

A method for representing periodic functions and enforcing exactly periodic boundary conditions with deep neural networks

Suchuan Dong*, Naxian Ni

Center for Computational and Applied Mathematics, Department of Mathematics, Purdue University, USA



ARTICLE INFO

Article history:

Available online 1 March 2021

Keywords:

Periodic function
Periodic boundary condition
Neural network
Deep neural network
Periodic deep neural network
Deep learning

ABSTRACT

We present a simple and effective method for representing periodic functions and enforcing exactly the periodic boundary conditions for solving differential equations with deep neural networks (DNN). The method stems from some simple properties about function compositions involving periodic functions. It essentially composes a DNN-represented arbitrary function with a set of independent periodic functions with adjustable (training) parameters. We distinguish two types of periodic conditions: those imposing the periodicity requirement on the function and all its derivatives (to infinite order), and those imposing periodicity on the function and its derivatives up to a finite order k ($k \geq 0$). The former will be referred to as C^∞ periodic conditions, and the latter C^k periodic conditions. We define operations that constitute a C^∞ periodic layer and a C^k periodic layer (for any $k \geq 0$). A deep neural network with a C^∞ (or C^k) periodic layer incorporated as the second layer automatically and exactly satisfies the C^∞ (or C^k) periodic conditions. We present extensive numerical experiments on ordinary and partial differential equations with C^∞ and C^k periodic boundary conditions to verify and demonstrate that the proposed method indeed enforces exactly, to the machine accuracy, the periodicity for the DNN solution and its derivatives.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

Deep neural networks (DNN) have emerged in the past few years as a promising alternative to the classical numerical methods (such as finite difference and finite element) for solving ordinary and partial differential equations (PDE). DNN-based solvers transform the PDE solution problem into an optimization problem. They typically represent the unknown field function in terms of a deep neural network, thanks to the universal approximation property of DNNs [18,19,4,23]. Then these methods compute the solution by minimizing a loss function that consist of residual norms of the governing equation and also possibly of the boundary and initial conditions in strong or weak forms (see e.g. [21,22,31,35,13,29,38,32], among others). DNN solutions are smooth analytical functions (depending on the activation function used) and, once the network is trained, can be evaluated for the function value and its derivatives exactly at any point inside the domain.

Boundary (and initial) conditions play a critical role in the solution of PDEs and make the problem well posed [17, 33, 9, 7, 11, 8, 27]. To enforce the boundary conditions (BC) in DNN solvers, an often-used approach is the penalty method,

* Corresponding author.

E-mail address: sdong@purdue.edu (S. Dong).

by incorporating a penalty term consisting of the residual norm of the boundary conditions into the loss function. The penalty method enforces the boundary conditions only approximately, and the penalty coefficient strongly influences the DNN training and convergence [3]. Choosing an appropriate or near-optimal penalty coefficient is largely an art, usually conducted by trial and error.

Enforcing the boundary conditions exactly with deep neural networks, if feasible, would be highly desirable. In this case, the DNN is constructed such that the boundary conditions are automatically and exactly satisfied. The constrained optimization problem for the PDE solution will then become less constrained or unconstrained, which will greatly facilitate the DNN training. Enforcing exactly the boundary conditions with deep neural networks is, however, highly non-trivial. For Dirichlet and Neumann type boundary conditions, several researchers have investigated the problem and promising techniques are available (see e.g. [21,22,25,1]). In [21] the unknown solution is decomposed into two components. One component satisfies the Dirichlet/Neumann boundary conditions and has no training parameters, while the other component vanishes on the boundary and is represented by a neural network in the domain. In [22] the authors decompose the unknown solution into a function represented by a deep neural network plus a linear combination of radial basis functions. The combination coefficients of the radial basis functions are determined by solving a linear system at every iteration of the DNN evaluation to satisfy the Dirichlet or Neumann boundary conditions. This process is understandably computationally expensive [22]. In [25] the authors rewrite the solution into two parts, similar to [21], with one part satisfying the Dirichlet/Neumann boundary conditions and the other part vanishing on the boundary but otherwise unconstrained. In order to deal with complex domain boundaries, the authors of [25] introduce a multiplicative length factor in front of the unconstrained part, which heuristically represents the distance of a point to the domain boundary. These ideas are further developed by [1], where the Dirichlet boundary condition is considered. In [1] the Dirichlet boundary data extension and the distance function to the boundary are both represented by low-capacity deep neural networks and pre-trained. This simplifies the implementation of the DNN solver. However, the enforcement of the boundary condition becomes only approximate. Similar ideas for the Dirichlet/Neumann type boundary conditions have also appeared in more recent works, see e.g. [30] for the DNN simulation of elastodynamic problems.

Periodic boundary conditions are widely encountered in computational science of various areas, especially when the physical domain involved in is infinite or homogeneous along one or more directions [10,6]. In such cases, usually only one cell will be computed in numerical simulations, and periodic boundary conditions are imposed on the cell boundaries. With classical numerical methods, another often-used technique for this type of problems is to express the unknown field function in terms of Fourier expansions, leading to what is known as the Fourier spectral or pseudo-spectral method [2,5,12]. While both are referred to as periodic conditions, the periodicity requirements imposed by the numerical method, when Fourier expansions are used and when they are not, are different. With the use of Fourier expansions, the method seeks a smooth periodic function as the solution to the governing equations, which automatically satisfies the periodicity for the solution value and *all* its derivatives (to infinite order) on the cell boundaries. On the other hand, when Fourier expansions are not used in the method, the periodicity needs to be imposed explicitly on the cell boundaries, and this can only be imposed for the solution value and its derivatives up to a certain finite order. We will distinguish these two types of periodic boundary conditions in this work. We refer to the former as the C^∞ periodic conditions, and the latter as the C^k periodic conditions, where $k \geq 0$ denotes the highest derivative the periodic condition imposes on.

The penalty method has been used to enforce the periodic boundary conditions (for up to the first derivative) with DNN-based PDE solvers in some recent studies; see e.g. [3,26]. In [3] a penalty term representing the residual norms of the periodic conditions for the function and its first derivative is included in the loss function. In [26] a sinusoidal activation function is employed with a shallow neural network (one hidden layer) to mimic the Fourier decompositions of the function, which is termed the Fourier neural network by some researchers (see e.g. [14,34,24,39]), and then the periodicity condition for the solution is imposed by a penalty term in the loss function. **The penalty method can only impose the periodic boundary conditions approximately.** It becomes more difficult, and perhaps impractical, to enforce C^k periodic conditions using the penalty for moderate or large k values. In particular, it is practically impossible to impose the C^∞ periodic conditions with the penalty method.

How to enforce exactly the C^∞ and C^k (for any $k \geq 0$) periodic conditions with deep neural networks is the focus of the current work. This problem seems to have barely been investigated before. To the best of the authors' knowledge, the only work close in theme to the current effort is perhaps [15], in which the authors enforce the periodicity condition for the solution value only, by adopting a similar idea to [21] and constructing a trial function with two parts. One part enforces the periodicity for the solution value, and the other part vanishes on the boundary and is represented by a DNN [15].

In the current paper we present a method for enforcing exactly the C^∞ and C^k (for any $k \geq 0$) periodic boundary conditions with deep neural networks. The DNN resulting from the current method, by design, automatically and exactly satisfies the C^∞ or C^k (for any prescribed k) periodic conditions. This method is based on some simple properties about function compositions involving periodic functions (Lemmas 2.1 and 2.2 in Section 2), and leverages the universal approximation power of deep neural networks. **It essentially composes a DNN-represented arbitrary function with a set of independent known periodic functions with adjustable parameters.** We consider the feed-forward neural network architecture [16] in this work, and define the operations that constitute a C^∞ periodic layer and a C^k periodic layer. To enforce the C^∞ periodic conditions, one only needs to set the second layer of the DNN (i.e. the first hidden layer) as a C^∞ periodic layer. To enforce the C^k periodic conditions, one only needs to set the second layer of the DNN as a C^k periodic layer. The C^∞ periodic layer constructs a set of independent C^∞ periodic functions with a user-prescribed period, based on sinusoidal functions, affine

mappings and nonlinear activation functions (such as “tanh” and “sigmoid”). The C^k periodic layer constructs a set of independent C^k periodic functions, based on the generalized Hermite interpolation polynomials, affine mappings and nonlinear activation functions. The output of the overall DNN, with the C^∞ (or C^k) periodic layer incorporated therein, automatically and exactly satisfies the C^∞ (or C^k) periodic conditions.

The operations involved in the C^∞ and C^k periodic layers can be implemented as user-defined Tensorflow/Keras (<http://www.tensorflow.org> and <http://keras.io>) layers, and incorporated into a DNN using Keras in a straightforward way. We present a number of numerical experiments with ordinary and partial differential equations to verify and demonstrate that the proposed method indeed enforces exactly, to the machine accuracy, the periodic boundary conditions as expected. All the numerical examples in the current paper are implemented and conducted based on Tensorflow, Keras and Python.

The contributions of this paper consist of two aspects: (i) the method for representing smooth periodic functions and exactly enforcing C^∞ periodic boundary conditions with deep neural networks; (ii) the method for exactly enforcing the C^k (for any $k \geq 0$) periodic boundary conditions with deep neural networks.

The rest of this paper is structured as follows. In Section 2 we define the operations that constitute the C^∞ periodic layer and the C^k periodic layer in one and higher dimensions, and establish that a deep neural network with these layers incorporated as the second layer exactly satisfies the C^∞ or C^k periodic boundary conditions for a given domain. In Section 3 we present extensive numerical experiments on periodic function approximations, and on solving the Helmholtz equations in one and two dimensions, the diffusion equation, and the wave equation, with C^∞ and C^k periodic boundary conditions. We demonstrate numerically that the proposed method enforces exactly the periodic boundary conditions, to the machine accuracy, for the DNN solution and its corresponding higher derivatives. Section 4 then concludes the presentation with some closing remarks.

2. Enforcing exact periodic conditions with DNN

2.1. C^∞ and C^k periodic conditions

Consider a smooth periodic function $f(x)$ with period L defined on the real axis,

$$f(x+L) = f(x), \quad \forall x \in (-\infty, \infty). \quad (1)$$

Now restrict $f(x)$ to a finite interval $[a, b]$, where $b - a = L$. Then f satisfies the following relations on the boundaries:

$$f(a) = f(b), \quad f'(a) = f'(b), \quad f''(a) = f''(b), \quad \dots, \quad f^{(m)}(a) = f^{(m)}(b), \quad \dots \quad (2)$$

We refer to the conditions in (2) as the C^∞ periodic conditions. Hereafter we will refer to a smooth function $f(x)$ satisfying these conditions as a C^∞ periodic function on $[a, b]$, or simply a periodic function.

In practice, the function may not be smooth and the conditions in (2) may only be required for the derivatives up to a finite order k ($k \geq 0$), i.e.

$$f^{(l)}(a) = f^{(l)}(b), \quad 0 \leq l \leq k, \quad (3)$$

where $f^{(0)}(x) = f(x)$ by convention. We refer to the $(k+1)$ conditions in (3) as the C^k periodic conditions. With a slight abuse of notation, we will refer to a function $f(x)$ satisfying the conditions (3) as a C^k periodic function on $[a, b]$.

Our goal here is to devise a method for representing C^∞ and C^k periodic functions with deep neural networks such that, by design, the output of the DNN automatically and exactly satisfies the C^∞ or C^k periodic conditions. Such neural networks will be referred to as C^∞ or C^k periodic deep neural networks. When solving a boundary value problem or initial/boundary value problem together with the C^∞ or C^k periodic boundary conditions, one can use the method developed herein to construct periodic DNNs as the trial functions that automatically take into account the periodic boundary conditions.

2.2. Enforcing exact C^∞ periodic conditions with DNN

We present a method below for representing C^∞ periodic functions and enforcing exactly the C^∞ periodic conditions with DNN. The method is based on the following property about function compositions involving periodic functions.

Lemma 2.1. Let $v(x)$ be a given smooth periodic function with period L on the real axis, i.e. $v(x+L) = v(x)$ for all $x \in (-\infty, \infty)$, and $f(x)$ denote an arbitrary smooth function. Define $u(x) = f(v) = f(v(x))$. Then

$$u(x+L) = u(x), \quad \forall x \in (-\infty, \infty); \quad (4a)$$

$$u^{(l)}(a) = u^{(l)}(b), \quad l = 0, 1, 2, \dots \quad (4b)$$

where a and b denote two real numbers with $b - a = L$.

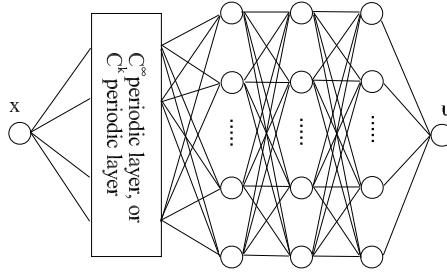


Fig. 1. Sketch of a feed-forward deep neural network with a C^∞ or C^k periodic layer incorporated as the second layer.

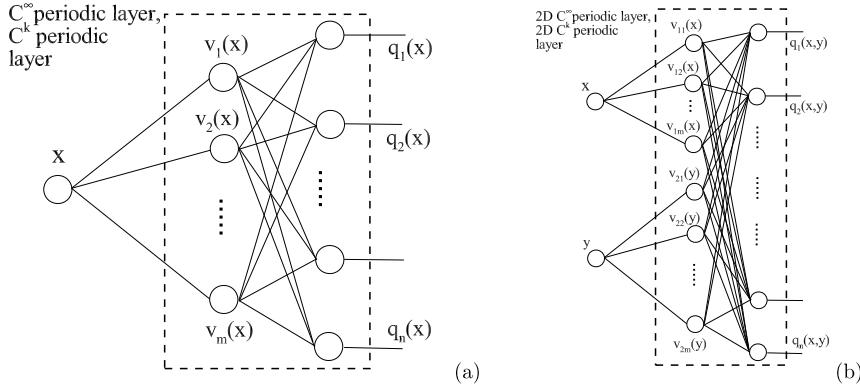


Fig. 2. Sketch illustrating the internal structures of (a) 1D, and (b) 2D C^∞ or C^k periodic layers.

This lemma can be proven by straightforward verifications.

We seek a DNN representation for an arbitrary smooth periodic function with a prescribed period L , such that the C^∞ periodic conditions are automatically satisfied. In light of Lemma 2.1, our basic idea for the representation is to compose an arbitrary DNN-represented function, together with a set of independent known periodic functions with period L and adjustable (training) parameters. Let us first use a single known periodic function with prescribed period L for illustration. We consider the sinusoidal functions,

$$p(x) = A \cos(\omega x + \phi) + c, \quad \text{with } \omega = \frac{2\pi}{L}, \quad (5)$$

where the constants A , c and ϕ are scalar adjustable (training) parameters. Here ω is a fixed constant as given above and ensures that $p(x)$ has a period L . Let $\sigma(\cdot)$ denote a nonlinear activation function (such as “tanh” or “sigmoid”). We define

$$v(x) = \sigma(p(x)) = \sigma(A \cos(\omega x + \phi) + c). \quad (6)$$

This step is crucial. The nonlinear function $\sigma(\cdot)$ will generate higher-frequency components in the output. So while $p(x)$ has a single frequency ω , $v(x)$ contains not only the frequency ω , but components with other and higher frequencies, all with a common period L . Finally, we consider an arbitrary smooth function $f(x)$ represented by a DNN, and define

$$u(x) = f_{\text{dnn}}(v(x)), \quad (7)$$

where f_{dnn} denotes the DNN-represented arbitrary smooth function. By Lemma 2.1, this $u(x)$ satisfies the C^∞ periodic conditions exactly.

In practice, we would like to compose the DNN-represented arbitrary function $f_{\text{dnn}}(\cdot)$ with a set of independent periodic functions $v(x)$ as defined above with adjustable parameters. This leads to the idea of a periodic layer with multiple nodes (neurons) within the layer. We consider the feed-forward deep neural network architecture [16] in the current work. Fig. 1 illustrates the idea of the current DNN with a sketch. Let x denote the input layer to the network, and $u(x)$ denote the output layer of the network. We use the second layer (behind the input x) to implement the set of independent known periodic functions (copies of $v(x)$) with period L and adjustable parameters, so that the output of the network $u(x)$ satisfies exactly the C^∞ periodic conditions (4). We refer to such a layer as a C^∞ periodic layer, or simply a periodic layer.

The operations within the C^∞ periodic layer are defined as follows (see Fig. 2(a)). Let $\mathcal{L}_p(m, n)$ denote the C^∞ periodic layer, where n denotes the number of nodes in the output of this layer and m denotes the size of the set of independent

periodic functions $v(x)$. Here both m and n are hyper-parameters of the C^∞ periodic layer $\mathcal{L}_p(m, n)$. The operations within $\mathcal{L}_p(m, n)$ are defined by:

$$v_i(x) = \sigma(A_i \cos(\omega x + \phi_i) + c_i), \quad 1 \leq i \leq m; \quad (8a)$$

$$q_j(x) = \sigma\left(\sum_{i=1}^m v_i(x) W_{ij} + B_j\right), \quad 1 \leq j \leq n. \quad (8b)$$

In these equations $q_j(x)$ ($1 \leq j \leq n$) are the output of this layer, and the fixed constant ω is given in (5) for a prescribed period L . $\sigma(\cdot)$ is the nonlinear activation function, and it is used twice in this layer. The training parameters of $\mathcal{L}_p(m, n)$ are the constants A_i , ϕ_i , c_i , W_{ij} and B_j , with $1 \leq i \leq m$ and $1 \leq j \leq n$.

In the current work, we have employed Tensorflow (<http://www.tensorflow.org>) and Keras (<http://keras.io>) to implement the operations of the C^∞ periodic layer as described above. The C^∞ periodic layer is implemented as a user-defined Tensorflow/Keras layer, and it can be used in the same way as the built-in core Keras layers.

Remark 2.1. The operations of the C^∞ periodic layer defined by (8) can be extended to two, three and higher dimensions (2D/3D) in a straightforward way. Here we outline the idea with two dimensions only (see Fig. 2(b)). Let (x_1, x_2) denote the coordinates in two dimensions, and $u(x_1, x_2)$ denote a smooth periodic function to be approximated, with properties

$$u(x_1 + L_1, x_2) = u(x_1, x_2), \quad u(x_1, x_2 + L_2) = u(x_1, x_2), \quad \forall x_1, x_2 \in (-\infty, \infty), \quad (9)$$

where L_1 and L_2 are the periods in the x_1 and x_2 directions, respectively. Equivalently, we can write the periodicity conditions in terms of a single periodic cell $[a_1, b_1] \times [a_2, b_2]$,

$$\begin{cases} \frac{\partial^\alpha}{\partial x_1^\alpha} u(a_1, x_2) = \frac{\partial^\alpha}{\partial x_1^\alpha} u(b_1, x_2), & \forall x_2 \in [a_2, b_2], \\ \frac{\partial^\alpha}{\partial x_2^\alpha} u(x_1, a_2) = \frac{\partial^\alpha}{\partial x_2^\alpha} u(x_1, b_2), & \forall x_1 \in [a_1, b_1], \quad \alpha = 0, 1, 2, \dots \end{cases} \quad (10)$$

where a_1 , b_1 , a_2 and b_2 are given constants satisfying $b_1 - a_1 = L_1$ and $b_2 - a_2 = L_2$. In this case, we define the 2D periodic layer, $\mathcal{L}_p^{2D}(m, n)$, with the following operations:

$$v_{1i}(x_1) = \sigma(A_{1i} \cos(\omega_1 x_1 + \phi_{1i}) + c_{1i}), \quad 1 \leq i \leq m; \quad (11a)$$

$$v_{2i}(x_2) = \sigma(A_{2i} \cos(\omega_2 x_2 + \phi_{2i}) + c_{2i}), \quad 1 \leq i \leq m; \quad (11b)$$

$$q_j(x_1, x_2) = \sigma\left(\sum_{i=1}^m v_{1i}(x_1) W_{ij}^{(1)} + \sum_{i=1}^m v_{2i}(x_2) W_{ij}^{(2)} + B_j\right), \quad 1 \leq j \leq n. \quad (11c)$$

In these equations, m and n are hyper-parameters of the layer \mathcal{L}_p^{2D} , $q_j(x_1, x_2)$ ($1 \leq j \leq n$) denote the output of this layer, and the constants ω_1 and ω_2 are defined by

$$\omega_1 = \frac{2\pi}{L_1}, \quad \omega_2 = \frac{2\pi}{L_2}, \quad (12)$$

with prescribed periods (L_1, L_2) . The training parameters of $\mathcal{L}_p^{2D}(m, n)$ consist of the constants:

$$A_{1i}, A_{2i}, \phi_{1i}, \phi_{2i}, c_{1i}, c_{2i}, W_{ij}^{(1)}, W_{ij}^{(2)}, B_j, \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

By composing an arbitrary DNN-represented function with the 2D C^∞ periodic layer defined above, we attain an overall DNN whose output automatically and exactly satisfies the 2D C^∞ periodic conditions. The C^∞ periodic layer for three and higher dimensions can be defined in a similar way.

Remark 2.2. In two or higher dimensions, if the C^∞ periodic conditions are imposed only in some (not all) directions, the C^∞ periodic layer as defined above can be modified in a simple way to accommodate the situation. For example, consider the 2D C^∞ periodic layer defined in (11) and suppose that the C^∞ periodic conditions are imposed only in the x_1 direction with period L_1 , but not in the x_2 direction. In this case, we can retain the equations (11a) and (11c), and replace (11b) by the following equation

$$v_{2i}(x_2) = \sigma(A_{2i} x_2 + c_{2i}), \quad 1 \leq i \leq m, \quad (13)$$

where the constants A_{2i} and c_{2i} are the training parameters. The modified 2D periodic layer consisting of equations (11a), (13) and (11c), when composed with a DNN-represented arbitrary function, will give rise to an overall DNN that automatically and exactly satisfies the C^∞ periodic conditions in the x_1 direction.

2.3. Enforcing exact C^k periodic conditions with DNN

We present in this subsection a method for representing C^k periodic functions and enforcing exactly the C^k periodic conditions (for any $k \geq 0$) with DNN. The method is based on the following simple property about function compositions involving C^k periodic functions:

Lemma 2.2. Let $v(x)$ ($x \in [a, b]$) denote a given function with continuous derivatives up to the order k and satisfying the following property,

$$v^{(l)}(a) = v^{(l)}(b), \quad 0 \leq l \leq k. \quad (14)$$

Let $f(x)$ denote an arbitrary function defined on the real axis with continuous derivatives up to the order k . Define $u(x) = f(v) = f(v(x))$ ($x \in [a, b]$). Then

$$u^{(l)}(a) = u^{(l)}(b), \quad 0 \leq l \leq k. \quad (15)$$

Proof. By induction one can show that $u^{(m)}(x) = g(v, v', v'', \dots, v^{(m)})$ for $0 \leq m \leq k$. In other words, $u^{(m)}(x)$ depends on x only through v and its derivatives. Equation (15) follows immediately from this relation and the conditions (14). \square

We seek a DNN representation for an arbitrary C^k periodic function on $[a, b]$, such that the C^k periodic conditions are automatically and exactly satisfied. In light of Lemma 2.2, our basic idea for the representation is to compose an arbitrary function represented by a DNN, together with a set of independent known C^k periodic functions with adjustable (training) parameters. To construct a C^k periodic function $v(x)$ on $[a, b]$ in Lemma 2.2, i.e. satisfying the conditions (14), we note that these conditions are reminiscent of the Hermite interpolation conditions. So the Hermite interpolation polynomial of degree at most $(2k+1)$ can be used to construct $v(x)$. Once $v(x)$ is obtained, we compose an arbitrary DNN-represented function f with $v(x)$, and the resultant function satisfies the C^k periodic conditions exactly.

Let us now use a single C^k periodic function $v(x)$ ($x \in [a, b]$) to illustrate the idea in some detail. Let s_i ($0 \leq i \leq k$) denote $(k+1)$ adjustable (training) parameters. Let $h(x)$ denote the unique Hermite interpolation polynomial of degree at most $(2k+1)$ that satisfies the following $(2k+2)$ interpolation conditions:

$$\begin{cases} h(a) = s_0, & h(b) = s_0; \\ h'(a) = s_1, & h'(b) = s_1; \\ \dots \\ h^{(k)}(a) = s_k, & h^{(k)}(b) = s_k. \end{cases} \quad (16)$$

The Newton form for $h(x)$ can be computed based on the divided differences, and the explicit Lagrange form for $h(x)$ is available in e.g. [36,37]. We then define

$$\begin{cases} p(x) = h(x) + (r_0 + r_1 x)(x - a)^{k+1}(x - b)^{k+1}, \\ v(x) = \sigma(p(x)) = \sigma\left(h(x) + (r_0 + r_1 x)(x - a)^{k+1}(x - b)^{k+1}\right), \end{cases} \quad (17)$$

where r_0 and r_1 are two additional adjustable parameters, and $\sigma(\cdot)$ is a nonlinear activation function (e.g. “tanh” or “sigmoid”). It is straightforward to verify that the $v(x)$ (and also $p(x)$) given in (17) satisfies the conditions (14). The explicit forms for $p(x)$ in (17) corresponding to $k = 0, 1$ and 2 are given by:

$$\begin{aligned} p(x) &= s_0 + (r_0 + r_1 x)(x - a)(x - b), \quad \text{for } k = 0; \\ p(x) &= s_0 + s_1(x - a)(b - x)(a + b - 2x) + (r_0 + r_1 x)(x - a)^2(x - b)^2, \quad \text{for } k = 1; \\ p(x) &= \xi^3(x - a)^3 \left[s_0 + (-3s_0\xi + s_1)(x - b) + \left(6s_0\xi^2 - 3s_1\xi + \frac{s_2}{2}\right)(x - b)^2 \right] \\ &\quad + \xi^3(b - x)^3 \left[s_0 + (3s_0\xi + s_1)(x - a) + \left(6s_0\xi^2 + 3s_1\xi + \frac{s_2}{2}\right)(x - a)^2 \right] \\ &\quad + (r_0 + r_1 x)(x - a)^3(x - b)^3, \quad \text{for } k = 2, \end{aligned}$$

where $\xi = \frac{1}{b-a}$. Let $f_{\text{dnn}}(x)$ denote an arbitrary function with continuous derivatives up to the order k represented by a deep neural network. With $v(x)$ given by (17), we finally define

$$u(x) = f_{\text{dnn}}(v(x)). \quad (18)$$

Then by Lemma 2.2 $u(x)$ satisfies exactly the C^k periodic conditions (15).

In practice, we would like to compose the DNN-represented arbitrary function $f_{dnn}(\cdot)$ with a set of independent C^k periodic functions $v(x)$ with adjustable parameters. This leads to the idea of a C^k periodic layer with multiple nodes (neurons) within the layer. Consider again the feed-forward neural network architecture, and let x denote the input and $u(x)$ denote the output of the network. Analogous to the C^∞ periodic layer in Section 2.2, we define a C^k periodic layer below, and use it as the second layer (behind the input x) of the network to implement the set of independent C^k periodic functions (copies of $v(x)$) and enforce the C^k periodic conditions. Fig. 1 sketches the DNN with a C^k periodic layer incorporated as the second layer.

The operations within the C^k periodic layer are defined as follows (see Fig. 2(a)). Let $\mathcal{L}_{C^k}(m, n)$ denote the C^k periodic layer, where n denotes the number of nodes in the output of this layer and m denotes the size of the set of independent C^k periodic functions $v(x)$. Both m and n are hyper-parameters of this layer. Given the input x , we compute the output $q_j(x)$ ($1 \leq j \leq n$) of the C^k periodic layer $\mathcal{L}_{C^k}(m, n)$ by:

$$v_i(x) = \sigma \left(h_i(x) + (r_{0i} + r_{1i}x)(x - a)^{k+1}(x - b)^{k+1} \right), \quad 1 \leq i \leq m; \quad (19a)$$

$$q_j(x) = \sigma \left(\sum_{i=1}^m v_i(x) W_{ij} + B_j \right), \quad 1 \leq j \leq n. \quad (19b)$$

In these equations $\sigma(\cdot)$ is the nonlinear activation function, and $h_i(x)$ ($1 \leq i \leq m$) are the Hermite interpolation polynomials of degree at most $(2k+1)$ satisfying the conditions

$$h_i^{(l)}(a) = s_{li}, \quad h_i^{(l)}(b) = s_{li}, \quad 0 \leq l \leq k, \quad 1 \leq i \leq m. \quad (20)$$

The constant parameters involved in these equations, r_{0i} , r_{1i} , W_{ij} , B_j , s_{li} , for $1 \leq i \leq m$, $1 \leq j \leq n$ and $0 \leq l \leq k$, are the training parameters of the C^k periodic layer $\mathcal{L}_{C^k}(m, n)$. By incorporating the C^k periodic layer defined by (19) as the second layer, the resultant DNN automatically and exactly satisfies the C^k periodic conditions with its output.

Remark 2.3. The operations of the C^k periodic layer $\mathcal{L}_{C^k}(m, n)$ defined by (19) can be extended to two, three and higher dimensions in a straightforward fashion. Here we use two dimensions only to illustrate the idea (see Fig. 2(b)). Let x_1 and x_2 ($x_1 \in [a_1, b_1]$, $x_2 \in [a_2, b_2]$) denote the coordinates in two dimensions, and $u(x_1, x_2)$ denote a 2D C^k periodic function, satisfying the C^k periodic conditions:

$$\begin{cases} \frac{\partial^\alpha}{\partial x_1^\alpha} u(a_1, x_2) = \frac{\partial^\alpha}{\partial x_1^\alpha} u(b_1, x_2), & \forall x_2 \in [a_2, b_2], \\ \frac{\partial^\alpha}{\partial x_2^\alpha} u(x_1, a_2) = \frac{\partial^\alpha}{\partial x_2^\alpha} u(x_1, b_2), & \forall x_1 \in [a_1, b_1], \quad \alpha = 0, 1, \dots, k. \end{cases} \quad (21)$$

In this case, we define the 2D C^k periodic layer $\mathcal{L}_{C^k}^{2D}(m, n)$ with the following operations:

$$v_{1i}(x_1) = \sigma \left(h_{1i}(x_1) + (r_{0i}^{(1)} + r_{1i}^{(1)}x_1)(x_1 - a_1)^{k+1}(x_1 - b_1)^{k+1} \right), \quad 1 \leq i \leq m; \quad (22a)$$

$$v_{2i}(x_2) = \sigma \left(h_{2i}(x_2) + (r_{0i}^{(2)} + r_{1i}^{(2)}x_2)(x_2 - a_2)^{k+1}(x_2 - b_2)^{k+1} \right), \quad 1 \leq i \leq m; \quad (22b)$$

$$q_j(x_1, x_2) = \sigma \left(\sum_{i=1}^m v_{1i}(x_1) W_{ij}^{(1)} + \sum_{i=1}^m v_{2i}(x_2) W_{ij}^{(2)} + B_j \right), \quad 1 \leq j \leq n. \quad (22c)$$

In the above equations, m and n are the hyper-parameters of this layer, $q_j(x_1, x_2)$ ($1 \leq j \leq n$) denote the output of this layer, and $h_{1i}(x_1)$ and $h_{2i}(x_2)$ are the Hermite interpolation polynomials of degree at most $(2k+1)$ satisfying the conditions:

$$\begin{cases} h_{1i}^{(l)}(a_1) = s_{li}^{(1)}, \quad h_{1i}^{(l)}(b_1) = s_{li}^{(1)}, & 0 \leq l \leq k, \quad 1 \leq i \leq m; \\ h_{2i}^{(l)}(a_2) = s_{li}^{(2)}, \quad h_{2i}^{(l)}(b_2) = s_{li}^{(2)}, & 0 \leq l \leq k, \quad 1 \leq i \leq m. \end{cases} \quad (23)$$

The constant parameters involved in the above equations,

$$s_{li}^{(1)}, s_{li}^{(2)}, r_{0i}^{(1)}, r_{0i}^{(2)}, r_{1i}^{(1)}, r_{1i}^{(2)}, W_{ij}^{(1)}, W_{ij}^{(2)}, B_j, \quad 0 \leq l \leq k, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n,$$

are the training parameters of the layer $\mathcal{L}_{C^k}^{2D}(m, n)$. By using the 2D C^k periodic layer $\mathcal{L}_{C^k}^{2D}(m, n)$ as the second layer of a DNN and with (x_1, x_2) as the input, the resultant DNN will automatically and exactly satisfy the 2D C^k periodic boundary conditions (21). The C^k periodic layer for three and higher dimensions can be defined in a similar way.

Remark 2.4. In two and higher dimensions, if the C^k periodic conditions are only imposed in some (not all) directions, the C^k periodic layer as defined above can be modified in a simple way to accommodate the situation. The modification is similar to what is discussed in Remark 2.2 for the modified C^∞ periodic layer. For illustration, let us consider the 2D C^k periodic layer defined by (22), and suppose that the C^k periodic conditions are imposed only in the x_1 direction, not in the x_2 direction. In this case, we can retain the equations (22a) and (22c), and replace equation (22b) by the following equation for $v_{2i}(x_2)$,

$$v_{2i}(x_2) = \sigma \left(r_{0i}^{(2)} + r_{1i}^{(2)} x_2 \right), \quad (24)$$

where the constants $r_{0i}^{(2)}$ and $r_{1i}^{(2)}$ are training parameters. The modified 2D C^k periodic layer consisting of equations (22a), (24) and (22c), when used as the second layer of a DNN, will impose exactly the C^k periodic conditions in the x_1 direction in the output of this DNN.

Remark 2.5. In the current work we use the Hermite interpolation polynomial (see equations (16) and (17)) to implement the known C^k periodic function. There are other possibilities. Since the summation (or product) of a C^k periodic function and a C^∞ periodic function produces a C^k periodic function, one can e.g. add a known C^∞ periodic function to the Hermite interpolation polynomial when constructing the known C^k periodic function. For example, one can add the sinusoidal function from Section 2.2 (see equation (5)) to the $p(x)$ in equation (17) to construct a new C^k periodic function with adjustable parameters.

Remark 2.6. To ensure the C^∞ or C^k periodicity with the deep neural network, the nonlinear activation functions $\sigma(\cdot)$ therein need to satisfy the corresponding regularity properties. With the C^∞ periodic layer, all the activation functions employed in the neural network need to be smooth functions (infinitely differentiable). With the C^k periodic layer, all the activation functions employed in the neural network need to have continuous derivatives up to the order of at least k .

3. Numerical examples

We present several numerical examples in what follows to demonstrate the effectiveness of the method presented in the previous section. We consider the approximation of periodic functions, and the solution of the Helmholtz equation, the unsteady diffusion equation and the wave equation, together with periodic boundary conditions, using deep neural networks. We employ a variant of the deep Galerkin method [35] for solving the differential equations with DNN and also for enforcing the initial conditions with unsteady problems. The periodic boundary conditions (BC) are dealt with based on the method from Section 2. Note that with the current method the periodic boundary conditions are satisfied automatically by the DNN. Therefore there is no need to account for the periodic boundary conditions in the loss function. The application codes for all the tests reported here are implemented using Tensorflow/Keras and Python. We employ the hyperbolic tangent (“tanh”) as the nonlinear activation function in all the tests. The neural network is trained using either the Adam [20] or the L-BFGS [28] optimizer. In the current implementation, for the Adam optimizer we employ the routine from the Keras library, and for the L-BFGS optimizer we employ the routine available from the Tensorflow-Probability library (<http://www.tensorflow.org/probability>).

3.1. Approximation of periodic functions

Let us look into the DNN approximation of periodic functions using the method developed in Section 2. We employ three different functions to illustrate the performance characteristics of the method: a C^∞ periodic function, a C^0 periodic function, and a non-periodic function. Note that in the case of the non-periodic function, we are essentially seeking a periodic function approximation of the non-periodic function.

Consider first the function

$$u_1(x) = \sin(2\pi x + 0.25\pi) + \cos(9\pi x - 0.1\pi) - 2 \sin(7\pi x + 0.33\pi) \quad (25)$$

on the domain $\Omega = \{x | 0 \leq x \leq 2\}$. This is a C^∞ periodic function on this domain. We would like to approximate $u_1(x)$ using a C^∞ periodic DNN, and using a C^k periodic DNN with $k = 0$ and $k = 1$.

To approximate $u_1(x)$, we employ a feed-forward neural network [16] as illustrated in Fig. 1. The input to the network is x , and the output is the approximation $u(x)$. We use 3 hidden layers in between, each with a width of 30 neurons. The hyperbolic tangent (“tanh”) function is used as the activation function for all the hidden layers, and no activation is applied on the output layer. As discussed in Section 2, we set the second layer of the network (i.e. the first hidden layer) as a C^∞ periodic layer for the C^∞ periodic approximation, and as a C^k periodic layer for the C^k periodic approximation. More specifically, we employ a $\mathcal{L}_p(m, n)$ with $m = 11$ and $n = 30$ for the C^∞ periodic layer, and a $\mathcal{L}_{C^k}(m, n)$ with $m = 11$ and $n = 30$ for the C^k -periodic layer with $k = 0$ and 1. In other words, a set of 11 independent periodic functions $v(x)$ has been

Table 1

Function approximation: DNN and simulation parameters.

parameter	value	parameter	value
hidden layers	depth=3, width=30	N_e	3
1st hidden layer	C^∞ periodic layer $\mathcal{L}_p(11, 30)$, or C^k periodic layer $\mathcal{L}_{C^k}(11, 30)$	Q	30 (C^∞ periodic DNN), or 40 (C^k periodic DNN)
maximum epochs	10000	optimizer	Adam
learning rate coefficient	1e - 3	activation	tanh
input data	x_i^e ($0 \leq e \leq N_e - 1, 0 \leq i \leq Q - 1$)	label data	$u_1(x_i^e)$, or $u_2(x_i^e)$, or $u_3(x_i^e)$
x_i^e	Gauss-Lobatto-Legendre quadrature	ω	$2\pi/V_\Omega$

used within the C^∞ periodic layer and the C^k periodic layer. For the C^∞ periodic layer $\mathcal{L}_p(m, n)$, the constant ω in equation (8a) is set to, according to equation (5),

$$\omega = \frac{2\pi}{V_\Omega} = \frac{2\pi}{2} = \pi, \quad (26)$$

where $V_\Omega = \int_\Omega dx = 2$ is the size of the domain Ω .

We minimize the following loss function with this DNN,

$$\begin{aligned} \text{Loss} &= \frac{1}{V_\Omega} \int_{\Omega} |u(x) - u_1(x)|^2 dx = \frac{1}{V_\Omega} \sum_{e=0}^{N_e-1} \int_{\Omega_e} |u(x) - u_1(x)|^2 dx \\ &= \frac{1}{V_\Omega} \sum_{e=0}^{N_e-1} \sum_{i=0}^{Q-1} |u(x_i^e) - u_1(x_i^e)|^2 J^e w_i, \end{aligned} \quad (27)$$

where N_e is the number of elements (i.e. sub-intervals) we have partitioned the domain Ω into in order to compute the integral, Q is the number of quadrature points within each element, Ω_e denotes the sub-interval occupied by the element e ($0 \leq e \leq N_e - 1$), x_i^e ($0 \leq i \leq Q - 1$) are the Gauss-Lobatto-Legendre quadrature points within Ω_e for $0 \leq e \leq N_e - 1$, J^e is the Jacobian of the element Ω_e with respect to the standard element $[-1, 1]$, and w_i ($0 \leq i \leq Q - 1$) are the quadrature weights associated with the Gauss-Lobatto-Legendre quadrature points. In the numerical experiments we have employed 3 elements ($N_e = 3$) to partition the domain Ω , with $\Omega_0 = [0, 0.7]$, $\Omega_1 = [0.7, 1.4]$ and $\Omega_2 = [1.4, 2]$. We employ 30 quadrature points ($Q = 30$) within each element for the C^∞ periodic DNN, and 40 quadrature points ($Q = 40$) within each element for the C^k ($k = 0, 1$) periodic DNN. The input data to the network consist of all the quadrature points x_i^e ($0 \leq i \leq Q - 1, 0 \leq e \leq N_e - 1$), and the label data consist of $u_1(x_i^e)$. The Adam optimizer [20] has been used to train the network for 10000 epochs for each case, with the learning rate coefficient (i.e. the parameter α in [20]) fixed at the default value 10^{-3} . The options of “early stopping” and “restore to best weight” have been used in Keras during the training process. The main parameters for the DNN and the simulations are summarized in Table 1.

Fig. 3 shows the training histories of the loss function corresponding to the C^∞ , C^0 and C^1 periodic DNN approximations. The loss function decreases rather slowly as the training begins. Then we observe a short stage when the loss function decreases sharply. After that, the reduction in the loss function slows down again, resulting in a long tail in the training history curve. These characteristics seem to be common to the DNN training for all the problems we have considered in this work. During the stage with slow reduction in the loss function (long tail), we observe that the loss value fluctuates from time to time during the training, resulting in a sequence of spikes in the training history curves (Fig. 3). This is in part due to the fixed learning rate coefficient employed here. Reducing the learning rate coefficient gradually as the training progresses will likely reduce the loss fluctuations. In spite of the spikes, one can observe the trend of decreasing loss as the training proceeds. Since we have turned on the options of “early stopping” and “restore to best weight” in Keras, the spikes in the training curves have no effect on the simulation results we have obtained here.

In Fig. 4 we compare the DNN approximation results of $u_1(x)$ (top row) obtained with C^∞ (plot (a)), C^0 (plot (b)), and C^1 (plot (c)) periodic conditions, together with the exact function $u_1(x)$. The distributions of the absolute error, $|u(x) - u_1(x)|$, corresponding to these approximations are shown in Figs. 4(d,e,f), respectively. It is observed that the DNN approximations computed with all three methods agree well with the exact function $u_1(x)$, and the approximation function curves overlap with the exact function curve.

In Table 2 we list the values of the function $u_1(x)$ and its derivatives $\frac{du_1}{dx}$ and $\frac{d^2u_1}{dx^2}$ on the domain boundaries $x = 0$ and $x = 2$, obtained from the C^∞ , C^0 and C^1 periodic DNN approximations, as well as from the exact $u_1(x)$ function given in (25). The function derivatives have been computed based on auto-differentiation. Once the DNN is trained, the derivatives computed in this way are exact values corresponding to the given DNN representation. We have shown 14 significant digits (double precision) for the values in this table. It is evident that the C^∞ periodic DNN enforces exactly, to the machine accuracy, the periodicity for the function as well as its derivatives. On the other hand, the C^0 periodic DNN enforces exactly the periodicity only for the function value, and the C^1 periodic DNN enforces exactly the periodicity only for the function value and the first derivative. These numerical results have verified the analyses about these methods in Section 2.

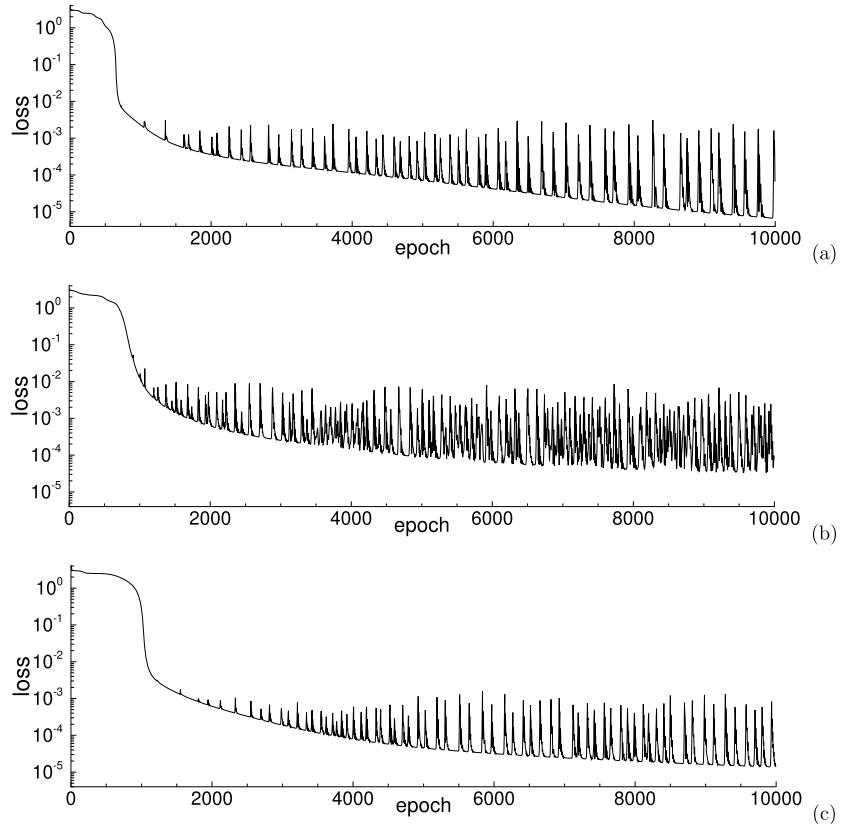


Fig. 3. Approximation of the periodic function $u_1(x)$: training loss histories corresponding to (a) C^∞ , (b) C^0 , and (c) C^1 periodic DNN approximations.

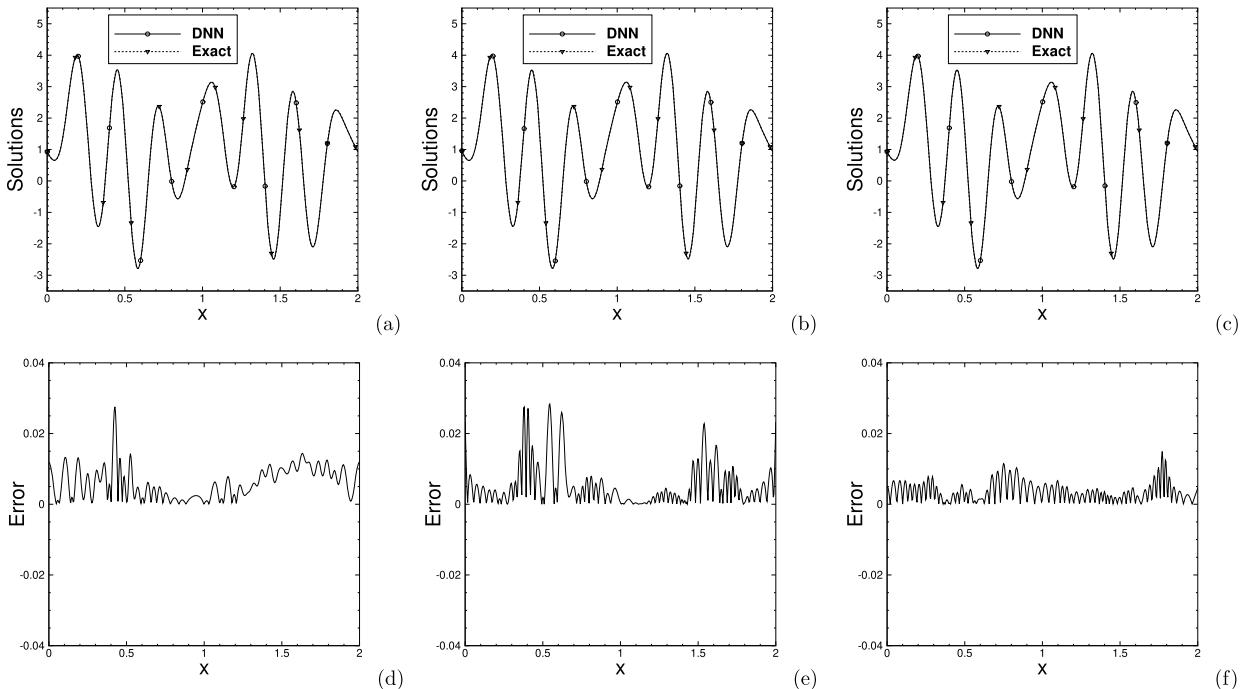


Fig. 4. Approximation of periodic function $u_1(x)$: comparison of DNN approximations (top row) and their absolute errors (bottom row) with (a,d) C^∞ , (b,e) C^0 , (c,f) C^1 periodic conditions. The exact function is shown in the plots (a,b,c) for reference.

Table 2

Approximations of the periodic function $u_1(x)$: Values of the function and its first and second derivatives at the left/right domain boundaries ($x = 0, 2$) from the DNN approximations with C^∞ , C^0 and C^1 periodic conditions and from the exact function. 14 significant digits (double precision) are listed to show that the current method enforces the periodic conditions exactly. The boxes highlight the values that mis-match on the boundaries.

	C^∞ periodic DNN	C^0 periodic DNN	C^1 periodic DNN	Exact value
$u_1(0)$	9.2478271387350e-01	9.5998617007276e-01	9.3202741230659e-01	9.3667924347381e-01
$u_1(2)$	9.2478271387350e-01	9.5998617007276e-01	9.3202741230659e-01	9.3667924347381e-01
$u'_1(0)$	-9.2010196979120e+00	-1.1103158994625e+01	-9.3797238469930e+00	-9.2086781968974e+00
$u'_1(2)$	-9.2010196979120e+00	-7.1832211211812e+00	-9.3797238469930e+00	-9.2086781968972e+00
$u''_1(0)$	6.0998633006437e+01	8.4031868694133e+01	1.0357079019194e+02	4.4301828505852e+01
$u''_1(2)$	6.0998633006437e+01	8.6538176666730e+01	3.5624115949487e+01	4.4301828505854e+01

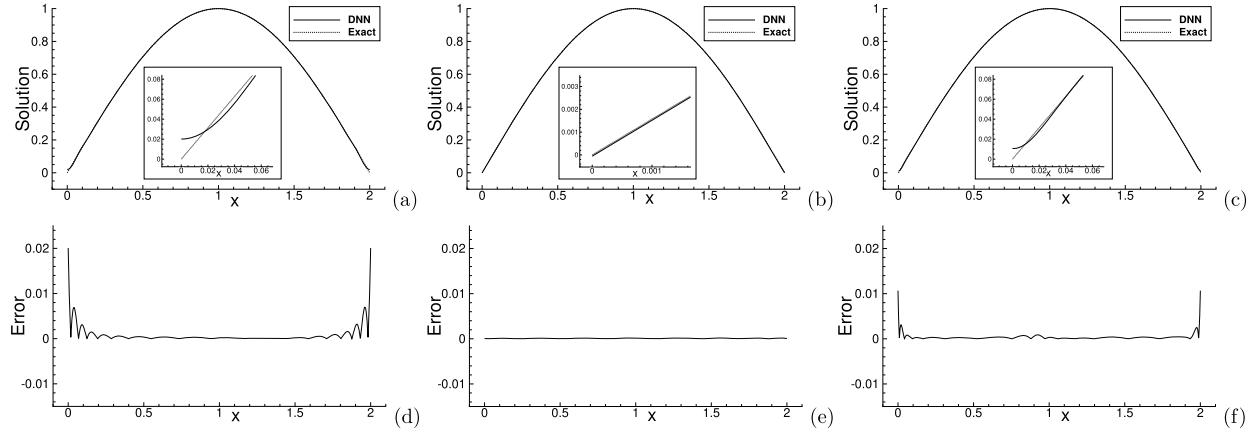


Fig. 5. Approximation of C^0 periodic function $u_2(x)$: Comparison of approximation results (top row) and their absolute errors (bottom row) obtained with (a,d) C^∞ , (b,e) C^0 , and (c,f) C^1 periodic DNNs. The exact function is included for comparison. The insets are the magnified views near $x = 0$.

Table 3

Approximation of C^0 periodic function $u_2(x)$: Values of the function and its derivatives at the left/right domain boundaries ($x = 0$ and $x = 2$) from the C^∞ , C^0 and C^1 periodic DNN approximations and from the exact function. 14 significant digits are listed to demonstrate that the current method enforces the periodic conditions exactly.

	C^∞ periodic DNN	C^0 periodic DNN	C^1 periodic DNN	Exact value
$u_2(0)$	2.0077934937638e-02	-5.6212180516792e-05	1.0675453567003e-02	0
$u_2(2)$	2.0077934937638e-02	-5.6212180516792e-05	1.0675453567003e-02	0
$u'_2(0)$	2.0046381097937e-03	1.5738634304121e+00	-8.9688048016955e-02	1.5707963267949e+00
$u'_2(2)$	2.0046381097902e-03	-1.5651374257149e+00	-8.9688048016955e-02	-1.5707963267949e+00
$u''_2(0)$	4.8956024103701e+01	-3.0860019634346e-02	1.1444431255354e+02	0
$u''_2(2)$	4.8956024103700e+01	1.7388612411815e-01	1.0788094853391e+02	0

We next consider the function

$$u_2(x) = \sin \frac{\pi x}{2} \quad (28)$$

on the domain $\Omega = \{x | 0 \leq x \leq 2\}$. This is a C^0 periodic function on this domain, with $u_2(0) = u_2(2)$ and $u'_2(0) \neq u'_2(2)$. We would like to approximate $u_2(x)$ with C^∞ and C^k ($k = 0$ and 1) periodic DNNs.

We employ the same DNN and simulation parameters to approximate $u_2(x)$ as for $u_1(x)$; see Table 1. The loss function is given by (27), with $u_1(x)$ replaced by $u_2(x)$. Fig. 5 is a comparison of the approximation results and their absolute errors obtained with C^∞ , C^0 and C^1 periodic DNNs. The exact function $u_2(x)$ has also been included for comparison. The C^0 periodic DNN produces results that are considerably more accurate than the other two methods, as expected. The C^∞ and C^1 periodic DNN approximations produce accurate results in the bulk of the domain, but exhibit larger errors near/at the domain boundaries. The C^∞ and C^1 periodic conditions appear to have the tendency of bending the function curve near the boundaries to achieve periodicity for the derivatives; see the insets of Figs. 5(a,b,c).

To verify the periodicity of the DNN approximations on the boundaries, we list in Table 3 the values of the approximated $u_2(x)$ and its derivatives (up to order two) on the domain boundaries $x = 0$ and 2 from different approximations and from the exact function $u_2(x)$. Again 14 significant digits have been shown for each value. It is observed that the current methods indeed enforce exactly the periodicity for the approximation function and its derivatives on the boundaries as expected. In the C^∞ periodic DNN approximation, the function and its derivatives (up to order 2 considered here) have identical values on the two boundaries. In contrast, with the C^0 periodic DNN approximation only the function value is identical on the

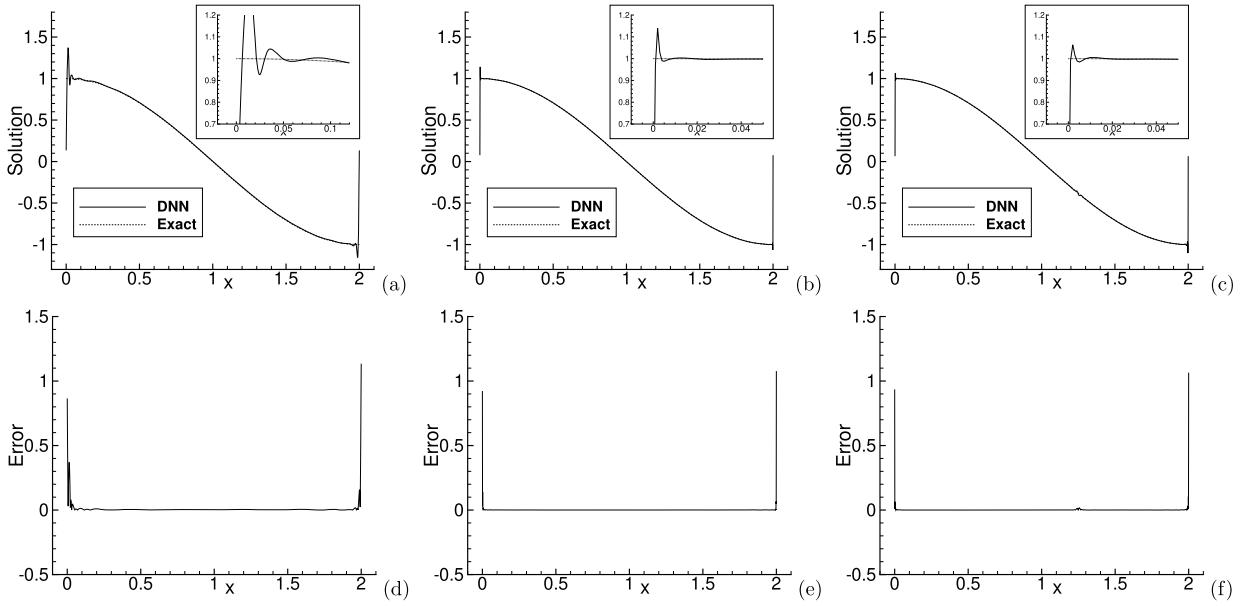


Fig. 6. Approximation of non-periodic function $u_3(x)$: Approximation results (top row) and their absolute errors (bottom row) obtained with (a,d) C^∞ , (b,e) C^0 , (c,f) C^1 , periodic DNNs. The exact profile of the function $u_3(x)$ is also included. The insets are magnified views near $x = 0$.

Table 4

Approximation of the non-periodic function $u_3(x)$: Values of the function and its derivatives at the left/right domain boundaries ($x = 0$ and $x = 2$) from the approximations with C^∞ , C^0 and C^1 periodic DNNs and from the exact function. 14 significant digits are listed to demonstrate that the current method enforces the periodic conditions exactly for the approximation function and its derivatives.

	C^∞ periodic DNN	C^0 periodic DNN	C^1 periodic DNN	Exact value
$u_3(0)$	1.3452550840065e-01	7.7219924514080e-02	6.4853245159018e-02	1
$u_3(2)$	1.3452550840063e-01	7.7219924514080e-02	6.4853245159018e-02	-1
$u'_3(0)$	1.7647534129428e+02	7.0097116541954e+02	1.5259262265715e+03	0
$u'_3(2)$	1.7647534129428e+02	1.0015322821138e+03	1.5259262265715e+03	0
$u''_3(0)$	-9.2430741908226e+03	8.4509087153852e+05	-4.6596618576551e+05	-2.4674011002723e+00
$u''_3(2)$	-9.2430741908221e+03	-9.4100497028271e+05	-4.5176214142043e+05	2.4674011002723e+00

boundaries, and with the C^1 periodic DNN approximation the function and the first derivative have identical values on the two boundaries.

We finally consider a non-periodic function,

$$u_3(x) = \cos \frac{\pi x}{2}, \quad (29)$$

on the domain $\Omega = \{x | 0 \leq x \leq 2\}$. We would like to approximate this function using C^∞ and C^k ($k = 0, 1$) periodic DNNs. The DNN and simulation parameter values employed here are the same as those for $u_1(x)$ and $u_2(x)$ (see Table 1), except for the number of quadrature points within each element. Here for $u_3(x)$ we employ $Q = 40$ with the C^∞ periodic DNN, and $Q = 50$ with the C^0 and C^1 periodic DNNs.

Fig. 6 shows the approximation functions $u_3(x)$ and their absolute errors obtained with C^∞ , C^0 and C^1 periodic DNNs. In the bulk of the domain the DNN approximations appear to be in good agreement with the exact function $u_3(x)$ with all three methods. In a region near the two boundaries, the periodic DNN approximations exhibit large errors, and one can observe fluctuations in the approximation functions (Gibbs phenomenon). Table 4 lists the values of the approximation function $u_3(x)$ and its derivatives on the two boundaries ($x = 0, 2$) obtained with the C^∞ , C^0 and C^1 periodic DNNs as well as the exact function $u_3(x)$. The results again demonstrate that the current methods enforce exactly the periodicity for the approximation function and its derivatives.

3.2. One-dimensional Helmholtz equation with periodic BCs

In this subsection we test the performance of the proposed method with the one-dimensional (1D) Helmholtz equation,

$$\frac{d^2u}{dx^2} - \lambda u = f(x), \quad (30)$$

on the domain $\Omega = \{x | a \leq x \leq b\}$, where $\lambda (\lambda \geq 0)$, a and b are given constants and $f(x)$ is a prescribed source term. We impose periodic boundary conditions (BC) on the domain boundaries, $x = a$ and b .

Specifically, we consider two types of periodic boundary conditions. The first type is the C^1 periodic condition,

$$u(a) = u(b), \quad u'(a) = u'(b). \quad (31)$$

The second type is the C^∞ periodic condition,

$$u(a) = u(b), \quad u'(a) = u'(b), \quad u''(a) = u''(b), \quad \dots, \quad u^{(m)}(a) = u^{(m)}(b), \quad \dots \quad (32)$$

Note that with the C^∞ periodic condition (32), we are effectively seeking a C^∞ periodic function, with the period $L = b - a$, on the infinite domain $x \in (-\infty, \infty)$ that solves the equation (30). Since the Helmholtz equation is a second-order equation, imposing the C^0 periodic condition only, i.e. $u(a) = u(b)$, does not lead to a unique solution to the problem.

For the numerical tests in this section we fix the problem parameters to the following values:

$$\lambda = 10, \quad a = 0, \quad b = 4, \quad L = b - a = 4. \quad (33)$$

We choose the source term $f(x)$ such that the Helmholtz equation (30) has an analytic solution

$$u(x) = \sin[3\pi(x + 0.05)] \cos[2\pi(x + 0.05)] + 2. \quad (34)$$

This is a periodic function with $L = 4$ as a period, and it satisfies the boundary conditions (31) and (32).

To simulate this problem, we employ a feed-forward neural network (Fig. 1) with 4 hidden layers, with 20 nodes in each layer, apart from the input and output layers. The input to the network is the coordinate x (1 node), and the output of the network is the solution to the Helmholtz equation u (1 node). The second layer of the network (or the first hidden layer) is set to be a C^∞ periodic layer $\mathcal{L}_p(m, n)$ with $m = 11$ and $n = 20$, in which we set the constant $\omega = \frac{2\pi}{L} = \frac{\pi}{2}$ in equation (8a), when the C^∞ periodic boundary conditions in (32) are imposed. When the C^1 periodic boundary conditions in (31) are imposed, we set the second layer of the network to be a C^1 periodic layer $\mathcal{L}_{C^1}(m, n)$ with $m = 11$ and $n = 20$, as detailed in Section 2.

We minimize the following loss function with this DNN,

$$\begin{aligned} \text{Loss} &= \frac{1}{L} \int_{\Omega} \left[\frac{d^2u}{dx^2} - \lambda u - f(x) \right]^2 dx = \frac{1}{L} \sum_{e=0}^{N_e-1} \int_{\Omega_e} \left[\frac{d^2u}{dx^2} - \lambda u - f(x) \right]^2 dx \\ &= \frac{1}{L} \sum_{e=0}^{N_e-1} \sum_{i=0}^{Q-1} \left[\frac{d^2u}{dx^2} \Big|_{x_i^e} - \lambda u(x_i^e) - f(x_i^e) \right]^2 J^e w_i. \end{aligned} \quad (35)$$

In this equation, N_e is the number of elements (sub-intervals) we have partitioned the domain Ω into in order to compute the integral, Ω_e ($0 \leq e \leq N_e - 1$) denotes the region of element e , Q is the number of quadrature points within each element, J^e is the Jacobian of Ω_e with respect to the standard element $[-1, 1]$, x_i^e ($0 \leq i \leq Q - 1$) are the Gauss-Lobatto-Legendre quadrature points within element e for $0 \leq e \leq N_e - 1$, and w_i ($0 \leq i \leq Q - 1$) are the weights associated with the Gauss-Lobatto-Legendre quadrature. The input data to the network consist of x_i^e ($0 \leq i \leq Q - 1, 0 \leq e \leq N_e - 1$), and the label data consist of $f(x_i^e)$ ($0 \leq i \leq Q - 1, 0 \leq e \leq N_e - 1$). In the expression (35), $u(x_i^e)$ can be obtained from the output of the DNN, and $\frac{d^2u}{dx^2} \Big|_{x_i^e}$ can be computed by auto-differentiation.

For the numerical experiments reported below, we have partitioned the domain into three elements ($N_e = 3$), with these elements being $\Omega_0 = [0, 1.3]$, $\Omega_1 = [1.3, 2.6]$ and $\Omega_2 = [2.6, 4.0]$. We employ 40 quadrature points ($Q = 40$) within each element. The activation functions for all the hidden layers are the hyperbolic tangent function ("tanh"), and no activation function is applied to the output layer. The DNN is trained using the Adam optimizer. For both C^∞ and C^1 periodic BCs, the neural network has been trained for 31,000 epochs, and the learning rate coefficient has been decreased linearly from 10^{-3} at the beginning to 10^{-5} at the end of the training process. The options of "early stopping" and "restore to best weight" have been used in Tensorflow/Keras when training the DNN. The training histories of the loss function for the C^∞ and C^1 periodic BCs are shown in Fig. 7. We observe characteristics in the loss histories similar to those observed in Section 3.1, such as the varied loss reduction rates at different stages and the fluctuations in the loss values.

Fig. 8 shows a comparison between the exact solution and the DNN solutions obtained with the C^∞ and C^1 periodic BCs enforced using the current method from Section 2 (left column), as well as the absolute errors of these DNN solutions against the exact solution (right column). The DNN solutions are observed to agree with the exact solution very well. The DNN solution curves almost exactly overlap with the exact-solution curve, and the maximum errors in the domain is on the order of 10^{-4} with both the C^∞ and C^1 periodic boundary conditions.

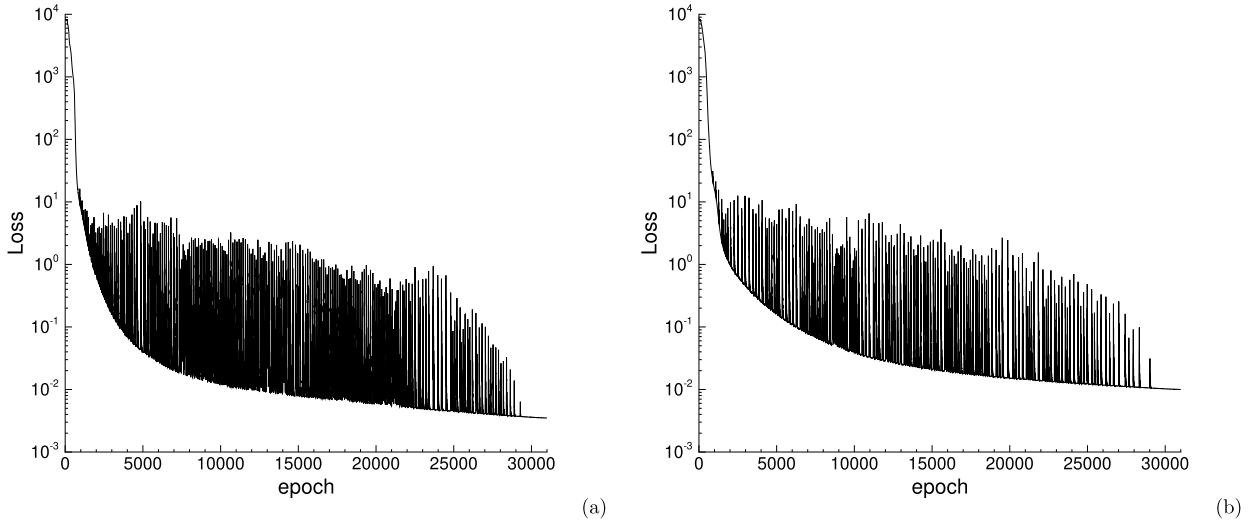


Fig. 7. 1D Helmholtz equation: training histories of the loss function with the (a) C^∞ and (b) C^1 periodic boundary conditions.

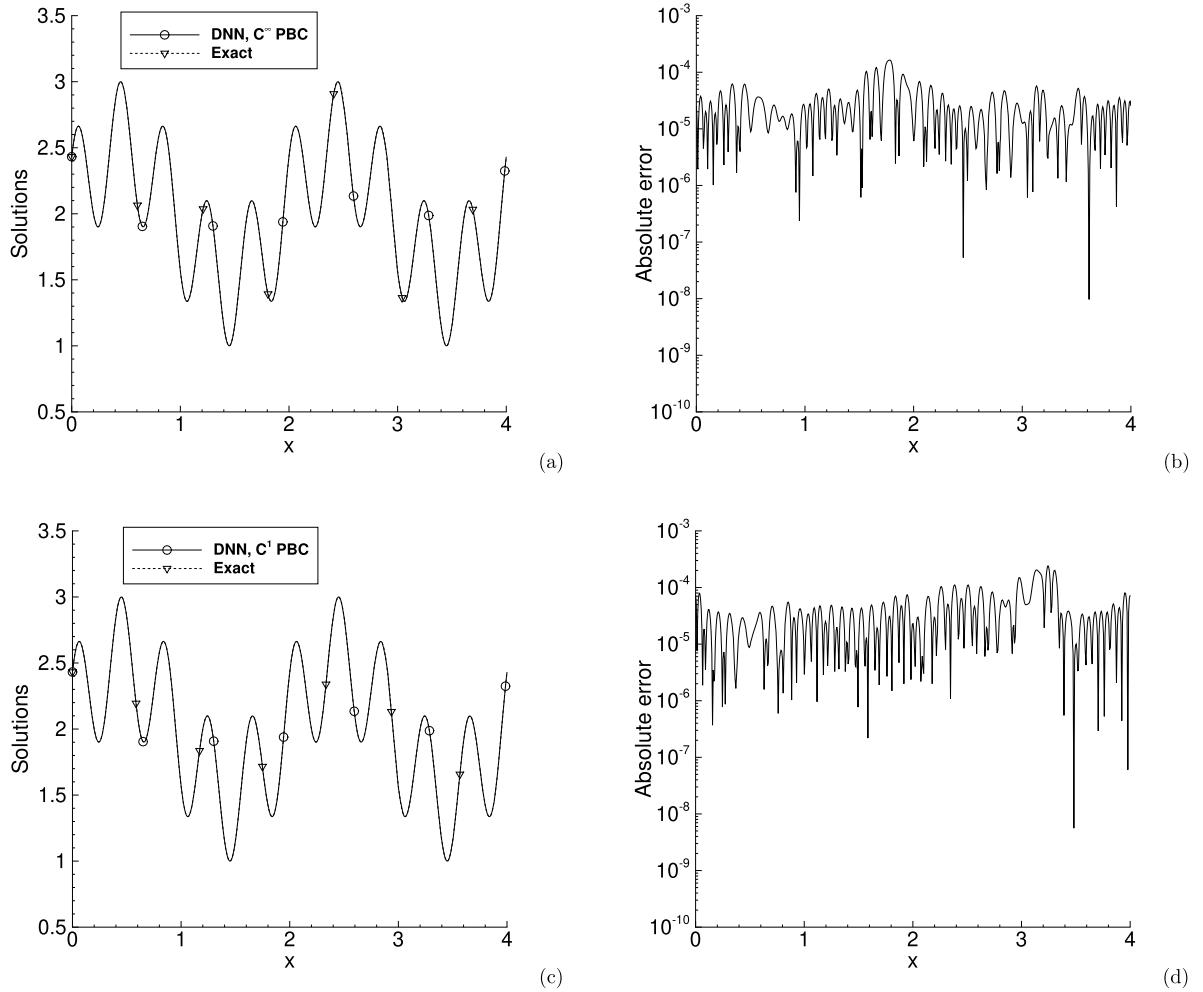


Fig. 8. 1D Helmholtz equation: (a,c), comparison between the exact and DNN solutions. (b,d), absolute errors of the DNN solutions. Results in (a,b) are obtained with the C^∞ periodic BCs, and those in (c,d) are obtained with the C^1 periodic BCs. Periodic BCs are enforced using the methods from Section 2.

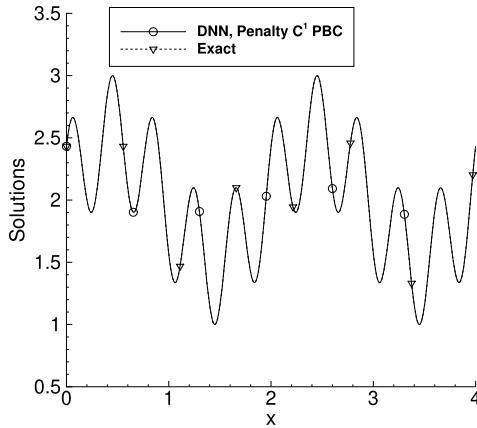


Fig. 9. 1D Helmholtz equation: DNN solution obtained with the C^1 periodic BC enforced using the penalty method.

Table 5

1D Helmholtz equation: Values of the solution and its first and second derivatives on the left/right domain boundaries from the exact solution and from the DNN solutions with C^∞ and C^1 periodic BCs enforced using the current method, and with the C^1 periodic BCs enforced using the penalty method.

	DNN C^∞ PBC (Current)	DNN C^1 PBC (Current)	DNN C^1 PBC (Penalty)	Exact solution
$u(0)$	2.4317959997357e+00	2.4318435955019e+00	2.4318606137503e+00	2.4317706231133e+00
$u(4)$	2.4317959997357e+00	2.4318435955019e+00	2.4318130678898e+00	2.4317706231133e+00
$u'(0)$	7.1038526507927e+00	7.1041119004958e+00	7.1047853002015e+00	7.1050608901229e+00
$u'(4)$	7.1038526507927e+00	7.1041119004958e+00	7.1048251209190e+00	7.1050608901229e+00
$u''(0)$	-8.8126699519835e+01	-8.9042115478782e+01	-8.7001193825524e+01	-8.8007775637807e+01
$u''(4)$	-8.8126699519835e+01	-8.867820891429e+01	-8.8802791765032e+01	-8.8007775637807e+01

For comparison, we have also computed this problem with the C^1 periodic BCs in another way, by enforcing the C^1 periodic BCs based on the penalty method. Fig. 9 shows the DNN solution computed using the penalty method. Here the DNN has the same parameters (4 hidden layers, with 20 nodes in each layer). The C^1 periodic BCs are enforced by including a penalty term in the loss function as follows,

$$\text{Loss} = \frac{1}{L} \int_{\Omega} \left[\frac{d^2 u}{dx^2} - \lambda u - f(x) \right]^2 dx + \theta_{bc} \left([u(a) - u(b)]^2 + \left[\frac{du}{dx} \Big|_{x=a} - \frac{du}{dx} \Big|_{x=b} \right]^2 \right), \quad (36)$$

where $\theta_{bc} = 10$ is the penalty coefficient in front of the boundary residual terms. The DNN has been trained with the Adam optimizer for 31,000 epochs, with the learning rate coefficient linearly decreasing from 10^{-3} at the beginning to 10^{-5} at the end of the training process. We observe that the DNN solution resulting from the penalty method also agrees well with the exact solution.

In Table 5 we list the values of the DNN solution and its first and second derivatives, with 14 significant digits shown, on the left and right domain boundaries obtained with the C^∞ and C^1 periodic boundary conditions enforced using the current method, together with those obtained with the C^1 periodic BCs enforced using the penalty method. The boundary values from the exact solution (34) are also included in the table for comparison. We observe that the current method enforces exactly, to the machine accuracy, the periodicity for the solution and its derivatives (up to order 2 shown here) with the C^∞ periodic BCs. With the C^1 periodic BCs, the current method enforces exactly the periodicity for the solution and its first derivative, but not for the second derivative. In contrast, the penalty method enforces the periodic condition for none of these quantities exactly.

In addition to the above tests, we have also considered L^2 regularization in the DNN simulation of the 1D Helmholtz equation. In this set of simulations we modify the loss function given in (35), by adding an L^2 regularization term

$$\gamma \sum_{i=1}^{N_p} |\chi_i|^2,$$

where the constant $\gamma \geq 0$ is a regularization coefficient, χ_i denotes a generic training parameter, and N_p is the total number of training parameters in the neural network, including those involved in the C^∞ or C^k periodic layers. This regularization term will tend to reduce the magnitudes of the training parameters. We vary the regularization coefficient γ in a range of values (between $\gamma = 10^{-5}$ and $\gamma = 0.1$). For each case we train the neural network for 31,000 epochs using the Adam optimizer, with the learning rate coefficient decreasing linearly from 10^{-3} at the beginning to 10^{-5} at the end of the

Table 6

1D Helmholtz equation: Effect of the L^2 regularization coefficient γ on the maximum and rms errors in the domain of the DNN solution computed using the C^∞ and C^1 periodic BCs.

Periodic BC	γ	maximum error	rms error
C^∞	0	1.64e - 4	2.59e - 5
	10 ⁻⁵	1.61e - 4	2.54e - 5
	10 ⁻⁴	1.52e - 4	2.46e - 5
	10 ⁻³	1.49e - 4	2.66e - 5
	10 ⁻²	4.03e - 4	9.41e - 5
	10 ⁻¹	2.16e - 3	5.88e - 4
C^1	0	2.41e - 4	3.97e - 5
	10 ⁻⁵	2.40e - 4	3.98e - 5
	10 ⁻⁴	2.37e - 4	3.92e - 5
	10 ⁻³	2.61e - 4	4.58e - 5
	10 ⁻²	3.31e - 4	7.26e - 5
	10 ⁻¹	2.19e - 3	4.38e - 4

Table 7

1D Helmholtz equation: Values of the DNN solution and its first/second derivatives on the left/right domain boundaries computed using C^∞ and C^1 periodic BCs, with different L^2 regularization coefficient values. These can be compared with Table 5 (columns two and three), which corresponds to no regularization.

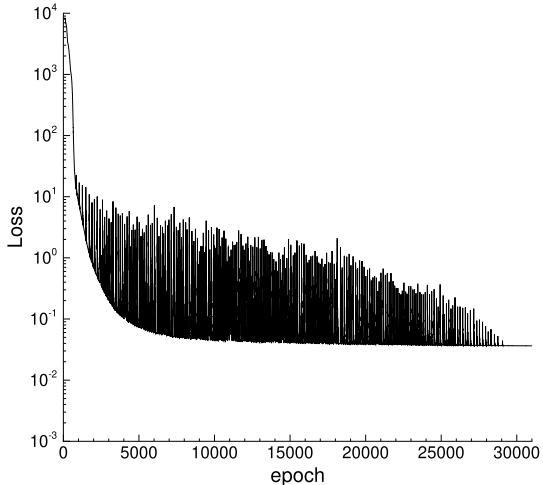
Periodic BC	γ	x	$u(x)$	$u'(x)$	$u''(x)$
C^∞	10 ⁻⁵	0	2.4317956200955e+00	7.1038405078644e+00	-8.8124787378696e+01
		4	2.4317956200955e+00	7.1038405078644e+00	-8.8124787378695e+01
	10 ⁻⁴	0	2.4317945868172e+00	7.1038451580625e+00	-8.8128268170738e+01
		4	2.4317945868172e+00	7.1038451580625e+00	-8.8128268170738e+01
	10 ⁻³	0	2.4317835455463e+00	7.1038350695957e+00	-8.8123605452033e+01
		4	2.4317835455463e+00	7.1038350695957e+00	-8.8123605452033e+01
	10 ⁻²	0	2.4316806415841e+00	7.1034826104860e+00	-8.8139635113550e+01
		4	2.4316806415841e+00	7.1034826104861e+00	-8.8139635113549e+01
	10 ⁻¹	0	2.4306387295347e+00	7.0984545906631e+00	-8.8292285064056e+01
		4	2.4306387295347e+00	7.0984545906631e+00	-8.8292285064056e+01
C^1	10 ⁻⁵	0	2.4318435801313e+00	7.1041221348866e+00	-8.9044025046988e+01
		4	2.4318435801313e+00	7.1041221348866e+00	-8.8677802019579e+01
	10 ⁻⁴	0	2.4318446061993e+00	7.1041204904255e+00	-8.9043608588260e+01
		4	2.4318446061993e+00	7.1041204904255e+00	-8.8680583719691e+01
	10 ⁻³	0	2.4318516299914e+00	7.1040912750379e+00	-8.9090042771614e+01
		4	2.4318516299914e+00	7.1040912750379e+00	-8.8705371438099e+01
	10 ⁻²	0	2.4318885280959e+00	7.1042513140124e+00	-8.9485778332921e+01
		4	2.4318885280959e+00	7.1042513140123e+00	-8.8652504850661e+01
	10 ⁻¹	0	2.4320845796666e+00	7.1044417448844e+00	-9.0794978595810e+01
		4	2.4320845796666e+00	7.1044417448844e+00	-8.8619367934323e+01

training. In this group of tests, all the weight and bias coefficients in the neural network are initialized with identical values as those in the DNN simulations without regularization discussed before.

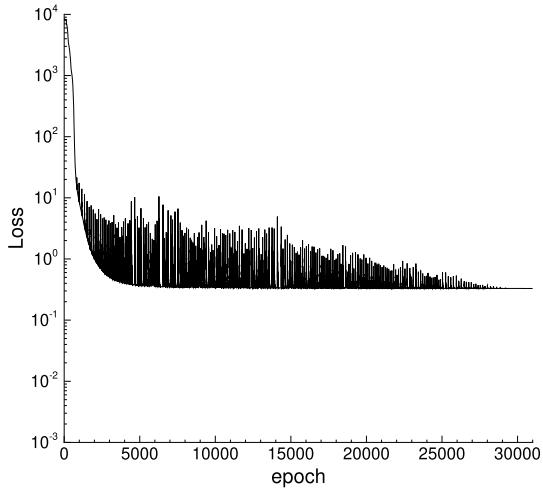
Fig. 10 shows the training histories of the loss function and the error distributions of the DNN solutions computed using the C^∞ periodic BC with regularization coefficients $\gamma = 10^{-4}$, 10^{-3} and 10^{-2} . These results can be compared with Fig. 7(a) and Fig. 8(b), which corresponds to the C^∞ periodic BC but with zero regularization in the loss function. With L^2 regularization we again observe the fluctuations in the loss histories. However, the fluctuations appear to become relatively weaker with a larger regularization coefficient. The loss histories signify that with $\gamma > 0$ the loss function asymptotically approaches some non-zero constant as the training progresses. This appears qualitatively to be somewhat different from the case without regularization (see Fig. 7(a)). It is also observed that, compared with the case of zero regularization, the error of the DNN result seems to be nearly the same with a small γ value (see Figs. 10(b) and 8(b)). On the other hand, a larger γ value can markedly alter the error magnitude and distribution (see Fig. 10(f)). The simulation error for this problem appears to become worse with a large regularization coefficient γ .

Table 6 illustrates the effect of L^2 regularization in some quantitative sense. Here we list the maximum errors and the root-mean-squares (rms) errors of the DNN solutions in the domain computed using the C^∞ and C^1 periodic BCs corresponding to different γ values in the L^2 regularization. The simulation errors are essentially the same as those in the case without regularization, when γ is below about 10^{-3} with the C^∞ periodic BC or below about 10^{-4} with the C^1 periodic BC. Beyond these values, the simulation errors increase with increasing γ .

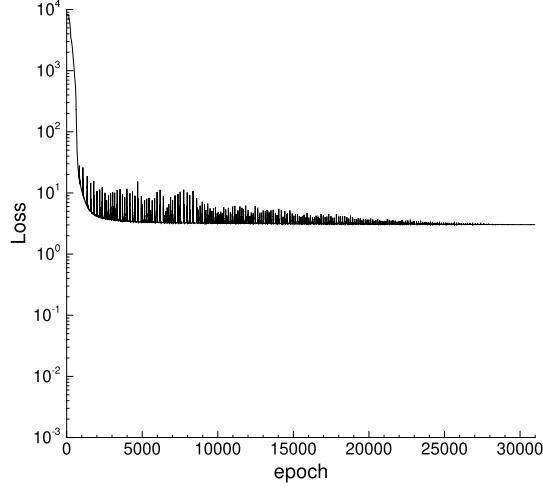
Regardless whether a regularization is imposed in the DNN simulation or not, the current method can exactly enforce the periodic conditions for the solution and its derivatives. This is demonstrated by Table 7 for the cases with L^2 regularization. Here we list the values of the DNN solution and its first and second derivatives, with 14 significant digits shown, on the domain boundaries $x = 0$ and $x = 4$ corresponding to different γ values, obtained using the C^∞ and C^1 periodic BCs. It is



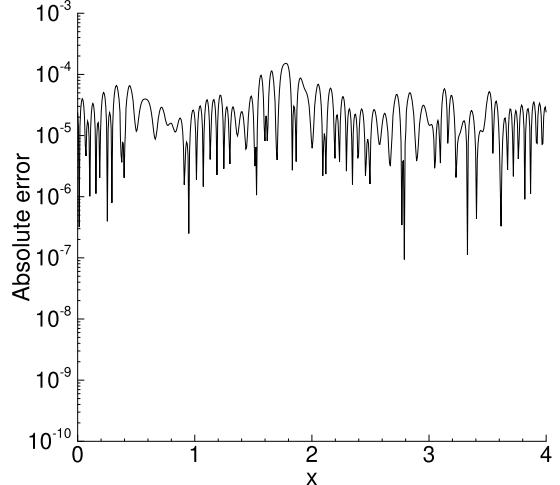
(a)



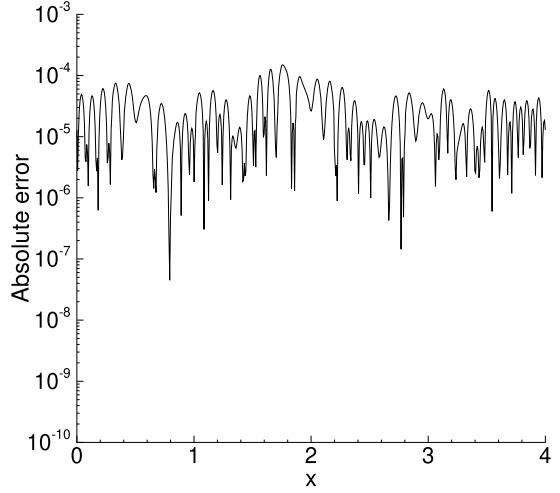
(c)



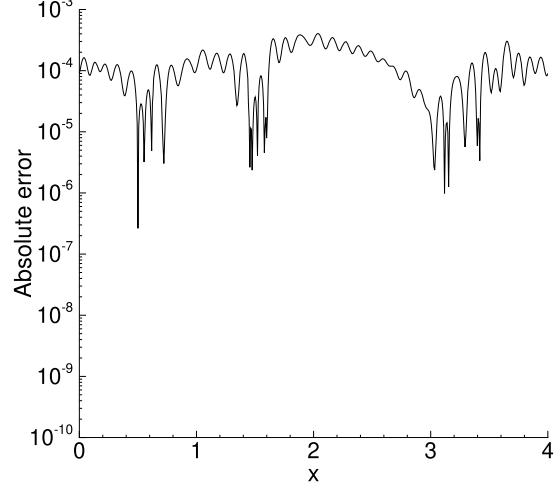
(e)



(b)



(d)



(f)

Fig. 10. 1D Helmholtz equation: The training histories (left column) and the absolute-error distributions (right column) of the DNN simulation using the C^∞ periodic BC, with L^2 regularization coefficients $\gamma = 10^{-4}$ (a,b), 10^{-3} (c,d), and 10^{-2} (e,f). These figures can be compared with Figs. 7(a) and 8(b), which are obtained using the C^∞ periodic BC with no regularization.

evident that the current method enforces the periodicity to the machine accuracy, for the solution and all its derivatives as shown here with the C^∞ periodic BC, and for the solution and its first derivative with the C^1 periodic BC.

3.3. Two-dimensional Helmholtz equation with periodic BCs

We next test the performance of the proposed methods using the Helmholtz equation in two dimensions (2D),

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \lambda u = f(x, y), \quad (37)$$

on a rectangular domain $\Omega = \{(x, y) | a_1 \leq x \leq b_1, a_2 \leq y \leq b_2\}$. Here λ, a_1, a_2, b_1 and b_2 are given constants, $u(x, y)$ is the unknown field function to be solved for, and $f(x, y)$ is a prescribed source term. We impose periodic boundary conditions in both the x and y directions.

Specifically, we consider C^1 and C^∞ periodic boundary conditions in 2D. The 2D C^1 periodic BC imposes the relations:

$$\begin{cases} u(a_1, y) = u(b_1, y), & \frac{\partial}{\partial x} u(a_1, y) = \frac{\partial}{\partial x} u(b_1, y), \quad \forall y \in [a_2, b_2]; \\ u(x, a_2) = u(x, b_2), & \frac{\partial}{\partial y} u(x, a_2) = \frac{\partial}{\partial y} u(x, b_2), \quad \forall x \in [a_1, b_1]. \end{cases} \quad (38)$$

The 2D C^∞ periodic BC imposes the relations:

$$\begin{cases} u(a_1, y) = u(b_1, y), & \frac{\partial}{\partial x} u(a_1, y) = \frac{\partial}{\partial x} u(b_1, y), & \frac{\partial^2}{\partial x^2} u(a_1, y) = \frac{\partial^2}{\partial x^2} u(b_1, y), \dots, \forall y \in [a_2, b_2]; \\ u(x, a_2) = u(x, b_2), & \frac{\partial}{\partial y} u(x, a_2) = \frac{\partial}{\partial y} u(x, b_2), & \frac{\partial^2}{\partial y^2} u(x, a_2) = \frac{\partial^2}{\partial y^2} u(x, b_2), \dots, \forall x \in [a_1, b_1]. \end{cases} \quad (39)$$

With the C^∞ periodic BC, we are effectively seeking a smooth periodic function $u(x, y)$ satisfying

$$u(x + L_1, y) = u(x, y), \quad u(x, y + L_2) = u(x, y), \quad \forall x, y \in (-\infty, \infty), \quad (40)$$

where $L_1 = b_1 - a_1$ and $L_2 = b_2 - a_2$.

We specifically consider the following parameter values for the numerical tests in this section:

$$\lambda = 10, \quad a_1 = a_2 = 0, \quad b_1 = b_2 = 4, \quad L_1 = L_2 = 4. \quad (41)$$

We choose the source term $f(x, y)$ such that the 2D Helmholtz equation (37) has the solution given by,

$$u(x, y) = -[1.5 \cos(\pi x + 0.4\pi) + 2 \cos(2\pi x - 0.2\pi)][1.5 \cos(\pi y + 0.4\pi) + 2 \cos(2\pi y - 0.2\pi)]. \quad (42)$$

This analytic solution satisfies the periodic boundary conditions (38) and (39).

To simulate this problem, we employ a feed-forward DNN with 2 nodes in the input layer, one node in the output layer, and 4 hidden layers in between. The input layer consists of the coordinates x and y , and the output layer is the solution to the 2D Helmholtz equation u . Each of the four hidden layers contains 20 nodes (neurons) in its output. For the C^∞ periodic BCs, the second layer of this DNN (i.e. the first hidden layer) is set to be a 2D C^∞ periodic layer $\mathcal{L}_p^{2D}(m, n)$ with $m = 12$ and $n = 20$ (see equation (11)), in which the constants ω_1 and ω_2 are set to

$$\omega_1 = \frac{2\pi}{L_1} = \frac{\pi}{2}, \quad \omega_2 = \frac{2\pi}{L_2} = \frac{\pi}{2}. \quad (43)$$

For the C^1 periodic BCs, the second layer of this DNN is set to be a 2D C^1 periodic layer $\mathcal{L}_{C^1}^{2D}(m, n)$ with $m = 12$ and $n = 20$ (see equation (22)).

We minimize the following loss function,

$$\begin{aligned} \text{Loss} &= \frac{1}{V_\Omega} \int_{\Omega} \left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \lambda u - f(x, y) \right]^2 d\Omega \\ &= \frac{1}{V_\Omega} \sum_{e=0}^{N_e-1} \sum_{i,j=0}^{Q-1} \left[\frac{\partial^2 u}{\partial x^2} \Big|_{(x_i^e, y_j^e)} + \frac{\partial^2 u}{\partial y^2} \Big|_{(x_i^e, y_j^e)} - \lambda u(x_i^e, y_j^e) - f(x_i^e, y_j^e) \right]^2 J^e w_{ij}, \end{aligned} \quad (44)$$

where $V_\Omega = \int_{\Omega} d\Omega = 16$ is the area of the domain, N_e is the number of elements (sub-domains) we have partitioned the domain Ω into for computing the integral, Q is the number of quadrature points in the x and y directions within

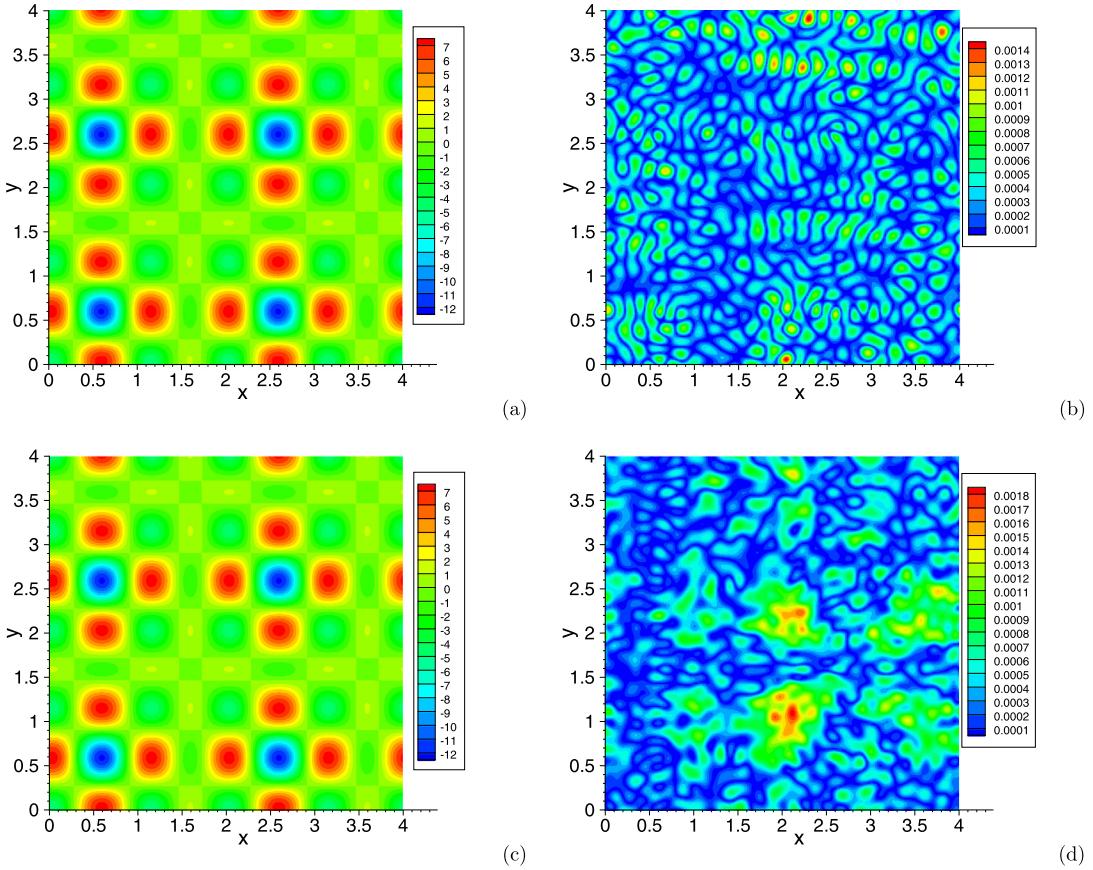


Fig. 11. 2D Helmholtz equation: Contours of DNN Solutions (left column) and their absolute errors against the exact solution (right column), obtained with C^∞ (top row) and C^1 (bottom row) periodic boundary conditions. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

each element, J^e is the Jacobian of the element e ($0 \leq e \leq N_e - 1$), (x_i^e, y_j^e) ($0 \leq i, j \leq Q - 1$) are the Gauss-Lobatto-Legendre quadrature points within the element e ($0 \leq e \leq N_e - 1$), and w_{ij} ($0 \leq i, j \leq Q - 1$) are the Gauss-Lobatto-Legendre quadrature weights associated with (x_i^e, y_j^e) . The input data to the DNN consist of (x_i^e, y_j^e) ($0 \leq i, j \leq Q - 1, 0 \leq e \leq N_e - 1$), and $f(x_i^e, y_j^e)$ ($0 \leq i, j \leq Q - 1, 0 \leq e \leq N_e - 1$) are passed to the DNN as the label data. In the loss expression, $u(x_i^e, y_j^e)$ can be obtained from the output of the DNN, and the derivatives $\frac{\partial^2 u}{\partial x^2}$ and $\frac{\partial^2 u}{\partial y^2}$ on (x_i^e, y_j^e) can be computed by auto-differentiation.

In the numerical tests below, we partition the domain into 4 elements ($N_e = 4$), with 2 uniform elements in both the x and y directions. We use 30 quadrature points ($Q = 30$) in each direction within each element. We employ the hyperbolic tangent (“tanh”) function as the activation function for each of the hidden layers, and no activation is applied to the output layer. The DNN is trained using the L-BFGS optimizer for 30,500 iterations with both the C^∞ and the C^1 periodic BCs. As mentioned before, we employ the L-BFGS routine “lbfgs_minimize()” from the Tensorflow-Probability library in our implementation, and use the default values for the parameters as specified in its documentation (e.g. the tolerance is 10^{-8}).

Fig. 11 shows contours of the DNN solutions (left column) and their absolute errors (right column) against the exact solution (42), computed with the C^∞ periodic boundary conditions (top row) and the C^1 periodic boundary conditions (bottom row). The distributions of the DNN solutions are qualitatively the same as that of the exact solution, and no difference can be discerned visually. The maximum absolute error of the DNN solution in the domain is about 1.5×10^{-3} with the C^∞ periodic BC and about 1.9×10^{-3} with the C^1 periodic BC. Fig. 12 provides a quantitative comparison between the DNN solutions and the exact solution. It shows the profiles of the DNN solutions obtained using C^∞ and C^1 periodic BCs, as well as the exact solution, along several horizontal lines across the domain located at $y = 0.5, 2$, and 3.5 . The error profiles of the DNN solutions along these lines are also shown in this figure. We observe that the DNN solutions with both the C^∞ and C^1 periodic BCs obtained using the current method agree very well with the exact solution.

To examine how well the current methods enforce the periodic boundary conditions for the 2D Helmholtz equation, we have extracted the values of the DNN solution and its partial derivatives (up to second order) on several corresponding points on the left and right boundaries, and on the top and bottom boundaries. Note that these derivatives are computed

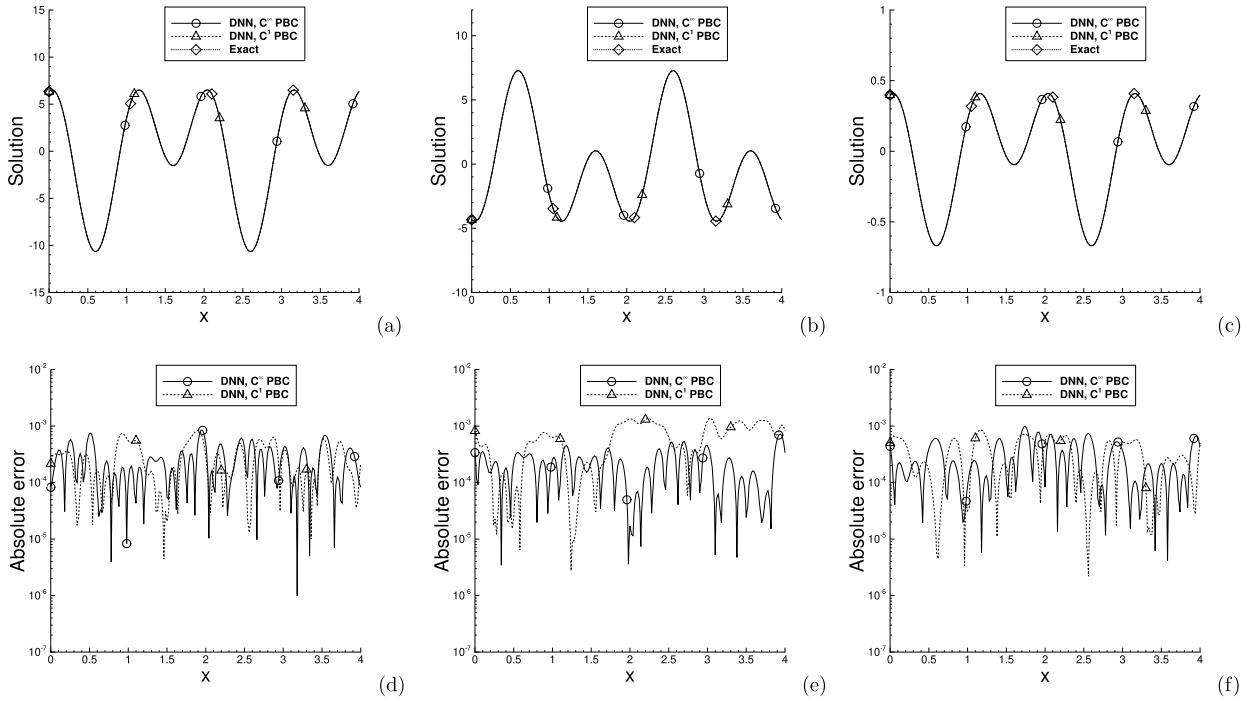


Fig. 12. 2D Helmholtz equation: comparison of profiles of the solution (top row) and its absolute error against the exact solution (bottom row) along several horizontal lines located at (a,d) $y = 0.5$, (b,e) $y = 2.0$, and (c,f) $y = 3.5$, from the DNN solutions with C^∞ and C^1 periodic boundary conditions. The profiles of the exact solution are also included for comparison.

by auto-differentiation, and they are the exact derivatives corresponding to the given DNN representation of the field. Table 8 lists the values of the DNN solutions and their partial derivatives on several corresponding points on the left/right boundaries and top/bottom boundaries. The second-order mixed derivatives and some first derivatives are not listed in the table, such as $\frac{\partial^2 u}{\partial x \partial y}$, $\frac{\partial u}{\partial y}$ on the left/right boundaries, and $\frac{\partial u}{\partial x}$ on the top/bottom boundaries. These unlisted values are exactly the same on the corresponding boundary points with both the C^∞ and C^1 periodic BCs. The boxed values in this table highlight the difference in the second partial derivatives on the corresponding boundary points of the DNN solution obtained with C^1 periodic BCs. As expected, the current methods have enforced the periodic boundary conditions exactly for the solution and the corresponding higher-order derivatives.

3.4. Diffusion equation with periodic BCs

The next test problem is the unsteady diffusion equation:

$$\frac{\partial u}{\partial t} - \nu \frac{\partial^2 u}{\partial x^2} = f(x, t), \quad (45)$$

where the constant $\nu > 0$ is the diffusion coefficient, $u(x, t)$ is the unknown field function to be solved for, $f(x, t)$ is a prescribed source term, x is the spatial coordinate, and t is time. We consider the spatial-temporal domain $\Omega = \{(x, t) | a \leq x \leq b, 0 \leq t \leq T\}$, where a, b, T are prescribed constants whose values are specified below. This equation is supplemented by the initial condition,

$$u(x, 0) = u_{in}(x), \quad (46)$$

where u_{in} denotes the initial distribution.

We impose periodic boundary conditions on the spatial boundaries $x = a$ and b . We specifically consider the C^∞ and C^1 periodic BCs. The C^∞ periodic BC requires,

$$u(a, t) = u(b, t), \quad \frac{\partial}{\partial x} u(a, t) = \frac{\partial}{\partial x} u(b, t), \quad \frac{\partial^2}{\partial x^2} u(a, t) = \frac{\partial^2}{\partial x^2} u(b, t), \quad \dots, \quad \forall t \in [0, T]. \quad (47)$$

The C^1 periodic BC requires,

$$u(a, t) = u(b, t), \quad \frac{\partial}{\partial x} u(a, t) = \frac{\partial}{\partial x} u(b, t), \quad \forall t \in [0, T]. \quad (48)$$

Table 8

2D Helmholtz equation: Values of the solution and its derivatives on selected corresponding points of the left/right and top/bottom boundaries, obtained from the DNN solutions with C^∞ and C^1 periodic BCs and from the exact solution. $u_x = \frac{\partial u}{\partial x}$, $u_y = \frac{\partial u}{\partial y}$, $u_{xx} = \frac{\partial^2 u}{\partial x^2}$, and $u_{yy} = \frac{\partial^2 u}{\partial y^2}$.

	DNN C^∞ PBC	DNN C^1 PBC	Exact solution
$u(0, 0.5)$	6.3376378742740e+00	6.3377710228780e+00	6.3375550504603e+00
$u(4, 0.5)$	6.3376378742740e+00	6.3377710228780e+00	6.3375550504603e+00
$u_x(0, 0.5)$	8.8435063231492e+00	8.8494566780609e+00	8.8433359504634e+00
$u_x(4, 0.5)$	8.8435063231492e+00	8.8494566780609e+00	8.8433359504634e+00
$u_{xx}(0, 0.5)$	-2.0831254187423e+02	-2.0869839735215e+02	-2.0841095826412e+02
$u_{xx}(4, 0.5)$	-2.0831254187423e+02	-2.0835671494578e+02	-2.0841095826412e+02
$u(0, 3.5)$	3.9895135572211e-01	3.9901841846678e-01	3.9851292703942e-01
$u(4, 3.5)$	3.9895135572206e-01	3.9901841846678e-01	3.9851292703942e-01
$u_x(0, 3.5)$	5.4847970782262e-01	5.6649740417208e-01	5.5607938177296e-01
$u_x(4, 3.5)$	5.4847970782265e-01	5.6649740417208e-01	5.5607938177297e-01
$u_{xx}(0, 3.5)$	-1.3183580874671e+01	-1.3617295595959e+01	-1.3105126558055e+01
$u_{xx}(4, 3.5)$	-1.3183580874671e+01	-1.3406939322481e+01	-1.3105126558055e+01
$u(1, 0)$	-2.4026513592099e+00	-2.4026445343439e+00	-2.4031781074217e+00
$u(1, 4)$	-2.4026513592099e+00	-2.4026445343439e+00	-2.4031781074217e+00
$u_y(1, 0)$	-3.3491228258428e+00	-3.3537672992540e+00	-3.3533612226667e+00
$u_y(1, 4)$	-3.3491228258428e+00	-3.3537672992540e+00	-3.3533612226667e+00
$u_{yy}(1, 0)$	7.8862068922108e+01	7.9273831286344e+01	7.9028686655862e+01
$u_{yy}(1, 4)$	7.8862068922108e+01	7.8873562549603e+01	7.9028686655862e+01
$u(3, 0)$	-2.4027241121576e+00	-2.4024234030698e+00	-2.4031781074217e+00
$u(3, 4)$	-2.4027241121576e+00	-2.4024234030698e+00	-2.4031781074217e+00
$u_y(3, 0)$	-3.3581682386601e+00	-3.3584974075468e+00	-3.3533612226667e+00
$u_y(3, 4)$	-3.3581682386600e+00	-3.3584974075468e+00	-3.3533612226667e+00
$u_{yy}(3, 0)$	7.8767439914223e+01	7.9141884228911e+01	7.9028686655862e+01
$u_{yy}(3, 4)$	7.8767439914222e+01	7.8931266277797e+01	7.9028686655862e+01

For the numerical tests in this subsection we employ the following parameter values,

$$\nu = 0.01, \quad a = 0, \quad b = 4, \quad T = 4.5, \quad L = b - a = 4. \quad (49)$$

We choose the source term $f(x, t)$ such that the function,

$$u(x, t) = (2 \cos(\pi x + 0.2\pi) + 1.5 \cos(2\pi x - 0.6\pi)) (2 \cos(\pi t + 0.2\pi) + 1.5 \cos(2\pi t - 0.6\pi)), \quad (50)$$

is a solution to the equation (45). We choose the initial distribution $u_{in}(x)$ by using the analytic expression (50) and setting $t = 0$. Note that the expression (50) satisfies the boundary conditions (48) and (47). So under the initial condition (46) and the periodic boundary conditions, the exact solution to the diffusion equation is given by (50).

We solve this initial/boundary value problem using DNN together with the method from Section 2 for enforcing the periodic BCs in x . We employ a DNN with two nodes in the input layer, which represent the spatial coordinate x and the time t , and one node in the output layer, which represents the unknown function $u(x, t)$ to be solved for. This DNN contains 3 hidden layers between the input and the output layers. Each of the hidden layers has an output consisting of 30 nodes. Note that periodic BCs are imposed only in the x direction, not in time. For the C^∞ periodic BCs, the second layer of this DNN (or the first hidden layer) is set to be a modified 2D C^∞ periodic layer as discussed in the Remark 2.2. For the periodic direction x , this modified 2D C^∞ periodic layer corresponds to a 1D C^∞ periodic layer $\mathcal{L}_p(m, n)$ with $m = 12$ and $n = 30$ (see equation (8a)), in which the constant ω is set to

$$\omega = \frac{2\pi}{L} = \frac{\pi}{2}. \quad (51)$$

For the C^1 periodic BCs, the second layer of this DNN is set to be a modified 2D C^1 periodic layer as discussed in Remark 2.4. For the periodic direction x , this modified 2D C^1 periodic layer corresponds to a 1D C^1 periodic layer $\mathcal{L}_{C^1}(m, n)$ with $m = 12$ and $n = 30$ (see equation (19)).

We minimize the following loss function,

$$\begin{aligned} \text{Loss} &= \theta_{eq} \frac{1}{V_\Omega} \int_{\Omega} \left[\frac{\partial u}{\partial t} - \nu \frac{\partial^2 u}{\partial x^2} - f(x, t) \right]^2 d\Omega + \theta_{ic} \frac{1}{b-a} \int_a^b [u(x, 0) - u_{in}(x)]^2 dx \\ &= \theta_{eq} \frac{1}{V_\Omega} \sum_{e=0}^{N_{el}-1} \int_{\Omega_e} \left[\frac{\partial u}{\partial t} - \nu \frac{\partial^2 u}{\partial x^2} - f(x, t) \right]^2 d\Omega + \theta_{ic} \frac{1}{b-a} \sum_{e=0}^{N_{el}-1} \int_{a_e}^{b_e} [u(x, 0) - u_{in}(x)]^2 dx \end{aligned}$$

$$\begin{aligned}
&= \frac{\theta_{eq}}{(b-a)T} \sum_{e=0}^{N_{el}-1} \sum_{i,j}^{Q-1} \left[\frac{\partial u}{\partial t} \Big|_{(x_i^e, t_j^e)} - \nu \frac{\partial^2 u}{\partial x^2} \Big|_{(x_i^e, t_j^e)} - f(x_i^e, t_j^e) \right]^2 J^e w_{ij} \\
&\quad + \frac{\theta_{ic}}{b-a} \sum_{e=0}^{N_{el}^x-1} \sum_{i=0}^{Q-1} [u(x_i^e, 0) - u_{in}(x_i^e)]^2 J^{ex} w_i,
\end{aligned} \tag{52}$$

where $V_\Omega = \int_\Omega d\Omega = (b-a)T$ is the volume of the spatial-temporal domain Ω , N_{el} is the number of spatial-temporal elements we have partitioned Ω into in order to compute the integral, Ω_e denotes the spatial-temporal element e for $0 \leq e \leq N_{el} - 1$, and Q is the number of Gauss-Lobatto-Legendre quadrature points in both the x and t directions within each spatial-temporal element. We use N_{el}^x to denote the number of elements in the spatial direction, and N_{el}^t to denote the number of elements in time, and then $N_{el} = N_{el}^x N_{el}^t$. The sub-interval $[a_e, b_e]$ denotes the spatial element e for $0 \leq e \leq N_{el}^x - 1$. The constants θ_{eq} and θ_{ic} are the penalty coefficients for the equation residual term and the initial condition residual term in (52). The Gauss-Lobatto-Legendre quadrature points within the spatial-temporal element Ω_e are denoted by (x_i^e, t_j^e) , with the associated quadrature weights w_{ij} . J^e is the Jacobian associated with the element Ω_e for $0 \leq e \leq N_{el} - 1$. J^{ex} is the Jacobian of the spatial element $[a_e, b_e]$ for $0 \leq e \leq N_{el}^x - 1$. w_i ($0 \leq i \leq Q - 1$) denote the quadrature weights associated with Gauss-Lobatto-Legendre quadrature points x_i^e in the spatial direction. The input data to the DNN consist of all the quadrature points within the domain, (x_i^e, t_j^e) , for $0 \leq i, j \leq Q - 1$ and $0 \leq e \leq N_{el} - 1$. The values of the source term on the quadrature points, $f(x_i^e, t_j^e)$, are passed to the DNN as the label data. The terms $\frac{\partial u}{\partial t} \Big|_{(x_i^e, t_j^e)}$ and $\frac{\partial^2 u}{\partial x^2} \Big|_{(x_i^e, t_j^e)}$ in (52) are computed based on auto-differentiation. It can be observed that here we are enforcing the initial condition by the penalty method.

In the numerical tests reported below, the domain Ω is partitioned into 9 spatial-temporal elements ($N_{el} = 9$), with 3 uniform elements in time ($N_{el}^t = 3$) and also 3 elements in the x direction ($N_{el}^x = 3$). Along the x direction, the two interior boundaries of the elements are located at $x = 1.3$ and 2.6 . We employ 20 quadrature points ($Q = 20$) in space and time within each spatial-temporal element. The penalty coefficients are set to be $\theta_{eq} = 0.9$ and $\theta_{ic} = 1 - \theta_{eq} = 0.1$. The “tanh” activation function has been employed for each hidden layer, and no activation function is applied to the output layer of the DNN. The L-BFGS optimizer has been employed to train the DNN for 35,000 iterations with the C^∞ and C^1 periodic BCs.

Fig. 13 shows contours in the spatial-temporal ($x - t$) plane of the DNN solutions (left column) and their absolute errors against the exact solution (right column), obtained using the C^∞ periodic BCs (top row) and the C^1 periodic BCs (bottom row). Qualitatively, no difference can be discerned between distributions of the DNN solutions and the exact solution given by (50). Fig. 14 provides a quantitative comparison between the DNN solutions and the exact solution. It shows profiles of the exact solution and the DNN solutions corresponding to C^∞ and C^1 periodic BCs at three time instants $t = 0.75, 2.25$ and 3.75 . The profiles of the absolute errors of the DNN solutions are shown in this figure as well. The DNN solution profiles corresponding to both the C^∞ and C^1 periodic BCs are observed to overlap with those of the exact solution at different time instants. It can be observed from the error profiles that the DNN with the C^∞ periodic BCs appears to result in generally smaller errors than that with the C^1 periodic BCs for this problem.

To assess how well the current method enforces the periodic boundary conditions, we have extracted the values of the solution and its partial derivatives (up to order two) on the spatial boundaries ($x = 0$ and 4) at several time instants from the exact solution and the DNN solutions obtained using the C^∞ and C^1 periodic boundary conditions. Table 9 lists these boundary values for the solution and its derivatives, with 14 significant digits shown. As expected, the current method has enforced exactly the periodicity for the solution and all its derivatives extracted here with the C^∞ periodic BCs, while with the C^1 periodic BCs it enforces exactly the periodicity only for the solution and its first derivative.

3.5. Wave equation with periodic BCs

In the last numerical example, we test the proposed method using the wave equation,

$$\frac{\partial u}{\partial t} - c \frac{\partial u}{\partial x} = 0, \tag{53}$$

where the prescribed constant c represents the wave speed, $u(x, t)$ is the unknown field function to be solved for, x is the spatial coordinate and t is the time. We consider the spatial-temporal domain $\Omega = \{(x, t) | a \leq x \leq b, 0 \leq t \leq T\}$ for this problem, where a, b and T are prescribed constants whose values are to be specified below. We consider the following initial condition,

$$u(x, 0) = u_{in}(x) = 2 \operatorname{sech} \frac{3(x - x_0)}{\delta}, \quad \forall x \in [a, b], \tag{54}$$

where $x_0 \in [a, b]$ and δ are prescribed constants whose values are specified below.

We impose the periodic boundary condition on the spatial boundaries of the domain, $x = a$ and b . Specifically, we consider C^∞, C^0, C^1 and C^2 periodic boundary conditions in this test. The C^k ($k = 0, 1, 2$) periodic BCs involve the conditions:

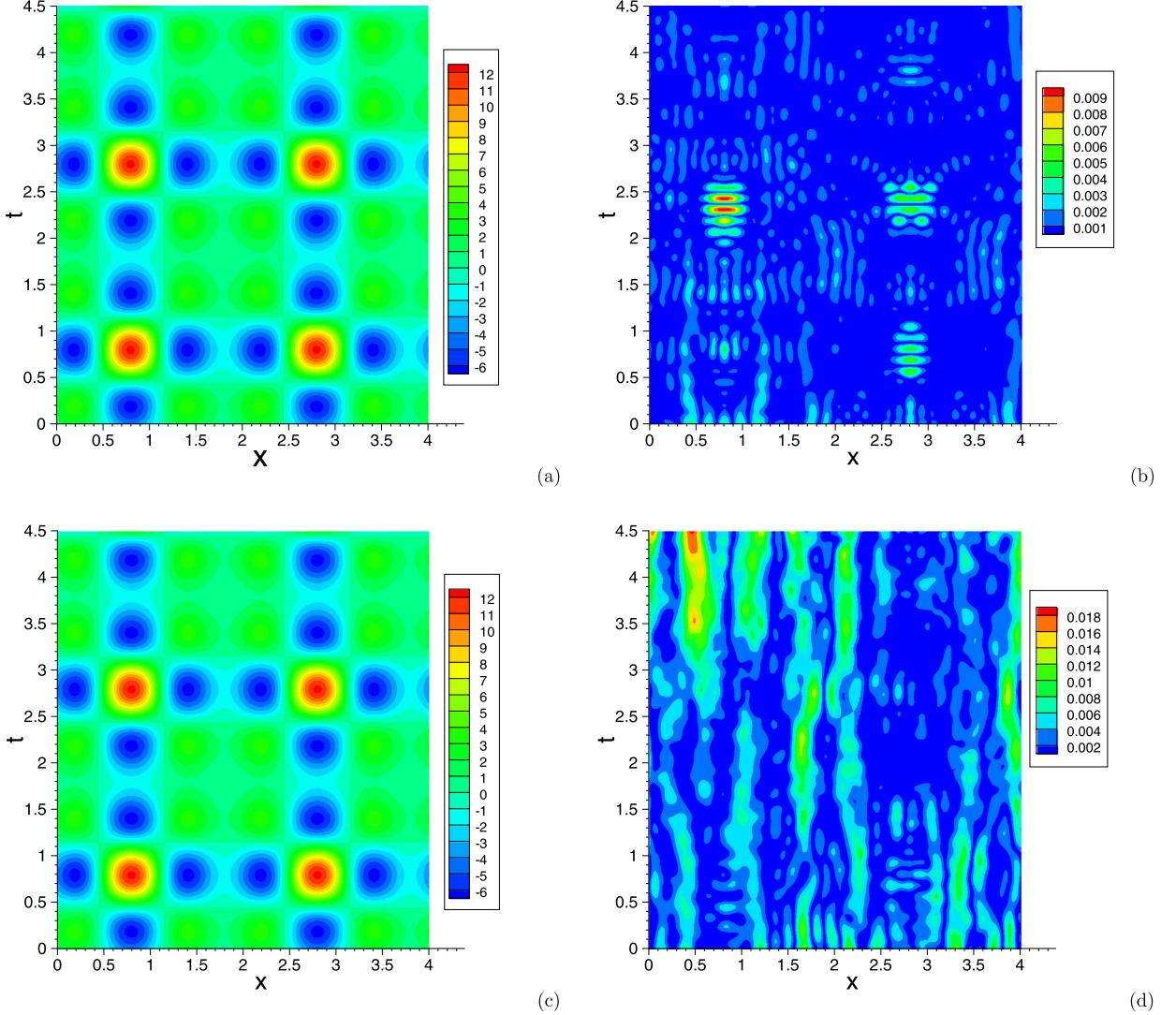


Fig. 13. Diffusion equation: DNN solutions (left column) and their absolute errors (right column) obtained with C^∞ (top row) and C^1 (bottom row) periodic boundary conditions in the x direction.

$$u(a, t) = u(b, t), \quad \forall t \in [0, T]; \quad (55)$$

$$\frac{\partial}{\partial x} u(a, t) = \frac{\partial}{\partial x} u(b, t), \quad \forall t \in [0, T]; \quad (56)$$

$$\frac{\partial^2}{\partial x^2} u(a, t) = \frac{\partial^2}{\partial x^2} u(b, t), \quad \forall t \in [0, T]. \quad (57)$$

The C^0 periodic BC imposes the condition (55). The C^1 periodic BC imposes the conditions (55) and (56). The C^2 periodic BC imposes the conditions (55)–(57). The C^∞ periodic BC imposes the conditions:

$$u(a, t) = u(b, t), \quad \frac{\partial}{\partial x} u(a, t) = \frac{\partial}{\partial x} u(b, t), \quad \dots, \quad \frac{\partial^m}{\partial x^m} u(a, t) = \frac{\partial^m}{\partial x^m} u(b, t), \quad \dots, \quad \forall t \in [0, T]. \quad (58)$$

This initial/boundary value problem has the solution,

$$u(x, t) = \begin{cases} u_{in}(x + ct) = 2 \operatorname{sech} \frac{3(x - x_0 + ct)}{\delta}, & \text{if } (x + ct) \in [a, b], \\ u(x \pm L, t), & \text{otherwise,} \end{cases} \quad \forall (x, t) \in \Omega, \quad (59)$$

where $L = b - a$. In the numerical tests reported below we have employed the following values for the parameters:

$$c = 2, \quad T = 4, \quad a = 0, \quad b = 4, \quad L = b - a = 4, \quad \delta = 1, \quad x_0 = 2. \quad (60)$$

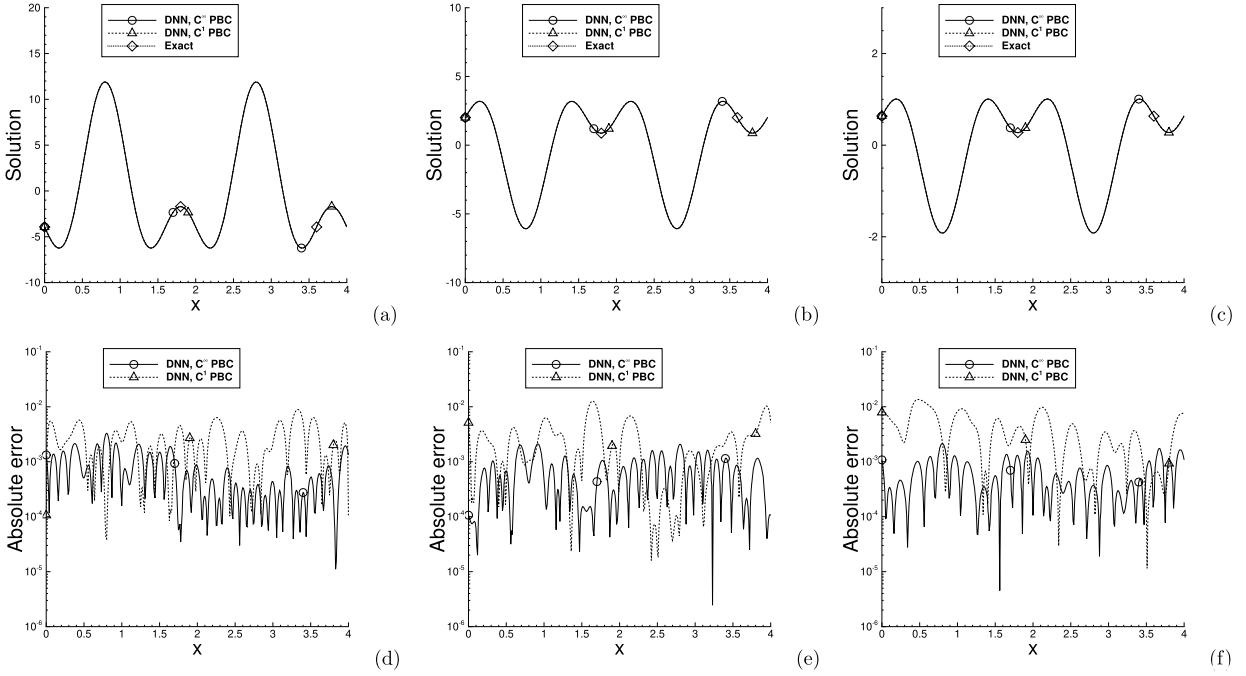


Fig. 14. Diffusion equation: comparison of profiles of the solution (top row) and the absolute error (bottom row) at time instants (a,d) $t = 0.75$, (b,e) $t = 2.25$, (c,f) $t = 3.75$ from the DNN solutions with C^∞ and C^1 periodic boundary conditions and from the exact solution.

Table 9

Diffusion equation: values of the solution and its derivatives on the spatial boundaries ($x = 0$ and 4) at several time instants (14 significant digits shown), obtained from the DNN solutions with C^∞ and C^1 periodic BCs and from the exact solution. The boxes highlight the differences in the obtained values for the second derivatives.

	DNN C^∞ PBC	DNN C^1 PBC	Exact solution
$u(0, 0.6)$	-2.4027959418736e+00	-2.4027209470767e+00	-2.4031781074217e+00
$u(4, 0.6)$	-2.4027959418736e+00	-2.4027209470767e+00	-2.4031781074217e+00
$u_x(0, 0.6)$	-1.0992666323436e+01	-1.0826214093012e+01	-1.0970511273440e+01
$u_x(4, 0.6)$	-1.0992666323436e+01	-1.0826214093012e+01	-1.0970511273440e+01
$u_{xx}(0, 0.6)$	-4.8267691835142e+00	-6.7181257407543e+00	-4.8498203327099e+00
$u_{xx}(4, 0.6)$	-4.8267691835143e+00	-4.7342701644443e+00	-4.8498203327100e+00
$u(0, 1.92)$	8.8578278309199e-01	8.8939474834158e-01	8.8446899507453e-01
$u(4, 1.92)$	8.8578278309200e-01	8.8939474834158e-01	8.8446899507453e-01
$u_x(0, 1.92)$	4.0439177602078e+00	3.9325284744288e+00	4.0376021450542e+00
$u_x(4, 1.92)$	4.0439177602078e+00	3.9325284744288e+00	4.0376021450542e+00
$u_{xx}(0, 1.92)$	1.3741473917761e+00	3.5766608911621e+00	1.7849345842144e+00
$u_{xx}(4, 1.92)$	1.3741473917761e+00	1.8868393439482e-03	1.7849345842144e+00
$u(0, 3.3)$	1.7324193446499e+00	1.7385194442865e+00	1.7317627457812e+00
$u(4, 3.3)$	1.7324193446499e+00	1.7385194442865e+00	1.7317627457812e+00
$u_x(0, 3.3)$	7.9201370767910e+00	7.7382513657243e+00	7.9054992498656e+00
$u_x(4, 3.3)$	7.9201370767909e+00	7.7382513657243e+00	7.9054992498656e+00
$u_{xx}(0, 3.3)$	3.2047702584285e+00	7.3508629515789e+00	3.4948463245323e+00
$u_{xx}(4, 3.3)$	3.2047702584286e+00	-1.0053353366089e+00	3.4948463245324e+00

To solve this initial/boundary value problem, we employ a feed-forward DNN together with the method from Section 2 for enforcing the periodic boundary conditions. The input layer of the DNN consists of two nodes, which represent the spatial coordinate x and the time t . The output layer of the DNN consists of one node, which represents the field function u to be solved for. We employ 3 hidden layers between the input and the output layers. Each hidden layer has an output with 30 nodes. Since the periodic BC is only imposed in the x direction, we employ the modified 2D periodic layers to enforce the periodic boundary conditions; see the Remarks 2.2 and 2.4. For the C^∞ periodic BCs, the second layer of this DNN (or the first hidden layer) is set to be a modified 2D C^∞ periodic layer as discussed in Remark 2.2. For the x direction, this modified periodic layer corresponds to the 1D C^∞ periodic layer $\mathcal{L}_p(m, n)$ with $m = 12$ and $n = 30$ (see equation (8)), in which the constant ω is set to $\omega = \frac{2\pi}{L} = \frac{\pi}{2}$. For the C^k ($k = 0, 1, 2$) periodic BCs, the second layer of this DNN is set to be

a modified 2D C^k periodic layer, which for the x direction corresponds to the 1D C^k periodic layer $\mathcal{L}_{C^k}(m, n)$ with $m = 12$ and $n = 30$ (see equation (19)).

We employ the following loss function for this DNN,

$$\begin{aligned}
 \text{Loss} &= \theta_{eq} \frac{1}{V_\Omega} \int_{\Omega} \left(\frac{\partial u}{\partial t} - c \frac{\partial u}{\partial x} \right)^2 d\Omega + \theta_{ic} \frac{1}{b-a} \int_a^b [u(x, 0) - u_{in}(x)]^2 dx \\
 &= \frac{\theta_{eq}}{V_\Omega} \sum_{e=0}^{N_{el}-1} \int_{\Omega_e} \left(\frac{\partial u}{\partial t} - c \frac{\partial u}{\partial x} \right)^2 d\Omega + \frac{\theta_{ic}}{b-a} \sum_{e=0}^{N_{el}^x-1} \int_{a_e}^{b_e} [u(x, 0) - u_{in}(x)]^2 dx \\
 &= \frac{\theta_{eq}}{(b-a)T} \sum_{e=0}^{N_{el}-1} \sum_{i,j=0}^{Q-1} \left[\frac{\partial}{\partial t} u(x_i^e, t_j^e) - c \frac{\partial}{\partial x} u(x_i^e, t_j^e) \right]^2 J^e w_{ij} \\
 &\quad + \frac{\theta_{ic}}{b-a} \sum_{e=0}^{N_{el}^x-1} \sum_{i=0}^{Q-1} [u(x_i^e, 0) - u_{in}(x_i^e)]^2 J^{ex} w_i.
 \end{aligned} \tag{61}$$

In the above expression, $V_\Omega = \int_{\Omega} d\Omega = (b-a)T$ is the volume of the spatial-temporal domain Ω , and the constants θ_{eq} and θ_{ic} are the penalty coefficients for the loss terms associated with the equation and the initial condition, respectively. In order to compute the integrals, we have partitioned the domain Ω into N_{el} spatial-temporal elements, with N_{el}^x elements in the spatial direction and N_{el}^t elements in time, leading to the relation $N_{el} = N_{el}^x N_{el}^t$. Ω_e denotes the region occupied by the spatial-temporal element e for $0 \leq e \leq N_{el} - 1$. The interval $[a_e, b_e]$ denotes the region of the spatial element e for $0 \leq e \leq N_{el}^x - 1$. Q is the number of quadrature points in both the spatial and temporal directions within each spatial-temporal element. (x_i^e, t_j^e) ($0 \leq i, j \leq Q - 1$) are the Gauss-Lobatto-Legendre quadrature points within the spatial-temporal element Ω_e , for $0 \leq e \leq N_{el} - 1$. J^e is the Jacobian of the spatial-temporal element Ω_e ($0 \leq e \leq N_{el} - 1$), and J^{ex} is the Jacobian of the spatial element $[a_e, b_e]$ ($0 \leq e \leq N_{el}^x - 1$). w_{ij} ($0 \leq i, j \leq Q - 1$) denote the weights associated with the Gauss-Lobatto-Legendre quadrature points (x_i^e, t_j^e) . w_i ($0 \leq i \leq Q - 1$) denote the weights associated with the spatial Gauss-Lobatto-Legendre quadrature point x_i^e . The input data to the DNN are the quadrature points (x_i^e, t_j^e) for $0 \leq i, j \leq Q - 1$ and $0 \leq e \leq N_{el} - 1$. In the loss function (61), $u(x_i^e, t_j^e)$ is obtained from the output of the DNN, and the terms $\frac{\partial u}{\partial t}|_{(x_i^e, t_j^e)}$ and $\frac{\partial u}{\partial x}|_{(x_i^e, t_j^e)}$ can be computed based on auto-differentiation. It can be observed that the initial condition is enforced by the penalty method.

For the numerical results reported below, we have partitioned the domain Ω into 4 spatial-temporal elements ($N_{el} = 4$), with 2 uniform elements along the spatial and temporal directions ($N_{el}^x = N_{el}^t = 2$). We employ 30 quadrature points in both space and time ($Q = 30$) within each spatial-temporal element. The penalty coefficients are set to be $\theta_{eq} = 0.9$ and $\theta_{ic} = 1 - \theta_{eq} = 0.1$. We use “tanh” as the activation functions for the hidden layers. No activation is applied to the output layer. The Adam optimizer has been employed to train the DNN for 60,000 epochs with the C^∞ and C^1 periodic BCs, for 90,000 epochs with the C^0 periodic BC, and for 80,000 epochs with the C^2 periodic BCs. The learning rate coefficient decreases gradually in a staircase fashion from 10^{-3} at the beginning to 10^{-4} at the end of the training process. The options for “early stopping” and “restore to best weight” in Tensorflow/Keras are employed during the training of the DNNs. The training histories of the loss function are illustrated in Fig. 15 with the C^∞ and C^1 periodic BCs.

Fig. 16 shows the distributions of the DNN solutions and their absolute errors in the spatial-temporal plane. The plots in the left column of this figure are contours of the DNN solutions in the spatial-temporal $(x-t)$ plane, obtained with the C^∞ and C^k ($k = 0, 1, 2$) periodic BCs on the spatial boundaries. The plots in the right column are contours of the absolute errors of these solutions against the exact solution (59). Qualitatively, no difference can be discerned of the distributions between the DNN solutions and the exact solution. The maximum errors in the domain are approximately on the order of magnitude 10^{-2} .

Fig. 17 shows a temporal sequence of snapshots of the wave form, obtained from the DNN solution with the C^∞ periodic boundary conditions. One can clearly observe the propagation of the wave form in the $-x$ direction at a constant speed. Because of the imposed periodic conditions, as soon as the wave exits the left boundary ($x = 0$), it re-enters the domain through the right boundary ($x = 4$) in a seamless and smooth fashion.

In Fig. 18 we compare the wave profiles from the exact solution (59) and from the DNN solutions with different types of periodic BCs at three time instants ($t = 0.5, 2$ and 3.5). The error profiles of these DNN solutions are also included in this figure (bottom row). It is observed that the wave profiles obtained from the DNN with various types of periodic BCs are in good agreement with that of the exact solution, and that the largest errors appear to coincide with the peak of the wave.

Table 10 provides a verification that the current method enforces the periodic conditions exactly for the wave equation as expected. Here we list the boundary values of the DNN solutions obtained with different types of periodic BCs and the exact solution, as well as their partial derivatives (up to order two), at several time instants ($t = 0.5, 2$ and 3.5). We have again included 14 significant digits for each value. It is evident that, with the C^∞ and C^2 periodic BCs, the current method

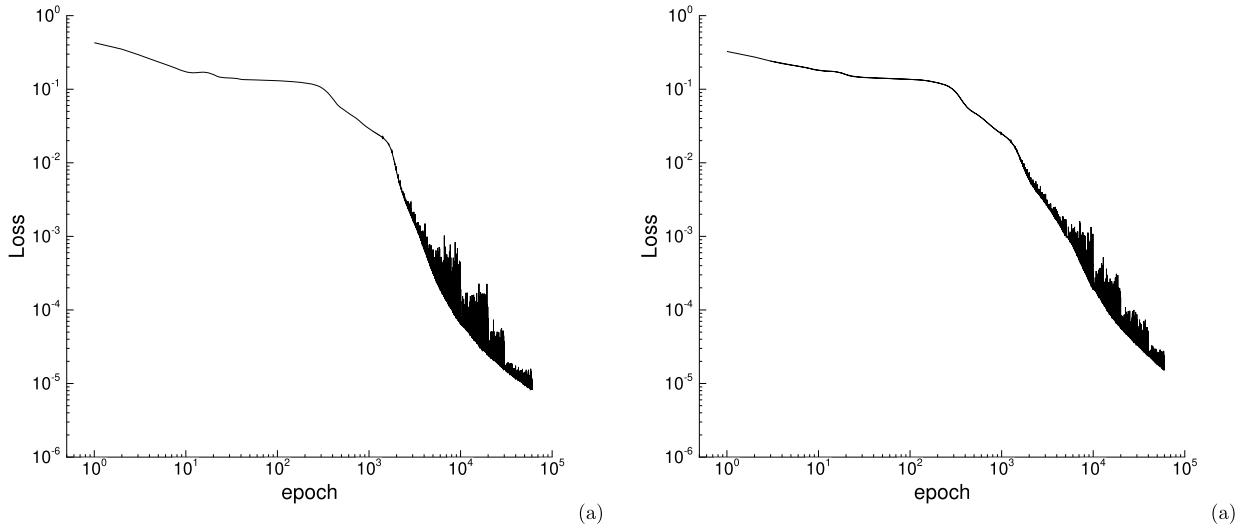


Fig. 15. Wave equation: training histories of the loss function with the (a) C^∞ and (b) C^1 periodic boundary conditions.

has enforced exactly the periodic conditions for the solution and its first and second derivatives. With the C^0 periodic BC, the method enforces exactly the periodic condition only for the solution. With the C^1 periodic BC, the method enforces exactly the periodic condition for the solution and its first derivative, but not for its second derivative.

4. Concluding remarks

In this paper we have presented a method for enforcing exactly the C^∞ and C^k (for any $k \geq 0$) periodic conditions with deep neural networks. The method stems from some simple properties about function compositions involving periodic functions. It essentially composes an arbitrary DNN-represented function with a set of independent known periodic functions with adjustable (training) parameters. More specifically, we have defined the operations that constitute a C^∞ periodic layer and a C^k periodic layer. The DNN with a C^∞ periodic layer incorporated as the second layer of the network (behind the input) automatically and exactly satisfies the C^∞ periodic boundary conditions in its output. The DNN with a C^k periodic layer incorporated as the second layer automatically and exactly satisfies the C^k periodic boundary conditions in its output. The C^∞ periodic layer comprises constructions of a set of independent C^∞ periodic functions with a prescribed period, based on sinusoidal functions, affine mappings, and nonlinear activation functions. The C^k periodic layer comprises constructions of a set of independent C^k periodic functions, based on the generalized Hermite interpolation polynomials, affine mappings, and nonlinear activation functions. We have tested the method in extensive numerical experiments with ordinary and partial differential equations involving C^∞ and C^k periodic boundary conditions. The numerical results demonstrate that the proposed method indeed enforces exactly, to the machine accuracy, the periodicity for the solution and its derivatives.

The proposed method can be implemented in a straightforward way. The C^∞ and C^k periodic layers defined herein can be implemented as user-defined Keras layers, and used in the same way as the built-in core Keras layers. All the numerical examples in the current work are implemented based on Tensorflow and Keras.

Periodic functions and periodic boundary conditions have widespread applications in computational science of various disciplines. The proposed method provides an effective tool, based on deep neural networks, for representing periodic functions and enforcing exactly the periodic boundary conditions. We anticipate that this method will be instrumental in expanding DNN-based techniques to new classes of applications that are unexplored or scarcely explored so far.

An outstanding question concerning the method developed herein is the following: Can an arbitrary periodic function of a certain regularity be represented by the current periodic DNNs to arbitrary accuracy? This is an important question and it is open at this point. Our extensive numerical experiments seem to suggest that the answer to this question is positive. The periodic DNNs from the current method are essentially compositions of an arbitrary DNN-represented function with a set of independent periodic functions with adjustable parameters. Can the universal approximation power of the original DNN carry over to the resultant periodic DNN as the set of independent periodic functions becomes sufficiently large? Can theoretical analysis establish an analogous universal approximation property for such periodic DNNs with respect to periodic functions? These are interesting questions that call for future research and should be pursued by the community.

CRediT authorship contribution statement

Suchuan Dong: Conceptualization, Methodology, Software, Data acquisition, Visualization, Writing of paper, Paper review, Funding acquisition.

Naxian Ni: Software, Data acquisition, Data curation, Visualization.

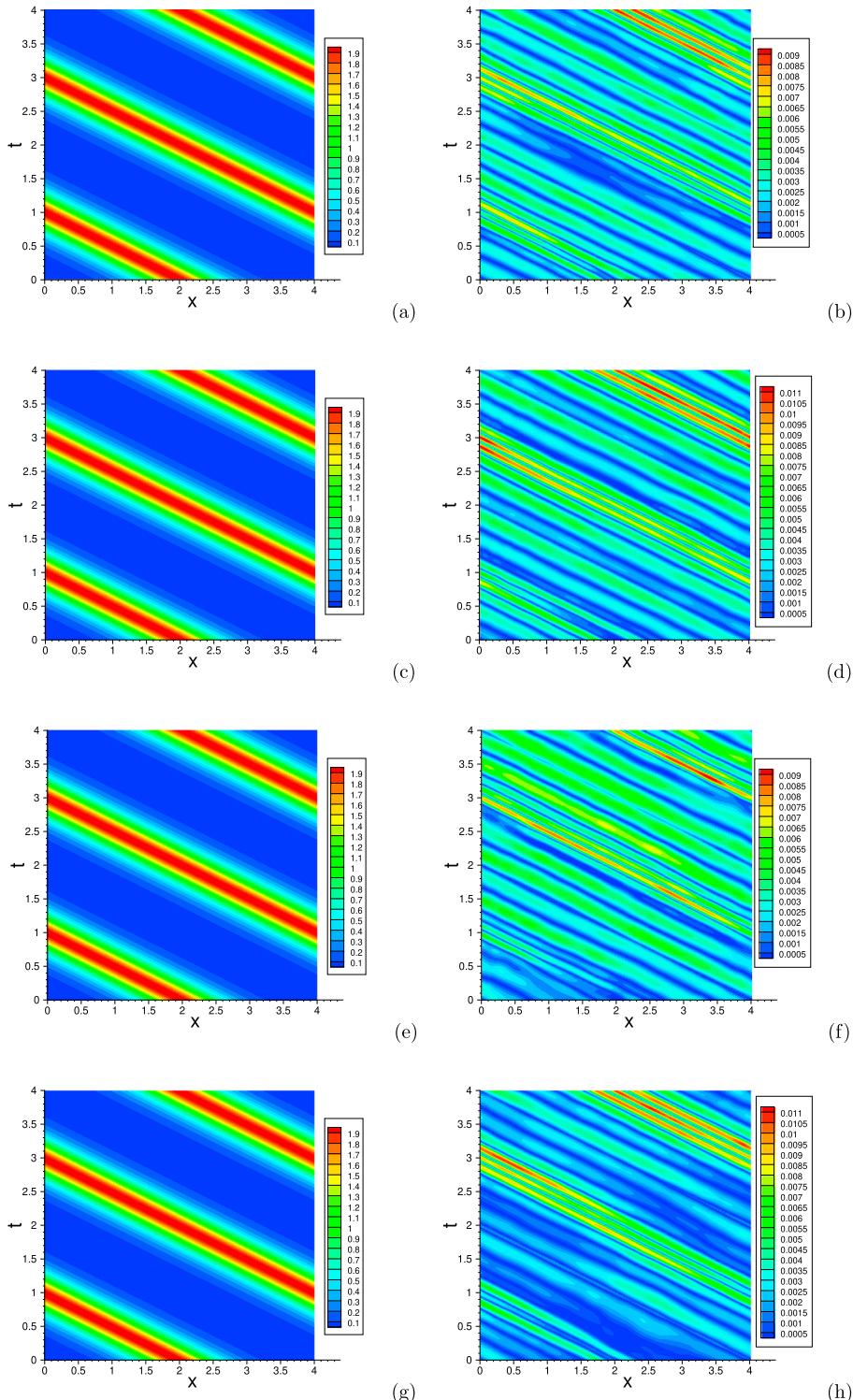


Fig. 16. Wave equation: Contours of the DNN solutions (left column) and their absolute errors (right column) obtained with the C^∞ (plots (a)-(b)), C^0 (plots (c)-(d)), C^1 (plots (e)-(f)), C^2 (plots (g)-(h)) periodic boundary conditions in the x direction.

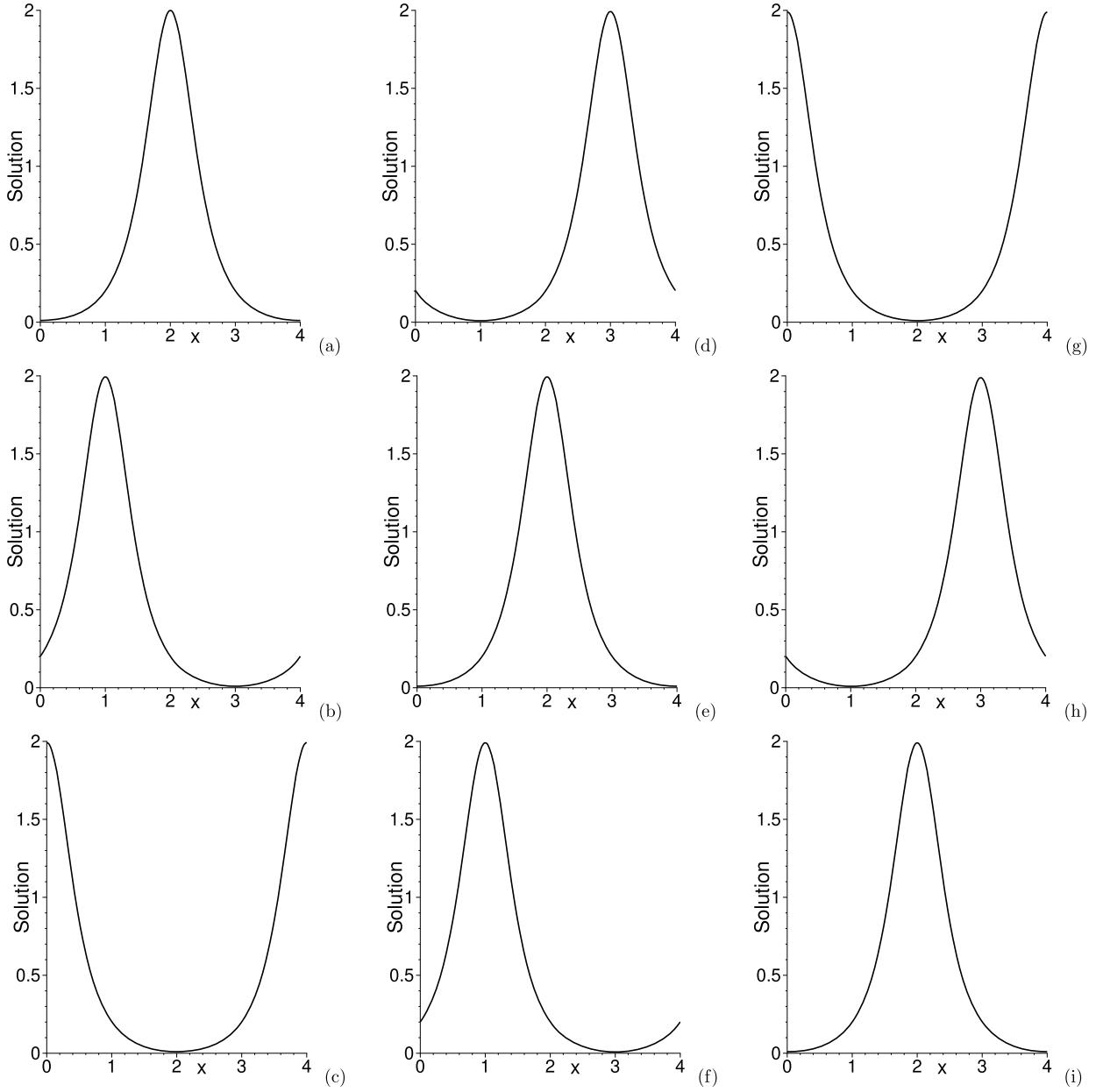


Fig. 17. Wave equation: Snapshots of the wave profiles obtained with the C^∞ periodic BC at time instants (a) $t = 0$, (b) $t = 0.5$, (c) $t = 1.0$, (d) $t = 1.5$, (e) $t = 2.0$, (f) $t = 2.5$, (g) $t = 3.0$, (h) $t = 3.5$, (i) $t = 4.0$.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was partially supported by NSF (DMS-2012415, DMS-1522537).

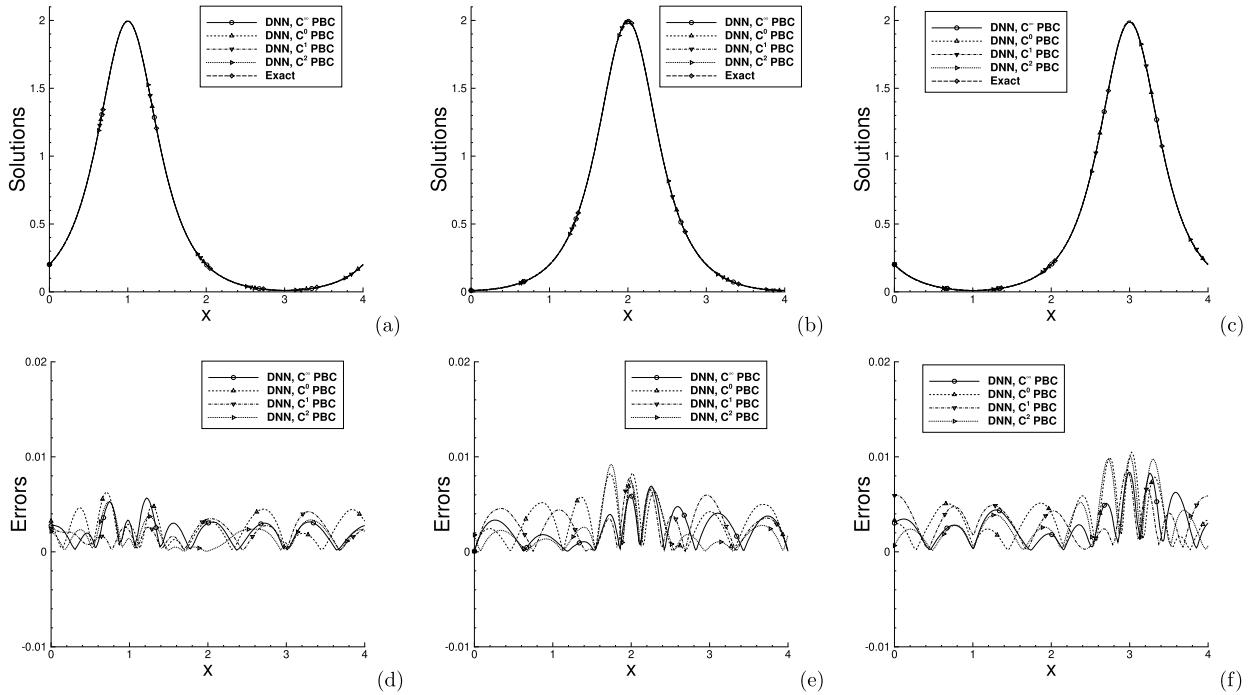


Fig. 18. Wave equation: comparison of profiles of the solution (top row) and the absolute error (bottom row) obtained with the C^∞ , C^0 , C^1 and C^2 periodic boundary conditions and with the exact solution at several time instants: (a, d) $t = 0.5$, (b, e) $t = 2.0$, (c, f) $t = 3.5$.

Table 10

Wave equation: values of the solution and its derivatives on the left/right boundaries at several time instants ($t = 0.5, 2$ and 3.5), obtained from the DNN solutions with C^∞ , C^0 , C^1 and C^2 periodic boundary conditions and from the exact solution.

	DNN C^∞ PBC	DNN C^0 PBC	DNN C^1 PBC	DNN C^2 PBC	Exact solution
$u(0, 0.5)$	2.0142799261625e-01	2.0190756733086e-01	2.0109869263559e-01	2.0083996838890e-01	1.9865585483886e-01
$u(4, 0.5)$	2.0142799261625e-01	2.0190756733086e-01	2.0109869263559e-01	2.0083996838891e-01	1.9865585483886e-01
$u_x(0, 0.5)$	5.9482715294468e-01	[5.7394660776802e-01]	5.9323728164392e-01	5.8746219179701e-01	5.9302035811534e-01
$u_x(4, 0.5)$	5.9482715294468e-01	[5.8248575768435e-01]	5.9323728164392e-01	5.8746219179701e-01	5.9302035811534e-01
$u_{xx}(0, 0.5)$	1.6980636297410e+00	[1.4696979666932e+00]	[1.5994875423905e+00]	1.6679798466138e+00	1.7526236647042e+00
$u_{xx}(4, 0.5)$	1.6980636297410e+00	[1.9395212811018e+00]	[1.5277304396975e+00]	1.6679798466138e+00	1.7526236647042e+00
$u(0, 2)$	9.8284051583224e-03	9.9241674367022e-03	9.9440310013471e-03	1.1696938690667e-02	9.9149477871207e-03
$u(4, 2)$	9.8284051583222e-03	9.9241674367022e-03	9.9440310013471e-03	1.1696938690667e-02	9.9149477871207e-03
$u_x(0, 2)$	1.5256678688244e-03	-4.8521014846574e-03	9.9336770382177e-04	2.0456109624086e-03	2.9744477846340e-02
$u_x(4, 2)$	1.5256678688247e-03	[2.3815674708337e-03]	9.9336770382177e-04	2.0456109624087e-03	2.9744477846340e-02
$u_{xx}(0, 2)$	2.4222108105432e-01	[4.2985279066538e-01]	[1.1927388753390e-01]	2.1278569005620e-01	8.9230143930769e-02
$u_{xx}(4, 2)$	2.4222108105432e-01	[2.2731343637548e-01]	[2.6553423293645e-01]	2.1278569005620e-01	8.9230143930769e-02
$u(0, 3.5)$	2.0170117943758e-01	2.0200326210422e-01	2.0458115518013e-01	1.9932351260077e-01	1.9865585483886e-01
$u(4, 3.5)$	2.0170117943758e-01	2.0200326210422e-01	2.0458115518013e-01	1.9932351260077e-01	1.9865585483886e-01
$u_x(0, 3.5)$	-5.8538598646821e-01	-6.1113064140580e-01	-5.8981323269116e-01	-5.7592303278591e-01	-5.9302035811534e-01
$u_x(4, 3.5)$	-5.8538598646821e-01	-5.9639323869798e-01	-5.8981323269116e-01	-5.7592303278591e-01	-5.9302035811534e-01
$u_{xx}(0, 3.5)$	1.6523878185422e+00	[2.3592894577992e+00]	[1.6262664810111e+00]	1.6631571855496e+00	1.7526236647042e+00
$u_{xx}(4, 3.5)$	1.6523878185422e+00	1.2305332426305e+00	1.8162774394782e+00	1.6631571855496e+00	1.7526236647042e+00

References

- [1] J. Berg, K. Nystrom, A unified deep artificial neural network approach to partial differential equations in complex geometries, Neurocomputing 317 (2018) 28–41.
- [2] C. Canuto, M.Y. Hussaini, A. Quarteroni, T. Zang, Spectral Methods in Fluid Dynamics, Springer, New York, 1988.
- [3] J. Chen, R. Du, K. Wu, A comprehensive study of boundary conditions when solving pdes by dnns, arXiv:2005.04554, 2020.
- [4] N.E. Cotter, The stone-weierstrass theorem and its application to neural networks, IEEE Trans. Neural Netw. 4 (1990) 290–295.
- [5] S. Dong, Direct numerical simulation of turbulent Taylor-Couette flow, J. Fluid Mech. 587 (2007) 373–393.
- [6] S. Dong, Turbulent flow between counter-rotating concentric cylinders: a direct numerical simulation study, J. Fluid Mech. 615 (2008) 371–399.
- [7] S. Dong, An outflow boundary condition and algorithm for incompressible two-phase flows with phase field approach, J. Comput. Phys. 266 (2014) 47–73.
- [8] S. Dong, A convective-like energy-stable open boundary condition for simulations of incompressible flows, J. Comput. Phys. 302 (2015) 300–328.

- [9] S. Dong, G.E. Karniadakis, C. Chryssostomidis, A robust and accurate outflow boundary condition for incompressible flow simulations on severely-truncated unbounded domains, *J. Comput. Phys.* 261 (2014) 83–105.
- [10] S. Dong, G.E. Karniadakis, A. Ekmekci, D. Rockwell, A combined direct numerical simulation-particle image velocimetry study of the turbulent near wake, *J. Fluid Mech.* 569 (2006) 185–207.
- [11] S. Dong, J. Shen, A pressure correction scheme for generalized form of energy-stable open boundary conditions for incompressible flows, *J. Comput. Phys.* 291 (2015) 254–278.
- [12] S. Dong, X. Zheng, Direct numerical simulation of spiral turbulence, *J. Fluid Mech.* 668 (2011) 150–173.
- [13] W. E, B. Yu, The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems, *Commun. Math. Stat.* 6 (2018) 1–12.
- [14] A.R. Gallant, H. White, There exists a neural network that does not make avoidable mistakes, in: Proceedings of the Second Annual IEEE Conference on Neural Networks, 1988.
- [15] F.S. Gokuzum, L.T.K. Nguyen, M.-A. Keip, An artificial neural network based solution scheme for periodic computational homogenization of electrostatic problems, *Math. Comput. Appl.* 24 (2019) 40.
- [16] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, The MIT Press, 2016.
- [17] P.M. Gresho, Incompressible fluid dynamics: some fundamental formulation issues, *Annu. Rev. Fluid Mech.* 23 (1991) 413–453.
- [18] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* 2 (1989) 359–366.
- [19] K. Hornik, M. Stinchcombe, H. White, Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks, *Neural Netw.* 3 (1990) 551–560.
- [20] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv:1412.6980, 2014.
- [21] I.E. Lagaris, A.C. Likas, D.I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Trans. Neural Netw.* 9 (1998) 987–1000.
- [22] I.E. Lagaris, A.C. Likas, D.G. Papageorgiou, Neural-network methods for boundary value problems with irregular boundaries, *IEEE Trans. Neural Netw.* 11 (2000) 1041–1049.
- [23] X. Li, Simultaneous approximations of multivariate functions and their derivatives by neural networks with one hidden layer, *Neurocomputing* 12 (1996) 327–343.
- [24] S. Liu, Fourier neural network for machine learning, in: Proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC), 2013.
- [25] K.S. McFall, J.R. Mahan, Artificial neural network method for solution of boundary value problems with exact satisfaction of arbitrary boundary conditions, *IEEE Trans. Neural Netw.* 20 (2009) 1221–1233.
- [26] M. Ngom, O. Marin, Approximating periodic functions and solving differential equations using a novel type of fourier neural networks, arXiv:2005.13100, 2020.
- [27] N. Ni, Z. Yang, S. Dong, Energy-stable boundary conditions based on a quadratic form: applications to outflow/open-boundary problems in incompressible flows, *J. Comput. Phys.* 391 (2019) 179–215.
- [28] J. Nocedal, S.J. Wright, *Numerical Optimization*, second edition, Springer, 2006.
- [29] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* 378 (2019) 686–707.
- [30] C. Rao, H. Sun, Y. Liu, Physics informed deep learning for computational elastodynamics without labeled data, arXiv:2006.0872, 2020.
- [31] K. Rudd, S. Ferrari, A constrained integration (CINT) approach to solving partial differential equations using artificial neural networks, *Neurocomputing* 155 (2015) 277–285.
- [32] E. Samaniego, C. Anitescu, S. Goswami, V.M. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, T. Rabczuk, An energy approach to the solution of partial differential equations in computational mechanics via machine learning: concepts, implementation and applications, *Comput. Methods Appl. Mech. Eng.* 362 (2020) 112790.
- [33] R.L. Sani, P.M. Gresho, Resume and remarks on the open boundary condition minisymposium, *Int. J. Numer. Methods Fluids* 18 (1994) 983–1008.
- [34] A. Silvscu, Fourier neural networks, in: Proceedings of the International Conference on Neural Networks (IJCNN99), 1999, pp. 488–491.
- [35] J. Sirignano, K. Spiliopoulos, DGM: a deep learning algorithm for solving partial differential equations, *J. Comput. Phys.* 375 (2018) 1339–1364.
- [36] A. Spitzbart, A generalization of Hermite's interpolation formula, *Am. Math. Mon.* 67 (1960) 42–46.
- [37] J.F. Traub, On lagrange-hermite interpolation, *J. Soc. Ind. Appl. Math.* 12 (1964) 886–891.
- [38] Y. Zang, G. Bao, X. Ye, H. Zhou, Weak adversarial networks for high-dimensional partial differential equations, *J. Comput. Phys.* 411 (2020) 109409.
- [39] A. Zhumekenov, M. Uteulyayeva, R. Takhanov, Z. Assylbekov, A.J. Castro, O. Kabdolov, Fourier neural networks: a comparative study, arXiv:1902.03011, 2019.