



Available online at www.sciencedirect.com

ScienceDirect

Comput. Methods Appl. Mech. Engrg. 389 (2022) 114333

**Computer methods
in applied
mechanics and
engineering**

www.elsevier.com/locate/cma

Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks

N. Sukumar^{a,*}, Ankit Srivastava^b

^a Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, USA

^b Department of Mechanical, Materials, and Aerospace Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA

Received 16 April 2021; received in revised form 21 October 2021; accepted 6 November 2021

Available online 2 December 2021

Abstract

In this paper, we introduce a new approach based on distance fields to exactly impose boundary conditions in physics-informed deep neural networks. The challenges in satisfying Dirichlet boundary conditions in meshfree and particle methods are well-known. This issue is also pertinent in the development of physics informed neural networks (PINN) for the solution of partial differential equations. We introduce *geometry-aware* trial functions in artificial neural networks to improve the training in deep learning for partial differential equations. To this end, we use concepts from constructive solid geometry (R-functions) and generalized barycentric coordinates (mean value potential fields) to construct $\phi(x)$, an approximate distance function to the boundary of a domain in \mathbb{R}^d . To exactly impose homogeneous Dirichlet boundary conditions, the trial function is taken as $\phi(x)$ multiplied by the PINN approximation, and its generalization via transfinite interpolation is used to a priori satisfy inhomogeneous Dirichlet (essential), Neumann (natural), and Robin boundary conditions on complex geometries. In doing so, we eliminate modeling error associated with the satisfaction of boundary conditions in a collocation method and ensure that kinematic admissibility is met pointwise in a Ritz method. With this new ansatz, the training for the neural network is simplified: sole contribution to the loss function is from the residual error at interior collocation points where the governing equation is required to be satisfied. Numerical solutions are computed using strong form collocation and Ritz minimization. To convey the main ideas and to assess the accuracy of the approach, we present numerical solutions for linear and nonlinear boundary-value problems over convex and nonconvex polygonal domains as well as over domains with curved boundaries. Benchmark problems in one dimension for linear elasticity, advection-diffusion, and beam bending, and in two dimensions for the steady-state heat equation, Laplace equation, biharmonic equation (Kirchhoff plate bending), and the nonlinear Eikonal equation are considered. The construction of approximate distance functions using R-functions extends to higher dimensions, and we showcase its use by solving a Poisson problem with homogeneous Dirichlet boundary conditions over the four-dimensional hypercube. The proposed approach consistently outperforms a standard PINN-based collocation method, which underscores the importance of exactly (a priori) satisfying the boundary condition when constructing a loss function in PINN. This study provides a pathway for meshfree analysis to be conducted on the exact geometry without domain discretization.

© 2021 Elsevier B.V. All rights reserved.

Keywords: Deep learning; Meshfree method; Distance function; R-function; Transfinite interpolation; Exact geometry

* Corresponding author.

E-mail address: nsukumar@ucdavis.edu (N. Sukumar).

1. Introduction

Machine learning algorithms based on supervised learning (deep neural networks) are relatively mature in fields such as computer vision, image processing, and speech recognition. Some of the earliest studies on physics-informed neural networks (PINN) to solve boundary-value problems can be traced to the contributions of Lagaris et al. [1–3], and those of McFall [4] and McFall and Mahan [5]. These studies have provided the impetus for the recent interest and advancement of the approach. Over the past 3–4 years there have been many new developments in a meshfree approach that is based on PINN to solve low- and high-order partial differential equations (PDEs). Some of the main contributions in this thread are based on: collocation [6–8], variational principle (deep Ritz) [9,10], and Petrov–Galerkin domain-decomposition [11,12]. Lu et al. [13] present an overview of PINN for the solution of PDEs. These developments have been made possible due to the advances in automatic differentiation tools that efficiently compute the derivatives of nonlinear composite functions, and in stochastic gradient descent algorithms that can deliver accurate solutions for nonlinear, nonconvex optimization problems. Furthermore, the availability of public-domain data analysis packages such as Tensorflow [14] and PyTorch [15] has also decreased the barriers to entry for newcomers to this field. Within the purview of solving PDEs with neural networks, it is instructive to highlight that unlike well-known computational methods (finite elements and its generalizations, finite volume, boundary elements, meshfree, and others), where function approximation is based on linear combinations of basis functions, which have to be chosen and defined a priori, function approximation in deep neural networks is through nonlinear function composition of an activation function $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}$, which yields the best approximation function via the solution procedure. A popular choice for σ is: $\sigma(x) := \text{ReLU}(x) = \max(0, x)$, which is known as the Rectified Linear Unit (ReLU) activation function. The collocation method using PINN is based on minimizing the least squares residual error (nonlinear and nonconvex mean squared error loss function) in order to satisfy the PDE and the boundary conditions at collocation points—whose solution yields the parameters that define the approximation function in the artificial neural network.

In PINN-based deep collocation [6] and deep Ritz [9] methods, the inexact imposition of boundary conditions adversely affects the training of the neural network as well as the accuracy of the method [16–18]. On complex geometries, this shortcoming becomes more acute. This issue has plagued meshfree Galerkin methods [19,20] since their inception, and a foolproof solution is still unavailable for arbitrary two- and three-dimensional geometries. Meshfree basis functions are also nonpolynomial, and hence an additional hurdle is a suitable cell-based numerical integration scheme that is consistent (patch test is passed) and stable (absence of spurious modes) for nonlinear simulations. Since solving for the approximating function is part of the minimization procedure in PINN, a simple (e.g., Monte Carlo) integration scheme for the entire domain can be used. Hence, if a reliable approach to impose boundary conditions on complex geometries in PINN is realized, it can lead to an accurate and robust meshfree method. In this paper, we tackle this problem by proposing a new approach in PINN to exactly impose boundary conditions, with an eye towards enabling meshfree simulations over complex geometries in \mathbb{R}^d ($d = 2, 3$). Our main contributions are as follows:

1. We introduce a new *geometry-aware* method in physics-informed neural networks that uses R-functions and transfinite interpolation to exactly impose boundary conditions over complex (affine, curved and multiply-connected) geometries. This geometry-aware approach, which is based on the construction of approximation distance functions (ADFs) to boundary sets, was first proposed by Kantorovich to satisfy Dirichlet boundary conditions [21], and has been extended using the theory of R-functions to exactly impose Dirichlet, Neumann and Robin boundary conditions in a meshfree method by initially Rvachev [22], and then subsequently by Rvachev and coworkers [23–25] and Shapiro and coworkers [26–31]. A related meshfree approach in the spirit of Kantorovich's method is that of Höllig et al. [32], who used web-splines that are multiplied by a weight function (approximates the distance function) to solve a Dirichlet problem. We also draw from a previous study [33], where R-functions are used to define smooth approximate distance fields over polygonal domains.
2. Approximate distance fields that stem from mean value potentials [34–36] are also used in PINN to solve PDEs.
3. On exactly satisfying boundary conditions in physics-informed neural networks, the training of the network is simplified, and this facilitates convergence and improved accuracy of the PINN approximation.
4. New application of R-functions in \mathbb{R}^4 : solution of the Poisson equation over the four-dimensional hypercube.

5. Solving PDEs on curved domains without discretizing the domain (mesh generation) is realized, which provides a pathway to conducting meshfree simulations on the exact geometry (isogeometric analysis) [37].

First, we present a few connections of prior work on finite elements and partition-of-unity meshfree methods to better understand and place the PINN approximation and its use to solve PDEs. Contributions in r -adaptivity with finite elements have a long history in computational mechanics. For instance, in finite-deformation simulations using finite elements, the optimal nodal locations and the solution coefficients have both been simultaneously treated as unknowns in the minimization of the potential energy functional [38]. Since the PINN approximation that is composed by the ReLU activation function can exactly represent piecewise affine functions (Delaunay basis functions) [39], one can view the ReLU network solution as a variational r -adaptive finite element solution procedure. Instead of refining elements in h -adaptive finite elements, adaptive solutions can be realized via a basis refinement strategy that has advantages (for example, “hanging nodes” are a nonissue), which was put forth by Grinspun [40], and a similar approximation refinement perspective can be associated with a multilayer neural network [41]. The connections between ReLU networks and hp -finite elements are studied in Opschoor et al. [42]. Initiated by Kansa [43,44], meshfree collocation schemes with positive-definite radial basis functions (RBFs) such as the Gaussian and Hardy’s inverse multiquadratics have been used to solve PDEs [45,46]. Schaback and Wendland [47] discuss the ties of kernel methods (for example, radial basis functions) to machine learning and meshfree methods. The choice on how to set the “shape parameter” in the Gaussian RBF (controls the support-width of the Gaussian) is an unsettled issue since it is problem dependent and must be carefully selected for boundary and interior points if exponential convergence is to be maintained without exacerbating the condition number. A related approach is the *local* maximum-entropy (max-ent) meshfree method [48] that yield compactly-supported basis functions of exponential form that constitute a partition-of-unity [49], possess linear completeness, and provide a smooth transition from Delaunay basis functions [50] to global maximum-entropy basis functions [51]. Consider a nodal set with nodes (centers) that are located at $\{\mathbf{x}_i\}_{i=1}^n$. When viewed through the lens of Gaussian weight functions [52–54], a single parameter $\{\beta_i\}_{i=1}^n$ controls the support-width of each Gaussian weight function. Rosolen et al. [55] proposed a variational adaptivity formulation to find optimal values of $\{\beta_i\}_{i=1}^n$ that minimize the potential energy functional for Poisson equation and nonlinear elasticity. Since RBFs can be represented in a neural network with a single hidden layer [56,57], the neural network solution optimizes the location of the centers $\{\mathbf{x}_i\}_{i=1}^n$ as well as the support-widths $\{\beta_i\}_{i=1}^n$ [56]. RBF-based partition-of-unity networks [58] for hp -approximation have been introduced, and numerical experiments have been conducted using sparse Gaussian networks to solve PDEs [59]. Lastly, we mention the recent work of Greco and Arroyo [60], who presented a collocation scheme for PDEs that is based on high-order max-ent approximants, which delivered accurate simulation results on domains with affine and curved boundaries. In Section 9.1.7, we present a one-dimensional example using Gaussian neural networks. Since PINN affords significant flexibility vis-à-vis existing meshfree (basis-set) methods, solving PDEs over complex geometries using collocation and Ritz methods with artificial neural networks holds significant promise.

In a strong collocation PINN method, the loss function consists of the residual errors from the interior of the domain, which is known as *interior or PDE loss*, and from the boundaries of the domain, which is referred to as *boundary (conditions) loss* [6]. There are three distinct contributions in the mean squared error loss function: (1) residual error at interior collocation points where the PDE must be satisfied; (2) residual error at boundary collocation points where the essential (Dirichlet) boundary condition must be satisfied; and (3) residual error at boundary collocation points where the Robin or natural (Neumann) boundary condition must be satisfied. Early approaches [1,3–5] had already recognized the importance of exact imposition of boundary conditions in artificial neural networks. Lagaris et al. [1] considered two terms in the trial function, with the first term being an analytical function that exactly imposed the boundary conditions and the second term was chosen as the product of the PINN approximation and a function that vanished on the boundary; for irregular boundaries, Lagaris et al. [3] used a RBF network in the first term to satisfy the boundary conditions at a collection of discrete points on the boundary. McFall et al. [4,5] and more recently Sheng and Yang [61] introduced a length factor (measure of the distance to the boundary) associated with the boundary to impose boundary conditions, and Berg and Nyström [7] approximated the distance function using a low-capacity neural network to impose boundary conditions over complex geometries. In many recent studies [16–18,62], the implications of imposing essential boundary conditions via the loss function in PINN have been studied, and numerical experiments have affirmed that the presence of the boundary residual terms compromises the convergence of the stochastic gradient descent algorithm and the accuracy of the method. To address this problem, remedies have been introduced in the PINN literature, such as using two neural networks, one

for the PDE and the other to satisfy the essential boundary condition [3,7,61–63], introduction of a penalty parameter via an augmented variational formulation to weakly impose the essential boundary conditions [9], and Nitsche's method to impose the essential boundary condition [64]. Some of these approaches mirror those previously pursued in meshfree and particle methods to satisfy essential boundary conditions [19,20,65]. In meshfree Galerkin methods, the choice of the space for Lagrange multipliers is delicate; penalty approach leads to a saddle-point problem and the inf-sup condition must be met; and though Nitsche's method is variationally consistent, the stabilization parameter in it must be judiciously chosen. For low-dimensional problems over complex geometries, an accurate and robust meshfree approach remains elusive. Since from the universal approximation theorem [66,67] we know that a neural network with one hidden layer can represent any L^2 function to arbitrary accuracy, it stands to choose an ansatz that satisfies the boundary conditions a priori, so that the loss function is expressed solely in terms of the residual error at only the interior collocation points where the PDE is required to be satisfied. If the essential boundary conditions are exactly met, then this precludes “variational crimes” in a Ritz method [68]. Lastly, and most importantly, in deep collocation [6], multiple terms (interior loss and boundary losses) that have to be individually minimized are incorporated within a single objective (loss) function. When this loss function is minimized, then the solution that is realized depends on the weight (equal weights is the unbiased choice) that is assigned to each objective function, which reflects the importance of each residual error contribution. Rohrhofer et al. [69] discuss network training in relation to the Pareto front that appears in multiobjective constrained optimization. In the NVIDIA SimNet™ toolkit [70], signed distance function weighting is used to dynamically assign the spatially varying weight functions for the PDE and boundary loss terms. Since these weights are problem dependent, they should not be fixed a priori, since then the magnitude of the training loss by itself is a misleading error measure. To establish the accuracy of the method, the error in u as well as in u' must be assessed. We present numerical results in Sections 9–11 that support this thesis, and which points to the merits of the new approach.

In this paper, we solve this problem of competing loss terms in PINN formulations by constructing a trial function for the neural network that a priori satisfies all boundary conditions in deep collocation, and meets kinematic admissibility when used in a deep Ritz method. This eliminates the boundary terms in the loss function. Our approach is based on constructing distance functions (exact or approximate) to the boundary of the domain, and it can treat essential (Dirichlet) as well as mixed (Dirichlet and Robin) boundary conditions over complex domains. We use the exact distance function whenever it is available and applicable. However, in general, we construct approximate distance functions using two different techniques: the theory of R-functions [23,30] and the theory of mean value potential fields [34–36]. These methods provide approximate distance functions that possess the desirable property of being zero on the boundary of the domain with unit (inward) normal directional derivative. In addition, they are smooth in the interior of the domain, a property that the exact distance function does not always possess. Functions whose sign solely depend on the sign of its arguments encode Boolean logic, and are known as R-functions. R-functions provide an implicit function representation for line segments, curves, and solid regions, and are composed by Boolean operations (negation, conjunction, disjunction, equivalence). Mean value potential fields are specific forms of a singular double-layer potential that yield L_p -distance fields [36]. For a domain in \mathbb{R}^2 , this singular potential is defined as the integral of the reciprocal of the p th power of the distance from its boundary. For a polygon with $p = 1$, closed-form expressions for the ADF are available [34], but for closed curves in \mathbb{R}^2 , numerical integration is required to compute the ADF. Once the approximate distance functions are formed, methods to impose essential and Robin boundary conditions are available that rely on transfinite interpolation [23,25]. R-functions with approximants such as B-splines [27,31] and RBFs [71] have been used in a meshfree Galerkin method to solve boundary-value problems.

The remainder of this paper is organized as follows. In Section 2, we discuss the properties of distance functions, and in Section 3, the essentials on R-functions and the construction of approximate distance functions are described. In particular, joining R-functions using R-equivalence composition is presented, which is used in this paper. The inverse of the normalizing weight function that appears in the expression for mean value coordinates (polygon) and transfinite mean value interpolant (closed curves), is an approximate distance field. These are particular instances ($p = 1$) of L_p -distance fields, and are discussed in Section 4. On using normalizing functions and solution structures in the R-function method [22,23], we describe in Section 5 the use of ADFs to construct an ansatz in PINN that exactly satisfies boundary conditions for second- and fourth-order problems. The construction of the trial function in a deep neural network is presented in Section 6, along with a summary of the feedforward neural network and the backpropagation (computation of the gradient of the loss function and use of stochastic gradient descent)

algorithm. The loss function for collocation and deep Ritz formulations are presented in Section 7. The numerical implementation is discussed in Section 8, where we also provide code snippets of some of the main functions. One- and two-dimensional numerical simulations are presented in Sections 9 and 10, where we apply this new approach to a broad suite of boundary-value problems (Poisson, harmonic coordinates, plate bending, Eikonal equation) on convex and nonconvex domains with affine and curved boundaries. In addition, the Poisson equation over the four-dimensional hypercube is solved in Section 11. These numerical results clearly demonstrate the benefits of exactly imposing the boundary conditions in PINN—it simplifies the training of the network and enhances the accuracy and robustness of the method. Finally, we conclude with Section 12, where we summarize the main developments in this paper and discuss some of the topics of future research.

2. Distance functions and their properties

The signed distance function is an implicit representation for curves and surfaces, and also provides fast evaluation of predicates for geometric objects. Let $S \subset \mathbb{R}^d$ denote a domain (open, bounded set) with boundary ∂S . The exact distance function $d(\mathbf{x})$ gives the shortest distance between any point $\mathbf{x} \in \mathbb{R}^d$ to ∂S . It is clear that $d(\mathbf{x})$ is identically zero on ∂S . Computing the exact distance function requires solving the Eikonal equation (see Section 10.5), which is computationally expensive. Therefore, it is desirable to construct an approximate distance function or ADF (formally represented by $\phi(\mathbf{x})$) that has a closed-form (non-iterative algorithm) expression. Furthermore, since the exact distance function may only be continuous and not continuously differentiable, it may not be suitable for use in a trial function to solve PDEs. Since our objective is to use ADFs in a collocation or Ritz method to solve boundary-value problems, their differential properties are important. If essential boundary conditions are imposed on the entire boundary ∂S , then the ADF must be zero on ∂S , positive in S , and its gradient must not vanish for any $\mathbf{x} \in \partial S$. In addition, since the exact distance function has derivative discontinuities on the medial axis of the domain, smooth approximations of the distance function must be used within the trial function for a collocation method with PINN. For a second-order problem, a C^0 distance function that has gradient discontinuities in the interior of the domain cannot be used in the collocation approach since the Laplacian of the distance function will be unbounded at a collocation point. These considerations are crucial when used to solve PDEs. For instance, positivity in S precludes the presence of singularities within the domain, which in general is difficult to construct as noted in McFall [4].

For clarity of exposition in this paper, we use $\mathbf{v} := \mathbf{v}(\mathbf{x})$ to denote the unit inward normal vector (appears in the theory related to R-functions) on the boundary ∂S , and $\mathbf{n} := \mathbf{n}(\mathbf{x})$ as the unit outward normal vector (used when defining Neumann or Robin boundary condition) on ∂S . It is noted that $\mathbf{n} = -\mathbf{v}$. If ∂S in \mathbb{R}^2 is composed of piecewise line segments and curves, then we use $\phi_i := \phi_i(\mathbf{x})$ to denote the ADF to each curve or line segment. For a point $\mathbf{x} \in \mathbb{R}^d$ on ∂S , it is essential that any approximation to the distance function satisfy $\phi = 0$. Furthermore, to mimic the exact distance function, the normal derivative with respect to \mathbf{v} on the boundary should be unity, $\partial d / \partial \mathbf{v} = \nabla d \cdot \mathbf{v} = 1$, and it is desirable that all higher order normal derivatives vanish. An m th order approximate distance function requires that the second- to m th-order normal derivatives vanish on all regular points (unit normal is well-defined) on ∂S [23]:

$$\frac{\partial \phi}{\partial \mathbf{v}} = 1, \quad \frac{\partial^k \phi}{\partial \mathbf{v}^k} = 0 \quad (k = 2, 3, \dots, m), \quad (1)$$

and such a function is said to be *normalized* to the m th order. For finite m , the normalized function matches the exact distance function only in the vicinity of the boundary; for points that are away from the boundary, it deviates from the exact distance. Apart from applications in solid modeling, mesh generation, real-time rendering and computer vision, where distance functions are used, normalized first-order distance functions are also a suitable choice for the initialization and assignment of the extension velocity at points away from the interface in the level set method [72]. As noted in Biswas and Shapiro [29], use of normalized distance functions mitigate the bulging phenomenon in the vicinity of where the segments or curves are joined [73], since undulations (presence of local extrema) are undesirable in the representation of the surface.

3. R-functions and approximate distance functions

The theory of R-functions can be used to construct a composite approximate distance function, $\phi(\mathbf{x})$, to any arbitrarily complex boundary ∂S , when approximate distance functions, $\phi_i(\mathbf{x})$, to the partitions of ∂S are known.

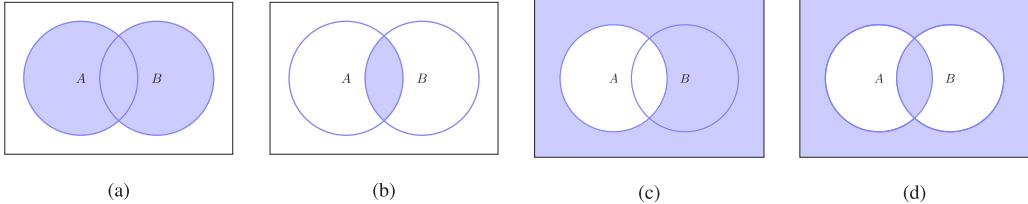


Fig. 1. Venn diagram for union, intersection, complement and equivalence in \mathbb{R}^2 . ω_A and ω_B are R-functions that are positive in Ω_A and Ω_B (open sets), respectively. (a) $A \cup B \equiv \omega_A \vee \omega_B$; (b) $A \cap B \equiv \omega_A \wedge \omega_B$; (c) $\bar{A} \equiv -\omega_A$; and (d) $(A \cap B) \cup (\bar{A} \cap \bar{B}) \equiv \omega_A \sim \omega_B$. Examples of \vee and \wedge operations using R-functions are given in (2) and (3).

Consider a real-valued function $F(\omega_1, \omega_2, \dots, \omega_q)$, where $\omega_i(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$ ($i = 1, \dots, q$) are also real-valued functions. If the sign of $F(\cdot)$ is solely determined by the signs of its arguments $\omega_i(\mathbf{x})$, then $F(\cdot)$ is known as an R-function [26,30]. R-functions were proposed by T. L. Rvachev in 1963 [26]. For example, $F_1(x, y) = 1 + x^2 + y^2$ and $F_2(x, y) = xy$ are R-functions in \mathbb{R}^2 , whereas $F_3(x, y) = \sqrt{x^2 + y^2} - 1$ and $F_4(x, y) = \sin xy$ are not. The important properties of R-functions are provided in Rvachev and Sheiko [23] and Shapiro [26]. On combining set-theoretic Boolean operations with such functions, the inverse problems of semi-analytic geometry (solid modeling) can be solved.

Consider a continuous function $\omega_i : \mathbb{R}^d \rightarrow \mathbb{R}$. Let $\Omega \subset \mathbb{R}^d$ be an open, bounded domain, $\bar{\Omega} = \Omega \cup \partial\Omega$ be the closure of Ω , and define Ω^c to be the complement of Ω ($\Omega \cup \Omega^c = \mathbb{R}^d$). If ω_i is strictly positive in Ω , identically equal to zero on $\partial\Omega$, and strictly negative in $\Omega^c \setminus \partial\Omega$, then it is evident that $F(\omega_i) = \omega_i$ is an R-function. Over the region $\bar{\Omega}$, we associate ω_i with the Boolean 1 (logical true) and over the region Ω^c we associate it with the Boolean 0 (logical false). Note that $\omega_i = 0$ (0 is assumed to be signed) is included in both sets so that it can be assigned to either the set of negative real values or the set of positive real values [30]. Hence, similar to Boolean functions, ω_i is closed under composition. Furthermore, just as Boolean functions are written using the symbols \neg , \vee , and \wedge , which correspond to complement, union and intersection in set theory, every R-function can be written as the composition of the corresponding elementary R-functions: R-negation ($-\omega$), R-disjunction ($\omega_1 \vee \omega_2$), and R-conjunction ($\omega_1 \wedge \omega_2$). On defining R-functions for regions in \mathbb{R}^d , a solid can then be composed using the set-theoretic operations of \neg , \vee , and \wedge . For the universal set $\mathbb{U} = \mathbb{R}^2$, Venn diagrams for some of the operations in set theory are shown in Fig. 1, and the corresponding operations using R-functions are indicated.

The simplest examples of R-functions are the R-disjunction (union) and the R-conjunction (intersection) functions. These are

$$\omega_1 \vee \omega_2 = \frac{\omega_1 + \omega_2}{2} + \frac{\sqrt{(\omega_1 - \omega_2)^2}}{2} = \max(\omega_1, \omega_2), \quad \omega_1 \wedge \omega_2 = \frac{\omega_1 + \omega_2}{2} - \frac{\sqrt{(\omega_1 - \omega_2)^2}}{2} = \min(\omega_1, \omega_2).$$

The generalization of the above R-functions is [30]:

$$R_\alpha(\omega_1, \omega_2) := \frac{1}{1 + \alpha} \left(\omega_1 + \omega_2 \pm \sqrt{\omega_1^2 + \omega_2^2 - 2\alpha\omega_1\omega_2} \right), \quad (2)$$

with (+) and (-) signs defining R-disjunction and R-conjunction, respectively. If ω_1 and ω_2 denote the sides of a triangle, then the triangle inequality is expressed in (2) with $-1 < \alpha < 1$ being the cosine of the angle between the two sides. For $\alpha = 1$, the max and min R-functions are recovered. If ω_1 and ω_2 are positive, then so are $\omega_1 \vee \omega_2$ and $\omega_1 \wedge \omega_2$. The R-functions defined in (2) are not analytic at points where $\omega_1 = \omega_2 = 0$. Smoothness can be obtained by defining the function ($\alpha = 0$ is selected) [30]

$$R_s(\omega_1, \omega_2) := \left[\omega_1 + \omega_2 \pm \sqrt{\omega_1^2 + \omega_2^2} \right] (\omega_1^2 + \omega_2^2)^{\frac{s}{2}}, \quad (3)$$

which renders these functions to be C^s -continuous at all points other than where $\omega_1 = \omega_2 = 0$.

3.1. Normalized functions for line segments and curves

Shapiro and Tsukanov [74] describe the representation of line segments and curves using R-functions and discuss their differential properties. Let us consider one line segment that joins $\mathbf{x}_1 := (x_1, y_1)$ and $\mathbf{x}_2 := (x_2, y_2)$. The center

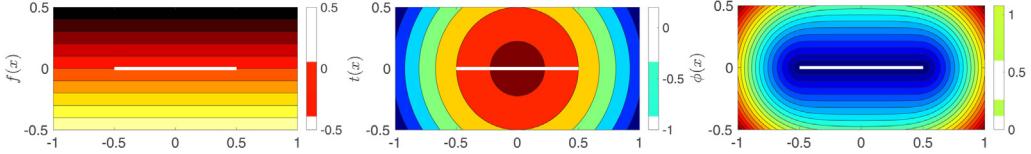


Fig. 2. Construction of the approximate distance function to a line segment. The leftmost plot depicts the signed distance function (4) to a line in \mathbb{R}^2 ; the middle plot shows the trimming function (5); and the rightmost plot displays the approximate distance function (6) to a line segment.

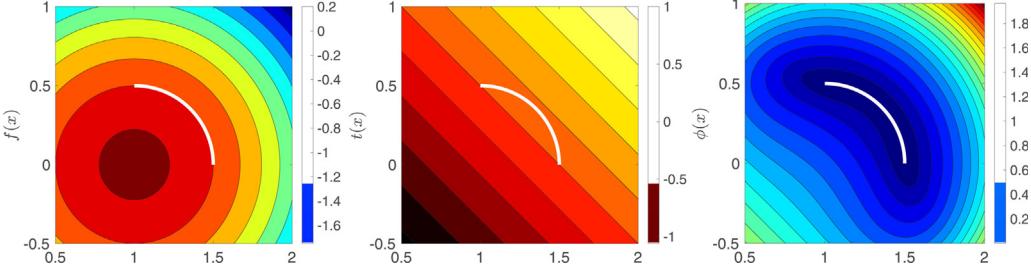


Fig. 3. Construction of the approximate distance function to a quarter-circular arc. The leftmost plot depicts f , the equation of the circular arc (normalized to first order); the trimming function is shown in the middle plot; and the rightmost plot displays the approximate distance function given by (6).

of this segment is denoted by $\mathbf{x}_c := (\mathbf{x}_1 + \mathbf{x}_2)/2$, and the length of the segment is: $L = \|\mathbf{x}_2 - \mathbf{x}_1\|$. Now, we define [25]

$$f := f(\mathbf{x}) = \frac{(x - x_1)(y_2 - y_1) - (y - y_1)(x_2 - x_1)}{L}, \quad (4)$$

which is the signed distance function from point \mathbf{x} to the line that passes through \mathbf{x}_1 and \mathbf{x}_2 .

Since the representation of the segment can be viewed as the intersection of an infinite line with a disk of radius $L/2$, we consider the following *trimming* function that is normalized to first order [25]:

$$t := t(\mathbf{x}) = \frac{1}{L} \left[\left(\frac{L}{2} \right)^2 - \|\mathbf{x} - \mathbf{x}_c\|^2 \right], \quad (5)$$

where $t \geq 0$ defines a disk with center at \mathbf{x}_c . Now, with $f(\mathbf{x})$ and $t(\mathbf{x})$ on-hand, we define a normalized function (up to first order) $\phi(\mathbf{x})$ that is C^2 at all points away from the line segment [25,29]:

$$\phi := \phi(\mathbf{x}) = \sqrt{f^2 + \left(\frac{\varphi - t}{2} \right)^2}, \quad \varphi = \sqrt{t^2 + f^4}, \quad (6)$$

which is an approximation of the distance function to the segment with end points \mathbf{x}_1 and \mathbf{x}_2 . The function ϕ in (6) is a modification of the form $\varphi = |t|$ [74], which has a derivative discontinuity at $t = 0$.

Fig. 2 provides a graphical illustration of f , t and ϕ for a line segment. For a quarter-circular arc, the functions f , t and ϕ are shown in Fig. 3.

In Fig. 4, the approximate distance function (normalized to order 1) to a circle and an ellipse are presented. The ADF to a circle of radius R and center located at $\mathbf{x}_c := (x_c, y_c)$ is given by

$$\phi(\mathbf{x}) = \frac{R^2 - (\mathbf{x} - \mathbf{x}_c) \cdot (\mathbf{x} - \mathbf{x}_c)}{2R}, \quad (7)$$

where $\phi(\mathbf{x})$ is a smooth (bivariate polynomial of degree 2) function. For an elliptical disk whose closure (interior and boundary) is given by the R-function $\omega(\mathbf{x}) \geq 0$, we construct an ADF that is normalized to order 1 using [23]

$$\phi(\mathbf{x}) = \frac{\omega(\mathbf{x})}{\sqrt{\omega^2(\mathbf{x}) + \|\nabla \omega(\mathbf{x})\|^2}}. \quad (8)$$

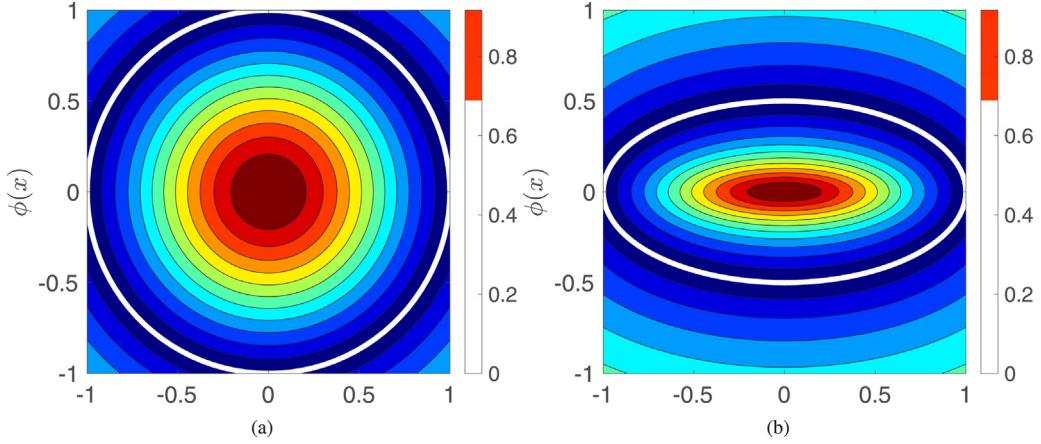


Fig. 4. Approximate distance function (normalized to order 1) to a (a) circle and an (b) ellipse.

In general for curves that are given in parametric form, such as Bézier and non-uniform rational B-spline (NURBS) curves, constructing ADFs require implicitization of the curve. These extensions are discussed and presented in Upreti et al. [75]. For some of the considerations and challenges in the representation (implicit and parametric) of curves in enriched computational methods, see Chin and Sukumar [76].

3.2. R-equivalence operation

Given the normalized distance functions ϕ_1 and ϕ_2 for two curves c_1 and c_2 , a distance field $\phi(\phi_1, \phi_2)$ for the union $c_1 \cup c_2$ must be zero when either $\phi_1 = 0$ or $\phi_2 = 0$ and positive otherwise. When c_1 and c_2 are line segments, the naive formula $\phi(\phi_1, \phi_2) = \phi_1\phi_2$ is no longer normalized at the regular points of the segments. An R-equivalence solution that preserves normalization up to order m of the distance function at all regular points (nonvertices for polygonal curves) is given by [29]:

$$\phi(\phi_1, \phi_2) := \phi_1 \sim \phi_2 = \frac{\phi_1\phi_2}{\sqrt[m]{\phi_1^m + \phi_2^m}} = \frac{1}{\sqrt[m]{\frac{1}{\phi_1^m} + \frac{1}{\phi_2^m}}}, \quad (9)$$

where $\lim_{m \rightarrow \infty} \phi(\phi_1, \phi_2) = \min(\phi_1, \phi_2)$. When ∂S (closed curve) is composed of n pieces, then a ϕ that is normalized up to order m is given by (see the proof in [33]):

$$\phi(\phi_1, \dots, \phi_n) := \phi_1 \sim \phi_2 \sim \dots \sim \phi_n = \frac{1}{\sqrt[m]{\frac{1}{\phi_1^m} + \frac{1}{\phi_2^m} + \dots + \frac{1}{\phi_n^m}}}. \quad (10)$$

The ϕ that is formed in (10) can be viewed as the reciprocal of the L_m -norm of inverse distance measures, which bears similarity to L_m -distance fields [36]. An alternative joining procedure is to consider the R-conjunction given by

$$\phi_s(\phi_1, \phi_2) := \phi_1 \wedge \phi_2 = \phi_1 + \phi_2 - \sqrt[s]{\phi_1^s + \phi_2^s}, \quad (11)$$

which is a function that is normalized to the $(s-1)$ -th order [29]. However, the joining operation is not associative, which makes this choice less desirable. The R-equivalence joining relation in (10) is associative.

The approximate distance function to two line segments is shown in Fig. 5, where the R-conjunction composition in (11) with $s = 2, 3$ and the R-equivalence relation in (9) with $m = 1, 2$ are compared. The bulging phenomenon [73] is noticeable in the vicinity of the joining point. In Fig. 6, we present the approximate distance function to a curved triangular region using R-equivalence for different orders of the normalizing parameter m . As m increases, the ADF approaches the exact distance, which is observed when inspecting the contours in the interior of the curved triangle. The ADFs for a triangle, square, hexagon, and an L-shaped polygon are shown in Fig. 7. The ADF for a square in Fig. 7b bears similarities to a superellipse, $|x|^m + |y|^m = 1$, which has rounded corners as m

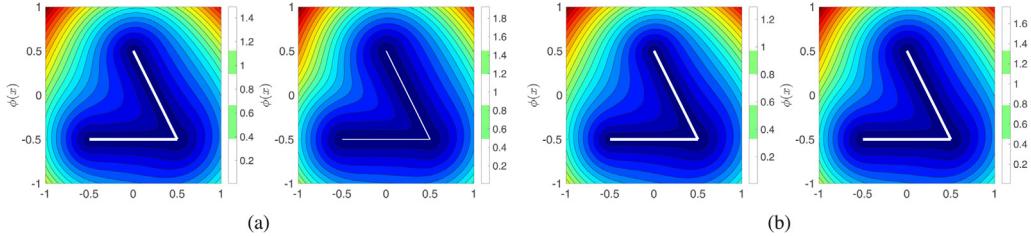


Fig. 5. Approximation of the distance function to two line segments. (a) R-conjunction composition with $s = 2, 3$ in (11), and (b) R-equivalence composition in (9) for the normalization parameter $m = 1, 2$. The ADFs are normalized to order $s - 1$ and m , respectively.

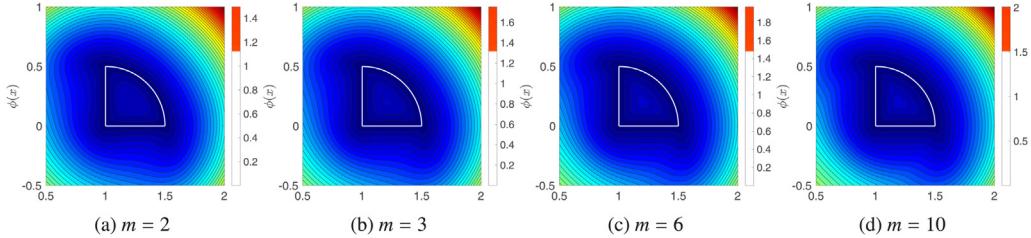


Fig. 6. Plots of the approximate distance function to a curved triangle using R-equivalence for different choices of the normalizing parameter ($m = 2, 3, 6, 10$).

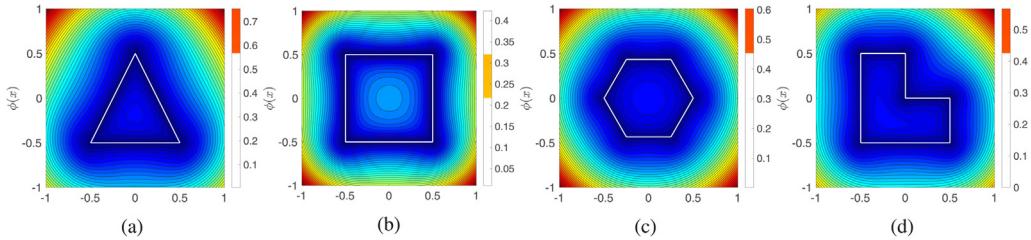


Fig. 7. Plots of the approximate distance function using R-equivalence ($m = 1$) for polygons. (a) triangle, (b) square, (c) regular hexagon, and (d) L-shaped (nonconvex) polygon.

increases. As the last example, we present the ADF for a complex polygonal domain. We consider the polygonalized map of Bhutan,¹ which has 291 boundary vertices. The contour plot for $\phi(\mathbf{x})$ is depicted in Fig. 8, and we observe that the contours are smooth in the interior and well-separated (ϕ is monotonic in Ω). Finally, we mention that a modified form of the R-equivalence relation (10) is also discussed in Biswas and Shapiro [29]. This modified form is constructed with an eye on better capturing the first-order normalization condition in the sector region between two line segments, where the normal is undefined and the closest point to either segment is the common vertex. With increase in m , the R-equivalence joining operation provides a better approximation to the exact distance away from the segments and also improved normalization properties in the vicinity of the segments; however, this comes at the expense of the function being higher order and hence its Laplacian will have greater undulations. Use of $m = 2$ has been adopted in prior computational studies [29,33,75], but herein, we adopt $m = 1$ in most of the numerical simulations that are presented in Sections 9–11.

4. Generalized mean value potentials and L_p -distance fields

In addition to the theory of R-functions, another approach to construct approximate distance fields is via the theory of mean value potential fields. This has connections to generalized barycentric coordinates and in particular to mean value interpolation over polygons and curved domains [77]. Generalized barycentric coordinates [78–80]

¹ Vectorized eps image obtained from <https://freevectormaps.com>.

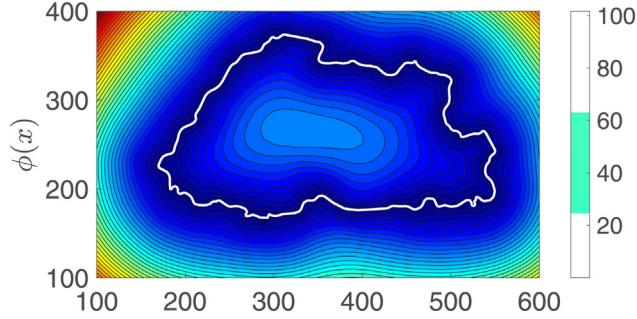


Fig. 8. Plot of the approximate distance function using R-equivalence ($m = 1$) for polygonalized map of Bhutan. The polygon has 291 vertices.

are an extension of barycentric coordinates over simplices to polygons and polyhedra. These coordinates (shape functions) have linear precision and are nonnegative over convex polygons. Transfinite barycentric interpolation over domains bounded by curves is the continuous counterpart of generalized barycentric coordinates over polygons [77]. Given a function $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ that assumes the function $g(\mathbf{x})$ on the boundary (curved) of a domain, a transfinite interpolant provides an approximation of $u(\mathbf{x})$ that matches $g(\mathbf{x})$ over the curved boundary of the domain. For a domain that is bounded by affine or curved boundaries, the reciprocal of the mean value normalization function is a smoothed approximation to the exact distance function [35,81,82], and is a specific instance ($p = 1$) of the reciprocal of a singular double-layer L_p -potential field [36]. We refer to the method that generates these smoothed distance functions by the acronym MVP, since they stem from (*generalized*) *mean value potential fields* [77]. The construction of $\phi(\mathbf{x})$ over polygons and curved domains is presented in Sections 4.1 and 4.2, respectively.

4.1. Approximate distance fields on arbitrary planar polygons

A popular generalized barycentric coordinate is Floater's mean value coordinates [34], which is derived using the circumferential mean value theorem for harmonic functions. This conception stemmed from the objective to approximate a harmonic map by a convex combination map (positive weights) over a triangulation, so that injectivity is preserved. Mean value coordinates have many remarkable properties: for instance, they are valid on arbitrary planar polygons, including nested polygons; C^∞ smooth in Ω with derivative discontinuities only at the vertices of the polygon; they reduce to piecewise affine functions on the edges of a polygon; are nonnegative in the kernel of the polygon; reciprocal of the normalizing weight function is a smoothed ADF; and they also have a smooth extension outside the polygon [81].

Consider the nonconvex polygon (n -gon) shown in Fig. 9, whose n vertices are defined in counterclockwise orientation. The coordinates of the vertices are $\{\mathbf{x}_i\}_{i=1}^n$, and \mathbf{x} is an arbitrary point in the interior of the polygon. The mean value coordinates, $\{\varphi_i(\mathbf{x})\}_{i=1}^n$, are defined as [34]:

$$\varphi_i(\mathbf{x}) = \frac{w_i(\mathbf{x})}{W(\mathbf{x})}, \quad w_i(\mathbf{x}) := \frac{\tan(\alpha_{i-1}/2) + \tan(\alpha_i/2)}{\|\mathbf{x}_i - \mathbf{x}\|}, \quad W(\mathbf{x}) = \sum_{j=1}^n w_j(\mathbf{x}), \quad (12)$$

where the angles α_{i-1} and α_i are shown in Fig. 9. Let $\mathbf{r}_i := \mathbf{x}_i - \mathbf{x}$ with $r_i = \|\mathbf{x}_i - \mathbf{x}\|$ represent the Euclidean distance between \mathbf{x} and \mathbf{x}_i . On noting the half-angle formula for $\tan(\cdot)$, we can define

$$\begin{aligned} t_i &:= \tan\left(\frac{\alpha_i}{2}\right) = \frac{\sin \alpha_i}{1 + \cos \alpha_i} = \frac{r_i r_{i+1} \sin \alpha_i}{r_i r_{i+1} + \mathbf{r}_i \cdot \mathbf{r}_{i+1}} = \frac{\det(\mathbf{r}_i, \mathbf{r}_{i+1})}{r_i r_{i+1} + \mathbf{r}_i \cdot \mathbf{r}_{i+1}}, \\ W(\mathbf{x}) &= \sum_{i=1}^n \left(\frac{1}{r_i} + \frac{1}{r_{i+1}} \right) t_i \quad (r_{n+1} := r_1), \end{aligned} \quad (13)$$

which is now valid for all points \mathbf{x} that are in the interior of a convex or nonconvex polygon. The denominator vanishes when $\alpha_i = \pi$, i.e., when \mathbf{x} lies on the boundary of the polygon, but there $\varphi_i(\mathbf{x})$ are known. The singularity of the weight function on the boundary is a property shared by nonnegative generalized barycentric coordinates.

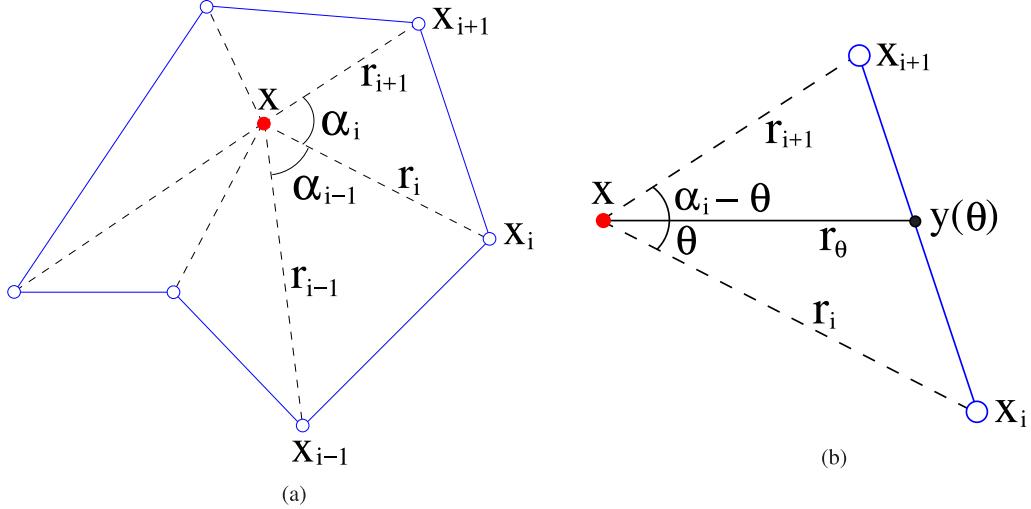


Fig. 9. Notation used in the definition of (a) mean value coordinates [34] and (b) generalized mean value potentials [77]. In (b), the parameters that are used to form $W(x)$ in (13) are shown.

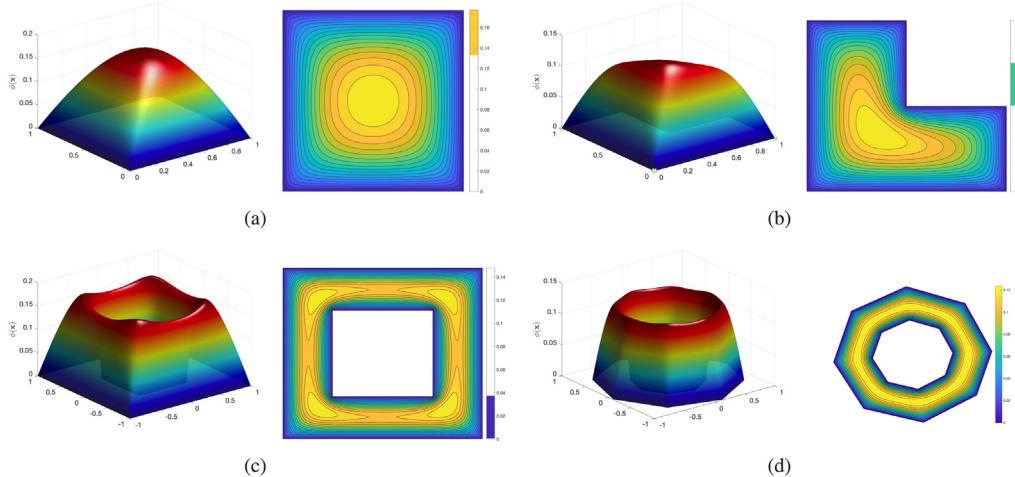


Fig. 10. Approximate distance fields for polygonal domains. (a) square, (b) L-shaped, (c) nested squares, and (d) nested octagons.

For a polygon with one interior (nested) m -gon, the vertices of the inner polygon are defined in clockwise orientation [81]. The contributions of $\{w_i(x)\}_{i=1}^{n+m}$ are used to form $W(x)$ in (12). Hormann and Floater [81] showed that $\phi(x) = 1/W(x)$ is an ADF to the boundary of the polygon, where its normal derivative is 1/2. On taking

$$\phi(x) = \frac{2}{W(x)}, \quad (14)$$

where the scaling factor is the volume of the unit sphere in \mathbb{R}^{d-1} [82] (equal to 2 when $d = 2$), the normal derivative becomes $\partial\phi/\partial\nu = 1$ on $\partial\Omega$. As we discuss in Section 4.2, the $W(x)$ that appears in (13) and (14) is a particular instance ($p = 1$) of the mean value potential field $W_p(x)$. We now have a smooth ADF for a polygon that is normalized to order 1. In Fig. 10, the surface and contour plots of $\phi(x)$ are shown for square and L-shaped domains, as well as nested squares and octagons. The ADF using R-equivalence for the polygonalized map of Bhutan is presented in Fig. 8. In Fig. 11, we show the approximate distance field (surface and contour plots) for the same polygonalized map.

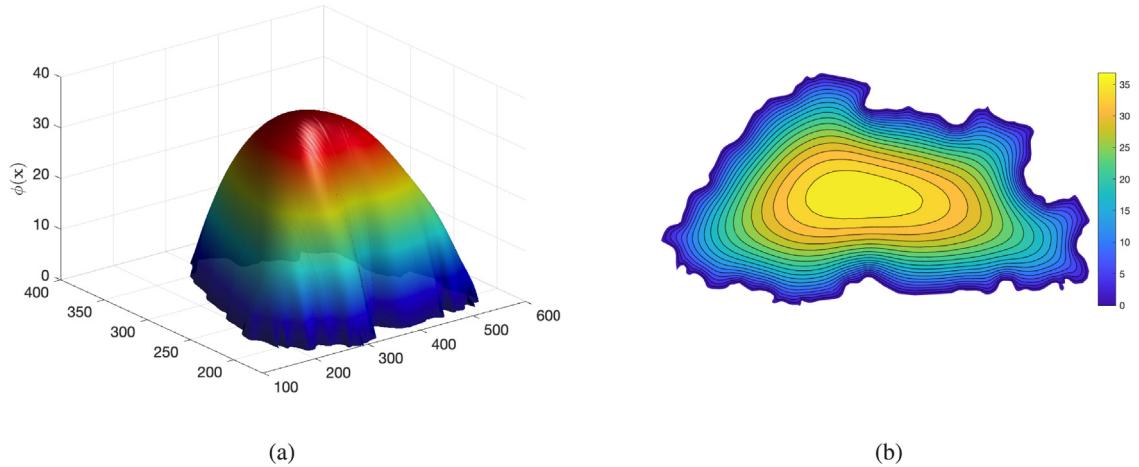


Fig. 11. Approximate distance field for the polygonalized map of Bhutan. Surface plot is shown in (a) and the contour plot in (b). The polygon has 291 vertices.

4.2. Approximate distance fields over curved domains

Consider an open, bounded nonconvex domain Ω with boundary $\Gamma = \partial\Omega$. Given a function $g : \Gamma \rightarrow \mathbb{R}$ that is prescribed on a curved boundary, the transfinite mean value interpolant $u : \Omega \rightarrow \mathbb{R}$ is defined as [35,83]

$$u(\mathbf{x}) = \frac{\int_{S_v} g(\mathbf{y}(\mathbf{x}, \mathbf{v})) K(\mathbf{x}, \mathbf{y}) dS_v}{W(\mathbf{x})}, \quad W(\mathbf{x}) = \int_{S_v} K(\mathbf{x}, \mathbf{v}) dS_v, \quad K(\mathbf{x}, \mathbf{v}) = \frac{1}{\|\mathbf{v} - \mathbf{x}\|}, \quad (15)$$

where $\mathbf{x} \in \Omega \setminus S_v$ and $\mathbf{v} \in S_v$, S_v is the unit circle that is centered at \mathbf{x} , the ray from \mathbf{x} that passes through \mathbf{v} intersects the boundary Γ at \mathbf{y} , and $K(\mathbf{x}, \mathbf{v})$ is a singular kernel function [35,83].

Similar to the behavior of the inverse weight in mean value coordinates on polygons, the function $\phi(\mathbf{x}) = 1/W(\mathbf{x})$ behaves like an approximate distance function to the boundary and its normal derivative on the boundary Γ is $1/2$ [35]. Belyaev et al. [36] introduced L_p -distance fields ($p \geq 1$), which approximates the exact distance function. These distance fields stem from a particular form of a singular double-layer potential, and hence the reference to them as generalized mean value potential fields. Consider the nonconvex domain shown in Fig. 12. The parametrization of the curved boundary $\Gamma : [0, 1] \rightarrow \mathbb{R}$ is $\mathbf{c}(t)$, and its tangent is $\mathbf{c}'(t)$. For a nonconvex domain, a ray from \mathbf{x} intersects the boundary at multiple points $\mathbf{c}(t_i) := \mathbf{y}_i(\mathbf{x}, \mathbf{v})$. On projecting the boundary curve onto the unit circle, an expression for $\phi(\mathbf{x})$ that is valid for convex as well as nonconvex domains is obtained in terms of the curve parameter $t \in [0, 1]$ [83]:

$$\phi(\mathbf{x}) = \left(\frac{1}{W_p(\mathbf{x})} \right)^{1/p}, \quad W_p(\mathbf{x}) = \int_0^1 \frac{(\mathbf{c}(t) - \mathbf{x}) \cdot \mathbf{c}'^\perp(t)}{\|\mathbf{c}(t) - \mathbf{x}\|^{2+p}} dt, \quad (16)$$

where $\mathbf{c}'^\perp(t) := \text{rot}(\mathbf{c}'(t))$ is obtained by rotating $\mathbf{c}'(t)$ through 90° in the clockwise direction. For $\mathbf{x} \in \Omega$, the integral in (16) is numerically integrated; if $\mathbf{x} \in \partial\Omega$ (integral is singular), we set $\phi(\mathbf{x}) = 0$. In (16), $W_p(\mathbf{x})$ is the generalized mean value potential field, which is used to form the approximate distance function $\phi(\mathbf{x})$. Eq. (16) is also applicable for polygons: on choosing $p = 1$ in (16), we recover the $W(\mathbf{x})$ that appears in (13). The approximate distance fields ($p = 1$) for an elliptical disk, annulus, hypocycloid, and a propeller-shaped domain are shown in Fig. 13. The distance function is smooth in the interior of the domain, $\phi \in C^\infty(\Omega)$, and it is C^k on the boundary for a C^k curve (derivative discontinuities occur at the vertices for a polygonal curve). Over curved two-dimensional domains, Dyken and Floater [35] assessed the approximation properties of the transfinite mean value interpolant as well as its use to solve the Poisson equation with web-splines [32], and Chin and Sukumar [84] have used it in verification tests of a cubature rule for numerical integration over curved regions.

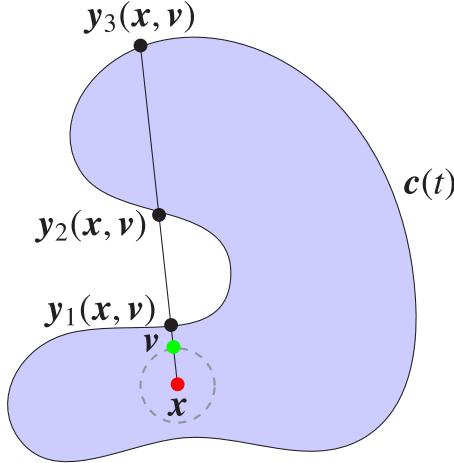


Fig. 12. Nonconvex domain bounded by the curve $c(t)$. The variables that appear in (15) and (16) are shown.

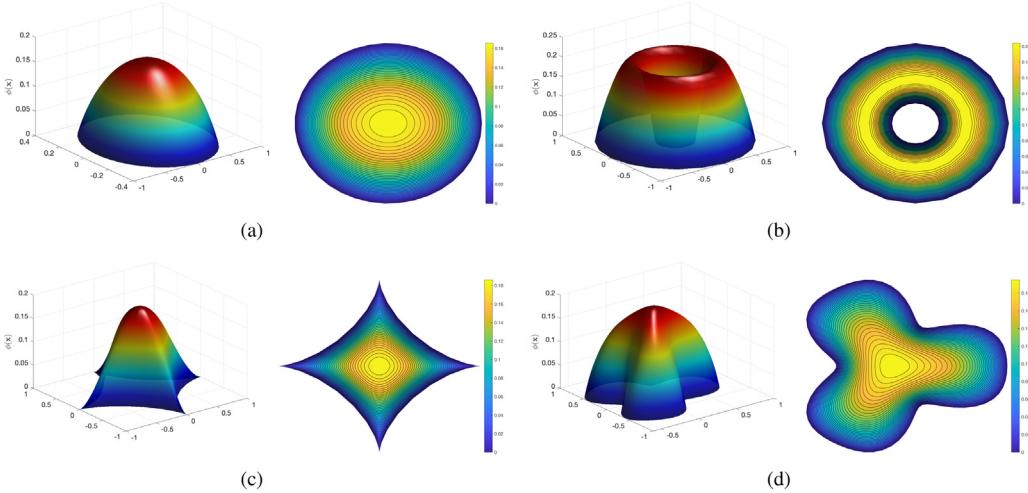


Fig. 13. Approximate distance fields on curved domains using $\phi(\mathbf{x}) = 2/W_1(\mathbf{x})$, with $W_1(\mathbf{x})$ given in (16). Surface and contour plots are shown for an (a) elliptical disk, (b) annulus, (c) hypocycloid, and (d) propeller [76].

5. Imposing boundary conditions in deep neural networks

Let $\tilde{u}_{\text{nn}}^{\text{bc}}(\mathbf{x}; \boldsymbol{\theta})$ denote the PINN trial function. We present the construction of $\tilde{u}_{\text{nn}}^{\text{bc}}(\mathbf{x}; \boldsymbol{\theta})$ so that it exactly satisfies all essential and Robin (natural boundary condition is a particular case) boundary conditions. The trial function includes the contribution from the neural network approximation, $\tilde{u}_{\text{nn}}^R(\mathbf{x}; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ contains the unknown parameters of the network. We defer the presentation of $\tilde{u}_{\text{nn}}^R(\mathbf{x}; \boldsymbol{\theta})$ to Section 6.

5.1. Normalizing functions and solution structures

Let $\Omega \subset \mathbb{R}^2$ be an open, bounded domain with boundary $\partial\Omega$. Consider a smooth function $u(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}$ and let $\phi(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}$ be an approximate distance function to $\partial\Omega$ that is normalized to the m th order (see Section 2). Rvachev defined the normalizers to u of the m th order with respect to $\phi(\mathbf{x})$ via the transformation [23,25]

$$u^*(\mathbf{x}) = u(\mathbf{x} - \phi \nabla \phi), \quad (17)$$

which leads to $u^*(\mathbf{x}) = u(\mathbf{x})$ on $\partial\Omega$ and $\partial^k u^*/\partial v^k = 0$ ($k = 1, 2, \dots, m$) on $\partial\Omega$. Since we are treating Dirichlet and Robin boundary conditions in this paper, we proceed to show that (17) is normalized up to order 1. If $u(\mathbf{x})$

is specified on $\partial\Omega$, then since $\phi = 0$ on $\partial\Omega$ it implies that $u^*(\mathbf{x}) = u(\mathbf{x})$ on $\partial\Omega$, which establishes zeroth-order normalization. The proof for first-order normalization follows.

Proof. Let $\mathbf{t} := \mathbf{x} - \phi\nabla\phi$. Then, by the chain rule we have

$$\nabla u^* = \nabla_t u^* \cdot \nabla \otimes \mathbf{t} = \nabla_t u^* \cdot [\mathbf{I} - \phi\nabla \otimes \nabla\phi - \nabla\phi \otimes \nabla\phi],$$

where \otimes is the dyadic (tensor) product and $\nabla_t(\cdot)$ is the gradient with respect to \mathbf{t} . On $\partial\Omega$, $\phi = 0$, $\mathbf{t} = \mathbf{x}$ and $\nabla\phi = \mathbf{v}$, since ϕ is normalized to the first order. Hence, we can write

$$\frac{\partial u^*}{\partial \nu} \Big|_{\partial\Omega} = [\nabla u^* \cdot \mathbf{v}]_{\partial\Omega} = [\nabla u^*]_{\partial\Omega} \cdot \left([\mathbf{I} - \phi\nabla \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{v}] \cdot \mathbf{v} \right)_{\partial\Omega} = [\nabla u^*]_{\partial\Omega} \cdot (\mathbf{v} - \mathbf{v}) = 0,$$

which is the desired result. \square

Note that this result also holds if we consider the unit outward normal vector, $\mathbf{n} = -\mathbf{v}$. By extension, if ϕ is normalized to the m th order, one can establish that u^* in (17) is normalized to the m th order, i.e.,

$$u^*(\mathbf{x}) = u(\mathbf{x}) \text{ on } \partial\Omega, \quad \frac{\partial^k u^*}{\partial \nu^k} = 0 \text{ on } \partial\Omega \quad (k = 1, 2, \dots, m). \quad (18)$$

Let u and its higher order normal derivatives along \mathbf{v} be prescribed on $\partial\Omega$, i.e.,

$$u(\mathbf{x}) = u_0(\mathbf{x}) \text{ on } \partial\Omega, \quad \frac{\partial^k u}{\partial \nu^k} = u_k(\mathbf{x}) \text{ on } \partial\Omega \quad (k = 1, 2, \dots, m). \quad (19)$$

Then, one can represent u in the vicinity of $\partial\Omega$ and in the direction (inward) normal to the boundary using a polynomial Taylor series expansion of u in terms of ϕ . Rvachev et al. [23,24] referred to this as a generalized Taylor series expansion, which takes the form:

$$u(\mathbf{x}) = u_0^*(\mathbf{x}) + \sum_{k=1}^m \frac{u_k^*(\mathbf{x})}{k!} \phi^k(\mathbf{x}) + \phi^{m+1}(\mathbf{x}) \Psi(\mathbf{x}), \quad (20)$$

where $u_k^*(\mathbf{x}) = u_k(\mathbf{x} - \phi\nabla\phi)$ and $\Psi(\mathbf{x})$ is an unknown function (approximation) in the remainder term. This equation resembles univariate Taylor series expansion about $x = 0$, where instead of evaluating derivatives that are scalar constants, $u_k^*(\mathbf{x})$ in (20) are scalar-valued functions. Eq. (20) is the general form of the *solution structure* for u that Rvachev introduced [23,24]. If a function $u \in C^k(\Omega)$, and its first m derivatives vanish on $\partial\Omega$, then the solution structure $\phi^{m+1}\Psi$ is sufficiently complete in the Hilbert space $H^k(\Omega)$ to approximate u and all its derivatives up to order k [24]. On using (20), the boundary conditions on $\partial\Omega$ that are given in (19) are exactly met. We now present the solution structure for three distinct sets of boundary conditions (Dirichlet, Neumann and Robin) that are imposed on $\partial\Omega$.

5.1.1. Solution structure for Dirichlet boundary condition

If $u = g$ is prescribed on $\partial\Omega$, then using $u_0^*(\mathbf{x}) = g(\mathbf{x} - \phi\nabla\phi)$ and $m = 0$ in (20), we can write:

$$\begin{aligned} u(\mathbf{x}) &= g(\mathbf{x} - \phi\nabla\phi) + \phi(\mathbf{x})\Psi(\mathbf{x}) = g(\mathbf{x}) - \phi(\mathbf{x}) \frac{\partial g(\mathbf{x})}{\partial \nu} \Big|_{\partial\Omega} + \phi(\mathbf{x})\Psi(\mathbf{x}) \\ &= g(\mathbf{x}) + \phi(\mathbf{x}) \frac{\partial g(\mathbf{x})}{\partial n} \Big|_{\partial\Omega} + \phi(\mathbf{x})\Psi(\mathbf{x}) \\ &= g(\mathbf{x}) + \phi(\mathbf{x})\tilde{u}, \end{aligned} \quad (21)$$

since $\mathbf{n} = -\mathbf{v}$ on $\partial\Omega$, and a first order linearization of $g(\cdot)$ is used. Therefore, the solution structure is:

$$u = g + \phi\tilde{u}, \quad (22)$$

where \tilde{u} is any suitable numerical approximation. So any trial function of the form given in (22) will exactly satisfy the essential boundary condition $u = g$ on $\partial\Omega$. In this paper, we use deep neural networks to construct \tilde{u} . For the homogeneous Dirichlet problem ($g = 0$) using Kantorovich's method, Babuška et al. [19] provide an a priori error estimate in the H^1 norm.

5.1.2. Solution structure for Neumann boundary condition

Since Neumann and higher order boundary conditions for PDEs are imposed in the direction of \mathbf{n} , the unit outward normal vector to $\partial\Omega$, we consider solution structures that are defined with respect to \mathbf{n} . This is a departure from the literature on the R-function method to solve PDEs, but is aligned with how boundary-value problems are posed.

For first order derivatives, linearizing $u^*(\mathbf{x}) := u(\mathbf{x} - \phi \nabla \phi)$ in the neighborhood of the boundary ($\phi = 0$) amounts to subtracting the variation in the normal direction \mathbf{v} , which leads to [23]

$$\begin{aligned} u^*(\mathbf{x}) &= u(\mathbf{x}) - [\phi(\mathbf{x}) \nabla u(\mathbf{x}) \cdot \mathbf{v}]_{\partial\Omega} + \phi^2(\mathbf{x}) \Psi(\mathbf{x}) = u(\mathbf{x}) - [\phi(\mathbf{x}) \nabla \phi(\mathbf{x}) \cdot \nabla u(\mathbf{x})]_{\partial\Omega} + \phi^2(\mathbf{x}) \Psi(\mathbf{x}) \\ &= (1 + \phi D_1^\phi)(u) + \phi^2 \Psi, \quad D_1^\phi(\cdot) := [-\nabla \phi \cdot \nabla(\cdot)]_{\partial\Omega} = \frac{\partial(\cdot)}{\partial n} \Big|_{\partial\Omega}, \end{aligned} \quad (23)$$

where $\mathbf{v} = -\mathbf{n} = \nabla \phi$ on $\partial\Omega$ and $D_1^\phi(\cdot)$ is a differential operator that acts in the outward normal direction to the boundary.

If $\partial u / \partial n = h$ is prescribed on $\partial\Omega$, then using (20) and (23), we can write

$$\begin{aligned} u(\mathbf{x}) &= u_0(\mathbf{x} - \phi(\mathbf{x}) \nabla \phi(\mathbf{x})) + \phi(\mathbf{x}) u_1(\mathbf{x} - \phi(\mathbf{x}) \nabla \phi(\mathbf{x})) + \phi^2(\mathbf{x}) \Psi_1(\mathbf{x}) \\ &= [1 + \phi D_1^\phi](u_0) + \phi u_1 + \phi^2 \Psi(\mathbf{x}) = [1 + \phi D_1^\phi](u_0) - \phi h + \phi^2 \Psi, \end{aligned}$$

since $\partial u / \partial \mathbf{v} = u_1 = -h$ on $\partial\Omega$, and therefore the most general solution structure can be written as:

$$u = [1 + \phi D_1^\phi](\tilde{u}_1) - \phi h + \phi^2 \tilde{u}_2, \quad (24)$$

where \tilde{u}_1 and \tilde{u}_2 are arbitrary approximation functions.

5.1.3. Solution structure for Robin boundary condition

If the Robin boundary condition, $\partial u / \partial n + cu = h$, is prescribed on $\partial\Omega$ ($c := c(\mathbf{x})$, $h := h(\mathbf{x})$), then following similar steps to that taken to obtain (22) and (24), we can write the most general solution structure for this boundary condition as [23]:

$$u = [1 + \phi(c + D_1^\phi)](\tilde{u}_1) - \phi h + \phi^2 \tilde{u}_2. \quad (25)$$

It is readily verified that the ansatz in (25) satisfies the boundary condition $\partial u / \partial n + cu = h$ on $\partial\Omega$.

5.2. Imposing inhomogeneous essential boundary conditions

Let us extend our analysis to the case when different inhomogeneous essential boundary conditions are imposed on distinct subsets of $\partial\Omega$. Now, let the boundary $\partial\Omega := \Gamma = \cup_{i=1}^N \Gamma_i$. The inhomogeneous essential boundary condition $u = g_i$ is imposed on Γ_i ($i = 1, 2, \dots, M$), and on Γ_i ($i = M+1, M+2, \dots, N$) we assume natural boundary conditions are imposed through the potential energy functional in the variational principle. Let ϕ_i be the ADF that is associated with Γ_i ($i = 1, 2, \dots, M$), and let ϕ be the approximate distance field that is composed either via R-equivalence using $\phi_1 \sim \phi_2 \dots \phi_M$ in (10) or using the mean value potential field, $W(\mathbf{x})$, given in (14) and (16). Transfinite interpolation is the generalization of scattered data interpolation to interpolation of functions over curves and surfaces. On using the singular inverse-distance based Shepard weight function [85], we can write the transfinite interpolant as [23,25]

$$g(\mathbf{x}) = \sum_{i=1}^M w_i(\mathbf{x}) g_i(\mathbf{x}), \quad w_i(\mathbf{x}) = \frac{\phi_i^{-\mu_i}}{\sum_{j=1}^M \phi_j^{-\mu_i}} = \frac{\prod_{j=1; j \neq i}^M \phi_j^{\mu_j}}{\sum_{k=1}^M \prod_{j=1; j \neq k}^M \phi_j^{\mu_j}}, \quad (26)$$

where the weights w_i form a partition-of-unity, and (26) interpolates g_i on the set Γ_i . In (26), $\mu_i \geq 1$ is a constant that controls the nature of interpolation that accrues. For $\mu_i = 1$, the function g_i is interpolated on Γ_i , whereas if $\mu_i = 2$, both g_i and $\partial g_i / \partial n$ are interpolated on Γ_i . Now, on using the solution structure for the Dirichlet problem given in (22) and following the work of Rvachev et al. [23,24], we can write the trial function in the deep Ritz method as:

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta}) = g(\mathbf{x}) + \phi(\mathbf{x}) \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}), \quad (27)$$

where ϕ can also be replaced by the product $\prod_{i=1}^M \phi_i$. Since $g(\mathbf{x}) = g_i(\mathbf{x})$ on Γ_i and $\phi(\mathbf{x})$ vanishes on $\cup_{i=1}^M \Gamma_i$, kinematic admissibility of (27) is verified.

5.3. Imposing inhomogeneous essential and Robin boundary conditions

Let us consider the case that the boundary $\partial\Omega := \Gamma = \Gamma_1 \cup \Gamma_2$ with $\Gamma_1 \cap \Gamma_2 = \emptyset$. The boundary conditions on Γ_1 and Γ_2 are:

$$u = g \text{ on } \Gamma_1, \quad \frac{\partial u}{\partial n} + cu = h \text{ on } \Gamma_2. \quad (28)$$

Let $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{x})$ be the approximate distance functions to the boundaries Γ_1 and Γ_2 , respectively. We use the R-equivalence ($m = 1$) relation in (9) to form

$$\phi(\mathbf{x}) = \frac{\phi_1(\mathbf{x})\phi_2(\mathbf{x})}{\phi_1(\mathbf{x}) + \phi_2(\mathbf{x})}, \quad (29)$$

which is the ADF to the boundary Γ . For the case when the boundary is partitioned into two disjoint sets, we consider two approaches to form a trial function. The first approach uses superposition of two solution structures and has a simple form. The second method, which is based on transfinite interpolation, is applicable in general when there are multiple boundaries on which essential and Robin boundary conditions are imposed. In this paper, we apply the first approach to solve a one-dimensional problem with mixed (Dirichlet and Neumann) boundary conditions, and adopt the second approach to solve a problem with Dirichlet and Robin boundary conditions in two dimensions.

Approach I: We form solutions structures u_1 and u_2 such that

$$u_1 = g \text{ on } \Gamma_1, \quad \frac{\partial u_1}{\partial n} + u_1 = 0 \text{ on } \Gamma_2, \quad u_2 = 0 \text{ on } \Gamma_1, \quad \frac{\partial u_2}{\partial n} + u_2 = h \text{ on } \Gamma_2, \quad (30)$$

and therefore the desired trial function is: $u(\mathbf{x}) = u_1(\mathbf{x}) + u_2(\mathbf{x})$. Now, on using (22), we know that the function $g + \phi_1 \tilde{u}_1$ satisfies the essential boundary condition on Γ_1 but does not meet the Robin boundary condition on Γ_2 . To satisfy the Robin boundary condition on Γ_2 , we normalize $u_1 + \phi \tilde{u}_1$ to the first order using $h = 0$ in (25) to obtain

$$u_1 = \left[1 + \phi(c + D_1^{\phi_2}) \right] (g + \phi_1 \tilde{u}_1) = \left[1 + \phi(D_1^{\phi_2}) \right] (g + \phi_1 \tilde{u}_1) + c\phi g + c\phi_1 \phi \tilde{u}_1, \quad (31a)$$

where

$$D_1^{\phi_2}(\cdot) = [-\nabla \phi_2 \cdot \nabla (\cdot)]_{\Gamma_2} = \left. \frac{\partial(\cdot)}{\partial n} \right|_{\Gamma_2} \quad (31b)$$

is the differential operator that acts in the outward normal direction on the boundary Γ_2 . Similarly, using (25), the minimal structure for u_2 is:

$$u_2 = \phi(\phi_2 \tilde{u}_2 - h). \quad (32)$$

Since $\phi = 0$ on Γ and $\phi_2 = 0$ on Γ_2 , the conditions on u_2 in (30) are satisfied. On choosing $\tilde{u}_1 = \tilde{u}_2 := \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})$ in (31) and (32) and adding them up, the ansatz $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ is:

$$\begin{aligned} \tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta}) &= \phi_1(\mathbf{x}) \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}) \\ &+ \phi(\mathbf{x}) \left[\left\{ \phi_2(\mathbf{x}) + c(\mathbf{x})\phi_1(\mathbf{x}) \right\} \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}) + D_1^{\phi_2} \left(\phi_1(\mathbf{x}) \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}) \right) \right. \\ &\quad \left. + D_1^{\phi_2}(g(\mathbf{x})) + c(\mathbf{x})g(\mathbf{x}) - h(\mathbf{x}) \right] + g(\mathbf{x}), \end{aligned} \quad (33)$$

where $\phi(\mathbf{x})$ is given in (29). For mixed (Dirichlet and Robin) boundary conditions, the form of the trial function in (33) appears in Rvachev and Sheiko [23].

Approach II: On using (22) and (25), we select the boundary functions u_1 and u_2 on Γ_1 and Γ_2 , respectively, as:

$$u_1 = g, \quad (34a)$$

$$u_2 = \left[1 + \phi_2(c + D_1^{\phi_2}) \right] (\tilde{u}) - \phi_2 h, \quad (34b)$$

with $\phi_1\phi_2^2\tilde{u}$ as the composite remainder term. From (26), we recall the transfinite interpolant, where we now choose $\mu_1 = 1$ and $\mu_2 = 2$. On carrying out a few algebraic simplifications and using (34) with $\tilde{u} := \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})$, we can write the trial function $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ that exactly imposes the mixed boundary conditions as:

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta}) = w_1(\mathbf{x})u_2(\mathbf{x}; \boldsymbol{\theta}) + w_2(\mathbf{x})u_1(\mathbf{x}) + \phi_1(\mathbf{x})\phi_2^2(\mathbf{x})\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}), \quad (35a)$$

where

$$w_1(\mathbf{x}) = \frac{\phi_1(\mathbf{x})}{\phi_1(\mathbf{x}) + \phi_2^2(\mathbf{x})}, \quad w_2(\mathbf{x}) = \frac{\phi_2^2(\mathbf{x})}{\phi_1(\mathbf{x}) + \phi_2^2(\mathbf{x})}, \quad (35b)$$

and

$$u_1(\mathbf{x}) = g(\mathbf{x}), \quad u_2(\mathbf{x}; \boldsymbol{\theta}) = \left[1 + \phi_2(\mathbf{x})(c(\mathbf{x}) + D_1^{\phi_2}) \right] (\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})) - \phi_2(\mathbf{x})h(\mathbf{x}). \quad (35c)$$

6. Approximation of trial functions in a deep neural network

In this paper, we exclusively use the densely connected neural network architecture, also known as the multi-layer perceptron (MLP), which has its origin in the early works of Rosenblatt [86]. MLPs consist of multiple layers of neurons, where each neuron has the task of converting its input to an output by generally passing it through a nonlinear function called activation. MLPs are characterized by an architecture where neurons in a given layer are connected densely to the neurons in the neighboring layers (Fig. 14). We note in passing, however, that the latest revolution in deep learning began with a different neural network architecture – convolutional neural networks (CNN) – applied to image classification tasks [87]. Independent of research in PINNs, some of these modern architectures have also been applied to mechanics problems [88]. In keeping with the broadly accepted definition, we consider any deep network to have two or more hidden layers.

In a standard collocation or Ritz method, the trial function is expanded as a linear combination of known basis functions. The point of departure in using deep neural networks is that the ansatz herein is represented by a nonlinear map that consists of unknown parameters. These parameters are obtained via the solution of a minimization problem, which in general but not necessarily, is a least squares optimization problem. Once the parameters are determined, one obtains the numerical solution to the boundary-value problem.

6.1. Feedforward neural network

Given a point $\mathbf{x} \in \mathbb{R}^d$, we use a multilayer feedforward deep neural network to construct $\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})$, which is then used to build $\tilde{u}_{nn}^{bc}(\mathbf{x})$, the approximation to $u(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$. The layers in between the input and output layer are known as hidden layers. Each hidden layer consists of neurons (hidden units), and each neuron in a hidden layer takes its input from the neurons in the preceding layer and computes its own activation. A network diagram of neural networks with one, two, and four hidden layers is shown in Fig. 14.

In this paper, we consider the following boundary-value problem:

$$\mathcal{L}u(\mathbf{x}) = f(\mathbf{x}) \text{ in } \Omega \subset \mathbb{R}^d, \quad (36a)$$

$$\mathcal{B}_u u(\mathbf{x}) = g(\mathbf{x}) \text{ on } \Gamma_u, \quad (36b)$$

$$\mathcal{B}_n u(\mathbf{x}) = h(\mathbf{x}) \text{ on } \Gamma_n, \quad (36c)$$

where \mathcal{L} is in general a differential operator plus the identity and d is the spatial dimension. In (36), Ω is an open bounded domain with boundary $\partial\Omega = \Gamma_u \cup \Gamma_n$ and $\Gamma_u \cap \Gamma_n = \emptyset$. Eq. (36b) represents essential boundary conditions for second- and fourth-order problems, and (36c) imposes natural and higher order boundary conditions. For a second-order problem, (36c) is a Robin boundary condition. As noted in Section 5, we form the trial function to approximate $u(\mathbf{x})$ as a combination of terms that involve the approximate distance functions ($\phi_k(\mathbf{x}), \phi(\mathbf{x})$) and the neural network approximation. Consider a neural network that consists of L hidden layers with \mathcal{N}_ℓ neurons in the hidden layer ℓ and activation function $\sigma : \mathbb{R} \rightarrow \mathbb{R}$. The size of the neural network is: $\mathcal{N} = \sum_{\ell=1}^L \mathcal{N}_\ell$. Let $\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})$ be the PINN approximation, where $\boldsymbol{\theta} := \{\mathbf{W}, \mathbf{b}\}$ is the unknown parameter vector, with weights $\mathbf{W}_\ell \in \mathbb{R}^{\mathcal{N}_\ell \times \mathcal{N}_{\ell-1}}$ and biases $\mathbf{b}_\ell \in \mathbb{R}^{\mathcal{N}_\ell}$. We write $\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})$ via the composition of $T^{(\ell)}$ ($\ell = 1, 2, \dots, L$) and a linear map \mathcal{G} as:

$$\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}) = \mathcal{G} \circ T^{(L)} \circ T^{(L-1)} \circ \dots \circ T^{(1)}(\mathbf{x}), \quad (37)$$

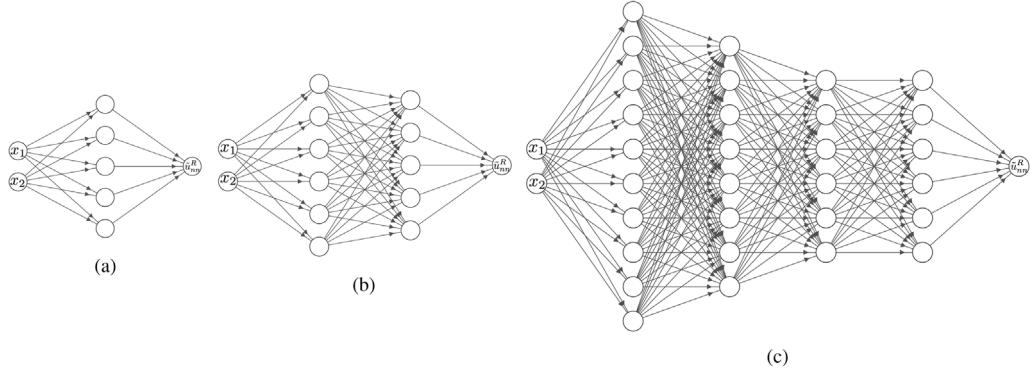


Fig. 14. Deep neural network with (a) one, (b) two, and (c) four hidden layers [90]. Input layer is a point $\mathbf{x} \in \mathbb{R}^2$, and the output layer is the PINN approximation, $\tilde{u}_{nn}^R(\mathbf{x}; \theta)$.

where $\mathcal{G} : \mathbb{R}^{\mathcal{N}_L} \rightarrow \mathbb{R}$ is the linear mapping for the output layer and in each hidden layer ($\ell = 1, 2, \dots, L$), the nonlinear mapping is:

$$T^{(\ell)}(\mathbf{z}) = \sigma(\mathbf{W}_\ell \cdot \mathbf{z} + \mathbf{b}_\ell), \quad (38)$$

where $\mathbf{z} \in \mathbb{R}^{\mathcal{N}_{\ell-1}}$. For a neural network with activation function σ and a single hidden layer that consists of \mathcal{N} neurons, the PINN approximation is:

$$\tilde{u}_{nn}^R(\mathbf{x}; \theta) = \sum_{i=1}^{\mathcal{N}} c_i \sigma(\mathbf{w}_i \cdot \mathbf{x} + b_i), \quad (39)$$

where $\mathbf{w}_i \in \mathbb{R}^d$ and $b_i, c_i \in \mathbb{R}$. It is known that a multilayer neural network is sufficiently rich to be able to approximate any L^2 function to arbitrary accuracy [66,67,89]. However, realization of this in practice hinges on the width and depth of the network, choice of the activation function and the computational algorithm to solve the optimization problem.

6.2. Backpropagation algorithm

Determination of the optimal parameters of the network is done through an iterative optimization process called backpropagation and a popular algorithm for backpropagation is stochastic gradient descent, which is the stochastic version of the gradient descent algorithm. An important step in this procedure is the efficient computation of the gradient of the loss function using automatic differentiation. In this paper, we exclusively use the Adam backpropagation algorithm whose details are given in [91]. Adam is an extension of the stochastic gradient descent algorithm and differs from it primarily in its implementation of per-parameter learning rates that are continuously tweaked during learning through both the first and second moments of the gradients.

7. Formulations

We now present the formulations for deep collocation and deep Ritz for second- and fourth-order problems. In the collocation approach, all boundary conditions are exactly satisfied; for the Ritz method, the essential boundary conditions are met.

7.1. Deep collocation

Let us consider a second-order boundary-value problem with mixed boundary conditions. We require the trial function given in (35) (meets all boundary conditions) to also satisfy the governing equation in (36a) at N_I interior

collocation points. We label these collocation points as $\{\mathbf{x}_k\}_{k=1}^{N_I}$. When substituted in (36a), this defines a residual error at each point \mathbf{x}_k and to determine the parameter $\boldsymbol{\theta}$, we minimize the mean squared residual error:

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} L_{\text{nn}}^{\text{bc}}(\boldsymbol{\theta}), \quad L_{\text{nn}}^{\text{bc}}(\boldsymbol{\theta}) := \|\mathcal{L}\tilde{u}_{\text{nn}}^{\text{bc}}(\mathbf{x}; \boldsymbol{\theta}) - f(\mathbf{x})\|_{\Omega, N_I}^2 = \frac{1}{N_I} \sum_{k=1}^{N_I} [\mathcal{L}\tilde{u}_{\text{nn}}^{\text{bc}}(\mathbf{x}_k; \boldsymbol{\theta}) - f(\mathbf{x}_k)]^2, \quad (40)$$

where $L_{\text{nn}}^{\text{bc}}(\boldsymbol{\theta})$ is known as the loss function, and $\|\cdot\|_{\Omega, N_I}$ denotes the mean discrete L_2 norm of its argument over the domain Ω that is discretized using N_I collocation points. The Adam algorithm [91] is used to solve (40).

When a standard PINN trial function, $\tilde{u}_{\text{nn}}(\mathbf{x}; \boldsymbol{\theta})$, is used, additional residual error contributions from the boundary conditions are present in the loss function. If only essential boundary conditions are imposed, with $u = g$ on $\partial\Omega$, then the solution for the parameter $\boldsymbol{\theta}$ is given by

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} L_{\text{nn}}(\boldsymbol{\theta}), \quad L_{\text{nn}}(\boldsymbol{\theta}) := \|\mathcal{L}\tilde{u}_{\text{nn}}(\mathbf{x}; \boldsymbol{\theta}) - f(\mathbf{x})\|_{\Omega, N_I}^2 + \|\tilde{u}_{\text{nn}}(\mathbf{x}; \boldsymbol{\theta}) - g(\mathbf{x})\|_{\partial\Omega, N_B}^2, \quad (41)$$

where $\|\cdot\|_{\partial\Omega, N_B}$ is the mean discrete L_2 norm of its argument over the boundary $\partial\Omega$, and N_B is the number of collocation points on $\partial\Omega$. As noted in Section 1, the first and second terms in $L_{\text{nn}}(\boldsymbol{\theta})$ are referred to as the interior (PDE) loss and the boundary loss, respectively.

7.2. Deep Ritz

E and Yu [9] introduced the deep Ritz method to solve low- and high-order boundary-value problems that have a variational structure. Samaniego et al. [92] have applied the Ritz approach to solve problems in computational solid mechanics. We consider second-order (Poisson) and fourth-order (plate bending) boundary-value problems. Essential and mixed boundary conditions are considered for the Poisson problem and clamped boundary conditions are imposed for the plate bending problem. We use a variational principle (minimization of the potential energy functional) to solve both problems. A trial function, $\tilde{u}_{\text{nn}}^{\text{bc}}(\mathbf{x}; \boldsymbol{\theta})$, from a finite-dimensional space is used in the variational principle, which forms the basis of the deep Ritz method.

7.2.1. Second-order problems

Referring to the model boundary-value problem in (36), we first consider a Poisson problem with Dirichlet boundary conditions:

$$-\nabla^2 u = f \quad \text{in } \Omega, \quad (42a)$$

$$u = g \quad \text{on } \partial\Omega. \quad (42b)$$

The variational principle for this problem is:

$$\min_{u \in \mathcal{S}} \left[\Pi[u] := \underbrace{\frac{1}{2} a(u, u)}_{W_{\text{int}}[u]} - \underbrace{\ell(u)}_{W_{\text{ext}}[u]}, \quad \mathcal{S} = \left\{ u : u \in H^1(\Omega), \quad u = g \text{ on } \partial\Omega \right\} \right], \quad (43a)$$

where

$$a(u, w) = \int_{\Omega} \nabla u \cdot \nabla w \, dx, \quad \ell(w) = \int_{\Omega} f w \, dx \quad (43b)$$

are a symmetric bilinear functional and a linear functional, respectively, and u and w are the trial and test functions, respectively. In (43a), $W_{\text{int}}[u]$ is the internal work (strain energy) and $W_{\text{ext}}[u]$ is the external work, and $H^k(\Omega)$ is the Sobolev space that consists of functions that have square integrable derivatives up to order k in Ω .

As a second problem, we consider a Poisson problem with mixed (Dirichlet and Robin) boundary conditions:

$$-\nabla^2 u = f \quad \text{in } \Omega, \quad (44a)$$

$$u = g \quad \text{on } \Gamma_u, \quad \frac{\partial u}{\partial n} + cu = h \quad \text{on } \Gamma_n, \quad (44b)$$

where \mathbf{n} is the unit outward normal on Γ_n , $c := c(\mathbf{x})$ and $h := h(\mathbf{x})$ are boundary data, and $\Gamma = \partial\Omega$ with $\Gamma = \Gamma_u \cup \Gamma_n$ and $\Gamma_u \cap \Gamma_n = \emptyset$. The variational principle for this problem is:

$$\min_{u \in \mathcal{S}} \left[\Pi[u] := \frac{1}{2} a(u, u) - \ell(u), \quad \mathcal{S} = \left\{ u : u \in H^1(\Omega), \quad u = g \text{ on } \partial\Omega \right\} \right], \quad (45a)$$

where

$$a(u, w) = \int_{\Omega} \nabla u \cdot \nabla w \, dx + \int_{\Gamma_n} cuw \, dS, \quad \ell(w) = \int_{\Omega} fw \, dx + \int_{\Gamma_n} hw \, dS. \quad (45b)$$

As indicated in (27), we choose a kinematically admissible trial function so that the essential boundary condition is satisfied. On substituting the PINN trial function in (43) or (45), the finite-dimensional minimization problem becomes

$$\min_{\theta} \left[\Pi[\tilde{u}_{nn}^{bc}] := \frac{1}{2} a(\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta), \tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)) - \ell(\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)) \right]. \quad (46)$$

To compute the potential energy functional, we use a Monte Carlo integration rule with points that are distributed (randomly or quasi-uniformly) in the domain with a constant weight that is attached to each point. To obtain the unknown parameters, we solve a discrete nonlinear minimization problem that is posed as:

$$\boldsymbol{\theta}^* = \arg \min_{\theta} L_{nn}^{bc}(\theta), \quad L_{nn}^{bc}(\theta) = \frac{1}{N_I} \sum_{k=1}^{N_I} \left[\frac{1}{2} a(\tilde{u}_{nn}^{bc}(\mathbf{x}_k; \theta), \tilde{u}_{nn}^{bc}(\mathbf{x}_k; \theta)) - \ell(\tilde{u}_{nn}^{bc}(\mathbf{x}_k; \theta)) \right], \quad (47)$$

where $L_{nn}^{bc}(\theta)$ is the loss function, and $a(\cdot, \cdot)$ and $\ell(\cdot)$ are given in (43b) and (45b) for the Poisson problems with Dirichlet and mixed boundary conditions, respectively. In general, the stochastic gradient descent algorithm or a variant of it is used to solve (47).

7.2.2. Fourth-order problem

We consider the fourth-order problem of Kirchhoff plate bending with clamped boundary conditions. The strong form of the boundary-value problem is:

$$\nabla^4 u = f \quad \text{in } \Omega \subset \mathbb{R}^2, \quad (48a)$$

$$u = 0 \quad \text{on } \partial\Omega, \quad u_{,n} := \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial\Omega, \quad (48b)$$

where $u := u(x, y)$ is the out-of-plane plate deflection, f is the transverse load per unit area, and \mathbf{n} is the unit outward normal on the boundary $\partial\Omega$. The flexural rigidity of the plate is assumed to be unity.

The variational principle that is associated with the strong form in (48) is:

$$\min_{u \in \mathcal{S}} \left[\Pi[u] := \frac{1}{2} a(u, u) - \ell(u), \quad \mathcal{S} = \left\{ u : u \in H^2(\Omega), \quad u = 0 \text{ on } \partial\Omega, \quad u_{,n} = 0 \text{ on } \partial\Omega \right\} \right], \quad (49a)$$

where

$$a(u, w) = \int_{\Omega} (\nabla^2 u)(\nabla^2 w) \, dx, \quad \ell(w) = \int_{\Omega} fw \, dx, \quad (49b)$$

and now both essential boundary conditions in (48b) must be met to satisfy kinematic admissibility. To meet this objective, the solution structure for the PINN trial function is chosen as:

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta) = [\phi(\mathbf{x})]^2 \tilde{u}_{nn}^R(\mathbf{x}; \theta), \quad (50)$$

where $\phi(\mathbf{x})$ is an approximate distance function to the boundary $\partial\Omega$.

On substituting the trial function in (49), the finite-dimensional minimization problem is of the same form as (43), where $a(\cdot, \cdot)$ and $\ell(\cdot)$ are now given by (49b). Similar to the deep Ritz approach for the Poisson problem, we use a Monte Carlo integration rule with equally-weighted points. Following the same steps, the loss function is of the same form as in (47) with the bilinear and linear functionals defined in (49b).

8. Numerical implementation

Numerical solutions using deep neural networks are presented for one-, two- and four-dimensional boundary-value problems. Both, second- and fourth-order problems are considered. The trial function, $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$, contains $\tilde{u}_{nn}^R(\mathbf{x}; \theta)$, the approximation that is composed by the neural network. The trial function in the standard PINN [6] is denoted by $\tilde{u}_{nn}(\mathbf{x}; \theta)$. Likewise, $L_{nn}^{bc}(\theta)$ is used to denote the loss function in our approach and $L_{nn}(\theta)$ is that for standard PINN. We refer to the numerical solution that is obtained (after training) by our approach as $\tilde{u}_{nn}^{bc}(\mathbf{x})$ and

that obtained (after training) using standard PINN [6] as $\tilde{u}_{nn}(\mathbf{x})$. Deep collocation and deep Ritz methods are used to solve boundary-value problems. For the Ritz method, either ReLU or cubic ReLU activation function is used, with the problem that is solved in Section 9.1.7 being the sole exception where the Gaussian activation function is used. Unless otherwise stated, R-equivalence (REQ) composition with $m = 1$ and mean value potential (MVP) with $p = 1$ are used to form approximate distance functions in $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$. For all problems that are considered in this paper, whenever REQ is indicated, we employ (10) to form $\phi(\mathbf{x})$; when MVP is mentioned, the expression for ϕ given in (14) or (16) applies. For collocation, all boundary conditions are exactly satisfied, and essential boundary conditions are exactly met in the Ritz method, which ensures that the trial function is kinematically admissible. All collocation points are considered in a single batch in the network; points are not sorted in bins and passed in batches, which in general is a more efficient approach.

The formulation described in Section 7 has been implemented using Google's JAX library in Python [93], which can automatically differentiate native Python and NumPy functions. As an example, the code listing below sets up the calculation for the approximate distance function over an arbitrary polygon through the R-equivalence operation in (10).

```
def dist(x1, y1, x2, y2):
    return jax.numpy.sqrt((x2-x1)**2+(y2-y1)**2)

def linseg(x, y, x1, y1, x2, y2):
    L = dist(x1, y1, x2, y2)
    xc = (x1+x2)/2.
    yc = (y1+y2)/2.
    f = (1/L)*((x-x1)*(y2-y1) - (y-y1)*(x2-x1))
    t = (1/L)*((L/2.)**2-dist(x, y, xc, yc)**2)
    varphi = jax.numpy.sqrt(t**2+f**4)
    phi = jax.numpy.sqrt(f**2 + (1/4.)*(varphi - t)**2)
    return phi

def phi(x, y, segments):
    m = 1.
    R = 0.
    for i in range(len(segments[:, 0])):
        phi = linseg(x, y, segments[i, 0], segments[i, 1], segments[i, 2], segments[i, 3])
        R = R + 1./phi**m
    R = 1/R***(1/m)
    return R
```

Here `segments` is a NumPy array, which contains the coordinates of the line segments that make up the polygon and `phi(x, y, segments)` returns an approximate distance function from the location (x, y) to the polygon. A multilayer perceptron neural network is created using the following set of functions:

```
def RePU3(x):
    return (jax.numpy.maximum(0, x**3))

def repu3_layer(params, x):
    return RePU3(jax.numpy.dot(params[0], x) + params[1])

def NN(params, x, y):
    """ Compute the forward pass for each example individually """
    activations = jax.numpy.array([x, y])
    # Loop over the RePU3 hidden layers
    for w, b in params[:-1]:
        activations = repu3_layer([w, b], activations)
```

```

final_w, final_b = params[-1]
final = jax.numpy.sum(jax.numpy.dot(final_w, activations)) + final_b
return (final[0])

```

In the above, the cubic ReLU function is used as the activation in the hidden layers, which is also an instance of the Rectified Power Unit (RePU) activation function $\text{RePU}_n(x) := [\max(0, x)]^n$ ($n = 3$ for cubic ReLU). A linear activation function is used in the output layer. Here, `params` is a NumPy array consisting of the optimizable parameters, (w, b) , of the network, and its shape is determined by the architecture of the network. A network architecture that has been used frequently in this paper for problems in \mathbb{R}^d is a 2 hidden layer network $d\text{-}\mathcal{N}\text{-}\mathcal{N}\text{-}1$, where \mathcal{N} is the number of neurons in each hidden layer and is typically 50. Finally, the following functions and their derivatives are used to construct the ansatz $\tilde{u}_{mn}^{bc}(x; \theta)$, which satisfies homogeneous essential boundary conditions that are imposed on the boundary of a polygonal domain:

```

def u(params, x, y):
    return phi(x, y, segments) * NN(params, x, y)

#Examples of first-order partial derivatives
gradx = grad(u, 1)
grady = grad(u, 2)

#Examples of second-order partial derivatives
gradxx = grad(grad(u, 1), 1)
gradyy = grad(grad(u, 2), 2)
gradxy = grad(grad(u, 1), 2)

```

Once the derivatives are formed, the appropriate loss functions are constructed in the interior of the domain to solve the problem through Ritz or collocation. For all problems that are solved in this paper, network training is done using Google's Colaboratory cloud platform [94]. Single-precision arithmetic is used in the computations. An important consideration is the generation of collocation points in the interior and on the boundary of the domain to evaluate the loss terms. When the domain is simple – for example, a square or a hypercube – then these points are generated on a uniform grid in \mathbb{R}^d . This is the case for the solution of the heat equation on square domains, Eikonal problems in Section 10.5, and for the Poisson problem over the 4-dimensional hypercube in Section 11. In other examples that involve more complicated domains, we use the Python library `dmsh` [95], which draws inspiration from `distmesh` [96], to create triangular meshes and we use the centroids of the generated triangles as the interior collocation points. Fig. 15 shows a few representative meshes generated by `dmsh` that are used in this paper. The corresponding interior collocations points are shown as dots. When needed, `dmsh` is also used to create collocation points on the boundaries of the domain.

9. Numerical examples in one dimension

We consider several second-order problems that involve essential and mixed boundary conditions. As a prototypical fourth-order problem, we solve the deflection for a clamped Euler–Bernoulli beam, and also investigate a single hidden layer meshfree RBF-network solution in Section 9.1.7. Our objective is to demonstrate the benefits of the new formulation vis-à-vis standard PINN [6] in which equal weights are chosen for the PDE and boundary loss terms. Hence, our emphasize is not on obtaining the most accurate PINN solution by finding the optimal hyperparameters nor comparing its performance versus the finite element method for forward problems. We compare our results with those obtained using standard PINN [6] and to the exact solution.

9.1. Deformation of a homogeneous elastic rod

Consider the boundary-value problem for the deformation of an elastic rod (Young's modulus and cross-sectional area are taken as unity):

$$u'' + b = 0 \quad \text{in } \Omega = (x_1, x_2) \tag{51a}$$

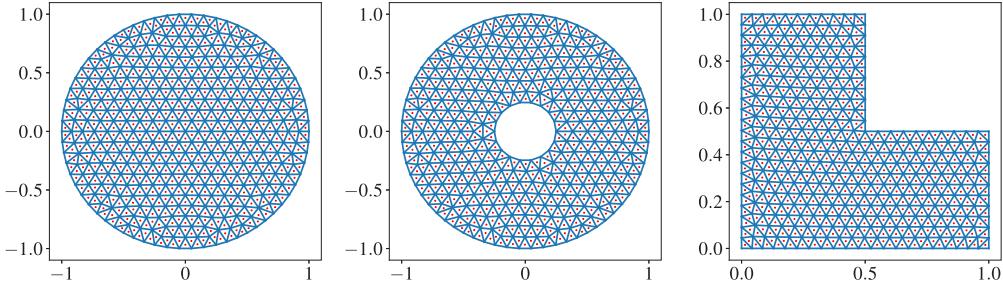


Fig. 15. Representative meshes generated from dmsh and the corresponding collocation points.

$$u(x_1) = g, \quad u'(x_2) + cu(x_2) = h, \quad (51b)$$

where $u := u(x)$ is the axial displacement field, $u'(x)$ is the strain field, $b := b(x)$ is the axial body force per unit length, and c, g and h are constants. The second boundary condition is a Robin boundary condition; if the bar is connected to a spring that is attached to a fixed end, then $h = 0$. We select test problems with different boundary conditions, and vary the regularity of $b(x)$ from it being a smooth function to a δ -function, and even choose $b(x)$ that has a singularity at the origin.

9.1.1. Example 1

As the first example, a Dirichlet problem in $\Omega = (0, 1)$ is selected with body force $b(x) = 1 - 2x + 10x^2$, and boundary conditions $u(0) = 1/2$ and $u(1) = -1/2$. The exact solution is: $u(x) = 1/2 - x^2/2 + x^3/3 - 10x^4/12$. The exact signed distance functions to $x = 0$ and $x = 1$ are $\phi_1(x) = x$ and $\phi_2(x) = 1 - x$, respectively. Now, we join these to form a smooth approximate distance function to the boundary $\partial\Omega = \{0, 1\}$. On using the product of ϕ_1 and ϕ_2 , and the R-conjunction ($\alpha = 0$) and R-equivalence ($m = 2$) relations in (2) and (9), respectively, we obtain the following combined distance functions:

$$\phi_A(x) = \phi_1(x)\phi_2(x), \quad \phi_B(x) = \phi_1(x) + \phi_2(x) - \sqrt{\phi_1^2(x) + \phi_2^2(x)}, \quad \phi_C(x) = \frac{\phi_1(x)\phi_2(x)}{\sqrt{\phi_1^2(x) + \phi_2^2(x)}}. \quad (52)$$

Note that in one dimension, the product formula is normalized to order 1, but this does not generalize to higher dimensions. Coincidentally, for any domain $\Omega = (x_1, x_2)$ in one dimension, the product formula scaled by $L := x_2 - x_1$ coincides with $m = 1$ in the R-equivalence relation.

In the numerical computations, the trial function is formed using (22):

$$\tilde{u}_{nn}^{bc}(x; \theta) = g(x) + \phi(x) \tilde{u}_{nn}^R(x; \theta), \quad g(x) = \frac{1 - 2x}{2}, \quad (53)$$

where $g(x)$ is formed using the transfinite interpolant in (26), $\tilde{u}_{nn}^R(x; \theta)$ is the neural network approximation, and $\phi(x)$ is chosen to be either $\phi_A(x)$, $\phi_B(x)$ or $\phi_C(x)$. Note that we did not solve the patch test ($b = 0$) since the exact solution, $u(x) = g(x)$, is already captured by the presence of $g(x)$ in (53). The network architecture 1-30-30-1 is used. We compute collocation solutions using ϕ_A , ϕ_B and ϕ_C in (53), and also for the standard PINN approximation, $\tilde{u}_{nn}(x; \theta)$ [6]. In Fig. 16, the numerical results are presented. All approaches are able to reach losses of the same order, and there is a good match between $\tilde{u}_{nn}^{bc}(x)$ and the exact displacement field. From Fig. 16d, we observe that the errors in the displacement and strain fields using \tilde{u}_{nn}^{bc} are uniformly smaller than those obtained using \tilde{u}_{nn} . This difference stems from the fact that in our approach the trial function is constructed with the exact satisfaction of the boundary conditions.

We emphasize that it is possible for both standard PINN and our formulation to deliver better accuracy if a larger network, more interior collocation points, and network training for a longer duration are chosen. This is realized at the expense of more computing time. So for just this example, we demonstrate the same. We now use a 1-50-50-1 network architecture with 300 interior collocation points. The training is conducted until 50,000 epochs. The solutions are presented in Fig. 17, which reveal that both approaches are now much more accurate than the

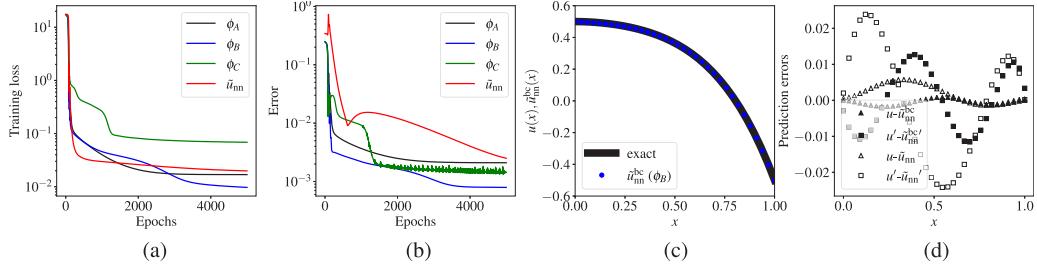


Fig. 16. Collocation solutions for a Dirichlet problem. The body force $b(x) = 1 - 2x + 10x^2$, and the essential boundary conditions are $u(0) = 1/2$ and $u(1) = -1/2$. The network architecture is 1-30-30-1. Numerical solutions, $\tilde{u}_{nn}^{bc}(x)$, are computed using different ADFs (ϕ_A , ϕ_B , ϕ_C), and are compared to the solution obtained using standard PINN, $\tilde{u}_{nn}(x)$. (a), (b) Training loss and normalized absolute error in the displacement field as a function of epochs. (c) Exact solution $u(x)$ and $\tilde{u}_{nn}^{bc}(x)$ (using ϕ_B) are shown, and (d) Errors in the displacement and strain fields are compared.

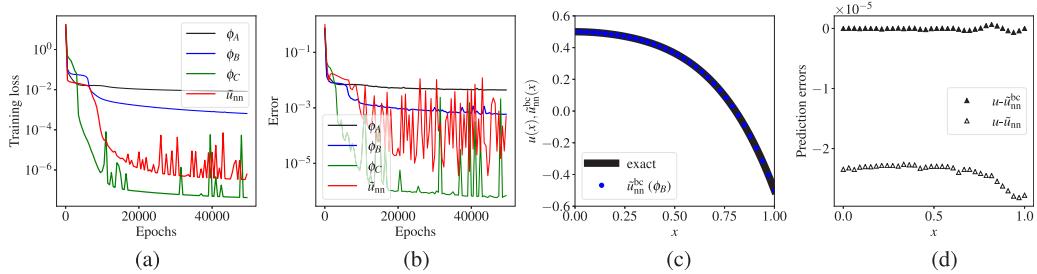


Fig. 17. Collocation solutions for the Dirichlet problem that is presented in Example 1. The network architecture is 1-50-50-1 with 300 interior collocation points. The captions for (a), (b), (c), (d) mirror those shown in Fig. 16.

results that are shown in Fig. 16. From Fig. 17, we observe that the solution obtained using standard PINN and our approach have relative errors of $\mathcal{O}(10^{-5})$ and $\mathcal{O}(10^{-6})$, respectively. So even at smaller relative errors, we note that our approach is more accurate than standard PINN for the same values of the hyperparameters.

9.1.2. Example 2

We reconsider the problem posed in Example 1, with Dirichlet boundary condition at $x = 0$ but homogeneous Neumann (traction-free) condition at $x = 1$. For the domain $\Omega = (0, 1)$, body force $b(x) = 1 - 2x + 10x^2$, and boundary conditions $u(0) = 1/2$ and $u'(1) = 0$, the exact solution is: $u(x) = 1/2 + 10x/3 - x^2/2 + x^3/3 - 10x^4/12$. The trial function, $\tilde{u}_{nn}^{bc}(x; \theta)$, is formed using (33), with $\phi_1(x) = x$, $\phi_2(x) = 1 - x$, $\phi_3(x) = x(1 - x)$, $g = 1/2$, and $c = h = 0$:

$$\tilde{u}_{nn}^{bc}(x; \theta) = x \tilde{u}_{nn}^R(x; \theta) + x(1 - x) \left[(1 - x) \tilde{u}_{nn}^R(x; \theta) + \{x \tilde{u}_{nn}^R(x; \theta)\}'_{x=1} \right] + \frac{1}{2}. \quad (54)$$

The network architecture 1-50-50-1 is used, since the network 1-30-30-1 did not converge for standard PINN. The collocation solutions $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$ are compared to the exact solution $u(x)$ in Fig. 18. We observe from Fig. 18a that the training loss for $\tilde{u}_{nn}^{bc}(x; \theta)$ is about two orders smaller than $\tilde{u}_{nn}(x; \theta)$, whereas in Fig. 18b, the corresponding normalized absolute error is one order smaller. The numerical solution $\tilde{u}_{nn}^{bc}(x)$ is in excellent agreement with the exact solution in Fig. 18c. Over the entire interval $x \in [0, 1]$, we find that the displacement and strain fields from $\tilde{u}_{nn}^{bc}(x)$ are markedly more accurate than those from $\tilde{u}_{nn}(x)$ (see Fig. 18d).

9.1.3. Example 3

For this example, we choose $\Omega = (0, 1)$, a sinusoidal body force $b(x) = -\sin(k\pi x)$ with varying k , and essential boundary conditions $u(-1) = u(1) = 0$. The exact solution is: $u(x) = -\sin(k\pi x)/\pi^2 k^2$. This example serves to reveal the spectral (low-frequency) bias [97,98] of neural network approximations. In Fig. 19, the numerical results are presented. For $k = 1$ and using standard PINN, we found that the normalized absolute error for the networks

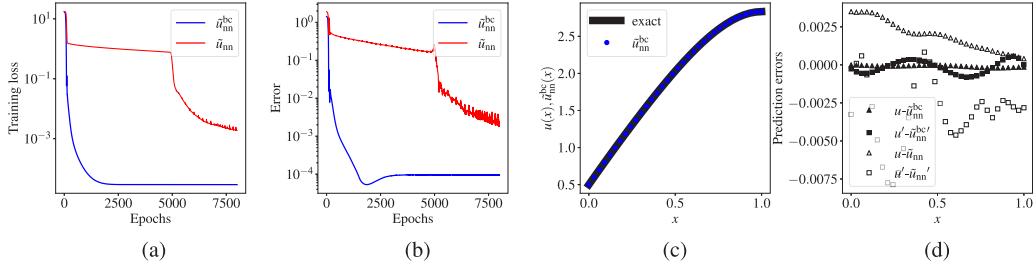


Fig. 18. Collocation solutions for a Neumann problem. The body force $b(x) = 1 - 2x + 10x^2$, and $u(0) = 1/2$ and $u'(1) = 0$. The network architecture is 1-50-50-1. Numerical solution using $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$ are compared. (a), (b) Training loss and normalized absolute error as a function of epochs. (c) $\tilde{u}_{nn}^{bc}(x)$ is compared to the exact solution. (d) Errors in the displacement and strain fields for $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$.

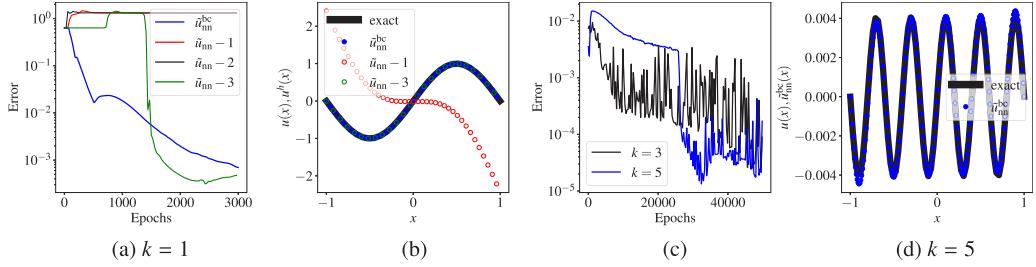


Fig. 19. Collocation solutions for a homogeneous Dirichlet problem with a sinusoidal body force $b(x) = -\sin(k\pi x)$ with varying k . (a) Normalized absolute error as a function of epochs for $k = 1$. $\tilde{u}_{nn}^{bc}(x; \theta)$ on a 1-30-30-1 architecture is compared to $\tilde{u}_{nn}(x; \theta)$. $\tilde{u}_{nn}-1$, $\tilde{u}_{nn}-2$ and $\tilde{u}_{nn}-3$ are solutions that are obtained on network architectures 1-30-30-1, 1-100-100-1 and 1-100-100-100-1, respectively. (b) $\tilde{u}_{nn}(x)$, $\tilde{u}_{nn}^{bc}(x)$, and $u(x)$ are plotted for $k = 1$. (c) Normalized absolute error during training using $\tilde{u}_{nn}^{bc}(x; \theta)$ for $k = 3$ and $k = 5$. Network architecture is 1-100-100-1. (d) For $k = 5$, error in the displacement field for $\tilde{u}_{nn}(x)$.

1-30-30-1 and even 1-100-100-1 did not converge; it took a network architecture of 1-100-100-100-1 for the error to be comparable to that obtained using $\tilde{u}_{nn}^{bc}(x; \theta)$ on a 1-30-30-1 architecture. The normalized absolute errors as a function of epochs is shown in Fig. 19a. The numerical solutions are compared to the exact solution $u(x)$ in Fig. 19b, and we notice the poor solution that is generated by $\tilde{u}_{nn}(x; \theta)$ on the 1-30-30-1 network. For $k = 3, 5$ on a 1-100-100-1 network, the normalized absolute error in the $\tilde{u}_{nn}^{bc}(x)$ displacement field is shown in Fig. 19c, and $\tilde{u}_{nn}^{bc}(x)$ and $u(x)$ are compared in Fig. 19d for $k = 5$. Good agreement between $\tilde{u}_{nn}^{bc}(x)$ and $u(x)$ is realized. We attribute the poor performance of the standard PINN approach to the fact that the boundary conditions are poorly approximated, whereas with $\tilde{u}_{nn}^{bc}(x; \theta)$ the boundary conditions are exactly satisfied. This observation is in broad agreement with the findings of Wang et al. [16], who further analyze the source of this discrepancy by drawing attention to the contributions from the boundary and interior terms in the loss function. For $k \geq 10$ with $\tilde{u}_{nn}^{bc}(x; \theta)$, accurate PINN solutions using a single neural network for the entire domain becomes infeasible due to the spectral bias of the neural network approximation; one can adopt a domain-decomposition strategy to obtain accurate numerical solutions for high-frequency problems.

9.1.4. Example 4

Consider an elastic rod that occupies $\Omega = (-1, 1)$ and is subjected to a discontinuous body force $b(x) = H(x)$, where $H(x)$ is the Heaviside function. Essential boundary conditions are imposed at both ends: $u(-1) = 0$ and $u(1) = -1/2$. The exact solution, $u(x) \in C^1(\Omega)$, is:

$$u(x) = \begin{cases} 0 & , -1 \leq x < 0 \\ -\frac{x^2}{2} & , 0 \leq x \leq 1. \end{cases} \quad (55)$$

Numerical computations are performed on a 1-50-50-1 network. Numerical results for \tilde{u}_{nn} and \tilde{u}_{nn}^{bc} are presented in Fig. 20. The training loss of $\tilde{u}_{nn}^{bc}(x; \theta)$ at 10,000 epochs is $\mathcal{O}(10^{-3})$ but $\tilde{u}_{nn}(x; \theta)$ stagnates to a value just below 0.1. These losses correspond to normalized absolute errors of $\mathcal{O}(10^{-5})$ and $\mathcal{O}(10^{-4})$ at the end of the training for

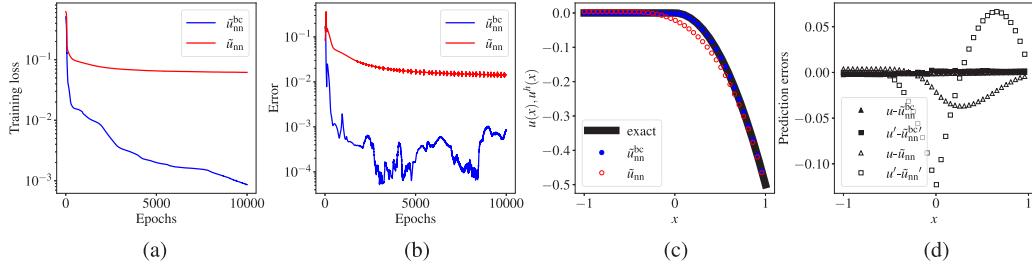


Fig. 20. Collocation solutions for a Dirichlet problem with a discontinuous body force, $b(x) = H(x)$. Essential boundary conditions are $u(-1) = 0$ and $u(1) = -1/2$, and the network architecture is 1-50-50-1. (a), (b) Training loss and normalized absolute error as a function of epochs for $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$. (c) Comparisons of $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$ with the exact solution. (d) Error in the displacement and strain fields.

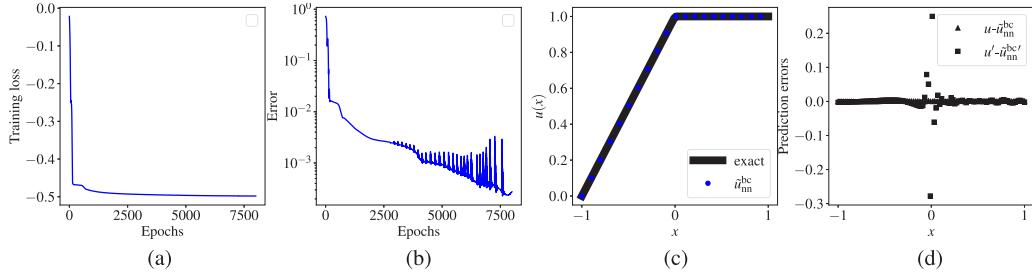


Fig. 21. Ritz solution for a point load, $b(x) = \delta(x)$. Homogeneous essential boundary condition is imposed at $x = -1$ and traction-free conditions prevail at $x = 1$. The network architecture is 1-50-50-1. (a), (b) Training (Ritz) loss and normalized absolute error as a function of epochs. (c) Comparison of $\tilde{u}_{nn}^{bc}(x)$ with the exact solution. (d) Errors in displacement and strain fields.

$\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$, respectively (see Fig. 20b). Fig. 20c and d reveal that $\tilde{u}_{nn}^{bc}(x)$ is in good agreement with $u(x)$, whereas the errors in $\tilde{u}_{nn}(x)$ are significant, and are especially pronounced in the vicinity of $x = 0$. The PDE loss is dominant over the boundary losses. We point out that if both boundary conditions are homogeneous then the accuracy of $\tilde{u}_{nn}(x)$ is comparable to that of $\tilde{u}_{nn}^{bc}(x)$.

9.1.5. Example 5

Let us consider an elastic rod that occupies $\Omega = (-1, 1)$ and is subjected to a unit point load at the origin, i.e., $b(x) = \delta(x)$, where $\delta(x)$ is the δ -function. Homogeneous essential boundary condition is prescribed at $x = -1$ and traction-free conditions prevail at $x = 1$. The exact solution, $u(x) \in C^0(\Omega)$, is:

$$u(x) = \begin{cases} 1 + x & , -1 \leq x < 0 \\ 1 & , 0 \leq x \leq 1, \end{cases} \quad (56)$$

which has a kink at $x = 0$. We use the deep Ritz method with $\tilde{u}_{nn}^{bc}(x; \theta)$ to solve this problem since it is not possible to solve this problem using the collocation method ($\delta(x)$ is a distribution that is defined over a zero measure). The network architecture is 1-50-50-1. Numerical results for \tilde{u}_{nn}^{bc} are presented in Fig. 21. From Fig. 21a, we observe that the training loss converges to a value close to -0.5 in a few epochs; this corresponds to a small normalized absolute error of $\mathcal{O}(10^{-3})$ (see Fig. 21b). On using u given in (56), we find that the potential energy of the exact solution is:

$$\Pi[u] = \frac{1}{2} \int_{-1}^1 u'^2 dx - u(0) = -\frac{1}{2},$$

which is the target loss that $\Pi[\tilde{u}_{nn}^{bc}]$ seeks to attain. Fig. 21c shows excellent agreement between $\tilde{u}_{nn}^{bc}(x)$ and the exact solution. The errors in the displacement and strain fields are presented in Fig. 21d. The errors in u are uniformly small for all $x \in [0, 1]$. The errors in the strain field follow the same trends, but have larger errors in the vicinity of the origin. This is not surprising since $\tilde{u}_{nn}^{bc}(x)$ is $C^2(\Omega)$ (cubic ReLU activation function), whereas the exact solution

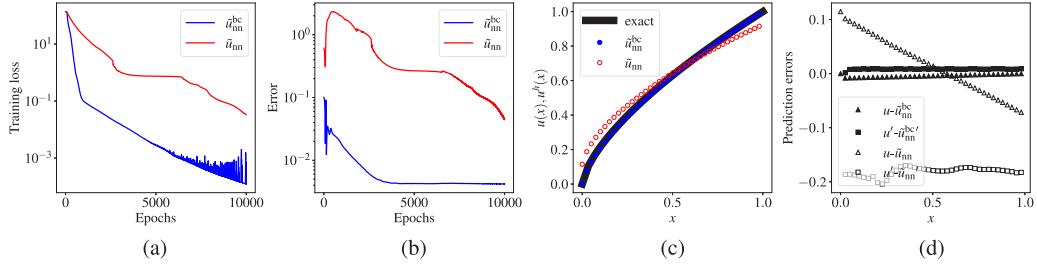


Fig. 22. Collocation solutions for a Dirichlet problem with a singular body force, $b(x) = 2x^{-4/3}/9$. Essential boundary conditions are $u(0) = 0$ and $u(1) = 1$, and the network architecture is 1-50-50-1. (a), (b) Training loss and normalized absolute error as a function of epochs for $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$. (c) Comparisons of $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$ with the exact solution. (d) Error in the displacement and strain fields.

$u \in C^0(\Omega)$, with $u'(x)$ being discontinuous at $x = 0$. This is also the source for the small discrepancy in the Ritz energy loss.

9.1.6. Example 6

To obtain a weakly singular solution for the elastic rod problem, we consider a body force that has a singularity at the origin. We choose $\Omega = (0, 1)$ with $b(x) = 2x^{-4/3}/9$ with essential boundary conditions $u(0) = 0$ and $u(1) = 1$. The exact solution is $u(x) = x^{2/3}$, and $u \in H^1(\Omega)$ is weakly singular. This problem is solved using the collocation approach on a 1-50-50-1 network architecture. Numerical results for $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$ are presented in Fig. 22. The training loss of $\tilde{u}_{nn}^{bc}(x; \theta)$ at 10,000 epochs is close to $\mathcal{O}(10^{-4})$ and $\tilde{u}_{nn}(x; \theta)$ is more than two orders larger. This same trend is observed in the normalized absolute error as a function of epochs (see Fig. 22b). Fig. 22c and d show that $\tilde{u}_{nn}^{bc}(x)$ is in fairly good agreement with $u(x)$, whereas the error in the displacement and strain fields of $\tilde{u}_{nn}(x)$ are appreciable. It appears that $\tilde{u}_{nn}(x)$ and $u(x)$ differ by close to an affine function, which one can infer as being present within $\tilde{u}_{nn}(x)$ to meet the essential boundary conditions.

9.1.7. Example 7

To draw connections to meshfree methods based on RBFs [46] and local maximum-entropy approximants [48], which are discussed in Section 1, we solve (51) with inhomogeneous Dirichlet boundary conditions. The domain $\Omega = (0, 1)$, and we choose the exact solution as:

$$u(x) = \sum_{i=1}^2 \exp[-\gamma_i(x - a_i)^2], \quad (57)$$

which is a sum of two Gaussian functions, and γ_i and a_i ($i = 1, 2$) are constants. The body force $b(x) = -u''(x)$. Essential boundary conditions are imposed at $x = 0$ and $x = 1$ that are consistent with the exact solution in (57).

In the numerical computations, we choose $a_1 = 1/4$, $a_2 = 6/10$, $\gamma_1 = 9$ and $\gamma_2 = 10$. The network architecture is 1-10-1 (1 hidden layer). For the hidden layer, we select a Gaussian activation function, $\sigma(x) = \exp(-x^2)$, and a linear activation function for the output layer. The centers of the Gaussian are chosen to be fixed, and only the support-widths of the Gaussians and the weights in the output layer are the unknown parameters in the network. The centers for the neurons are chosen as $\mathbf{b} = \{0, 1/9, 2/9, \dots, 1\}$. In Fig. 23, the numerical results for $\tilde{u}_{nn}(x; \theta)$ using collocation and Ritz are presented. By the end of the training, the loss for the collocation and Ritz solutions are $\mathcal{O}(10^{-1})$ or better. Note that the loss measures are distinct for the collocation and Ritz solutions; PDE loss is shown for the former whereas it is the (Ritz) energy loss that is presented for the latter. The collocation solution is in good agreement with the exact solution (Fig. 23b). From Fig. 23c, we observe that the errors in the displacement field for both methods are small, but the errors in the derivative (strain fields) are appreciable. With more number of neurons in the hidden layer, it is expected that the accuracy in the strain field will substantially improve. This example reveals the flexibility that the PINN affords in that variational adaptive solutions can be captured by a Gaussian neural network with a single hidden layer. Realizing this is much more difficult using meshfree basis functions, since the underlying Ritz formulation becomes a nonlinear, nonconvex minimization problem.

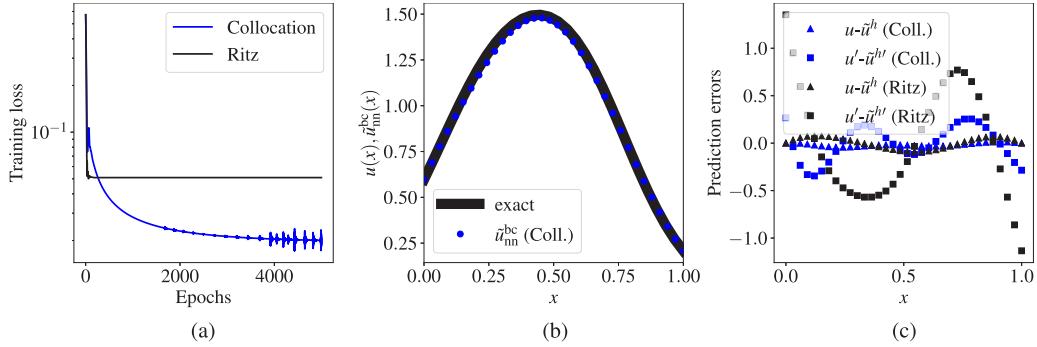


Fig. 23. Collocation and Ritz solutions using a Gaussian activation function for a Dirichlet problem with the exact solution as the sum of two Gaussians. Network architecture is 1-10-1. (a) Training loss as a function of epochs for $\tilde{u}_{nn}^{bc}(x; \theta)$ (collocation and Ritz methods). (b) Comparisons of $\tilde{u}_{nn}^{bc}(x)$ (collocation) with the exact solution. (c) Error in the displacement and strain fields for collocation and Ritz methods.

9.2. Longitudinal vibrations of a homogeneous elastic rod

The eigenproblem for the longitudinal vibrations of an elastic bar that is fixed at both ends is:

$$u'' + \omega^2 u = 0 \quad \text{in } \Omega = (0, 1), \quad (58a)$$

$$u(0) = u(1) = 0. \quad (58b)$$

The exact eigenfunctions are: $u_n(x) = \sin(\omega_n x)$, where $\omega_n = n\pi$ ($n \in \mathbb{N}$) are the natural frequencies. The eigenvalue $\lambda_n = \omega_n^2$ corresponds to the eigenfunction $u_n(x)$.

We use the Ritz method to solve this problem using $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$. The Rayleigh quotient minimization problem for the smallest eigenvalue (lowest mode) is [9]:

$$\min_{u \in \mathcal{S}} \frac{\int_0^1 u'^2 dx}{\int_0^1 u^2 dx}, \quad \text{subject to } \int_0^1 u^2 dx = 1, \quad (59)$$

where $\mathcal{S} = \{u : u \in H^1(\Omega), u(0) = u(1) = 0\}$ and the normalization constraint on the eigenfunction appears in (59). For the trial function $\tilde{u}_{nn}^{bc}(x; \theta)$, the loss function is:

$$L_{nn}^{bc}(\theta) = \frac{\sum_{k=1}^{N_I} (\tilde{u}_{nn}^{bc}(x_k; \theta))^2}{\sum_{k=1}^{N_I} (\tilde{u}_{nn}(x_k; \theta))^2} + \left[\left\{ \frac{1}{N_I} \sum_{k=1}^{N_I} (\tilde{u}_{nn}^{bc}(x_k; \theta))^2 \right\} - 1 \right]^2, \quad (60)$$

where N_I is the number of interior integration points. Note that apart from the PDE loss term, we have an additional loss term due to the normalization constraint in (59). The loss function for $\tilde{u}_{nn}(x; \theta)$ consists of the two contributions that appears in (60), and in addition it will include two boundary loss terms to impose the essential boundary conditions.

The network architecture 1-50-50-50-1 is used. In Fig. 24, the Ritz solutions for $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$ are presented. The loss function for \tilde{u}_{nn}^{bc} and \tilde{u}_{nn} saturate to values of 1 and 10, respectively. The error in the natural frequency for $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$ are $\mathcal{O}(10^{-4})$ and $\mathcal{O}(1)$ at 10,000 epochs (see Fig. 24b). In Fig. 24c, we compare the lowest mode (eigenfunction) from the numerical solutions to the exact solution: $\tilde{u}_{nn}^{bc}(x)$ is in much better agreement with the exact solution than $\tilde{u}_{nn}(x)$. The exact mode shape is well-captured by $\tilde{u}_{nn}^{bc}(x)$ but it is has not been exactly normalized, which leads to the observed discrepancy in the maximum amplitude.

9.3. Advection-diffusion problem

$$u'' = \alpha u' \quad \text{in } \Omega = (0, 1) \quad (61a)$$

$$u(0) = 0, \quad u(1) = 1, \quad (61b)$$

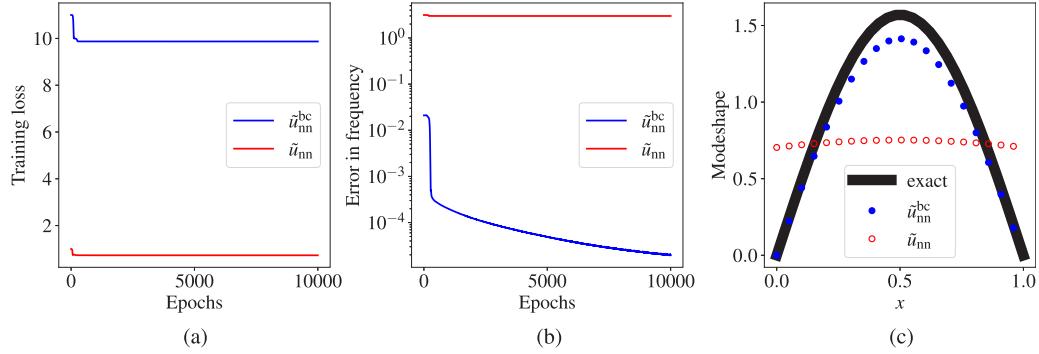


Fig. 24. Ritz solutions for the longitudinal vibrations (lowest mode) of a homogeneous elastic rod. The network architecture is 1-50-50-50-1. (a), (b) Training loss and error in lowest natural frequency as a function of epochs for $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$. (c) Comparisons of $\tilde{u}_{nn}^{bc}(x)$ and $\tilde{u}_{nn}(x)$ for the lowest mode (eigenfunction) that corresponds to the lowest natural frequency $\omega_1 = \pi$.

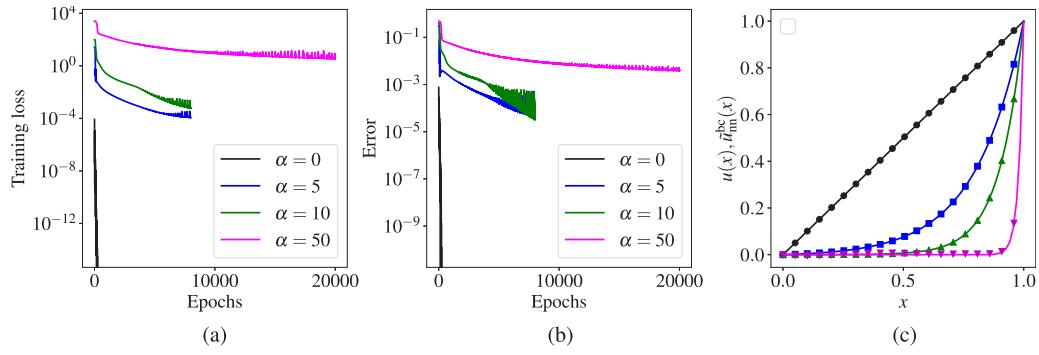


Fig. 25. Collocation solutions using $\tilde{u}_{nn}^{bc}(x; \theta)$ for the advection-diffusion problem. For $\alpha = 1, 5, 10$, the network architecture 1-50-50-1 is used, and for $\alpha = 50$, the architecture is 1-50-50-50-1. (a), (b) Training loss and normalized absolute error as a function of epochs. (c) Comparisons of $\tilde{u}_{nn}^{bc}(x)$ to the exact solutions for different α . The solid lines are for the exact solutions and the markers (colors are consistent with those shown in (b)) represent the numerical solutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where α is the Peclet number, which measures the ratio of the advective rate to the diffusion rate. The exact solution of the problem posed in (61) is:

$$u(x) = \frac{e^{\alpha x} - 1}{e^\alpha - 1}. \quad (62)$$

We choose $\alpha = 0, 5, 10, 50$ in this study (pure diffusion for $\alpha = 0$ to strongly advective flow for $\alpha = 50$) and run collocation simulations using $\tilde{u}_{nn}^{bc}(x; \theta)$. For $\alpha = 1, 5, 10$, the network architecture 1-50-50-1 is used, and for $\alpha = 50$, the architecture is 1-50-50-50-1. In Fig. 25, the simulation results are presented. For all cases shown in Fig. 25b, we observe that the normalized absolute error during the training is $\mathcal{O}(10^{-3})$ or less. For all α that are selected, Fig. 25c shows an excellent match between $\tilde{u}_{nn}^{bc}(x)$ and the exact solutions.

9.4. Euler-Bernoulli Beam bending

Consider the boundary-value problem for the deflection of a cantilever (Euler-Bernoulli) beam of unit length that is clamped at both ends and is subjected to a distributed load $q(x)$:

$$EIv''' = q \quad \text{in } \Omega = (0, 1), \quad (63a)$$

$$v(0) = v'(0) = v(1) = v'(1) = 0, \quad (63b)$$

where EI is the flexural rigidity of the beam.

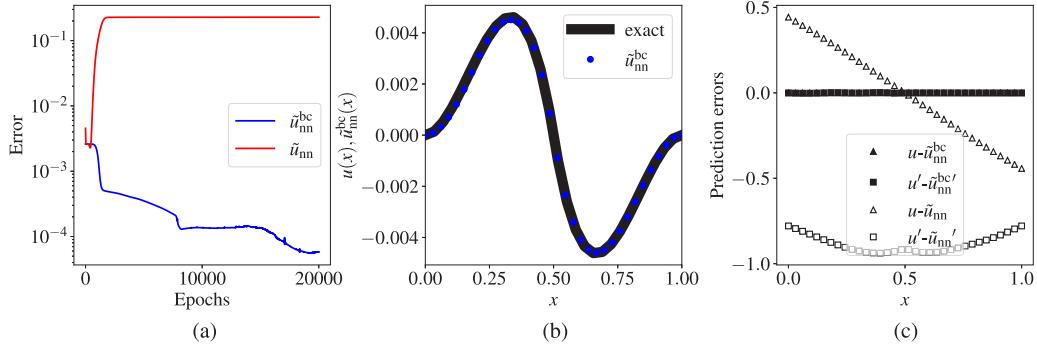


Fig. 26. Deep Ritz solutions using $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$ for the Euler–Bernoulli beam problem. The network architecture 1-50-50-1 is used. (a) Normalized absolute error as a function of epochs. (b) Comparison of $\tilde{u}_{nn}^{bc}(x)$ with the exact solution. (c) Errors in the deflection and rotation fields.

We apply a point moment (clockwise orientation) of magnitude M_0 at $x = 1/2$ so that $q(x) = M_0\delta'(x - 1/2)$. For this $q(x)$, the **variational principle** that is associated with the strong form in (63) is:

$$\min_{v \in S} \left[\Pi[v] = \frac{1}{2} \int_0^1 EI(v'')^2 dx + M_0 v'(1/2), \quad S = \left\{ v : v \in H^1(0, 1), \quad v(0) = v'(0) = v(1) = v'(1) = 0 \right\} \right]. \quad (64)$$

For this problem, all homogeneous boundary conditions associated with v and v' that appear in (63b) are essential boundary conditions, and hence a kinematically admissible PINN trial function is given by (50):

$$\tilde{v}_{nn}^{bc}(x; \theta) = [\phi(x)]^2 \tilde{v}_{nn}^R(x; \theta), \quad (65)$$

where $\phi(x)$ is an ADF (normalized to order 1) that vanishes at $x = 0$ and $x = 1$, and its normal derivative has unit magnitude on the boundary.

For the computations, we choose $EI = 1$ and $M_0 = 1$. The exact solution of (63) is:

$$v(x) = \frac{[\text{ReLU}(x - 1/2)]^2}{2} + \frac{x^2}{8} - \frac{x^3}{4}, \quad x \in \Omega, \quad (66)$$

which is a $C^1(\Omega)$ (piecewise cubic) function with a moment discontinuity of unit magnitude at $x = 1/2$. Deep Ritz solutions $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$ are computed on the network architecture 1-50-50-1 using the cubic ReLU activation function, and the results are presented in Fig. 26. From Fig. 26a, we observe that $\tilde{u}_{nn}^{bc}(x; \theta)$ converges to a loss of $\mathcal{O}(10^{-4})$ at 20,000 epochs but the loss for $\tilde{u}_{nn}(x; \theta)$ remains at about 0.1. Fig. 26b shows that the numerical solution $\tilde{u}_{nn}^{bc}(x)$ is in excellent agreement with the exact solution. In Fig. 26c, the error fields are shown: the deflection and rotation fields using $\tilde{u}_{nn}^{bc}(x)$ are accurate, whereas both fields have large errors for $\tilde{u}_{nn}(x)$. We attribute the poor performance of $\tilde{u}_{nn}(x; \theta)$ due to scaling issues of the interior and boundary terms in the loss function. If the ratio of the weights assigned to the boundary loss term and the interior loss terms is set to $10^3 : 1$, we find that the results improve but they are still worse than those obtained using $\tilde{u}_{nn}^{bc}(x)$. It requires a weight ratio of 10^4 for the two solutions to have comparable accuracy.

10. Numerical examples in two dimensions

The promise of PINN is intriguing for inverse and parameteric (design) problems. But this rests on its accuracy, robustness and reliability on solving the forward problem, which is the emphasis in this contribution. To this end, we focus on the performance of the PINN formulation with exact imposition of boundary conditions versus the standard PINN [6] with equally-weighted loss terms. For the two-dimensional problems, we consider polygonal domains and also domains with curved boundaries. We consider four distinct types of problems: steady-state heat conduction; computation of harmonic (Laplace equation) coordinates [99], which is an instance of generalized barycentric coordinates [80]; clamped Kirchhoff plate bending (fourth-order PDE); and the Eikonal equation to

compute the signed distance function to a boundary. For these problems, we present our solutions and compare them to either the exact solution (if available) or a reference finite element solution, and to deep collocation [6]. In addition, we identify key differential properties of the approximate distance function and bring to fore the issue of exact satisfaction of boundary conditions and its implications in the training of the network and the accuracy that the PINN approximation delivers.

Prior to presenting the numerical examples, it is instructive to understand the properties and behavior of approximate distance functions that are obtained by either R-functions with R-equivalence as presented in Section 3.2 or via mean value potential fields that are discussed in Section 4. Since these functions are used in the PINN ansatz $\tilde{u}_{mn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ that is presented in Section 5, one must consider the regularity of these functions when used in a deep collocation or a deep Ritz method.

10.1. Laplacian of approximate distance fields

For Poisson or Laplace boundary-value problems that involve the Laplace operator, we must understand the behavior of the Laplacian of the ADFs that stem from R-functions and mean value potentials. Let us consider the unit square, $\Omega = (0, 1)^2$. The boundary $\partial\Omega$ consists of four line segments. On using either (10) or (14), we can construct an approximate distance function to the boundary, $\phi(\mathbf{x})$, which is normalized to order 1. Let us refer to these functions as $\phi_R(\mathbf{x})$ (REQ) and $\phi_M(\mathbf{x})$ (MVP). In Fig. 27, $\phi(\mathbf{x})$ and its Laplacian over the unit square are presented. We observe that both ϕ_R and ϕ_M are zero on the entire boundary and monotonic (concave) inside the domain. This property of these functions on $\partial\Omega$ is used to impose essential boundary conditions, as described in Section 5.2. From Fig. 27c and d, we observe that the Laplacians, $\nabla^2\phi_R$ and $\nabla^2\phi_M$, blow up at the vertices of the square. In fact, it is known that both $\nabla^2\phi_R$ and $\nabla^2\phi_M$ are singular at the vertices of a polygon, and therefore very large in magnitude near any of its vertices. Therefore, in a collocation-based approach to solve the Poisson equation, which involves the Laplacian, the contributions to the total loss from regions near the vertices can be very large. This inference does not influence Ritz-based solutions of the Poisson equation since the highest derivative in the variational principle is of order 1, and both ϕ_R and ϕ_M and its first-order derivatives are well-behaved (bounded) over the entire domain. There are two possible remedies to address this issue. The first involves modifying the ϕ_i that are obtained from R-equivalence and the ADFs that stem from mean value potential fields, so that the corresponding Laplacians are bounded in the domain. The second is to consider collocation points inside the domain that are not very close to the vertices. We leave the first route as part of future work, and proceed in this paper with the second choice. For instance, we show in Section 10.2.2 that if all interior collocation points are located in $\Omega_\delta = [\delta, 1 - \delta]^2$ ($\delta = 0.01$), which is a subset of the unit square, then both methods perform well. Finally, we reemphasize that it is imperative that in most instances ϕ be smooth in the interior of a computational domain; otherwise, $\nabla^2\phi$ will blow up at an interior collocation point and then one cannot use a trial function that uses ϕ in a collocation-based PINN method. So in most instances in 2D or 3D, this precludes the use of exact distance functions in the ansatz, and hence *approximate distance functions* should be used. To show this, let us consider the exact signed distance to the unit disk, $\phi(\mathbf{x}) = 1 - \sqrt{x^2 + y^2}$. We can write

$$\frac{\partial^2\phi}{\partial x^2} = \frac{x^2}{(x^2 + y^2)^{3/2}} - \frac{1}{\sqrt{x^2 + y^2}},$$

and since the second term is unbounded at the origin, the Laplacian of ϕ blows up at the origin. There are exceptions when the exact distance function can be used. It is a suitable choice when the medial axis of a domain is not part of the computational domain. For example, when solving a boundary-value problem over an annulus in 2D (see the problem solved in Section 10.2.3) or a hollow cylinder in 3D, then the exact distance function can be used since the origin (where the exact distance function has derivative discontinuities) lies outside the computational domain.

10.2. Steady-state heat conduction

Let us consider the following model problem for isotropic steady-state heat conduction:

$$-\nabla^2 u = f \quad \text{in } \Omega \subset \mathbb{R}^2 \tag{67a}$$

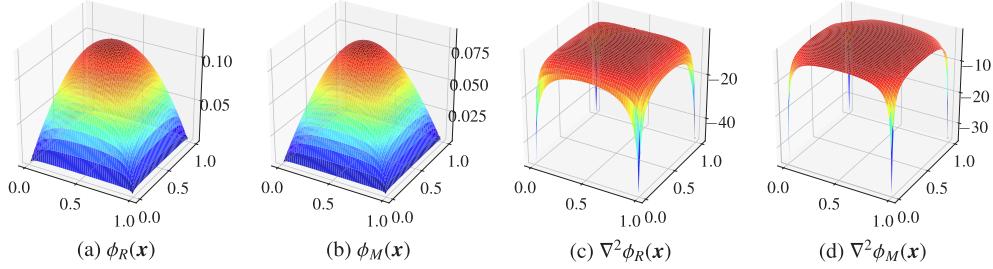


Fig. 27. Computation of $\phi(\mathbf{x})$ and $\nabla^2\phi(\mathbf{x})$ over the unit square for ADFs constructed from R-equivalence and mean value potential fields.

$$u = g \text{ on } \Gamma_u, \quad \frac{\partial u}{\partial n} + cu = h \text{ on } \Gamma_n \quad (67b)$$

where $u(\mathbf{x})$ is the temperature field and $f(\mathbf{x})$ is the heat source. The boundary $\partial\Omega = \Gamma_u \cup \Gamma_n$ is partitioned into two parts, with $\Gamma_u \cap \Gamma_n = \emptyset$. The temperature field $g(\mathbf{x})$ is imposed on the essential boundary Γ_u , and boundary data $h(\mathbf{x})$ that is associated with a Robin boundary condition is prescribed on Γ_n (c is in general a spatially varying field).

10.2.1. Essential boundary conditions

As the first example, we consider the biunit square, $\Omega = (-1, 1)^2$, with $u = g = 0$ prescribed on $\partial\Omega$. If $k \in \mathbb{N}$ and $f(x, y) = \sin(k\pi x)\sin(k\pi y)$ is the forcing function, then the exact solution for this problem is:

$$u(\mathbf{x}) = \frac{\sin(k\pi x)\sin(k\pi y)}{2k^2\pi^2}.$$

In the numerical computations, we consider two distinct forms of trial functions in the neural network. The first ansatz is the standard PINN that is given by $\tilde{u}_{nn}(\mathbf{x}; \theta)$, which does not a priori satisfy the boundary condition. The second form consists of trial functions that are given by $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta) = \phi(\mathbf{x})\tilde{u}_{nn}^R(\mathbf{x}; \theta)$, where $\phi(\mathbf{x})$ is a function that is zero on $\partial\Omega$. This property of $\phi(\mathbf{x})$ ensures that $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ automatically satisfies the essential boundary conditions. While an obvious choice for $\phi(\mathbf{x})$ is $(1 - x^2)(1 - y^2)$, here we consider $\phi(\mathbf{x})$ that are constructed using R-functions with R-equivalence composition (see Section 3.2) and by mean value potential fields (see Section 4) as they readily generalize to more complex domains. When needed, we use the acronyms REQ and MVP to distinguish the numerical solutions, $\tilde{u}_{nn}^{bc}(\mathbf{x})$, which are obtained using these two methods. The plot of $\phi(\mathbf{x})$ using REQ and MVP over a square are shown in Figs. 7b and 10a, and it can be observed that ϕ is zero on the boundary of the domain. For the collocation scheme, we randomly sample N_I number of points in Ω and N_B number of points on $\partial\Omega$. To solve the problem using standard PINN, we minimize the loss $L_{nn}(\theta)$ given in (41), which is reproduced below (using $g = 0$):

$$L_{nn}(\theta) = \|\nabla^2\tilde{u}_{nn}(\mathbf{x}; \theta) + f(\mathbf{x})\|_{\Omega, N_I}^2 + \|\tilde{u}_{nn}(\mathbf{x}; \theta)\|_{\Omega, N_B}^2,$$

where $\|\cdot\|_{\Omega, N_I}$ and $\|\cdot\|_{\partial\Omega, N_B}$ are defined in (40) and (41). Since $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ automatically satisfies the boundary condition, the parameters in this ansatz are found by minimizing the loss given in (40), which for this problem is:

$$L_{nn}^{bc}(\theta) = \|\nabla^2\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta) + f(\mathbf{x})\|_{\Omega, N_I}^2.$$

Fig. 28 shows the training loss and normalized absolute error as functions of the training epochs for $k = 1, 2$ ($N_I = 5,000$, $N_B = 400$). We observe from Fig. 28a and c that the training loss for $\tilde{u}_{nn}(\mathbf{x}; \theta)$ is either comparable to or less than the loss for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ over the same number of epochs. However, this does not translate into better prediction accuracy. Fig. 28b and d show the normalized absolute errors for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ (REQ and MVP) as well as $\tilde{u}_{nn}(\mathbf{x}; \theta)$. These plots reveal that both REQ- and MVP-based schemes deliver an order of magnitude more accurate solutions compared to $\tilde{u}_{nn}(\mathbf{x}; \theta)$.

It is interesting to note that $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ produces smaller normalized absolute errors during training than $\tilde{u}_{nn}(\mathbf{x}; \theta)$ even though it has larger losses. This observation is noticed in almost all cases, and it deserves some comments here. We mention that there is no reason to assume that the losses for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ and $\tilde{u}_{nn}(\mathbf{x}; \theta)$ should be comparable. For

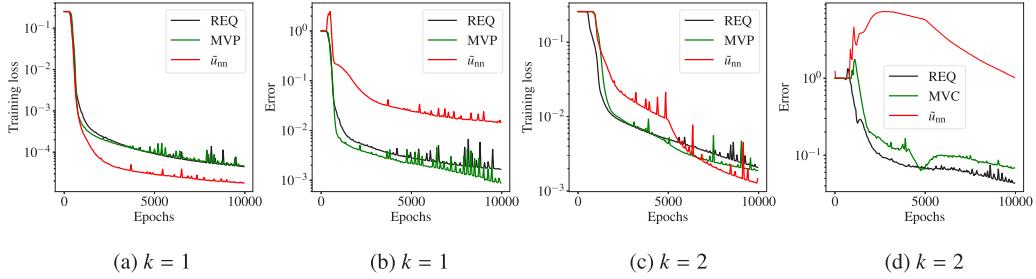


Fig. 28. Training loss and normalized absolute errors for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ and MVP) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ in the heat conduction problem with homogeneous Dirichlet boundary conditions and forcing function $\sin k\pi x$. (a), (b) $k = 1$, and (c), (d) $k = 2$.

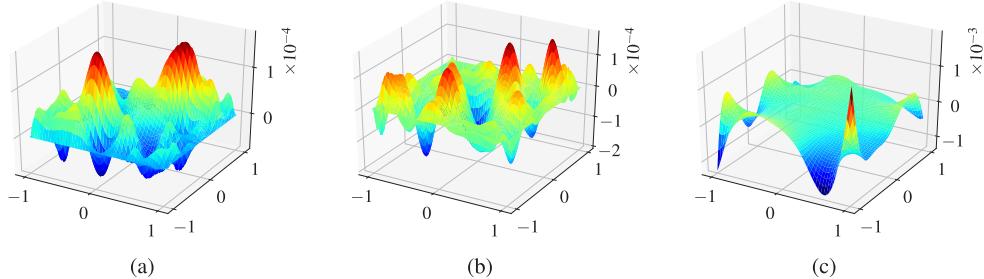


Fig. 29. Prediction errors for $k = 1$ in Example 1. (a) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (REQ), (b) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (MVP), and (c) $\tilde{u}_{nn}(\mathbf{x})$.

the problem under consideration, $L_{nn}^{bc}(\boldsymbol{\theta})$ comprises of terms that involve the derivatives of $\phi(\mathbf{x})$ and these terms are not present in $L_{nn}(\boldsymbol{\theta})$. A further issue with $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ is the relative scaling of the losses in $L_{nn}(\boldsymbol{\theta})$, which comprises of a loss on $\nabla^2 \tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ and a loss on $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. In this problem, the norm of u is much smaller than the norm of its Laplacian, and therefore the $\nabla^2 \tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ term dominates in $L_{nn}(\boldsymbol{\theta})$. However, the optimizer can drive $L_{nn}(\boldsymbol{\theta})$ to very small values without adequately addressing the boundary loss term. We can see that this is indeed the issue if we compare the prediction errors for $\tilde{u}_{nn}(\mathbf{x})$ over the domain (see Fig. 29). We see that the prediction errors for REQ and MVP schemes are much smaller than $\tilde{u}_{nn}(\mathbf{x})$, and furthermore, the prediction errors for $\tilde{u}_{nn}(\mathbf{x})$ are large near the boundary of the domain. This suggests that $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ is undervaluing the boundary loss term in $L_{nn}(\boldsymbol{\theta})$. In principle, it is possible to improve the results for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ by assigning a larger weight to the boundary loss term in $L_{nn}(\boldsymbol{\theta})$ [16], but this is an ad-hoc remedy, which is not needed in our approach since the boundary condition is exactly met. To see that this is indeed the case, one can assume the loss function for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ to be a convex combination of the interior and boundary loss terms:

$$L_{nn}(\boldsymbol{\theta}) = \underbrace{w \|\nabla^2 \tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta}) + f(\mathbf{x})\|_{\Omega, N_I}^2}_{\text{PDE loss}} + \underbrace{(1-w) \|\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})\|_{\partial\Omega, N_B}^2}_{\text{Boundary loss}},$$

where $w \in [0, 1]$ is a scalar that can be used to tune the relative weights of the two losses.

Fig. 30a–c show the evolution of the PDE loss and the boundary loss for different values of w ($k = 1$). The value $w = 0.1$ weighs the boundary loss term 9 times more than the PDE loss term and it achieves smaller error than both $w = 0.5, 0.9$ (Fig. 30d). This suggests that there likely is a sweet spot for w that results in very low errors. However, the exact regime for this solution is likely dependent on the problem under consideration and on the boundary conditions, the determination of which may be impossible to ascertain in problems where the exact solution is unknown. A distinguishing attribute of our approach is that no such tuning of relative weights is needed.

10.2.2. Example 2

We consider the Laplace equation ($f = 0$) over the unit square, $\Omega = (0, 1)^2$, with boundary conditions

$$u(\mathbf{x}) = 0 \quad \text{on } \Gamma_1, \Gamma_2, \Gamma_3, \quad u(\mathbf{x}) = g_4(\mathbf{x}) = \sin \pi x \quad \text{on } \Gamma_4, \quad (68)$$

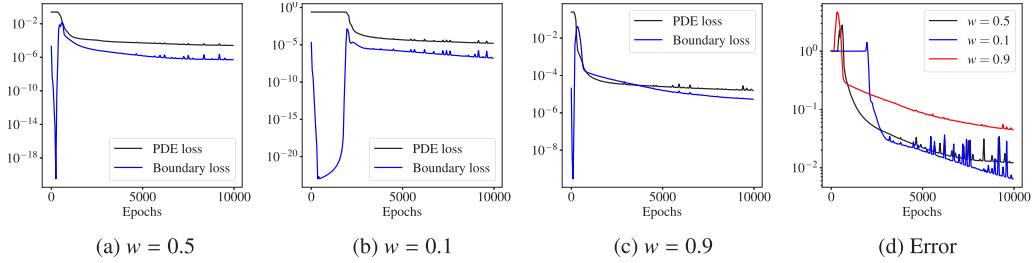


Fig. 30. Loss function for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ in Example 1 is a convex combination of PDE loss and boundary loss terms ($k = 1$). Evolution of loss function for (a) $w = 0.5$, (b) $w = 0.1$ and (c) $w = 0.9$. (d) Evolution of normalized absolute error during training for $w = 0.5, 0.1, 0.9$.

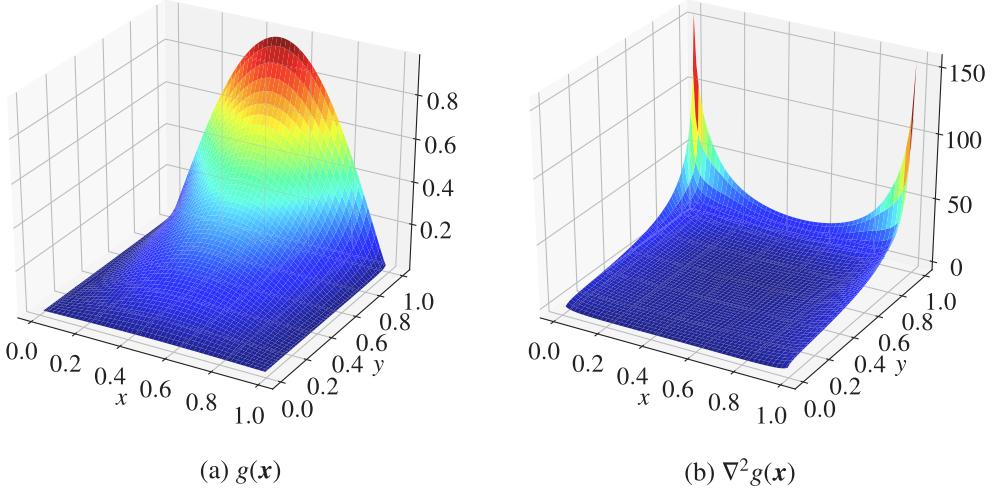


Fig. 31. Plots of (a) $g(\mathbf{x})$ and (b) $\nabla^2 g(\mathbf{x})$ over the unit square.

where $\Gamma_1 = \{(x, y) : x = 0, 0 \leq y \leq 1\}$, $\Gamma_2 = \{(x, y) : 0 \leq x \leq 1, y = 0\}$, $\Gamma_3 = \{(x, y) : x = 1, 0 \leq y \leq 1\}$, and $\Gamma_4 = \{(x, y) : 0 \leq x \leq 1, y = 1\}$ are the boundary edges. The exact solution for this problem is:

$$u(\mathbf{x}) = \frac{(e^{-\pi y} + e^{\pi y}) \sin \pi x}{e^{-\pi} + e^{\pi}}.$$

We chose this problem to demonstrate how to exactly satisfy nonzero essential boundary conditions on different subsets of the boundary through the use of transfinite interpolation. To construct a trial solution that satisfies the boundary conditions, we first create a composite approximate distance function, $\phi(\mathbf{x})$, to $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$. This ADF can either be formed by the joining operation via R-equivalence, or directly via (14) that uses mean value potential fields on polygons. The resultant $\phi(\mathbf{x})$ is similar to the $\phi(\mathbf{x})$ used in Example 1 (see Fig. 27). We combine the Dirichlet boundary data into one function $g(\mathbf{x})$ by using the transfinite interpolant in (26).

In Fig. 31, the function $g(\mathbf{x})$ and its Laplacian are plotted over the unit square. We observe that $g(\mathbf{x})$ is zero on Γ_α ($\alpha = 1, 2, 3$), and it is equal to $\sin \pi x$ on Γ_4 . Referring to (27), the trial function for PINN is:

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta}) = g(\mathbf{x}) + \phi(\mathbf{x}) \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}).$$

Since $g(\mathbf{x})$ satisfies the boundary conditions in (68) and $\phi = 0$ on Γ , $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ satisfies the Dirichlet boundary conditions on all edges of the boundary. From Fig. 31b, we observe that the Laplacian of $g(\mathbf{x})$ is singular at two of the four vertices on the boundary and again this is handled by performing collocation over the smaller square $[0.01, 0.99]^2$.

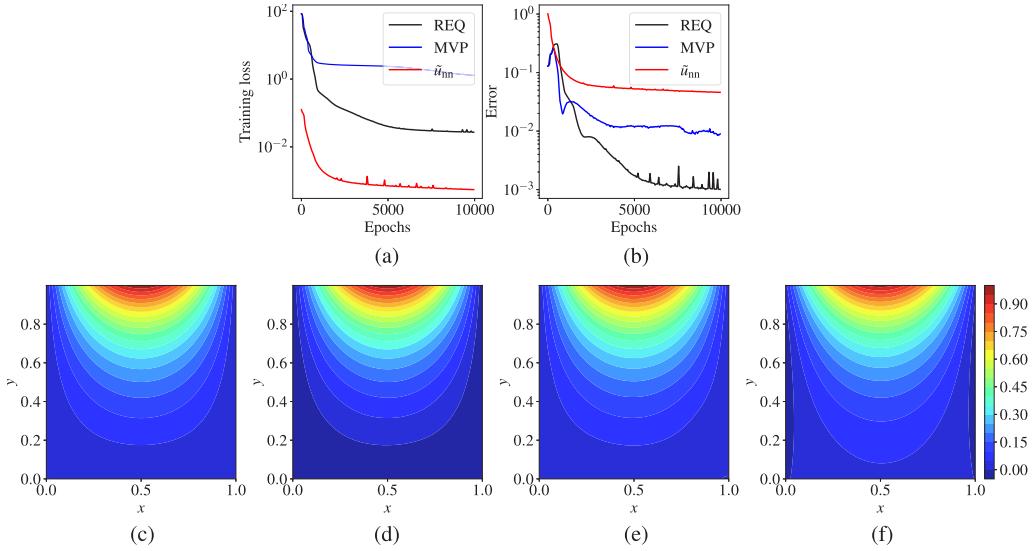


Fig. 32. Numerical results for the Laplace problem on the unit square with nonzero essential boundary conditions. (a), (b) Training loss and normalized absolute errors for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ and MVP) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. Contour plots over the unit square of the (c) exact solution, (d) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (REQ), (e) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (MVP), and (f) $\tilde{u}_{nn}(\mathbf{x})$.

We determine the parameters of this network by minimizing $L_{nn}^{bc}(\boldsymbol{\theta})$ as described earlier (with f set to zero). For standard PINN, the parameters of $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ are determined by minimizing the following loss function:

$$L_{nn}(\boldsymbol{\theta}) = \|\nabla^2 \tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})\|_{\Omega, N_I}^2 + \sum_{\alpha=1}^4 \|\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta}) - g_\alpha(\mathbf{x})\|_{\Gamma_\alpha, N_B^\alpha}^2,$$

where N_B^α is the number of collocation points on Γ_α ($\alpha = 1, 2, 3, 4$) and $g_\alpha(\mathbf{x}) = 0$ ($\alpha = 1, 2, 3, 4$) in this problem. In the computations, we select a total of 400 boundary collocation points and 5000 interior collocation within $[.01, .99]^2$. The numerical results are shown in Fig. 32, where the training loss, the normalized absolute error, and exact and approximate solutions are presented for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. As in the previous example, here also we notice that the training loss for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ is orders of magnitude smaller than both REQ and MVP (Fig. 32a). The losses for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ start at relatively high values during the initial stages of the training, which is due to the large contributions from the Laplacian in the vicinity of the vertices. However, the network is quickly able to optimize and bring the losses down by almost two orders of magnitude in a couple of thousand training epochs. Still, the losses for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ remain several orders of magnitude larger than $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ at the end of the training. However, this example also reveals that the absolute value of the loss in itself is not very meaningful. The normalized absolute errors of the three schemes are presented in Fig. 32b as a function of the training epochs. It is evident from this plot that the errors in $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ are orders of magnitude smaller than the error in $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. The error achieved by $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ) is almost an order of magnitude smaller than $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (MVP), and almost two orders of magnitude better than $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. Contour plots of the exact and approximate solutions appear in Fig. 32c–f.

The predicted errors of the three schemes are displayed in Fig. 33. It can be seen that the boundary errors in both REQ and MVP are precisely zero. This is expected, since $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ has been designed to satisfy the boundary conditions. On the other hand, $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ has large errors on the boundary of the domain. The errors from $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ are roughly an order of magnitude larger than MVP and two orders of magnitude larger than REQ. As in the previous example, one can improve the results of $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ by weighing the boundary loss more than the PDE loss in $L_{nn}(\boldsymbol{\theta})$.

This problem can also be solved using a Ritz scheme, which is appealing since only first-order derivatives are required in the loss function. We form a trial function that satisfies the essential conditions using (27):

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta}) = g(\mathbf{x}) + \phi(\mathbf{x}) \tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta}).$$

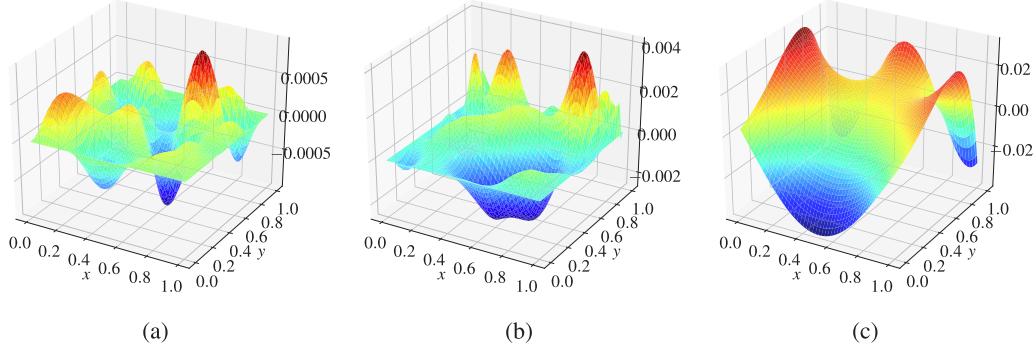


Fig. 33. Surface plots of the errors in the numerical solutions for the Laplace problem with nonzero essential boundary condition. (a) $\tilde{u}_{nn}^{bc}(x)$ (REQ), (b) $\tilde{u}_{nn}^{bc}(x)$ (MVP), and (c) $\tilde{u}_{nn}(x)$.

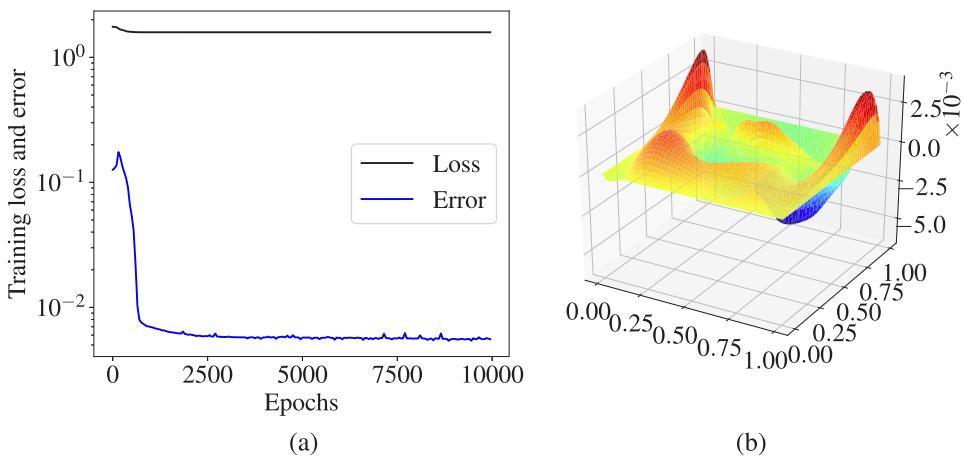


Fig. 34. Ritz solution using $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ) for the Laplace problem in Example 2. (a) Evolution of training loss and normalized absolute error, and (b) Prediction errors over the domain after training.

The parameters of the network can now be found by minimizing the loss in (47). To numerically evaluate the integral, we divide the square into a uniform grid with N_I number of interior points. Since the Ritz loss does not involve second derivatives of $\phi(\mathbf{x})$ or $g(\mathbf{x})$, all terms in the loss are well-defined and bounded even arbitrarily close to the boundaries of the domain. We select 5000 interior points on the square $[0.0001, 0.9999]^2$. For a Dirichlet problem, it is especially important to sample close to the boundaries, because in the absence of doing so, the loss may be trivially minimized by a $u(\mathbf{x})$ that is a constant. Sampling close to the boundaries informs the algorithm that a constant $u(\mathbf{x})$ leads to large errors near the boundaries, which are manifested in the loss term as large gradients. Numerical results for the Ritz method using $\tilde{u}_{nn}(\mathbf{x}; \theta)$ with REQ are presented in Fig. 34. The training loss and normalized absolute errors as a function of epochs are shown in Fig. 34a, which reveal that the error reduces to $\mathcal{O}(10^{-2})$ in less than 2000 epochs. In Fig. 34b, the prediction errors over the square are displayed. Compared to Fig. 33, we find that the errors in the Ritz scheme are smaller than MVP-based collocation but larger than REQ-based collocation.

10.2.3. Curved domain

For a problem with a curved domain, we consider the Laplace equation, $\nabla^2 u = 0$, in an annulus that is bounded between circles of radii, $R_1 = 1$ (boundary Γ_1) and $R_2 = 1/4$ (boundary Γ_2) [71]. The essential boundary conditions are:

$$u = 1 \quad \text{on } \Gamma_1, \quad u = 2 \quad \text{on } \Gamma_2.$$

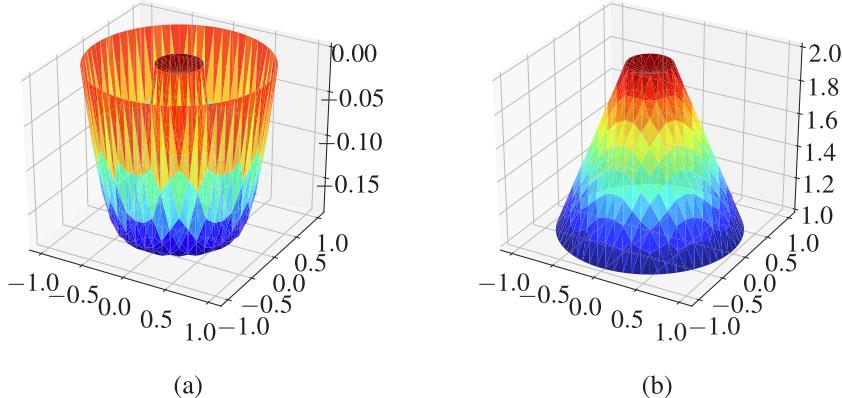


Fig. 35. Approximate distance function to the boundaries of the annulus using R-equivalence composition of the exact distance functions to Γ_1 and Γ_2 . Plots of (a) $-\phi(\mathbf{x})$ and (b) $g(\mathbf{x})$ that interpolate essential boundary data on Γ_1 and Γ_2 .

The exact solution for this problem is:

$$u(\mathbf{x}) = 1 - \frac{\ln \sqrt{x^2 + y^2}}{\ln 4}. \quad (69)$$

To impose the boundary conditions, we need a distance function to both Γ_1 and Γ_2 , and also a composite boundary data function $g(\mathbf{x})$. In this case, on using the exact distance functions to the two circles, we have $\phi_1(\mathbf{x}) = 1 - \sqrt{x^2 + y^2}$ (positive in the interior of the larger disk), and $\phi_2(\mathbf{x}) = \sqrt{x^2 + y^2} - 1/4$ (positive outside the smaller disk). Since the origin is not part of the computational domain, here we can use the exact distance functions to form ϕ_1 and ϕ_2 . Now, on combining these two ADFs using the R-equivalence operation ($m = 1$) in (9), we obtain $\phi(\mathbf{x}) = \phi_1 \sim \phi_2$ (positive in the annulus). Finally, we use the transfinite formula (26) to construct the composite boundary data as

$$g(\mathbf{x}) = \frac{2\phi_1 + \phi_2}{\phi_1 + \phi_2} = \frac{7 - 4\sqrt{x^2 + y^2}}{3}.$$

In Fig. 35, $\phi(\mathbf{x})$ and $g(\mathbf{x})$ are plotted over the annulus. To clearly see that $\phi(\mathbf{x})$ is zero on Γ_1 and Γ_2 , we show $-\phi(\mathbf{x})$ in Fig. 35a. From the plot in Fig. 35b, we observe that $g(\mathbf{x})$ matches the imposed boundary data on Γ_1 and Γ_2 . An ansatz that exactly satisfies the boundary conditions is: $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta}) = g(\mathbf{x}) + \phi(\mathbf{x})\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})$. We compare the performance of $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ with a standard PINN trial function, $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$, where the boundary conditions need to be enforced through the loss function. The loss function for the two cases are similar to those discussed in previous examples, the only difference being that now we sample the boundary data at N_B points on the curved boundary $\Gamma = \Gamma_1 \cup \Gamma_2$ and the interior collocation data at N_I points in $\Omega_1 \cap \Omega_2$. We use dmsh [95] to triangulate the annulus and choose the centroid of the triangles as interior collocation points and the center of the edges on the boundary as the boundary collocation points. For this problem, we pick $N_I = 612$ points in the interior of the domain, 66 points on Γ_1 and 30 points on Γ_2 for a total of $N_B = 96$ points on $\Gamma_1 \cup \Gamma_2$. A representative mesh that display the interior collocation points over the annulus is shown in Fig. 15. We use the Adam optimizer with a learning rate (step size) of 10^{-3} for training both $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$, and the training is stopped at 10,000 epochs for both networks in order to perform a fair comparison.

In Fig. 36, the results of the training as well as the approximate solutions produced by $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ are presented. It is pertinent to mention here that $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ required a much larger network architecture (2-150-150-1) compared to $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$, which only required a 2-50-50-1 network in order to converge to acceptable results. However, even with a much larger network, as observable in Fig. 36a and b, the error in $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ by the end of the training is two orders of magnitude larger than $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$. It can be seen from Fig. 36e that $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ does not satisfy the boundary conditions. The approximation errors of $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (REQ) and $\tilde{u}_{nn}(\mathbf{x})$ appear in Fig. 36f and g. The errors in $\tilde{u}_{nn}(\mathbf{x})$ are especially large on the boundaries of the domain.

We now use the Laplace problem over the annulus to also demonstrate how to solve the problem using mixed boundary conditions. To this end, we retain the essential boundary condition on Γ_1 and convert the boundary

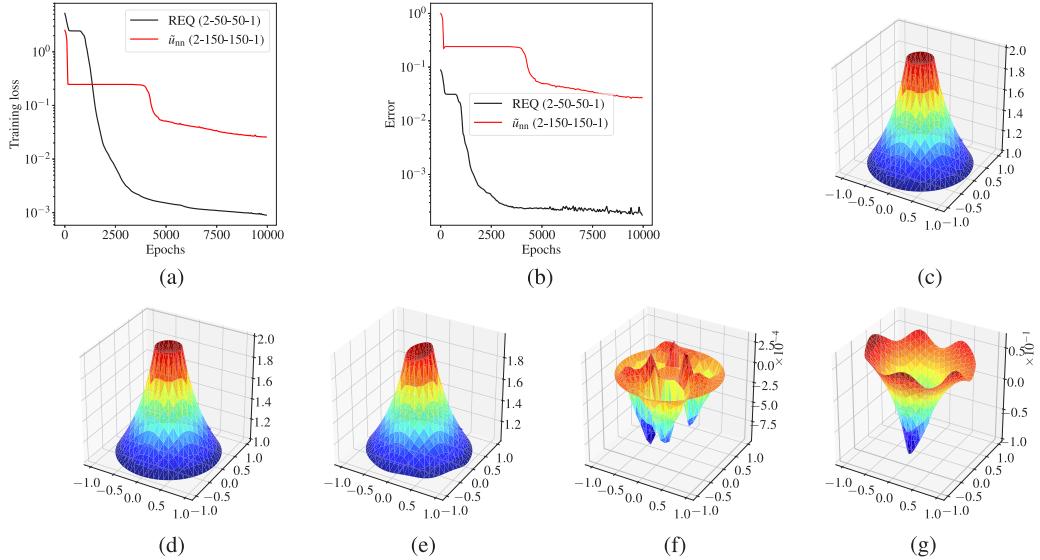


Fig. 36. Numerical solutions for the Laplace problem on an annulus with Dirichlet boundary conditions. (a), (b) Training loss and normalized absolute errors of $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. Surface plots of the (c) exact solution, (d) $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ using 2-50-50-1 network, and (e) $\tilde{u}_{nn}(\mathbf{x})$ using 2-150-150-1 network. Surface plots of the error for the numerical solutions (f) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ and (g) $\tilde{u}_{nn}(\mathbf{x})$.

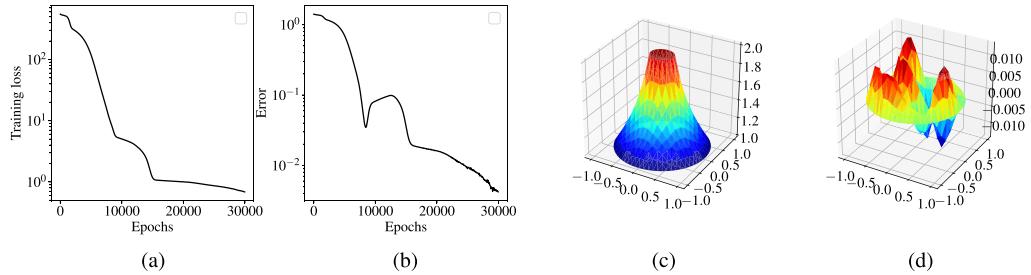


Fig. 37. Numerical solution using $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ) for the Laplace problem in an annulus with mixed boundary conditions. (a), (b) Training loss and normalized absolute errors as a function of epochs. Surface plots of (c) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ and (d) error in $\tilde{u}_{nn}^{bc}(\mathbf{x})$.

condition on Γ_2 to a Robin boundary condition. For this problem, we find that $\partial u / \partial n = 4 / \ln 4$ on the inner boundary. We use the following mixed boundary conditions:

$$u = 1 \quad \text{on } \Gamma_1, \quad \frac{\partial u}{\partial n} + u = 2 + \frac{4}{\ln 4} =: h \quad \text{on } \Gamma_2. \quad (70)$$

The exact solution remains unchanged and is given in (69). To create an ansatz for this problem with mixed boundary conditions, we follow the formulation in Section 5.3. We form ϕ_1, ϕ_2 , which remain unchanged from the previous case when essential boundary conditions are imposed on Γ_1 and Γ_2 . Since $g = 1$, referring to (35), we can write

$$u_1(\mathbf{x}) = g(\mathbf{x}) = 1, \quad u_2(\mathbf{x}) = \left[1 + \phi_2(1 + D_1^{\phi_2}) \right] (\tilde{u}_{nn}^R(\mathbf{x}; \boldsymbol{\theta})) - \phi_2 h, \quad (71)$$

where h is given in (70), and then on using transfinite interpolation given in (35a) and (35b), we form the trial function $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$. So we now have an ansatz that satisfies both the Dirichlet and Robin boundary conditions. The training loss and normalized absolute errors for the numerical solution using $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ are presented in Fig. 37a and b. On comparing $\tilde{u}_{nn}^{bc}(\mathbf{x})$ in Fig. 37c to the exact solution shown in Fig. 36, we see that the boundary conditions have been exactly satisfied. The error plot in Fig. 37d reveals that the numerical solution is within 1 percent of the exact solution over the domain.

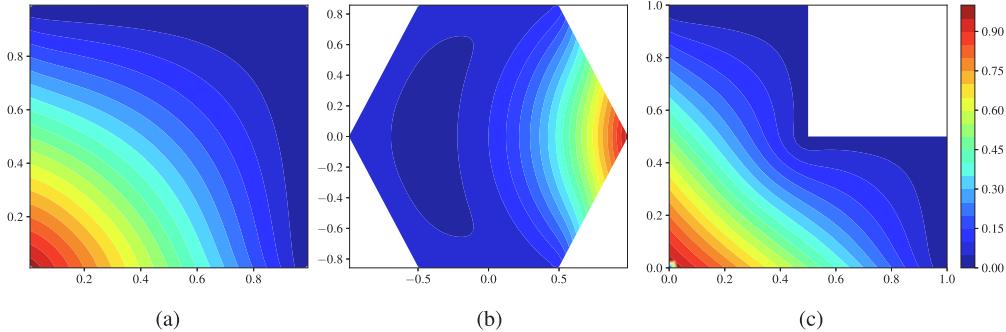


Fig. 38. Computation of harmonic coordinate for a vertex in a polygon. Contour plots of $g(\mathbf{x})$ over a (a) square (vertex at the origin), (b) regular hexagon (rightmost vertex), and (c) L-shaped polygon (vertex at the origin). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

10.3. Generalized barycentric coordinates over polygons

Consider a planar polygon with n vertices (nodal coordinate $\{\mathbf{x}_i\}_{i=1}^n$) that are in counterclockwise orientation. On an n -gon, harmonic coordinates [99] are one of the instances of generalized barycentric coordinates [79]. Each coordinate (shape function), $\varphi_i := \varphi_i(\mathbf{x})$, is associated with vertex i and is obtained by solving the Laplace equation with piecewise affine Dirichlet boundary conditions. The boundary-value problem for harmonic coordinates is: find $\varphi_i \geq 0$ ($i = 1, 2, \dots, n$) that solves

$$\nabla^2 \varphi_i = 0 \quad \text{in } \Omega, \tag{72a}$$

$$\varphi_i = g_i \quad \text{on } \partial\Omega, \tag{72b}$$

where $g_i := g_i(\mathbf{x})$ is a piecewise affine (hat) function that is unity at \mathbf{x}_i and is zero at all other vertices, i.e., $g_i(\mathbf{x}_j) = \delta_{ij}$, where δ_{ij} is the Kronecker-delta. By virtue of the maximum principle for the Laplace equation, $\varphi_i > 0$ in the interior of the polygon.

Here we will solve the harmonic coordinate problem on two representative polygons: a square and an L-shaped (nonconvex) polygon. For both examples, the network architecture 2-50-50-1 is used. In both examples, an important step is to assemble the boundary data into a function $g(\mathbf{x})$ through transfinite interpolation. Fig. 38 shows the function $g(\mathbf{x})$ for specific choices of the vertex i . If the vertices of the square are numbered 1-2-3-4 (counterclockwise), starting at vertex 1 that is at $(0, 0)$, then the harmonic coordinate $u(\mathbf{x})$ must satisfy $u(0, 0) = 1$, $u(1, 0) = u(1, 1) = u(0, 1) = 0$. So the boundary conditions are affine along edges 1-2 and 4-1 and identically zero along edges 2-3 and 3-4. All these boundary conditions are simultaneously captured in the $g(\mathbf{x})$ function. It can be seen from the colormaps that $g(\mathbf{x})$ appropriately interpolates the boundary data in all cases.

10.3.1. Harmonic coordinates on a square

We first consider the case of computing the harmonic coordinate over a square. On a square, there exists an exact solution for this problem—harmonic coordinates coincide with bilinear finite element shape functions. So the solution that is associated with vertex 1 is: $u(x, y) = (1 - x)(1 - y)$. To compute the harmonic coordinates, we adopt the Ritz method to determine the approximate solutions. Since harmonic coordinates minimize the Dirichlet energy, use of the Ritz formulation is natural. As done earlier, we consider the approximations $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. For $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$, the loss to be minimized is given in (47), whereas for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$, it is supplemented with an additional term $\sum_{\alpha=1}^n \|\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta}) - g_\alpha(\mathbf{x})\|_{\partial\Omega_\alpha, N_R^\alpha}^2$ to impose the essential boundary conditions. Here, n is the number of boundary segments over which the Dirichlet boundary conditions are specified.

In Fig. 39, numerical results using $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ are presented. In Fig. 39a and b, the training loss and normalized absolute error as a function of epochs are presented. The exact solution along with the numerical solutions are displayed in Fig. 39c–e. Surface plots of the errors in $\tilde{u}_{nn}^{bc}(\mathbf{x})$ and $\tilde{u}_{nn}(\mathbf{x})$ appear in Fig. 39f and g. Once again, and consistent with prior findings, $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ exactly satisfies the essential boundary conditions and has far smaller errors than $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. The large errors in $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ are particularly noticeable in Fig. 39g. The issue again

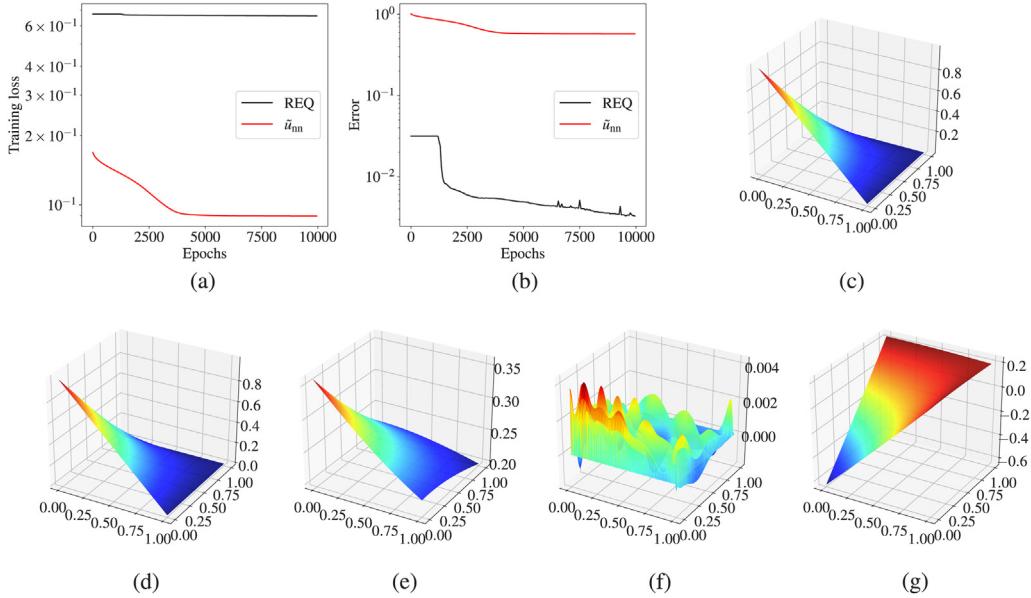


Fig. 39. Computation of harmonic coordinates on the unit square for the vertex at the origin. (a), (b) Training loss and errors for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. Surface plots of the (c) exact solution, (d) $\tilde{u}_{nn}^{bc}(\mathbf{x})$, and (e) $\tilde{u}_{nn}(\mathbf{x})$. Surface plot of the error for the numerical solutions (f) $\tilde{u}_{nn}^{bc}(\mathbf{x})$ and (g) $\tilde{u}_{nn}(\mathbf{x})$.

has to do with the relative scaling of the PDE loss versus the boundary loss. As done earlier, the numerical results from $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ can be improved by considering a loss of the form:

$$L_{nn}(\boldsymbol{\theta}) = w \left\{ \frac{1}{N_I} \sum_{k=1}^{N_I} \left[\frac{1}{2} a \left(\tilde{u}_{nn}(\mathbf{x}_k; \boldsymbol{\theta}), \tilde{u}_{nn}(\mathbf{x}_k; \boldsymbol{\theta}) \right) - \ell \left(\tilde{u}_{nn}(\mathbf{x}_k; \boldsymbol{\theta}) \right) \right] \right\} + (1-w) \left\{ \sum_{\alpha=1}^n \| \tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta}) - g_\alpha(\mathbf{x}) \|_{\partial \Omega_\alpha, N_B^\alpha}^2 \right\},$$

and then tuning the weight $w \in [0, 1]$. This tunes the relative importance of the boundary loss term with respect to the PDE loss term. In fact, if one considers $w = 10^{-3}$, $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ achieves errors that are comparable to $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$. In other words, the boundary loss in $L_{nn}(\boldsymbol{\theta})$ has to be weighed a thousand times more than the PDE loss in order for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ to produce results comparable to $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$.

10.3.2. Harmonic coordinates on an L-shaped polygon

We repeat the computations for the square over an L-shaped polygon. Here we consider $\phi(\mathbf{x})$ that is formed using REQ and MVP. The plot of $\phi(\mathbf{x})$ for REQ and MVP over the L-shaped polygon are shown in Figs. 7b and 10b. The function $g(\mathbf{x})$ over the L-shaped polygon is shown in Fig. 38c. We only present numerical simulation results for REQ and MVP. As shown for the case of the square, the loss terms have to be weighed judiciously to obtain acceptable accuracy for $\tilde{u}_{nn}(\mathbf{x})$. Harmonic coordinates associated with the vertex at the origin and for the vertex at the reentrant corner are computed. Since an exact analytical solution for this problem is not available, we compute an accurate finite element solution that we use as the reference solution. This finite element solution is used to compute the errors in $\tilde{u}_{nn}^{bc}(\mathbf{x})$. A Delaunay triangular mesh is created using the mesh generation package Triangle [100]: mesh has 13,952 elements with very small elements in the vicinity of the reentrant corner and larger elements near other vertices. The mesh size, $h = 10^{-3}$, is used near the reentrant corner to capture the weakly singular behavior of the Laplace equation at the reentrant corner.

A representative mesh that displays the interior collocation points in the L-shaped polygon is shown in Fig. 15. In Figs. 40 and 41, numerical solutions obtained from $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ and MVP) are presented for the computation of u that is associated with vertices at $(0, 0)$ and $(1/2, 1/2)$, respectively. In Fig. 40a, we observe that the training

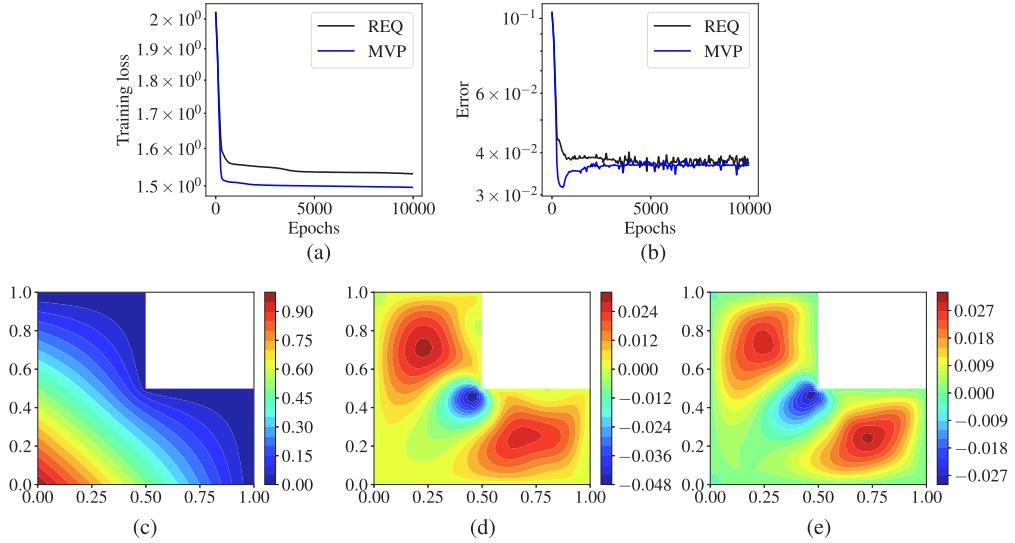


Fig. 40. Computation of harmonic coordinates (vertex at (0, 0)) on an L-shaped polygon. (a), (b) Training loss and normalized absolute error for $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ and MVP). (c) Reference finite element solution. Contour plots of the error for (d) $\tilde{u}_{nn}^{bc}(x)$ (REQ) and (e) $\tilde{u}_{nn}^{bc}(x)$ (MVP).

loss stabilizes within a few thousand epochs to about 1.5 for REQ and MVP, and this correspond to a normalized absolute error (see Fig. 40b) of $\mathcal{O}(10^{-2})$. The reference finite element solution is presented in Fig. 40c, and the error in $\tilde{u}_{nn}^{bc}(x)$ (REQ and MVP) are displayed in Fig. 40d and e. Numerical solutions using REQ and MVP have maximum errors of about 3 percent. From Fig. 41a, we observe that the training loss stabilizes within a few thousand epochs to about 8 for REQ and MVP, and this correspond to a normalized absolute error (see Fig. 41b) of $\mathcal{O}(10^{-1})$. The reference finite element solution is depicted in Fig. 41c, which display sharp gradients near the reentrant corner. The error in $\tilde{u}_{nn}^{bc}(x)$ (REQ and MVP) are shown in Fig. 41d and e, with maximum errors near the singularity on the order of 20 percent. Compared to the errors in Fig. 40d and e, this is a 10-fold increase in the maximum error (due to the presence of the derivative singularity at the reentrant corner).

10.4. Clamped circular Kirchhoff plate

We consider the boundary-value problem for a clamped plate that is given in (48). For a clamped circular plate of unit radius and transverse load $f = 1$, the boundary-value problem is:

$$\nabla^4 u = 1 \quad \text{in } \Omega = \{(x, y) : x^2 + y^2 < 1\}, \quad (73a)$$

$$u = 0 \quad \text{on } \partial\Omega, \quad u, n := \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial\Omega. \quad (73b)$$

The exact solution for this problem in polar coordinates is [101]:

$$u(r) = \frac{(1 - r^2)^2}{64}. \quad (74)$$

Given that both u and $\partial u / \partial n$ are specified on the boundary, this problem illustrates the use of a different solution structure than the ones considered until now. To impose the boundary conditions, we first create a distance function ϕ to the circular boundary. In this case, the distance function can be exactly determined as $\phi(\mathbf{x}) = 1 - \sqrt{x^2 + y^2}$; however, this exact distance function has derivative singularities at the origin. Therefore, we use an approximate distance function to a unit circle that is given in (7), which is reproduced below:

$$\phi(\mathbf{x}) = \frac{1 - \mathbf{x} \cdot \mathbf{x}}{2},$$

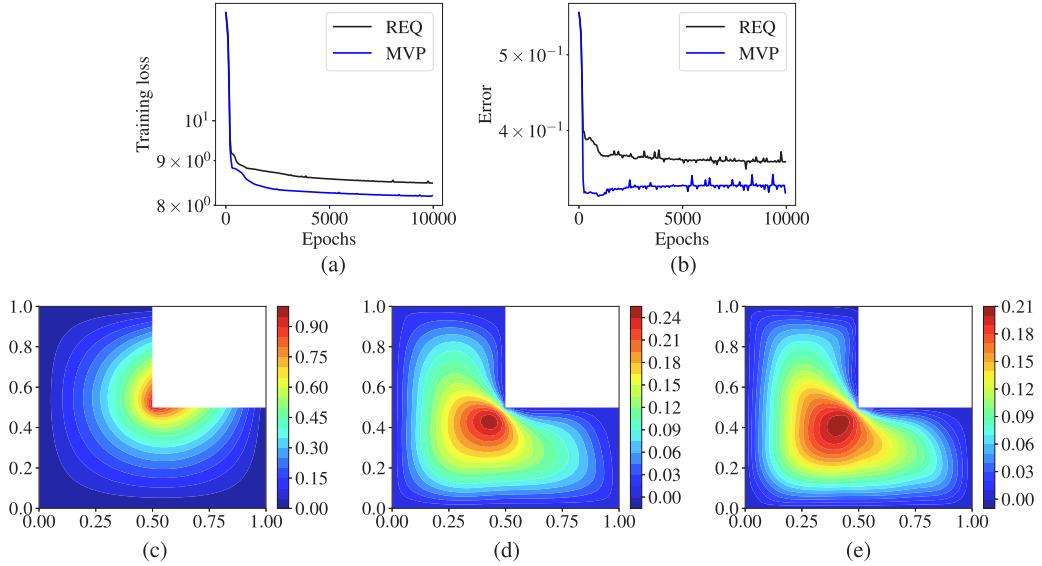


Fig. 41. Computation of harmonic coordinates (vertex at $(1/2, 1/2)$) on an L-shaped polygon. (a), (b) Training loss and normalized absolute error for $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ and MVP). (c) Reference finite element solution. Contour plots of the error for (d) $\tilde{u}_{nn}^{bc}(x)$ (REQ) and (e) $\tilde{u}_{nn}^{bc}(x)$ (MVP).

which is a bivariate polynomial. Now we construct an ansatz that satisfies both essential boundary conditions using (50): $\tilde{u}_{nn}^{bc}(x; \theta) = \phi^2 \tilde{u}(x; \theta)$. As with previous problems, this problem can be solved using either the collocation approach (see Guo et al. [102]) or the Ritz method. For collocation, we minimize the loss function

$$L_{nn}^{bc}(\theta) = \|\nabla^4 \tilde{u}_{nn}^{bc}(x; \theta) - 1\|_{\Omega, N_f}^2,$$

whereas for the Ritz approach, we minimize the loss function that is presented in Section 7.2.2.

Numerical results for collocation and Ritz using $\tilde{u}_{nn}^{bc}(x; \theta)$ are presented in Fig. 42. In the computations, 2800 interior points are used for both methods. For both collocation and Ritz we use the cubic ReLU activation function. Note that for collocation with standard PINN one cannot use the cubic ReLU since it is a biharmonic function. The trial function $\tilde{u}_{nn}^{bc}(x; \theta)$ has other terms that are present in it, and hence in general it is not biharmonic ($\nabla^4 \tilde{u}_{nn}^{bc} \neq 0$). A representative mesh that displays the interior points over a disk is shown in Fig. 15. The network architecture is 2-50-50-1. The training loss as a function of epochs for the collocation and Ritz methods are shown in Fig. 42a. The exact solution is shown in Fig. 42b, and the collocation and Ritz solution for $\tilde{u}_{nn}^{bc}(x)$ after 10,000 epochs are plotted in Fig. 42c and d. The surface plots of the error in the Ritz and collocation solutions are presented in Fig. 42e and f. We observe that the unnormalized error in the Ritz and collocation methods are $\mathcal{O}(10^{-5})$; the latter is consistent with the accuracy reported in Guo et al. [102].

10.5. Eikonal equation

We consider the Eikonal equation, which is a first-order nonlinear hyperbolic PDE. The boundary-value problem of the Eikonal equation is:

$$\|\nabla u\| = \frac{1}{f} \quad \text{in } \Omega \subset \mathbb{R}^2, \tag{75a}$$

$$u = 0 \quad \text{on } \Gamma, \tag{75b}$$

where Γ is an interface in two dimensions and $f(x) > 0$ is the speed on the interface. If $f(x) = 1$, then $u(x)$ is the shortest (signed) distance from x to the boundary $\partial\Omega$. If the zero level curve of $u(x)$ represents the initial location of the interface, then $u^{-1}(t)$ yields the location of the interface at time t . Hence, $u(x)$ represents the shortest time (arrival time) that is required to travel from the boundary Γ to x . For monotonically advancing fronts ($f > 0$), the

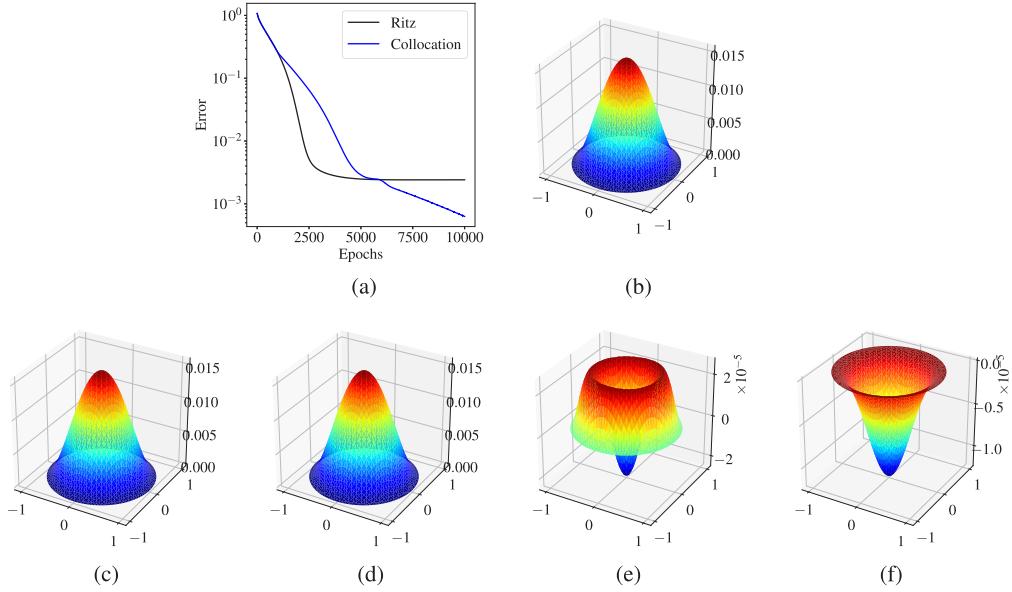


Fig. 42. Numerical solutions for clamped circular Kirchhoff plate bending problem using $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (Ritz and collocation). Network architecture used is 2-50-50-1. (a) Normalized absolute error in training. Surface plots of the (b) exact solution, and (c), (d) Ritz and collocation solutions. Surface plots of (e) error in Ritz solution and (f) error in collocation solution.

fast marching method [72] and the fast sweeping method [103] are highly efficient methods to solve (75). When upwind finite-differences are used to solve (75), it implies a causality: value of $u(\mathbf{x})$ only depends on values of $u(\mathbf{y})$ for which $u(\mathbf{x}) > u(\mathbf{y})$. No such restriction is used in PINN during training—we use a collocation method, where the exact satisfaction of the Dirichlet condition on Γ is met by constructing ADFs that use R-equivalence from Section 3.2 and the generalized mean value potential from Section 4.1.

We solve (75) with $f = 1$ to compute the signed distance function using PINN. The cubic ReLU activation function, which is a C^2 function, is used for all problems. The closed interface Γ is embedded within the biunit square, $\Omega_0 = (-1, 1)^2$, and as benchmark problems we consider affine (polygonal) and curved interfaces for Γ . The first problem that we consider is the computation of the signed distance function to the boundary of a smaller square, $\Omega = (-1/2, 1/2)^2$. The Dirichlet boundary condition $u = 0$ is imposed on $\Gamma = \partial\Omega$. The network architecture used is 2-30-30-30-1 for \tilde{u}_{nn}^{bc} and 2-70-70-1 for \tilde{u}_{nn} . For collocation, 10,000 points are used in the interior of the biunit square, and 400 points on the boundary Γ for $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$. Numerical results are presented in Fig. 43. The training loss and normalized absolute errors as a function of epochs for $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ and MVP) and $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ are shown in Fig. 43a and b. Once again, we notice that while $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$ attains the lowest loss among the three schemes, it has the highest error. This problem has an exact solution, which is shown as a contour plot in Fig. 43c. The exact distance, $u(\mathbf{x})$, achieves its maximum value of $1/\sqrt{2}$ at the corners of the biunit square and its minimum value of $-1/2$ at the center. The error in the numerical solutions are plotted in Fig. 43d-f. While $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (REQ and MVP) satisfy the boundary condition exactly, MVP results in a slightly more accurate solution in the entire domain. The L_∞ norm of the error using REQ, MVP and $\tilde{u}_{nn}(\mathbf{x})$ are 0.03, 0.026, and 1.34, respectively. The standard PINN, $\tilde{u}_{nn}(\mathbf{x}; \boldsymbol{\theta})$, does poorly on satisfying the boundary condition, which in turn leads to larger pointwise errors over the whole domain.

As the next problem, we consider the signed distance function to the boundary Γ of an ellipse that is centered at the origin and with semi-major and semi-minor axes of 0.25 and 0.15, respectively. The approximate distance function to the ellipse, $\phi(\mathbf{x})$, is computed using (8) (see Fig. 4(b)) for REQ and using (16) for MVP (see Fig. 13(a)) for MVP. The network architecture used is 2-50-50-1. The results that $\tilde{u}_{nn}(\mathbf{x})$ produced are very poor, and hence are not included. Numerical results using $\tilde{u}_{nn}^{bc}(\mathbf{x}; \boldsymbol{\theta})$ (REQ and MVP) are presented in Fig. 44. The training loss and the normalized error as a function of epochs is shown in Fig. 44a and b. The exact distance function (computed numerically) is presented in Fig. 44c, and the error plots for the numerical solutions $\tilde{u}_{nn}^{bc}(\mathbf{x})$ (REQ and MVP) are

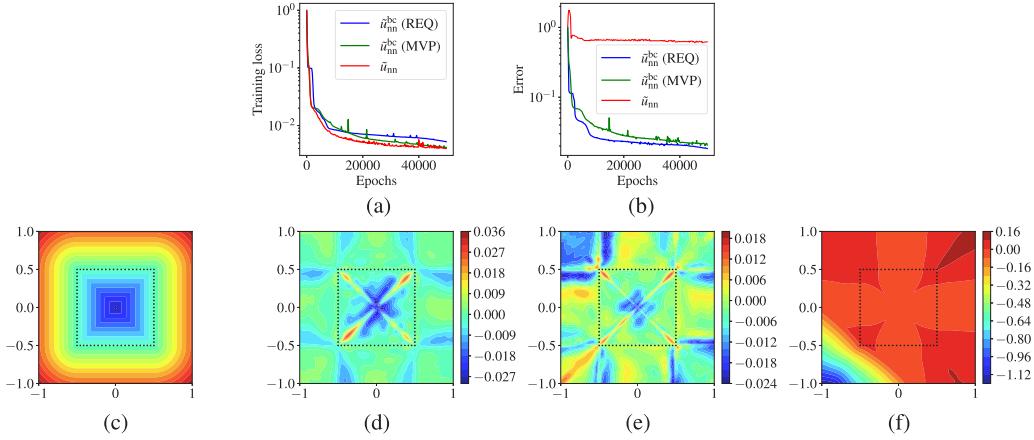


Fig. 43. Solving the Eikonal equation using $\tilde{u}_{nn}^{bc}(x; \theta)$ and $\tilde{u}_{nn}(x; \theta)$ to compute the signed distance to the boundary of a square. The network architecture used is 2-30-30-30-1 for \tilde{u}_{nn}^{bc} and 2-70-70-1 for \tilde{u}_{nn} . (a), (b) Training loss and normalized absolute error as a function of epochs for for $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ and MVP) and $\tilde{u}_{nn}(x; \theta)$. (c) Exact signed distance function. Contour plots of the error for (d) $\tilde{u}_{nn}^{bc}(x)$ (REQ), (e) $\tilde{u}_{nn}^{bc}(x)$ (MVP), and (f) $\tilde{u}_{nn}(x)$.

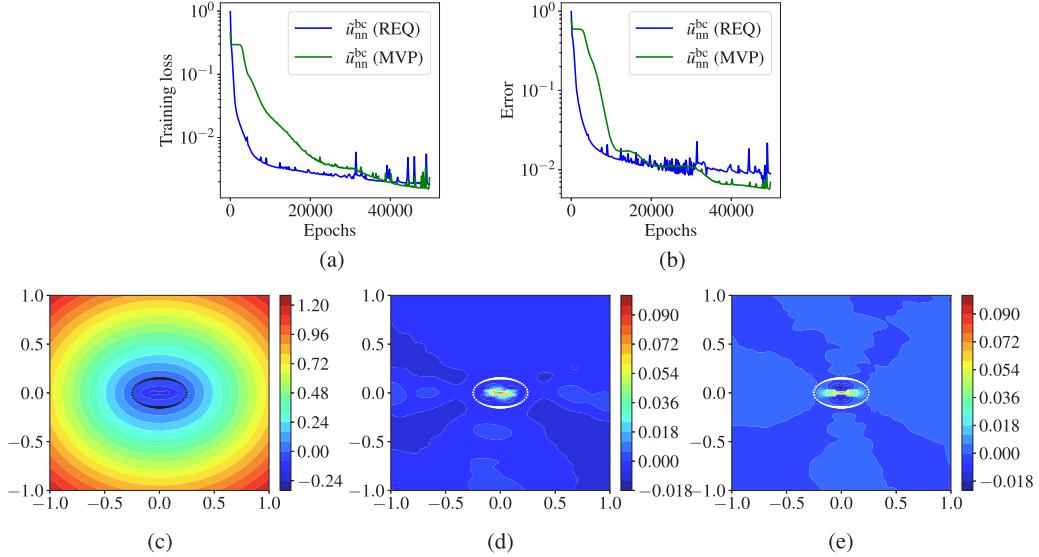


Fig. 44. Solving the Eikonal equation using $\tilde{u}_{nn}^{bc}(x; \theta)$ to compute the signed distance to the boundary of an ellipse. The network architecture used in 2-50-50-1. (a), (b) Training loss and normalized absolute error in training as a function of epochs for $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ and MVP) . (c) Exact signed distance function. Contour plots of the error for (d) $\tilde{u}_{nn}^{bc}(x)$ (REQ) and (e) $\tilde{u}_{nn}^{bc}(x)$ (MVP).

shown in Fig. 44d and e, respectively. Larger errors are concentrated in the region that is close to the center of the ellipse; away from the center the errors are less than 1 percent. The exact distance function is C^0 (derivative discontinuities at the center of the ellipse), whereas the numerical solution is C^2 smooth.

Lastly, we consider the signed distance function to the boundary Γ of the polygonalized map of Bhutan; see plots of the approximate distance functions to Γ using REQ and MVP that are presented in Figs. 8 and 11a, respectively. The network architecture used is 2-50-50-1. Here too the results of $\tilde{u}_{nn}(x)$ are not included since they are very poor. Numerical results using $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ and MVP) are presented in Fig. 45. The training loss and the normalized absolute error as a function of epochs are shown in Fig. 45a and b. At 20,000 epochs, the absolute normalized error is $\mathcal{O}(10^{-1})$. The exact distance function is presented in Fig. 45c, and the error plots for the numerical solutions $\tilde{u}_{nn}^{bc}(x)$ (REQ and MVP) are presented in Fig. 45d and e, respectively. The maximum error is about 4 percent.

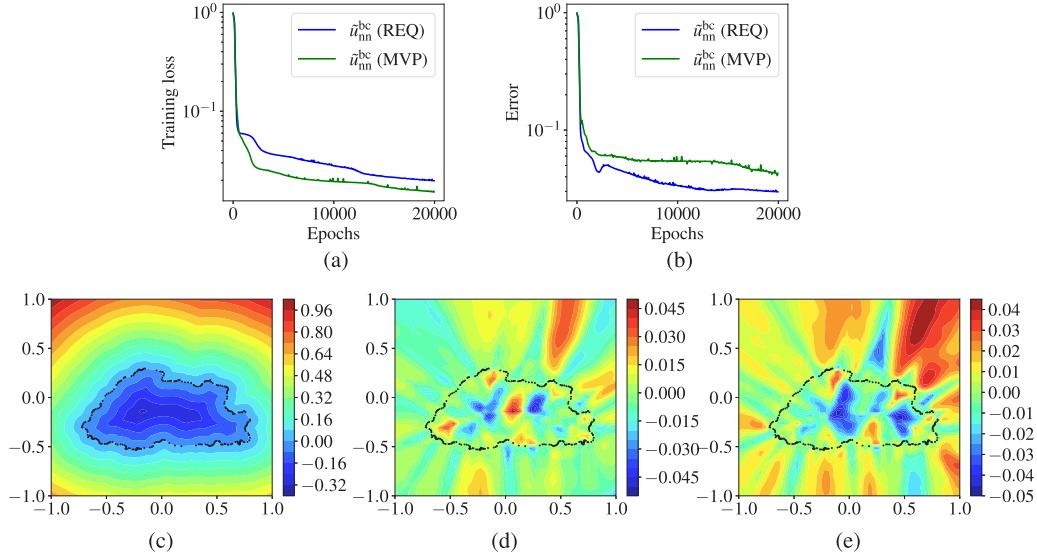


Fig. 45. Solving the Eikonal equation using $\tilde{u}_{nn}^{bc}(x; \theta)$ to compute the signed distance to the boundary of the polygonal map of Bhutan. The network architecture is 2-50-50-1. (a), (b) Training loss and normalized absolute error in training as a function of epochs for $\tilde{u}_{nn}^{bc}(x; \theta)$ (REQ and MVP). (c) Exact signed distance function. Contour plots of the error for (d) $\tilde{u}_{nn}^{bc}(x)$ (REQ) and (e) $\tilde{u}_{nn}^{bc}(x)$ (MVP).

11. Poisson boundary-value problem over the four-dimensional hypercube

As the last problem, we consider the model isotropic steady-state heat conduction (Poisson) problem over the 4-dimensional hypercube to show that essential boundary conditions can be readily imposed in higher dimensions as well. Consider the following Poisson problem with homogeneous Dirichlet boundary conditions:

$$-\nabla^2 u = f \quad \text{in } \Omega = (-1, 1)^4 \quad (76a)$$

$$u = 0 \quad \text{on } \partial\Omega, \quad (76b)$$

where $u(\mathbf{x}) : \mathbb{R}^4 \rightarrow \mathbb{R}$ is sought and $f(\mathbf{x})$ is the forcing function. Let $\mathbf{x} := (x_1, x_2, x_3, x_4) \in \Omega$ denote a point in the hypercube. We choose $f(\mathbf{x}) = \prod_{i=1}^4 \sin(\pi x_i)$, so that the exact solution is:

$$u(\mathbf{x}) = \frac{\prod_{i=1}^4 \sin(\pi x_i)}{4\pi^2}.$$

As noted in prior high-dimensional studies using PINN [8,9], numerical solutions for problems in high-dimensions are challenging since there is no easy way to mesh the domain and they are also subject to the curse of dimensionality. Among meshfree methods, since construction of radial basis functions is dimension-independent, RBF-based meshfree methods have had success in solving high-dimensional problems [104]. It is in these problems that the power and potential of a meshfree method such as PINN becomes most apparent. For this problem, we only consider trial functions, $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$, which exactly enforce the homogeneous Dirichlet boundary condition on $\partial\Omega$. One way of enforcing the boundary conditions is to assume \tilde{u}_{nn}^{bc} to be of the form

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta) = \left(\prod_{i=1}^4 (1 - x_i^2) \right) \tilde{u}_{nn}^R(\mathbf{x}; \theta),$$

which we refer to as the “product method” and note that while this is an obvious approach for the present problem, it leads to very small numbers inside the domain and away from the boundaries. In this case, the multiplicative factor scales as $\mathcal{O}(x^8)$ inside the biunit hypercube, and therefore the network parameters have to compensate for this highly nonlinear behavior during training. It is preferable to have a multiplicative factor that is better behaved in order to aid the training. To this end, we construct $\phi(\mathbf{x})$ using R-equivalence in (10), which seamlessly extends to higher dimensions. For this choice, the trial function $\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta)$ is of the form:

$$\tilde{u}_{nn}^{bc}(\mathbf{x}; \theta) = \phi(\mathbf{x}) \tilde{u}_{nn}^R(\mathbf{x}; \theta), \quad (77)$$

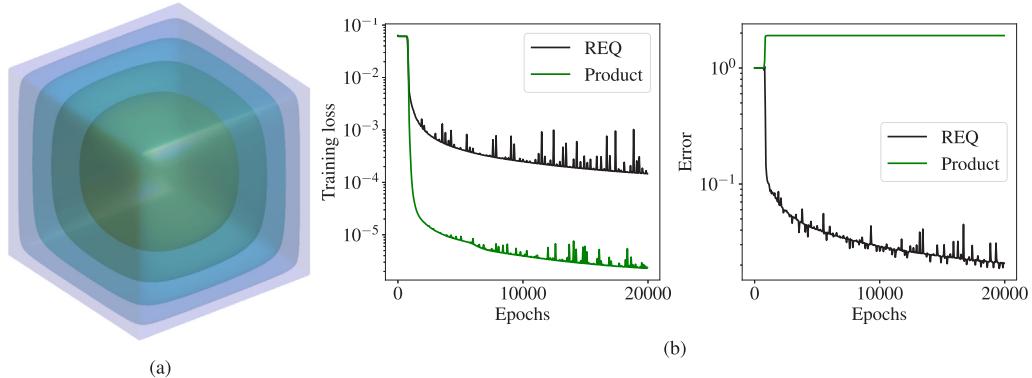


Fig. 46. Numerical solution of the Poisson problem over the four-dimensional hypercube. The network architecture is 2-100-100-1. The trial function, $\tilde{u}_{nn}^{bc}(x)$, is constructed using R-equivalence (REQ) and the product method. (a) Isosurface plot of the approximate distance function using REQ, $\phi(x)$, over the 3-dimensional biunit cube ($x_4 = 0$), with $\phi = 0$ being satisfied on the boundary of the cube. (b) Training loss and normalized absolute error for REQ and the product method as a function of epochs.

where $\phi(x)$ consists of R-equivalence (REQ) operations on $\phi_i(x)$, where $\phi_i(x)$ is the R-function for the region (strip) bounded by the hyperplanes $1 - x_i$ and $x_i - 1$. We form $\phi_i(x) = (1 - x_i^2)/2$, which is an ADF that is normalized to order 1. On using the REQ composition in (10), we write

$$\phi(x) = \phi_1(x) \sim \phi_2(x) \sim \phi_3(x) \sim \phi_4(x), \quad (78)$$

which generalizes to the hypercube in \mathbb{R}^d . Note that $\phi(x)$ only scales as x_i^2 in each coordinate direction, and therefore is much better behaved. For this problem, another choice for $\phi(x)$ is to define two ϕ_i 's in each dimension, i.e., $\phi_{2i-1} = 1 + x_i$ and $\phi_{2i} = 1 - x_i$ and then define $\phi(x) = \phi_1 \sim \phi_2 \dots \phi_8$ for the 4-dimensional hypercube. For $m = 1$, this construction coincides with the expression for ϕ in (78).

In Fig. 46, we present the numerical solutions using the product and REQ ($m = 1$) forms of the trial function. For both choices, we consider 5000 randomly generated interior points in Ω for training and compute the normalized error at a separate set of 5000 interior points. The network architecture for both choices is 4-100-100-1. The isosurface plot of $\phi(x)$ ($x_4 = 0$ plane) using REQ is shown in Fig. 46a. In Fig. 46b, the evolution of the training loss and normalized absolute error is presented for the product and REQ trial functions. We observe that while the REQ method is able to reach error levels of about 1 percent, the product method does not converge. The $\tilde{u}_{nn}^{bc}(x)$ solution with REQ yields $\mathcal{O}(10^{-2})$ error, whereas the product method has $\mathcal{O}(1)$ error even though it has a much smaller PDE loss of $\mathcal{O}(10^{-5})$.

12. Conclusions

Starting from the seminal works of Lagaris et al. [1–3] and the recent extensions and major advancements by Raissi et al. [6] and E and Yu [9], there has been a surge in the development and application of physics-informed neural networks to solve partial differential equations. In this paper, we have introduced a new approach based on distance fields to construct *geometry-aware approximations* in physics-informed neural networks by ensuring that the necessary boundary conditions are met *a priori*: all boundary conditions in a collocation method, and the essential boundary conditions (kinematic admissibility) in a Ritz method. Our approach relied on the theory of R-functions [23,24] to construct approximate distance fields and their use within a meshfree method to exactly impose boundary conditions to solve PDEs [27,30]. Apart from R-functions, we also showed that mean value potential fields [34–36] can be used to construct suitable distance field to solve PDEs over domains with affine as well as curved boundaries.

We presented several numerical examples to reveal the benefits of exactly imposing the boundary conditions versus the current state-of-the-art in deep collocation and deep Ritz methods for physics-informed neural networks. Notably, requiring only the interior residual error contribution in the loss function simplifies the training of the network and leads to more accurate numerical solutions. This was shown through several verification tests on benchmark one- and two-dimensional boundary-value problems—and consistently revealed the pitfalls of being

guided by the magnitude of the loss function in standard PINN-based collocation approaches [6]. In PINN methods, when there are multiple terms (PDE loss and boundary losses) that are present in the loss function and the loss weights are fixed *a priori*, the magnitude of the loss function at the end of the training does not provide a measure of the accuracy of the approximation. There can be a many-fold difference in the two error measures, which is revealed in our simulations. This is not surprising since the weights associated with each loss term is not known *a priori* since it depends on the PDE, boundary conditions, and the training. There is no clear rationale way to set these weights in order to ensure that the approach is robust and guaranteed to lead to reliable results. This study has reinforced that it is important that the PDE loss stands on its own, and boundary conditions are enforced via the ansatz. One approach is to construct a separate neural network to meet the boundary conditions as is done by Berg and Nyström [7], but the inherent inaccuracy in the satisfaction of the boundary conditions then propagates when training for the PDE is conducted. Moreover, when the geometry is complicated and different types of boundary conditions are imposed on different subset of the boundary, then this approach may soon become impractical. Our approach ensures exact satisfaction of all necessary boundary condition, which makes it appealing—so that training of the network (training loss and PDE loss coincide) is more efficient and accurate solutions can be realized.

This study has provided a method to perform meshfree analysis – solving PDEs without domain discretization – on complex two-dimensional geometries using physics-informed neural networks. This was made possible on using approximate distance functions, which were based on R-functions (R-equivalence composition) and generalized mean value potentials, in conjunction with transfinite inverse-distance based interpolation to exactly satisfy the boundary conditions. For the problems that we considered, both ADFs delivered the same order of accuracy with the constant being smaller for R-equivalence in most cases (Eikonal equation was the exception). If $u = 0$ is imposed on the boundary of a polygon with many (tens to hundreds) edges, then the ADF using mean value potential is more efficient since it results in an ADF by construction, whereas R-equivalence requires first constructing an ADF to each edge and then the joining operation to be performed. However, if different boundary conditions are imposed on distinct boundary segments then only R-equivalence is viable since this cannot be realized using the ADF based on mean value potential fields. While we have used R-functions and the mean value potential to construct smooth approximate distance functions, fields such as constructive geometric modeling [105] with implicit functions [73,106–108], PDE-based solutions for distance computations [109–111], and deep learning in computer vision [112] are rapidly advancing and may offer other attractive alternatives to construct smooth distance fields for use in deep neural networks to solve PDEs. A separate and more in depth investigation is needed to explore if there are other network architectures and optimizers for network training that are better-suited for the PINN ansatz with approximate distance functions, and to also quantify the accuracy that is obtained using different ADFs. Lastly, the formulation permitted exact modeling of affine and curved boundaries, thereby providing a pathway to conducting simulations on the exact geometry (isogeometric analysis) [37].

The ideas herein can be extended for higher dimensional problems, since R-equivalence composition for the approximate distance function is additive and does not suffer from the curse of dimensionality. This was demonstrated in Section 11, where we obtained an accurate PINN solution for a Poisson problem over the 4-dimensional hypercube. Extending our formulation to complex geometries in 3D, and the development of a deep Petrov–Galerkin domain-decomposition method are topics that we plan to pursue.

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.

Acknowledgments

AS acknowledges support from the NSF CAREER grant #1554033 to the Illinois Institute of Technology. NS thanks Anand Reddy, Eric Chin and Kai Hormann for many helpful discussions.

References

- [1] I.E. Lagaris, A. Likas, D.I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Trans. Neural Netw.* 9 (5) (1998) 987–1000.

- [2] I.E. Lagaris, A. Likas, D.I. Fotiadis, Artificial neural network methods in quantum mechanics, *Comput. Phys. Comm.* 104 (1997) 1–14, Ch. 3.
- [3] I.E. Lagaris, A.C. Likas, D.G. Papageorgiou, Neural-network methods for boundary value problems with irregular boundaries, *IEEE Trans. Neural Netw.* 11 (5) (2000) 1041–1049.
- [4] K.S. McFall, An artificial neural network method for solving boundary value problems with arbitrary irregular boundaries (Ph.D. thesis), Georgia Institute of Technology, Atlanta, GA, USA, 2006.
- [5] 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 (8) (2009) 1221–1233.
- [6] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* 378 (2019) 686–707.
- [7] J. Berg, K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, *Neuralcomputing* 317 (2018) 28–41.
- [8] J. Sirignano, K. Spiliopoulos, DGM: A deep learning algorithm for solving partial differential equations, *J. Comput. Phys.* 375 (2018) 1339–1364.
- [9] W. E, B. Yu, The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems, *Commun. Math. Stat.* 6 (1) (2018) 1–12.
- [10] J. Han, A. Jentzen, W. E, Solving high-dimensional partial differential equations using deep learning, *Proc. Natl. Acad. Sci.* 115 (34) (2018) 8505–8510.
- [11] E. Kharazmi, Z. Zhang, G.E. Karniadakis, Variational physics-informed neural networks for solving partial differential equations, 2019, arXiv:1912.00873.
- [12] E. Kharazmi, Z. Zhang, G.E. Karniadakis, hp-VPINNs: Variational physics-informed neural networks with domain decomposition, *Comput. Methods Appl. Mech. Engrg.* 374 (2020) 113547.
- [13] L. Lu, X. Meng, Z. Mao, G.E. Karniadakis, DeepXDE: A deep learning library for solving differential equations, *SIAM Rev.* 63 (1) (2021) 208–228.
- [14] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: A system for large-scale machine learning, in: 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), 2016, pp. 265–283.
- [15] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, R. Garnett (Eds.), *Advances in Neural Information Processing Systems*, 32, Curran Associates, Inc., 2019.
- [16] S. Wang, Y. Teng, P. Perdikaris, Understanding and mitigating gradient pathologies in physics-informed neural networks, 2020, arXiv:v2001.04536.
- [17] J. Chen, R. Du, K. Wu, A comparison study of deep Galerkin method and deep Ritz method for elliptic problems with different boundary conditions, *Commun. Math. Res.* 36 (3) (2020) 354–376.
- [18] L. Lyu, K. Wu, R. Du, J. Chen, Enforcing exact boundary and initial conditions in the deep mixed residual method, 2020, arXiv:2008.01491.
- [19] I. Babuška, U. Banerjee, J.E. Osborn, Survey of meshless and generalized finite element methods: a unified approach, *Acta Numer.* 12 (2003) 1–125.
- [20] A. Huerta, T. Belytschko, S. Fernández-Méndez, T. Rabcuk, X. Zhuang, M. Arroyo, Meshfree methods, second ed., in: *Encycl. Comput. Mech.*, vol. 2, Wiley, 2017, pp. 1–38.
- [21] L.V. Kantorovich, V.I. Krylov, *Approximate Methods of Higher Analysis*, Interscience, New York, NY, USA, 1958.
- [22] V.L. Rvachev, *Theory of R-Functions and Some Applications*, Naukova Dumka, Kiev. In Russian, 1982.
- [23] V.L. Rvachev, T.I. Sheiko, R-functions in boundary value problems in mechanics, *Appl. Mech. Rev.* 48 (4) (1995) 151–188.
- [24] V.L. Rvachev, T.I. Sheiko, V. Shapiro, I. Tsukanov, On completeness of RFM solution structures, *Comput. Mech.* 25 (2000) 305–3163.
- [25] V.L. Rvachev, T.I. Sheiko, V. Shapiro, I. Tsukanov, Transfinite interpolation over implicitly defined sets, *Comput. Aided Geom. Design* 18 (2001) 195–220.
- [26] V. Shapiro, Theory of R-functions and applications: A primer, Tech. Rep. CPA88-3, Cornell Programmable Automation, Sibley School of Mechanical Engineering, Ithaca, NY 14853, USA, 1991.
- [27] V. Shapiro, I. Tsukanov, Meshfree simulation of deforming domains, *Comput. Aided Des.* 31 (7) (1999) 459–471.
- [28] V. Shapiro, I. Tsukanov, The architecture of SAGE—a meshfree system based on RFM, *Eng. Comput.* 18 (4) (2002) 295–311.
- [29] A. Biswas, V. Shapiro, Approximate distance fields with non-vanishing gradients, *Graph. Models* 66 (3) (2004) 133–159.
- [30] V. Shapiro, Semi-analytic geometry with R-functions, *Acta Numer.* 16 (2007) 239–303.
- [31] M. Freytag, V. Shapiro, I. Tsukanov, Finite element analysis in situ, *Finite Elem. Anal. Des.* 47 (9) (2011) 957–972.
- [32] K. Höllig, U. Reif, J. Wipper, Weighted extended B-spline approximation of Dirichlet problems, *SIAM J. Numer. Anal.* 39 (2) (2001) 442–462.
- [33] D. Millán, N. Sukumar, M. Arroyo, Cell-based maximum-entropy approximants, *Comput. Methods Appl. Mech. Engrg.* 284 (2015) 712–731.
- [34] M.S. Floater, Mean value coordinates, *Comput. Aided Geom. Design* 20 (1) (2003) 19–27.
- [35] C. Dyken, M.S. Floater, Transfinite mean value interpolation, *Comput. Aided Geom. Design* 26 (1) (2009) 117–134.
- [36] A. Belyaev, P.-A. Fayolle, A. Pasko, Signed L_p -distance fields, *Comput. Aided Des.* 45 (2) (2013) 523–528.
- [37] T.J.R. Hughes, J.A. Cottrell, Y. Bazilevs, Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement, *Comput. Methods Appl. Mech. Engrg.* 194 (39–41) (2005) 4135–4195.

- [38] P. Thoutireddy, M. Ortiz, A variational r -adaption and shape-optimization method for finite-deformation elasticity, *Internat. J. Numer. Methods Engrg.* 61 (1) (2004) 1–21.
- [39] J. He, L. Li, J. Xu, C. Zheng, ReLU deep neural networks and linear finite elements, 2018, arXiv:1807.03973.
- [40] E. Grinpun, The basis refinement method (Ph.D. thesis), California Institute of Technology, Pasadena, CA, USA, 2003.
- [41] E.C. Cyr, M.A. Gulian, R.G. Patel, M. Perego, N.A. Trask, Robust training and initialization of deep neural networks: An adaptive basis viewpoint, in: Mathematical and Scientific Machine Learning, PMLR, 2020, pp. 512–536.
- [42] J.A.A. Opschoor, P.C. Petersen, C. Schwab, Deep ReLU networks and high-order finite element methods, *Anal. Appl.* 18 (05) (2020) 715–770.
- [43] E.J. Kansa, Multiquadratics—A scattered data approximation scheme for applications to computational fluid-dynamics. 1. Surface approximations and partial derivative estimates, *Comput. Math. Appl.* 19 (8/9) (1990) 127–145.
- [44] E.J. Kansa, Multiquadratics—A scattered data approximation scheme for applications to computational fluid-dynamics. 2. Solutions to parabolic, hyperboloc and elliptic partial-differential equations, *Comput. Math. Appl.* 19 (8/9) (1990) 147–161.
- [45] M.D. Buhmann, Radial Basis Functions: Theory and Implementations, Cambridge University Press, Cambridge, UK, 2003.
- [46] G. Fasshauer, Meshfree Approximation Methods in MATLAB, Interdisciplinary Mathematical Sciences – Vol. 6, World Scientific Publishers, Singapore, 2007.
- [47] R. Schaback, H. Wendland, Kernel techniques: from machine learning to meshless methods, *Acta Numer.* 15 (2006) 543.
- [48] M. Arroyo, M. Ortiz, Local *maximum-entropy* approximation schemes: a seamless bridge between finite elements and meshfree methods, *Internat. J. Numer. Methods Engrg.* 65 (13) (2006) 2167–2202.
- [49] I. Babuška, J.M. Melenk, The partition of unity method, *Internat. J. Numer. Methods Engrg.* 40 (1997) 727–758.
- [50] V. Rajan, Optimality of the Delaunay triangulation in \mathbb{R}^d , *Discrete Comput. Geom.* 12 (1) (1994) 189–202.
- [51] N. Sukumar, Construction of polygonal interpolants: a maximum entropy approach, *Internat. J. Numer. Methods Engrg.* 61 (12) (2004) 2159–2181.
- [52] N. Sukumar, Maximum entropy approximation, in: K.H. Knuth, A.E. Abbas, R.D. Morris, J.P. Castle (Eds.), AIP Conf. Proc. 803 (1) (2005) 337–344.
- [53] M. Arroyo, M. Ortiz, Local maximum-entropy approximation schemes, in: M. Griebel, M.A. Schweitzer (Eds.), Meshfree Methods for Partial Differential Equations III, in: Lecture Notes in Computational Science and Engineering, vol. 57, Springer, Berlin, Germany, 2007, pp. 1–16.
- [54] N. Sukumar, R.W. Wright, Overview and construction of meshfree basis functions: From moving least squares to entropy approximants, *Internat. J. Numer. Methods Engrg.* 70 (2) (2007) 181–205.
- [55] A. Rosolen, D. Millán, M. Arroyo, On the optimum support size in meshfree methods: a variational adaptivity approach with maximum entropy approximants, *Internat. J. Numer. Methods Engrg.* 82 (7) (2010) 868–895.
- [56] J. Park, I.W. Sandberg, Universal approximation using radial-basis-function networks, *Neural Comput.* 3 (1991) 246–257.
- [57] H.N. Mhaskar, Neural networks for optimal approximation of smooth and analytic functions, *Neural Comput.* 8 (1) (1996) 164–177.
- [58] K. Lee, N.A. Trask, R.G. Patel, M.A. Gulian, E.C. Cyr, Partition of unity networks: deep hp-approximation, 2021, arXiv:2101.11256.
- [59] A.A. Ramabathiran, P. Ramachandran, SPINN: Sparse, physics-based, and partially interpretable neural networks for PDEs, *J. Comput. Phys.* 445 (2021) 110600.
- [60] F. Greco, M. Arroyo, High-order maximum-entropy collocation methods, *Comput. Methods Appl. Mech. Engrg.* 367 (2020) 113115.
- [61] H. Sheng, C. Yang, PFNN: A penalty-free neural network method for solving a class of second-order boundary-value problems on complex geometries, *J. Comput. Phys.* 428 (2021) 110085.
- [62] V. Dwivedi, B. Srinivasan, Physics informed extreme learning machine (PIELM)—a rapid method for the numerical solution of partial differential equations, *Neuralcomputing* 391 (2020) 96–118.
- [63] V. Dwivedi, B. Srinivasan, Solution of biharmonic equation in complicated geometries with physics informed extreme learning machine, *J. Comput. Inf. Sci. Eng.* 20 (6) (2020).
- [64] Y. Liao, P. Ming, Deep Nitsche method: Deep Ritz method with essential boundary conditions, 2019, arXiv:1912.01309.
- [65] S. Li, W.K. Liu, Meshfree Particle Methods, Springer-Verlag, New York, NY, USA, 2004.
- [66] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* 2 (1989) 359–366.
- [67] K. Hornik, Approximation capabilities of multilayer perceptrons, *Neural Netw.* 4 (1991) 251–257.
- [68] G. Strang, G.J. Fix, An Analysis of the Finite Element Method, Prentice-Hall, New York, NY, USA, 1973.
- [69] F.M. Rohrhofer, S. Posch, B.C. Geiger, On the Pareto front of physics-informed neural networks, 2021, arXiv:2105.00862.
- [70] O. Hennigh, S. Narasimhan, M.A. Nabian, A. Subramaniam, K. Tangsali, Z. Fang, M. Rietmann, W. Byeon, S. Choudhry, NVIDIA SimNetTM: An AI-accelerated multi-physics simulation framework, in: International Conference on Computational Science, Springer, 2021, pp. 447–461.
- [71] I. Tsukanov, S.R. Posireddy, Hybrid method of engineering analysis: Combining meshfree method with distance fields and collocation technique, *J. Comput. Inf. Sci. Eng.* 11 (3) (2011).
- [72] J.A. Sethian, Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, Cambridge University Press, Cambridge, U.K., 1999.
- [73] J. Bloomenthal, Bulge elimination in convolution surfaces, *Comput. Graph. Forum* 16 (1) (1997) 31–41.
- [74] V. Shapiro, I. Tsukanov, Implicit functions with guaranteed differential properties, in: Proceedings of the Fifth ACM Symposium on Solid Modeling and Applications, 1999, pp. 258–269.
- [75] K. Upadhyay, T. Song, A. Tambat, G. Subbarayan, Algebraic distance estimations for enriched isogeometric analysis, *Comput. Methods Appl. Mech. Engrg.* 280 (2014) 28–56.

- [76] E.B. Chin, N. Sukumar, Modeling curved interfaces without element-partitioning in the extended finite element method, *Internat. J. Numer. Methods Engrg.* 120 (5) (2019) 607–649.
- [77] A.G. Belyaev, P.-A. Fayolle, in: K. Hormann, N. Sukumar (Eds.), *Transfinite barycentric coordinates*, CRC Press, New York, NY, 2017, pp. 43–62.
- [78] M.S. Floater, Generalized barycentric coordinates and applications, *Acta Numer.* 24 (2015) 161–214.
- [79] D. Anisimov, in: K. Hormann, N. Sukumar (Eds.), *Barycentric coordinates and their properties*, CRC Press, New York, NY, 2017, pp. 3–22.
- [80] K. Hormann, N. Sukumar (Eds.), *Generalized barycentric coordinates in computer graphics and computational mechanics*, CRC Press, New York, NY, 2017.
- [81] K. Hormann, M.S. Floater, Mean value coordinates for arbitrary planar polygons, *ACM Trans. Graph.* 25 (4) (2006) 1424–1441.
- [82] S. Bruvoll, M.S. Floater, Transfinite mean value interpolation in general dimension, *J. Comput. Appl. Math.* 233 (7) (2010) 1631–1639.
- [83] T. Ju, S. Schaefer, J. Warren, Mean value coordinates for closed triangular meshes, *ACM Trans. Graph.* 24 (3) (2005) 561–566.
- [84] E.B. Chin, N. Sukumar, Scaled boundary cubature scheme for numerical integration over planar regions with affine and curved boundaries, *Comput. Methods Appl. Mech. Engrg.* 380 (2021) 113791.
- [85] D. Shepard, A two-dimensional interpolation function for irregularly-spaced data, in: *Proceedings of the 23rd ACM National Conference, Association for Computing Machinery*, New York, New York, 1968, pp. 517–524.
- [86] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain., *Psychol. Rev.* 65 (6) (1958) 386.
- [87] Y. LeCun, The MNIST database of handwritten digits, 1998, <http://yann.lecun.com/exdb/mnist/>.
- [88] D. Finol, Y. Lu, V. Mahadevan, A. Srivastava, Deep convolutional neural networks for eigenvalue problems in mechanics, *Internat. J. Numer. Methods Engrg.* 118 (5) (2019) 258–275.
- [89] K. Hornik, Some new results on neural network approximation, *Neural Netw.* 6 (1993) 1069–1072.
- [90] A. Le Nail, Nn-svg: Publication-ready neural network architecture schematics, *J. Open Sour. Softw.* 4 (33) (2019) 747.
- [91] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [92] 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. Engrg.* 362 (2020) 112790.
- [93] J. Bradbury, R. Frostig, P. Hawkins, M.J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, Q. Zhang, JAX: composable transformations of Python+NumPy programs, 2018, URL [http://github.com/google/jax](https://github.com/google/jax).
- [94] E. Bisong, Google colaboratory, in: *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Springer, 2019, pp. 59–64.
- [95] N. Schlömer, J. Hariharan, Dmsh, 2020, Available at <https://github.com/nschloe/dmsh>. Accessed on April 1, 2021.
- [96] P.-O. Persson, G. Strang, A simple mesh generator in MATLAB, *SIAM Rev.* 46 (2) (2004) 329–345.
- [97] N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, A. Courville, On the spectral bias of neural networks, in: *International Conference on Machine Learning*, PMLR, 2019, pp. 5301–5310.
- [98] S. Wang, H. Wang, P. Perdikaris, On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks, *Comput. Methods Appl. Mech. Engrg.* 384 (2021) 113938.
- [99] J. Pushkar, M. Meyer, T. DeRose, B. Green, T. Sanocki, Harmonic coordinates for character articulation, *ACM Trans. Graph.* 26 (3) (2007) 71.
- [100] J.R. Shewchuk, Triangle: Engineering a 2D quality mesh generator and delaunay triangulator, in: M.C. Lin, D. Manocha (Eds.), *Applied Computational Geometry: Towards Geometric Engineering*, in: *Lecture Notes in Computer Science*, vol. 1148, Springer-Verlag, 1996, pp. 203–222.
- [101] S.P. Timoshenko, S. Woinowsky-Krieger, *Theory of Plates and Shells*, second ed., McGraw Hill, New York, NY, USA, 1959.
- [102] H. Guo, X. Zhuang, T. Rabczuk, A deep collocation method for the bending analysis of Kirchhoff plate, 2021, [arXiv:2102.02617](https://arxiv.org/abs/2102.02617).
- [103] H. Zhao, A fast sweeping method for Eikonal equations, *Math. Comp.* 74 (250) (2005) 603–627.
- [104] T. Cecil, J. Qian, S. Osher, Numerical methods for high dimensional Hamilton–Jacobi equations using radial basis functions, *J. Comput. Phys.* 196 (1) (2004) 327–347.
- [105] A. Ricci, A constructive geometry for computer graphics, *Comput. J.* 16 (2) (1973) 157–160.
- [106] A. Sherstyuk, Kernel functions in convolution surfaces: a comparative analysis, *Vis. Comput.* 15 (4) (1999) 171–182.
- [107] L. Barthe, B. Wyvill, E. De Groot, Controllable binary CSG operators for “soft objects”, *Int. J. Shape Model.* 10 (02) (2004) 135–154.
- [108] O. Gourmet, L. Barthe, M.-P. Cani, B. Wyvill, A. Bernhardt, M. Paulin, H. Grasberger, A gradient-based implicit blend, *ACM Trans. Graph.* 32 (2) (2013) 12:1–12:12.
- [109] A. Belyaev, P.-A. Fayolle, On variational and PDE-based distance function approximations, *Comput. Graph. Forum* 34 (8) (2015) 104–118.
- [110] K. Crane, C. Weischedel, M. Wardetzky, The heat method for distance computation, *Commun. ACM* 60 (11) (2017) 90–99.
- [111] A.G. Belyaev, P.-A. Fayolle, A variational method for accurate distance function estimation, in: *Numerical Geometry, Grid Generation and Scientific Computing*, Springer, Cham, 2019, pp. 175–181.
- [112] J.J. Park, P. Florence, J. Straub, R. Newcombe, S. Lovegrove, DeepSDF: Learning continuous signed distance functions for shape representation, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 165–174.