

Kolmogorov-Arnold PointNet: Deep learning for prediction of fluid fields on irregular geometries

Ali Kashefi^{a,*}

^aDepartment of Civil and Environmental Engineering, Stanford University, Stanford, 94305, CA, USA

ARTICLE INFO

Keywords:

PointNet

Kolmogorov-Arnold networks

Irregular geometries

Jacobi polynomials

Steady-state incompressible flow

ABSTRACT

We present Kolmogorov-Arnold PointNet (KA-PointNet) as a novel supervised deep learning framework for the prediction of incompressible steady-state fluid flow fields in irregular domains, where the predicted fields are a function of the geometry of the domains. In KA-PointNet, we implement shared Kolmogorov-Arnold Networks (KANs) in the segmentation branch of the PointNet architecture. We utilize Jacobi polynomials to construct shared KANs. As a benchmark test case, we consider incompressible laminar steady-state flow over a cylinder, where the geometry of its cross-section varies over the data set. We investigate the performance of Jacobi polynomials with different degrees as well as special cases of Jacobi polynomials such as Legendre polynomials, Chebyshev polynomials of the first and second kinds, and Gegenbauer polynomials, in terms of the computational cost of training and accuracy of prediction of the test set. Additionally, we compare the performance of PointNet with shared KANs (i.e., KA-PointNet) and PointNet with shared Multilayer Perceptrons (MLPs). It is observed that when the number of trainable parameters is approximately equal, PointNet with shared KANs (i.e., KA-PointNet) outperforms PointNet with shared MLPs.

1. Introduction and motivation

Kolmogorov-Arnold Networks (KANs) [1] were recently proposed as an alternative modeling approach to Multilayer Perceptrons (MLPs) [2, 3, 4]. KANs are primarily based on the principles of the Kolmogorov-Arnold representation theorem [5, 6, 7, 8, 9, 10, 11, 12]. The core idea of KANs is to propose a neural network where the objective is to train activation functions rather than training weights and biases with fixed activation functions, as in MLPs [1]. The efficiency and capability of KANs have been demonstrated in various areas and applications, such as physics-informed machine learning [13, 14, 15, 16, 17, 18, 19], deep operator networks [20, 14], neural ordinary differential equations [21], image classification [22, 23, 24, 25, 26, 27, 28, 29, 30, 31], image segmentation [32, 33], image detection [34], audio classification [29], and many other scientific and industrial fields [35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54]. Additionally, KANs have been integrated into Convolutional Neural Networks (CNNs) [22, 55] and graph neural networks [56, 57, 58, 59]. In the current study, our focus is on the field of computational physics and mechanics.

One of the most important applications of supervised deep learning in computational mechanics is to accelerate the investigation of geometric parameters for device optimization. Supervised deep learning models are first trained on labeled data obtained from numerical simulations or lab experiments, and then used to predict quantities of interest for unseen geometries. For this specific application, geometric deep learning models such as CNNs and their extensions (e.g., see Refs. [60, 61, 62, 63, 64, 65]), graph neural networks (e.g., see Refs. [66, 67, 68, 69, 70]), and point cloud-based neural networks such as PointNet and its derived versions (e.g., see Refs. [71, 72, 73, 74, 75, 76, 77]) are particularly useful. Focusing on this aspect, the goal of the current study is to investigate the capability of KANs, embedded in a geometric deep learning framework, for rapid predictions of quantities of interest in computational mechanics when the geometry of the domains varies across the data set. In these types of networks, the geometry of a domain is learned in a latent space, making the network's prediction a function of the geometry.

Graph neural networks [78] and PointNet [71] have advantages for handling complicated geometries, which are common in industrial design. These benefits have been extensively discussed in the literature (e.g., see Ref. [74]), and a summary is presented here. Using graph neural networks or PointNet, the geometry of the domain of interest or the

*Corresponding author

✉ kashefi@stanford.edu (A. Kashefi)

ORCID(s): 0000-0003-0014-9051 (A. Kashefi)

object under investigation for shape optimization can be represented without pixelation, unlike CNN-based models. This approach avoids artifacts in the representation of object or domain boundaries (e.g., see Fig. 1 in Ref. [73]). Moreover, because the geometry of the target object is precisely illustrated in graph neural networks or PointNet, the network's predictions are highly responsive to small variations in the geometry from one labeled data to another, such as minor adjustments in the angle of attack of an airfoil. Additionally, graph neural networks or PointNet allow for adaptive variation in the spatial distribution of points from fine to coarse scales (e.g., see Figs. 7–10 in Ref. [73]), a feature similar to unstructured grids. This optimizes the computational cost of training, rather than imposing high computational costs due to uniform pixels across the entire domain, as in CNN-based models. Finally, the spatial dimensions of the domain under study can vary from one labeled data to another, and are not constrained to a fixed size, unlike in CNN-based networks (e.g., see Figs. 7–11 in Ref. [73]). Due to these advantages, we embed KANs (instead of MLPs) into PointNet to create a new geometric deep learning framework. For the rest of the article, we will refer to this framework as KA-PointNet.

PointNet [71] was initially introduced in 2017 for the classification and segmentation of three-dimensional objects modeled as three-dimensional points. PointNet [71] and its advanced versions [72, 79] have received remarkable attention in the areas of computer graphics and computer vision (e.g., see Refs. [80, 81, 82, 83]). Historically, Kashefi et al. [73] were the first to use PointNet, with necessary adaptations and adjustments, in the area of computational physics, specifically for steady-state incompressible flow in two dimensions. Subsequently, other researchers employed PointNet and its advanced versions, such as PointNet++ [72], in other areas such as compressible flow [76], elasticity [77], and others [84, 85, 86]. Kashefi et al. [74] also introduced physics-informed PointNet (PIPNet) in 2022 to solve forward and inverse problems involving incompressible flows [74], thermal fields [74], linear elasticity [75], and flow transport in porous media [87]. In simple terms, PIPNet combines PointNet with the concept of physics-informed deep learning introduced by Raissi et al. [88], integrating the residual of governing equations associated with the physics of a problem into the loss function of PIPNet.

To build KA-PointNet, we fundamentally use the segmentation branch of PointNet [71]. We implement shared KANs in all layers of PointNet instead of shared MLPs. Furthermore, we utilize Jacobi polynomials in shared KANs. Jacobi polynomials have already been used for KAN implementations (e.g., see Refs. [14, 24, 41, 89, 90]). Moreover, we investigate the capability of special cases of Jacobi polynomials, such as Legendre polynomials, Chebyshev polynomials of the first and second kinds, and Gegenbauer polynomials. Additionally, we assess the performance of KA-PointNet for different degrees of Jacobi polynomials. Flow past a cylinder with different cross-sectional shapes is considered as a representative benchmark problem. To construct point cloud data readable for feeding into PointNet [71], grid vertices of finite volume meshes of computational domains are viewed as points. A comparison between the outcomes of PointNet with shared KANs (i.e., KA-PointNet) and PointNet with shared MLPs is made in terms of prediction accuracy, training time, and the number of trainable parameters. The Adam optimizer [91], along with the mean squared error loss function, is used to perform the gradient descent optimization. It is noteworthy that, to the best of the author's knowledge, this is the first time a geometric deep learning model combined with KANs is used to predict solutions for computational fields over domains with distinct geometries. Although KANs have already been integrated into CNNs and graph neural networks, they have been utilized for other applications in computer vision [22, 55, 56, 57, 58, 59].

The rest of this article is organized as follows. Governing equations and data generation are explained in Sect. 2. The architecture of KA-PointNet with shared KANs and its training on the desired data set are elaborated in Sect. 3. Analysis of the performance of KA-PointNet and the effect of different parameters on the accuracy of its predictions are given in Sect. 4. Additionally, a comparison between the performance of KA-PointNet and PointNet with shared MLPs is provided in Sect. 4 as well. Finally, a summary of the research article and potential future directions are explained in Sect. 5.

2. Problem formulation

In the current article, we propose a supervised version of KA-PointNet for predicting the velocity and pressure fields of steady-state incompressible flow. To train the deep learning framework and perform error analysis on the predictions of KA-PointNet for unseen data, we require labeled data. We use the classical problem of steady-state flow past a cylinder as a benchmark test case. Accordingly, the input of KA-PointNet is the geometry of the domain, and the output is the prediction of velocity and pressure fields. We describe the governing equations of the problem and the procedure of data generation in the subsections of 2.1 and 2.2, respectively.

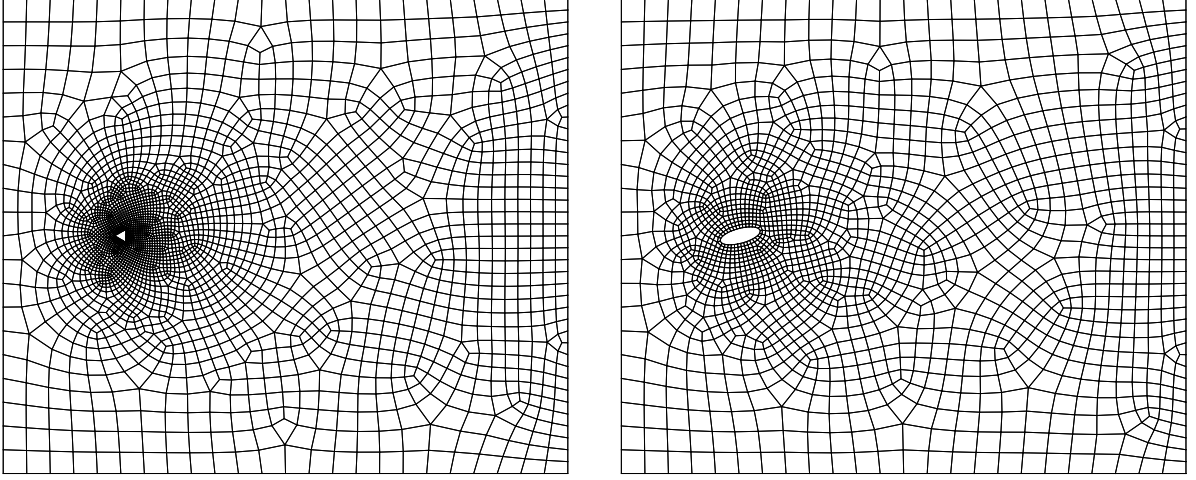


Figure 1: Finite volume meshes used for the numerical simulation of flow over a cylinder with a triangular cross section, with 2775 vertices on the left panel, and an elliptical cross section, with 2672 vertices on the right panel.

2.1. Governing equations of interest

The behavior of incompressible viscous Newtonian fluid flow is described by the conservation of mass and conservation of momentum, along with the associated boundary conditions, as follows:

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } V, \quad (1)$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) - \mu \Delta \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } V, \quad (2)$$

$$\mathbf{u} = \mathbf{u}_{\Gamma_D} \quad \text{on } \Gamma_D, \quad (3)$$

$$-pn + \mu \nabla \mathbf{u} \cdot \mathbf{n} = \mathbf{t}_{\Gamma_{Ne}} \quad \text{on } \Gamma_{Ne}, \quad (4)$$









where \mathbf{u} and p are the velocity vector and absolute pressure of the fluid in the domain V , respectively. Fluid density and dynamic viscosity are denoted by ρ and μ , respectively. The external body force is indicated by \mathbf{f} . The Dirichlet and Neumann boundaries of domain V are represented by Γ_D and Γ_{Ne} , respectively, with no overlap between Γ_D and Γ_{Ne} . The stress vector acting on Γ_{Ne} is denoted by $\mathbf{t}_{\Gamma_{Ne}}$, where \mathbf{n} indicates the outward unit normal vector to Γ_{Ne} . The components of the velocity vector \mathbf{u} are demonstrated by u and v in the x and y directions, respectively.

We are interested in solving Eqs. (1)–(2) for flow over an infinite cylinder with different cross-sectional shapes. To achieve this goal, we consider a rectangular domain $V = [0, 38 \text{ m}] \times [0, 32 \text{ m}]$. The cross-section of the cylinder is represented by a two-dimensional object with its center of mass at the point (8 m, 16 m) in the domain V . The cylinder is rigid, and we impose no-slip conditions on its surfaces. The free stream velocity, with a magnitude of u_∞ and parallel to the x axis, is applied at the inflow, bottom, and top boundaries of the domain V . To satisfy far-field assumptions [92, 93, 94, 95, 96], the outflow velocity is described by the Neumann stress-free conditions (i.e., $\mathbf{t}_{\Gamma_{Ne}} = \mathbf{0}$) as follows:

$$-pn + \mu \nabla \mathbf{u} \cdot \mathbf{n} = \mathbf{0}. \quad (5)$$

Working in the International System of Units, the fluid density (ρ), free stream velocity (u_∞), and dynamic viscosity (μ) are set to 1.00, 1.00, and 0.05, respectively.

Table 1
Details of the generated data

Shape	Schematic figure	Variation in orientation	Variation in length scale	Number of data
Circle		-	$a = 1 \text{ m}$	1
Equilateral hexagon		$3^\circ, 6^\circ, \dots, 60^\circ$	$a = 1 \text{ m}$	20
Equilateral pentagon		$3^\circ, 6^\circ, \dots, 72^\circ$	$a = 1 \text{ m}$	24
Square		$3^\circ, 6^\circ, \dots, 90^\circ$	$a = 1 \text{ m}$	30
Equilateral triangle		$3^\circ, 6^\circ, \dots, 180^\circ$	$a = 1 \text{ m}$	60
Rectangle		$3^\circ, 6^\circ, \dots, 180^\circ$	$a = 1 \text{ m}; b/a = 1.2, 1.4, \dots, 3.6$	780
Ellipse		$3^\circ, 6^\circ, \dots, 180^\circ$	$a = 1 \text{ m}; b/a = 1.2, 1.4, \dots, 3.8$	840
Triangle		$3^\circ, 6^\circ, \dots, 360^\circ$	$a = 1 \text{ m}; b/a = 1.5, 1.75$ $\gamma = 60^\circ, 80^\circ$	480

2.2. Data generation

We employ Gmsh [97] to discretize the domain V using unstructured finite volume meshes. Figure 1 exhibits two examples of generated meshes for cylinders with triangular and elliptical cross sections. We utilize the OpenFOAM solver [98] with the Semi-Implicit Method for Pressure Linked Equations (SIMPLE) [99] to numerically solve Eqs. (1)–(2) under the conditions and assumptions described above. Numerical simulations are conducted until the L^2 norm of residuals of the discretized Eqs. (1)–(2) reaches 10^{-3} . A solution is considered a steady state once this criterion is met. The solution of the velocity and pressure fields in Eqs. (1)–(2) is a function of the geometry of the cylinder's cross-section. Hence, to generate the labeled data, we consider seven different classical geometries for the cross-section, with varying sizes and rotations, as depicted and described in Table 1. In our computational setting, the cross-sectional geometry of the cylinder has a characteristic length L . Note that the Reynolds number can be expressed as

$$\text{Re} = \frac{\rho L u_\infty}{\mu}, \quad (6)$$

and thus, it is a function of L . As illustrated in Table 1, the data set includes cross-sections with the shapes of circle [93, 100, 101], square [102], triangle [103], rectangle [104], ellipse [105], pentagon [106], and hexagon [106], each with different length scales and orientations. The range of the Reynolds number in the data set varies from 20.0 to 76.0. The total number of labeled data is 2235. A similar data set was generated and used by Kashefi et al. [73]. It is noteworthy that the process of data generation, starting from defining the domain geometry, mesh generation, executing the numerical solver, and storing the velocity and pressure fields, is fully automated using C++ codes, Python scripts, journal files, and batch files.

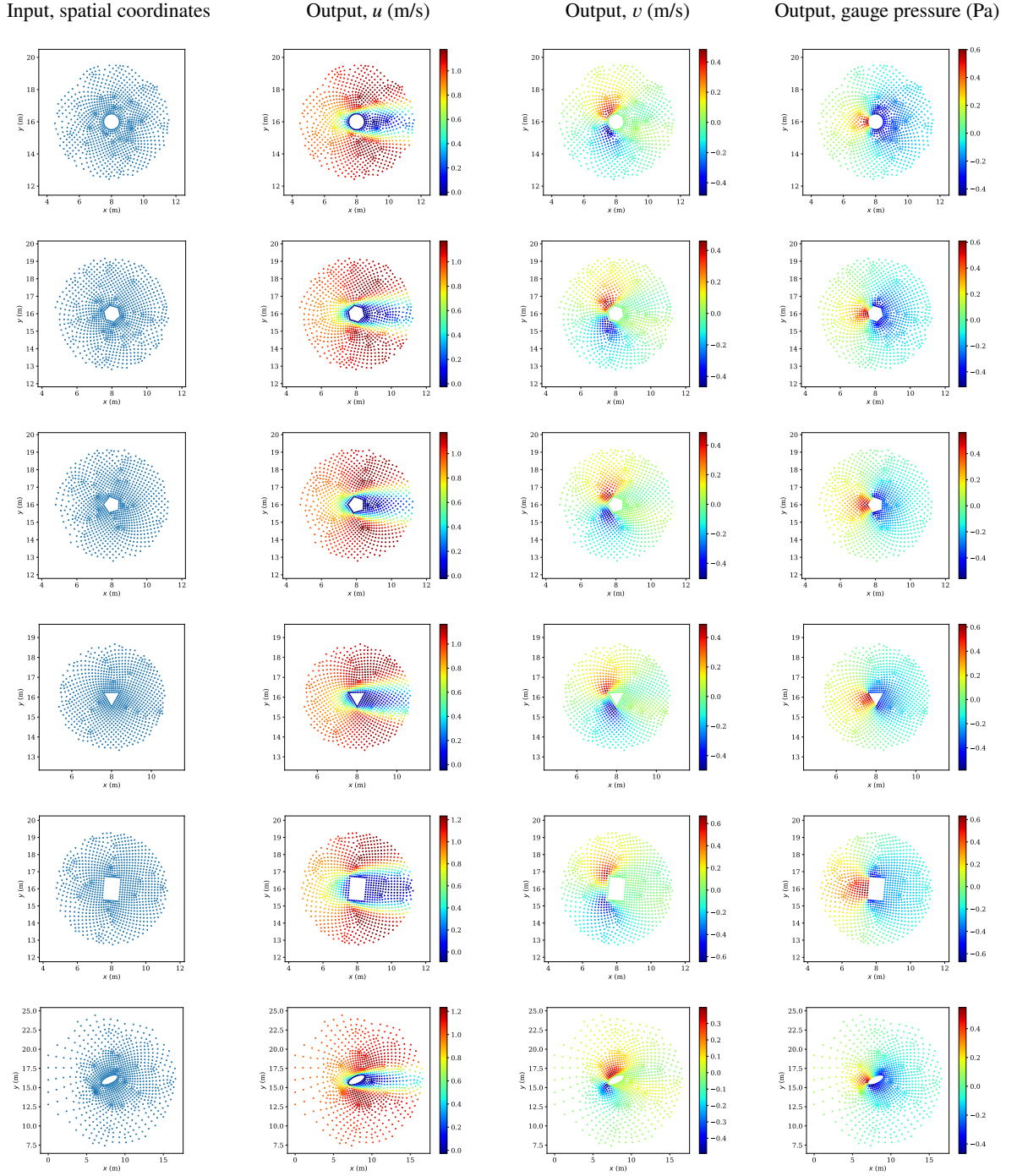


Figure 2: Examples of input and output of the generated dataset in the form of point clouds

In the next step, we introduce u^* , v^* , and p^* as dimensionless variables associated respectively with u , v , and p . They are defined as follows:

$$u^* = \frac{u}{u_\infty}, \quad (7)$$

$$v^* = \frac{v}{u_\infty}, \quad (8)$$

$$p^* = \frac{p - p_0}{\rho u_\infty^2}, \quad (9)$$

where the atmospheric pressure is indicated by p_0 . As the next stage, we scale the spatial coordinates as well as the velocity and pressure fields of the training set to a range of $[-1, 1]$ using the following formulation:

$$\{\phi'\} = 2 \left(\frac{\{\phi\} - \min(\{\phi\})}{\max(\{\phi\}) - \min(\{\phi\})} \right) - 1, \quad (10)$$

where the set $\{\phi\}$ includes $\{x^*\}$, $\{y^*\}$, $\{u^*\}$, $\{v^*\}$, and $\{p^*\}$. According to Eq. (10), the scaled input and output data are represented by $\{x'\}$, $\{y'\}$, $\{u'\}$, $\{v'\}$, and $\{p'\}$. Note that scaling the input spatial coordinates to the range of $[-1, 1]$ is critical for KA-PointNet; otherwise, the training loss diverges. The prediction of KA-PointNet will be scaled back to the physical domain using a similar process.

The input of KA-PointNet is a set of spatial points, and the output is the corresponding velocity and pressure values at those points. To represent the generated data set as point clouds suitable for feeding into KA-PointNet, we select the grid vertices of finite volume meshes as points. Note that a similar practice has been conducted by Kashefi et al. [73]. Additionally, since we are more interested in the solution of the velocity and pressure fields around the cylinder and in the wake region compared to the far field, we select the first N closest grid vertices to the center of mass of the cylinder to establish the point clouds. In this study, we set $N = 1024$, so each point cloud contains 1024 points. Because the distribution of grid vertices in the unstructured finite volume meshes is nonuniform, the range of spatial coordinates of points in each point cloud varies. In the described data set, $x_{\min} \in [0, 5.46 \text{ m}]$, $x_{\max} \in [10.55 \text{ m}, 22.14 \text{ m}]$, $y_{\min} \in [0, 13.41 \text{ m}]$, and $y_{\max} \in [18.57 \text{ m}, 32 \text{ m}]$. A few examples of input and output of the labeled data set are shown in Fig. 2. Note that these types of flexibility in input domain definition are the advantage of point-cloud and graph-based neural networks compared to traditional CNNs (e.g., see Ref [64]). We randomly split the data into three categories of the training set (1772 data), validation set (241 data), and test set (222 data).

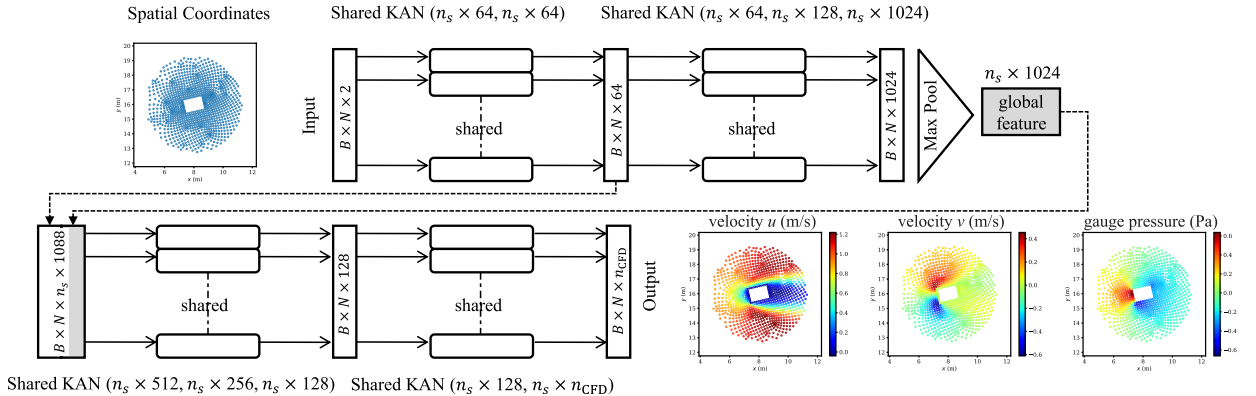


Figure 3: Architecture of the Kolmogorov-Arnold PointNet. Shared KANs with the labels (A_1, A_2) and (A_1, A_2, A_3) are explained in the text. n_{CFD} denotes the number of CFD variables. N is the number of points in the point clouds. B represents the batch size. n_s is the global scaling parameter used to control the network size.

3. Kolmogorov-Arnold PointNet

3.1. Kolmogorov-Arnold network layers with Jacobi polynomials

In this subsection, we explain the concept of Kolmogorov-Arnold network layers. For simplicity, let us consider a Kolmogorov-Arnold network with one hidden layer, where the input of the network is a vector \mathbf{r} of size d_{input} and the

output of the network is a vector \mathbf{s} of size d_{output} . In this setup, the one-layer Kolmogorov-Arnold network relates the input to the output as

$$\mathbf{s}_{d_{\text{output}} \times 1} = \mathbf{\Phi}_{d_{\text{output}} \times d_{\text{input}}} \mathbf{r}_{d_{\text{input}} \times 1}, \quad (11)$$

where the matrix $\mathbf{\Phi}_{d_{\text{output}} \times d_{\text{input}}}$ is expressed as

$$\mathbf{\Phi}_{d_{\text{output}} \times d_{\text{input}}} = \begin{bmatrix} \psi_{1,1}(\cdot) & \psi_{1,2}(\cdot) & \cdots & \psi_{1,d_{\text{input}}}(\cdot) \\ \psi_{2,1}(\cdot) & \psi_{2,2}(\cdot) & \cdots & \psi_{2,d_{\text{input}}}(\cdot) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{d_{\text{output}},1}(\cdot) & \psi_{d_{\text{output}},2}(\cdot) & \cdots & \psi_{d_{\text{output}},d_{\text{input}}}(\cdot) \end{bmatrix}, \quad (12)$$

where $\psi(z)$ is defined as

$$\psi(z) = \sum_{i=0}^n \Lambda_i P_i^{(\alpha,\beta)}(z), \quad (13)$$

where $P_i^{(\alpha,\beta)}(z)$ is the Jacobi polynomial of order i , n is the order of the polynomial ψ , and Λ_i are trainable parameters. Hence, the number of trainable parameters (i.e., Λ_i) of this KAN layer is equal to $(n+1) \times d_{\text{output}} \times d_{\text{input}}$. We implement $P_n^{(\alpha,\beta)}(z)$ recursively, using the following relationship [107]

$$P_n^{(\alpha,\beta)}(z) = (A_n z + B_n) P_{n-1}^{(\alpha,\beta)}(z) + C_n P_{n-2}^{(\alpha,\beta)}(z), \quad (14)$$

where A_n , B_n , and C_n coefficients are defined as follow

$$A_n = \frac{(2n + \alpha + \beta - 1)(2n + \alpha + \beta)}{2n(n + \alpha + \beta)}, \quad (15)$$

$$B_n = \frac{(2n + \alpha + \beta - 1)(\alpha^2 - \beta^2)}{2n(n + \alpha + \beta)(2n + \alpha + \beta - 2)}, \quad (16)$$

$$C_n = \frac{-2(n + \alpha - 1)(n + \beta - 1)(2n + \alpha + \beta)}{2n(n + \alpha + \beta)(2n + \alpha + \beta - 2)}, \quad (17)$$

along with the following two initial conditions:

$$P_0^{(\alpha,\beta)}(z) = 1, \quad (18)$$

$$P_1^{(\alpha,\beta)}(z) = \frac{1}{2}(\alpha + \beta + 2)z + \frac{1}{2}(\alpha - \beta). \quad (19)$$

Since $P_n^{(\alpha,\beta)}(z)$ is constructed recursively, $P_i^{(\alpha,\beta)}(z)$, for $(0 \leq i \leq n)$, are constructed along the way as well. Moreover, because the input of Jacobi polynomials must fall in the interval $[-1, 1]$, the input vector \mathbf{r} should be scaled in this interval before being fed into the KAN layer. To handle this situation, we use the hyperbolic tangent function defined as follows

$$\tanh(z) = \frac{e^{2z} - 1}{e^{2z} + 1}. \quad (20)$$

Other researchers have also used this strategy (e.g., see Refs. [14, 90, 41]). Other reasonable strategies might be used for this scaling. For example, one option is to use the following formula:

$$\{z\} = 2 \left(\frac{\{z\} - \min(\{z\})}{\max(\{z\}) - \min(\{z\})} \right) - 1. \quad (21)$$

The motivation for this choice is that since we scale both input and output data using the formula given in Eq. (10), a similar scaling scheme for the intermediate layers of KA-PointNet might be beneficial. However, our machine learning experiments show that using the hyperbolic tangent function is more efficient in terms of prediction accuracy. Therefore, we use it in this study.

Several sequential KAN layers establish a KAN component. Specifically, we implement shared KAN components in PointNet. We denote a shared KAN component with two layers of sizes A_1 and A_2 as (A_1, A_2) . Similarly, we define a shared KAN component with three layers as (A_1, A_2, A_3) . After each KAN layer, we implement batch normalization [108]. This implementation is mandatory; otherwise, the training loss will diverge.

At the end of this subsection, we note that with the choice of $\alpha = \beta = 0$, the Legendre polynomial is obtained [109, 107]. The Chebyshev polynomial of the first and second kinds are respectively covered with the choice of $\alpha = \beta = -0.5$ and $\alpha = \beta = 0.5$ [109, 107]. Additionally, the Gegenbauer polynomial (or ultraspherical polynomials) is obtained if $\alpha = \beta$ [107].

3.2. Architecture

The fundamental concept of KA-PointNet is to use shared KANs instead of shared MLPs. To explain the architecture of KA-PointNet, we first discuss the general aspects and then delve into the details and formulation of shared KANs. More specifically, we will explain why it is essential to implement KANs as “shared” layers.

Figure 3 depicts the architecture of KA-PointNet. To explain the architecture, let us first denote the number of labeled data in a desired set by m . The desired set can be the training set, validation set, or test set. In a data set with m labeled data, each domain V_i ($1 \leq i \leq m$) in the set is represented by a point cloud \mathcal{X}_i with N points, where $\mathcal{X}_i = \{\mathbf{x}_j \in \mathbb{R}^d\}_{j=1}^N$. The spatial dimension of V_i and consequently \mathcal{X}_i is denoted by d . Because our focus is on two-dimensional problems in this study, we set $d = 2$. Accordingly, \mathbf{x}_j ($1 \leq j \leq N$) is the vector representing the spatial coordinates of each point in the point cloud \mathcal{X}_i . We indicate the x and y components of each \mathbf{x}_j , respectively, by x'_j and y'_j . The task of KA-PointNet is to provide an end-to-end mapping from \mathcal{X}_i to \mathcal{Y}_i , where $\mathcal{Y}_i = \{\mathbf{y}_j \in \mathbb{R}^{n_{\text{CFD}}}\}_{j=1}^N$. We demonstrate the number of desired predicted fields by n_{CFD} . In this study, we are particularly interested in the prediction of the velocity vector (in two dimensions) and the pressure field. Hence, $n_{\text{CFD}} = 3$. Based on the given definition, \mathbf{y}_j is the vector corresponding to the predicted fields at the spatial point \mathbf{x}_j , and thus, has three components of u'_j , v'_j , and p'_j . This procedure can be mathematically formulated as:

$$(u'_j, v'_j, p'_j) = f\left(\left(x'_j, y'_j\right), g\left(\mathcal{X}_i\right)\right); \quad \forall \left(x'_j, y'_j\right) \in \mathcal{X}_i \text{ and } \forall \left(u'_j, v'_j, p'_j\right) \in \mathcal{Y}_i \text{ with } 1 \leq i \leq m \text{ and } 1 \leq j \leq N, \quad (22)$$

where f is the mapping function representing KA-PointNet. KA-PointNet is invariant to any of the $N!$ permutations of the input vector. In other words, if the input vector (\mathcal{X}_i) is randomly permuted, the geometry of the domain remains unchanged, and thus the solution (\mathcal{Y}_i) should also remain unchanged. KA-PointNet achieves this permutation invariance through a symmetric function and the use of shared KANs. In this sense, g is a symmetric function encoding the geometric features of the point cloud \mathcal{X}_i . Among the possible options for the symmetric function g , we follow the original function introduced in Ref. [71], which is the maximum function, such that

$$g\left(\mathcal{X}_i\right) = \max\left(h\left(x'_1, y'_1\right), \dots, h\left(x'_N, y'_N\right)\right); \quad \forall \left(x'_j, y'_j\right) \in \mathcal{X}_i \text{ with } 1 \leq i \leq m \text{ and } 1 \leq j \leq N, \quad (23)$$

where h is a function representing two shared KAN layers in the first branch of KA-PointNet (see Fig. 3). From a computer science perspective, $g(\mathcal{X}_i)$ is regarded as the global feature in the KA-PointNet architecture, as shown in Fig. 3. The key idea behind using KA-PointNet, as a geometric deep learning model, is that the predicted fields at each spatial point depend on both the spatial coordinates of that specific point and the overall geometry of the domain constructed by all the points, including the specific point. This fact can be seen in Eqs. (22)–(23).

We introduce n_s as a global scaling variable that controls the size of KA-PointNet. Furthermore, We denote the batch size by B , referring to the number of point clouds (i.e., \mathcal{X}_i and \mathcal{Y}_i pairs) fed into KA-PointNet at each epoch. As can be seen from Fig. 3, the input of KA-PointNet is a three-dimensional tensor of size $B \times N \times 2$. Following this, two sequential shared KANs are applied, with sizes $(n_s \times 64, n_s \times 64)$ and $(n_s \times 64, n_s \times 128, n_s \times 1024)$, as shown in Fig. 3. The maximum function then generates the global feature with a size of $n_s \times 1024$. As illustrated in Fig. 3, this global feature is concatenated with an intermediate feature tensor of size $B \times N \times (n_s \times 64)$, resulting in a new tensor with dimension $B \times N \times (n_s \times 1088)$. Next, two additional sequential shared KANs are applied within the KA-PointNet architecture, with sizes $(n_s \times 512, n_s \times 256, n_s \times 128)$ and $(n_s \times 128, n_{\text{CFD}})$, respectively. The result of the previous step is a tensor with dimensions $B \times N \times n_{\text{CFD}}$, as depicted in Fig. 3. It is important to note that suitable values for n_s should lead to positive integers for the size of shared KANs.

3.3. Shared Kolmogorov-Arnold networks

The idea of shared KANs is distinct from regular fully connected layers. Here, we illustrate the concept of shared KANs with a straightforward example. Consider the initial layer in the shared KAN given in the first branch of KA-PointNet, which has a size of $(n_s \times 64, n_s \times 64)$, as shown in Fig. 3. Assume $n_s = 1$ for simplicity. Particularly, we focus on the first shared KAN layer with a size of 64. The transposed input vector \mathcal{X}_i can be expressed as

$$\mathcal{X}_i^{\text{tr}} = \begin{bmatrix} x'_1 & x'_2 & \cdots & x'_N \\ y'_1 & y'_2 & \cdots & y'_N \end{bmatrix}. \quad (24)$$

After processing \mathcal{X}_i through the first shared KAN layer, the result is a matrix of size $64 \times N$ and can be described as:

$$\begin{bmatrix} \mathbf{s}_{64 \times 1}^{(1)} & \mathbf{s}_{64 \times 1}^{(2)} & \cdots & \mathbf{s}_{64 \times 1}^{(N)} \end{bmatrix}, \quad (25)$$

where $\mathbf{s}_{64 \times 1}^{(1)}, \mathbf{s}_{64 \times 1}^{(2)}, \dots, \mathbf{s}_{64 \times 1}^{(N)}$ are vectors, computed as follows:

$$\begin{aligned} \mathbf{s}_{64 \times 1}^{(1)} &= \Phi_{64 \times 2} \begin{bmatrix} x'_1 \\ y'_1 \end{bmatrix}, \\ \mathbf{s}_{64 \times 1}^{(2)} &= \Phi_{64 \times 2} \begin{bmatrix} x'_2 \\ y'_2 \end{bmatrix}, \\ &\vdots \\ \mathbf{s}_{64 \times 1}^{(N)} &= \Phi_{64 \times 2} \begin{bmatrix} x'_N \\ y'_N \end{bmatrix}, \end{aligned} \quad (26)$$

where $\Phi_{64 \times 2}$ is the shared KAN layer. As can be observed in Eq. (26), the same (shared) $\Phi_{64 \times 2}$ used for each spatial point in the domain, corresponding to the vector $[x'_j, y'_j]^{\text{tr}}$, where $1 \leq j \leq N$. This is the reason it is called shared KANs. This procedure is similarly applied to the remaining layers. With this strategy, it becomes evident that each point is independently processed within the KA-PointNet framework. The only instance when the points interact is when the global feature is determined (see Eq. (23)).

3.4. Training

Mean squared error is used as the loss function, expressed as

$$\mathcal{L} = \frac{1}{3 \times N} \left(\sum_{i=1}^N [(u'_i - \tilde{u}'_i)^2 + (v'_i - \tilde{v}'_i)^2 + (p'_i - \tilde{p}'_i)^2] \right), \quad (27)$$

where \tilde{u}' , \tilde{v}' , and \tilde{p}' are the predicted velocity and pressure fields, respectively. Note that for error analysis and visualization, predicted variables are scaled back into the physical domain. The Adam optimizer [91] is used with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\hat{\epsilon} = 10^{-8}$. We refer the audience to Ref. [91] for the mathematical definitions of β_1 , β_2 , and $\hat{\epsilon}$ in the Adam optimizer [91]. For a fair comparison between different setups for KA-PointNet, we perform all the training procedures on an NVIDIA A100 Tensor Core GPU with 80 gigabytes of RAM. We use mini-batch gradient descent with a batch size of 128 and a constant learning rate of 5×10^{-4} . To avoid overfitting, the evolution of the loss function for both the training and validation sets is monitored during the training procedure. We discuss the number of trainable parameters, the history of the loss function, and the computational cost of training in detail in Sect. 4.

Table 2

Computational cost and error analysis of the velocity and pressure fields predicted by Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the test set containing 222 unseen geometries for different degrees of the Jacobi polynomial. Here, $n_s = 1$ is set. In the Jacobi polynomial, $\alpha = \beta = 1$ is set. $\|\dots\|$ indicates the L^2 norm.

Polynomial degree	2	3	4	5	6
Average $\ \tilde{u} - u\ /\ u\ $	1.43377E-2	1.73537E-2	1.67633E-2	1.17467E-2	1.07525E-2
Maximum $\ \tilde{u} - u\ /\ u\ $	1.66709E-1	1.40088E-1	1.37067E-1	1.40614E-1	1.68387E-1
Minimum $\ \tilde{u} - u\ /\ u\ $	5.40694E-3	7.22247E-3	6.70142E-3	5.01482E-3	4.82974E-3
Average $\ \tilde{v} - v\ /\ v\ $	5.28306E-2	5.75906E-2	5.38728E-2	4.81333E-2	5.08094E-2
Maximum $\ \tilde{v} - v\ /\ v\ $	5.25991E-1	4.42245E-1	4.48167E-1	4.61926E-1	5.44426E-1
Minimum $\ \tilde{v} - v\ /\ v\ $	2.79826E-2	2.05299E-2	2.51025E-2	2.11034E-2	2.36179E-2
Average $\ \tilde{p} - p\ /\ p\ $	3.49754E-2	4.75586E-2	4.67659E-2	3.30555E-2	3.47443E-2
Maximum $\ \tilde{p} - p\ /\ p\ $	1.73144E-1	1.55694E-1	1.69091E-1	1.64282E-1	1.90192E-1
Minimum $\ \tilde{p} - p\ /\ p\ $	1.71494E-2	1.75673E-2	1.93152E-2	1.50220E-2	1.44066E-2
Training time per epoch (s)	3.71432	6.52910	10.12893	14.28542	19.26667
Number of trainable parameters	2660480	3545728	4430976	5316224	6201472

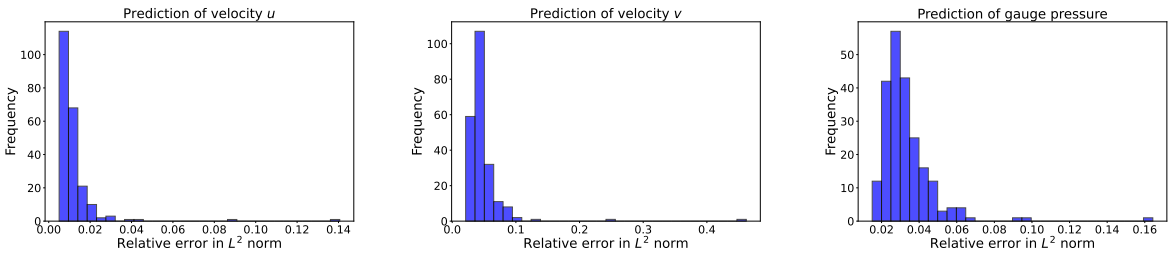


Figure 4: Histograms of the relative pointwise error in L^2 norm for the velocity and pressure fields predicted by Kolmogorov-Arnold PointNet (i.e., KA-PointNet). The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

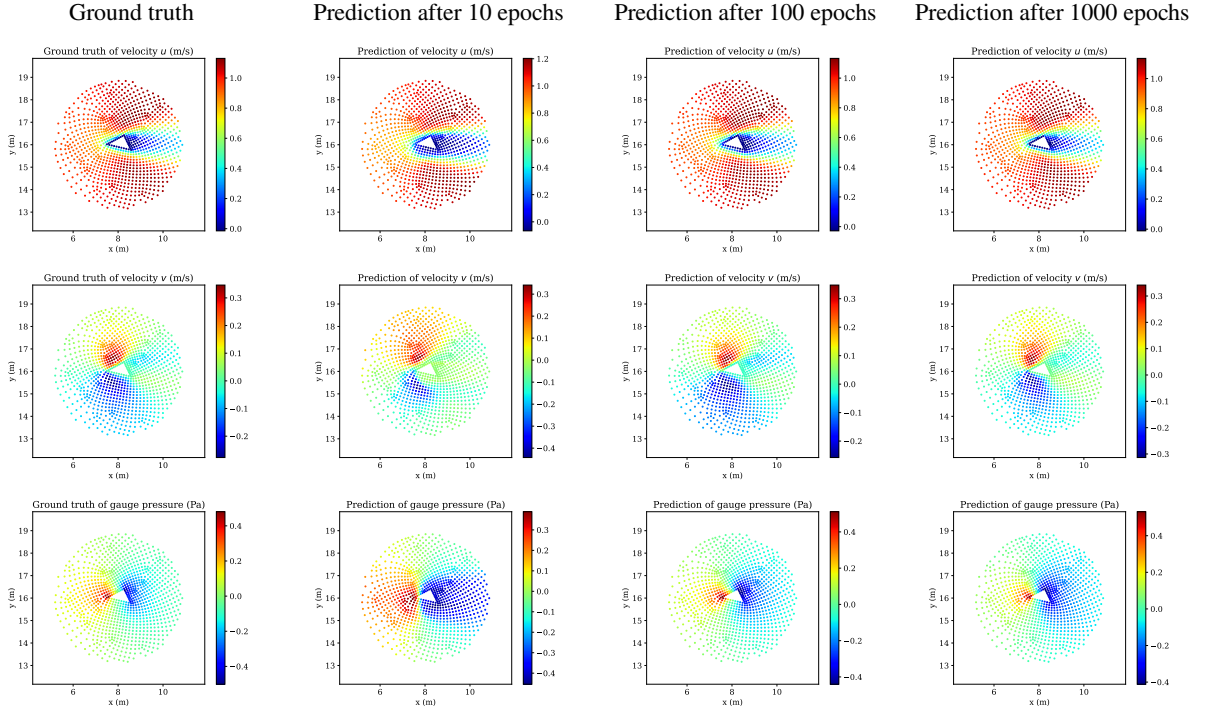


Figure 5: A comparison between the ground truth and prediction of the Kolmogorov-Arnold PointNet for the velocity and pressure fields after 10, 100, and 1000 epochs. The Jacobian polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

Table 3

Error analysis of the velocity and pressure fields predicted by Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the test set containing 222 unseen geometries for different values of α and β in Jacobi polynomials. The degree of the Jacobi polynomials used is 3. Here, $n_s = 1$ is set. $\|\dots\|$ indicates the L^2 norm.

	$\alpha = \beta = 0$	$\alpha = \beta = -0.5$	$\alpha = \beta = 0.5$	$\alpha = \beta = 1$	$2\alpha = \beta = 2$	$\alpha = 2\beta = 2$
Average $\ \tilde{u} - u\ /\ u\ $	1.36404E-2	1.21807E-2	1.24601E-2	1.73537E-2	1.65403E-2	2.98503E-2
Maximum $\ \tilde{u} - u\ /\ u\ $	1.37743E-1	1.43526E-1	1.42853E-1	1.40088E-1	1.51214E-1	1.54898E-1
Minimum $\ \tilde{u} - u\ /\ u\ $	5.46195E-3	6.51784E-3	6.28078E-3	7.22247E-3	7.76342E-3	1.00267E-2
Average $\ \tilde{v} - v\ /\ v\ $	6.38935E-2	4.70539E-2	4.32044E-2	5.75906E-2	5.26547E-2	9.47980E-2
Maximum $\ \tilde{v} - v\ /\ v\ $	4.34468E-1	4.40415E-1	4.76479E-1	4.42245E-1	5.06639E-1	4.42623E-1
Minimum $\ \tilde{v} - v\ /\ v\ $	2.76012E-2	1.94763E-2	1.92582E-2	2.05299E-2	2.09325E-2	3.83024E-2
Average $\ \tilde{p} - p\ /\ p\ $	4.13612E-2	2.49465E-2	3.55946E-2	4.75586E-2	3.70187E-2	6.39906E-2
Maximum $\ \tilde{p} - p\ /\ p\ $	1.56200E-1	1.66064E-1	1.62299E-1	1.55694E-1	1.68458E-1	1.61651E-1
Minimum $\ \tilde{p} - p\ /\ p\ $	1.53982E-2	9.89027E-3	1.36161E-2	1.75673E-2	1.13910E-2	2.14740E-2

4. Results and discussion

4.1. General analysis

To analyze the performance of KA-PointNet, we first focus on a specific case where the degree of the Jacobi polynomial (see Eqs. (13)–(14)) is set to 5, $\alpha = \beta = 1$ (leading to the Gegenbauer polynomial), and $n_s = 1$. All the information provided in this paragraph and the next three paragraphs is associated with this setup for KA-PointNet. Analysis of the relative pointwise error (L^2 norm) of the predicted velocity and pressure fields for 222 geometries of the test set is tabulated in Table 2 (see the fifth column). Based on the information collected in Table 2, the average relative error over all these 222 geometries for u , v , and p variables are approximately 1.18%, 4.82%, and 3.31%, respectively,

Table 4

Computational cost and error analysis of the velocity and pressure fields predicted by Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the test set containing 222 unseen geometries for different choices of n_s . The Jacobi polynomial used has a degree of 3, with $\alpha = \beta = 1$. $\|\dots\|$ indicates the L^2 norm.

n_s	0.5	0.75	1	1.25	1.5	2
Average $\ \tilde{u} - u\ /\ u\ $	2.63597E-2	2.57360E-2	1.73537E-2	1.82822E-2	1.68094E-2	1.69208E-2
Maximum $\ \tilde{u} - u\ /\ u\ $	1.30750E-1	1.56881E-1	1.40088E-1	1.49949E-1	1.47257E-1	1.50143E-1
Minimum $\ \tilde{u} - u\ /\ u\ $	1.30994E-2	1.04247E-2	7.22247E-3	7.63074E-3	6.98780E-3	1.08372E-2
Average $\ \tilde{v} - v\ /\ v\ $	1.00292E-1	8.05285E-2	5.75906E-2	5.34776E-2	5.66996E-2	5.48352E-2
Maximum $\ \tilde{v} - v\ /\ v\ $	4.19706E-1	4.83207E-1	4.42245E-1	5.15754E-1	4.59807E-1	4.85161E-1
Minimum $\ \tilde{v} - v\ /\ v\ $	5.58101E-2	3.13250E-2	2.05299E-2	1.95743E-2	2.54641E-2	2.49093E-2
Average $\ \tilde{p} - p\ /\ p\ $	9.86696E-2	6.35050E-2	4.75586E-2	4.21600E-2	3.63245E-2	4.25412E-2
Maximum $\ \tilde{p} - p\ /\ p\ $	2.88919E-1	1.82941E-1	1.55694E-1	1.60423E-1	1.76526E-1	1.63125E-1
Minimum $\ \tilde{p} - p\ /\ p\ $	4.67306E-2	3.69896E-2	1.75673E-2	1.75982E-2	1.53689E-2	1.74415E-2
Training time per epoch (s)	2.83741	4.53013	6.52910	8.85618	11.15280	16.69120
Number of trainable parameters	888128	1995744	3545728	5538080	7972800	14169344

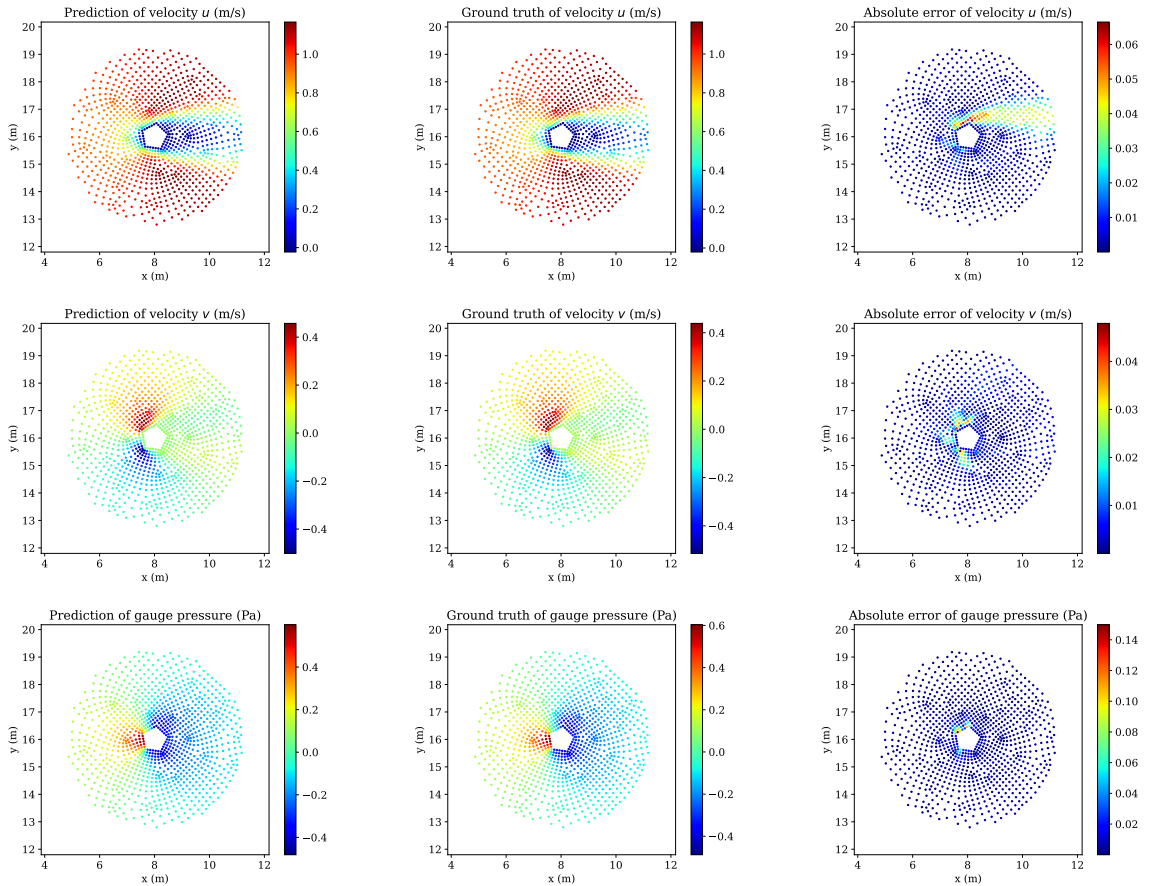


Figure 6: The first set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

Kolmogorov-Arnold PointNet

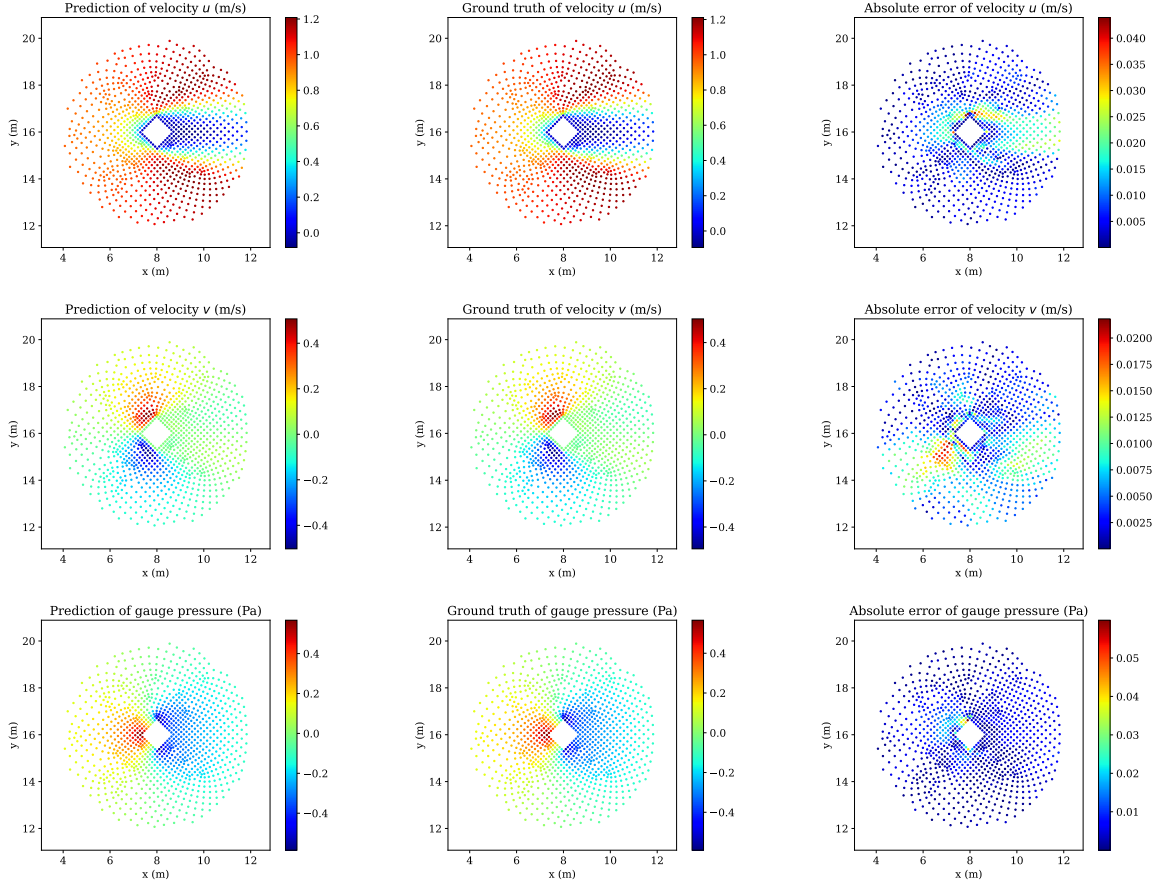


Figure 7: The second set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

showing a strong performance of KA-PointNet. The histogram of the distribution of the relative errors (L^2 norm) is shown in Fig. 4. In general, we observed that the y component of the velocity vector (v) experiences a higher level of errors compared to the x component of the velocity vector (u) and the pressure field (p).

Figure 5 visually compares the ground truth and the predictions of velocity and pressure fields by KA-PointNet (with the Jacobi polynomial of degree 5, $\alpha = \beta = 1$, and $n_s = 1$) after 10, 100, and 1000 epochs for one of the geometries of the test set. As shown in Fig. 5, KA-PointNet accurately predicts the general structure of flow fields after just 10 epochs. After 100 epochs, the prediction accuracy improves, particularly near the object surface, where the no-slip velocity boundary condition is better satisfied. Additionally, the scale of predicted values becomes more accurate. It is observed that after just 100 epochs, the prediction of the u variable achieves a higher level of accuracy compared to the predictions of the v and p variables. This observation can explain why, in general, the relative error of the u variable is lower than that of the v and p variables over the test set, as shown in the histogram in Fig. 4 as well as the data in Table 2. Comparing the predictions at 1000 epochs with those at 100 epochs, the velocity vector in the y direction (v) and the pressure field (p) experience further modifications and increased accuracy, while the velocity vector in the x direction (u) remains approximately unchanged.

Figures 6–15 present a comparison between the ground truth and the KA-PointNet predictions, along with the absolute pointwise error for the velocity and pressure fields, for ten cylinders with different cross-sectional geometries from the test set. As observed in Figs. 6–15, there is excellent agreement between the ground truth and the predictions, despite the non-uniform distributions of the point clouds and varying spatial coordinate ranges within the domains of the test set. As seen in Fig. 11, KA-PointNet successfully predicts the flow separation (i.e., when the u variable

Kolmogorov-Arnold PointNet

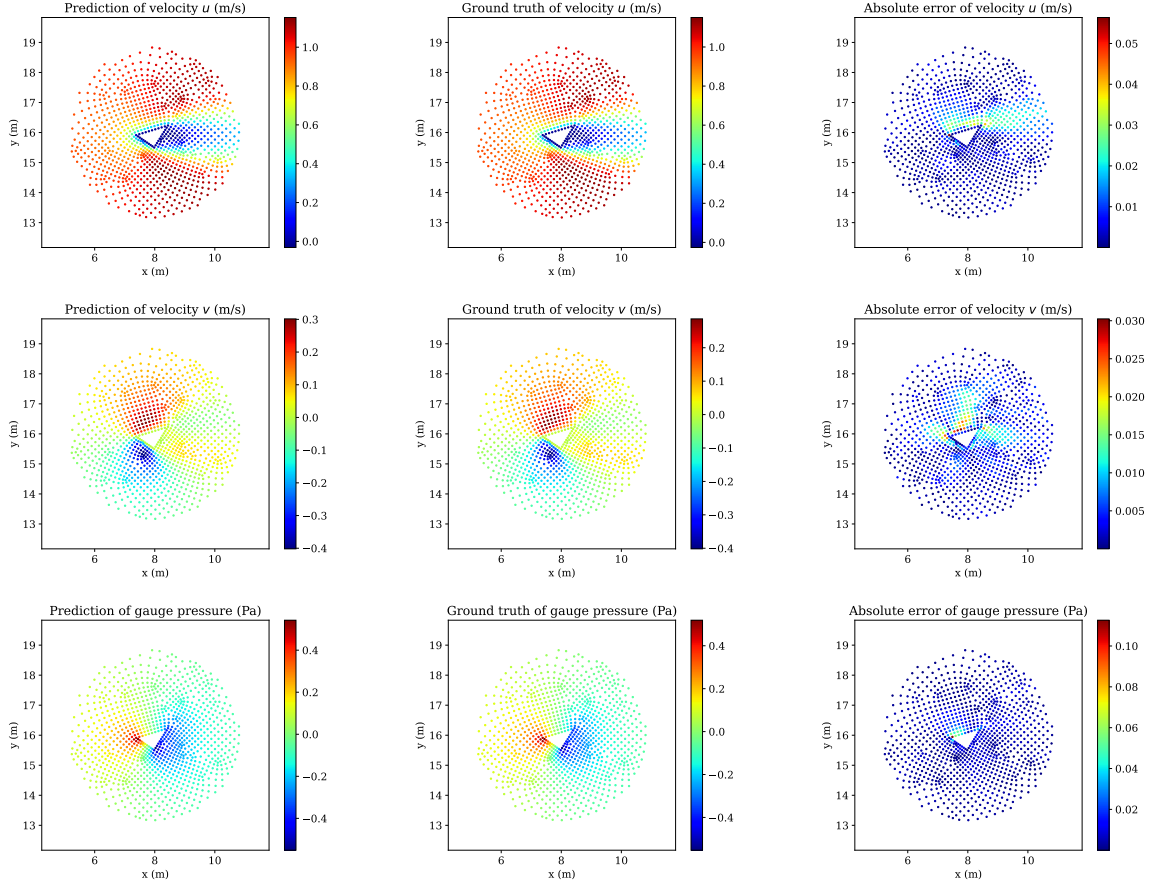


Figure 8: The third set of examples comparing the ground truth to predictions of the Kolmogorov-Arnold PointNet for velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

becomes negative). As a general observation in Figs. 6–15, the maximum absolute pointwise error typically occurs at or near the surface of the cylinders, where there are the highest variations in the geometry of the cylinders across different data in the test set.

Distributions of absolute pointwise errors are shown in Fig. 16 for geometries from the test set where the relative pointwise error (L^2 norm) for the velocity vector components (u and v) and the pressure field (p) becomes maximum (displayed in the first row of Fig. 16) and minimum (displayed in the second row of Fig. 16). The numerical values of these errors are listed in Table 2. From Fig. 16, it is observed that the maximum relative pointwise error (L^2 norm) for the velocity and pressure fields occurs in a single domain with an elliptical cross-section for the cylinder, whereas the minimum relative pointwise error (L^2 norm) for the velocity and pressure fields occurs in three different domains, each with a distinct geometry for the cross-section of the cylinder. Comparing geometries with maximum and minimum relative pointwise error (L^2 norm), it is recognized that the maximum error occurs when the spatial coordinates of the domain (as a point cloud) reach their maximum possible value compared to other domains in the test set. Note that from the histogram shown in Fig. 4, we see that the geometry with the maximum error is indeed a single outlier and falls outside the relative errors (L^2 norm) distribution of the histogram.

Returning to Table 2, with $n_s = 1$ and $\alpha = \beta = 1.0$, we analyze the performance of KA-PointNet as a function of the degree of the Jacobi polynomial. Based on the relative pointwise error (L^2 norm) tabulated in Table 2, the average error across the test set (222 unseen data) varies approximately from 1.07% to 1.74% for the u variable, 4.82% to 5.76% for the v variable, and 3.31% to 4.76% for the p variable. Additionally, we observe that increasing the polynomial degree does not necessarily improve prediction accuracy. For instance, when comparing the average relative errors from a

Kolmogorov-Arnold PointNet

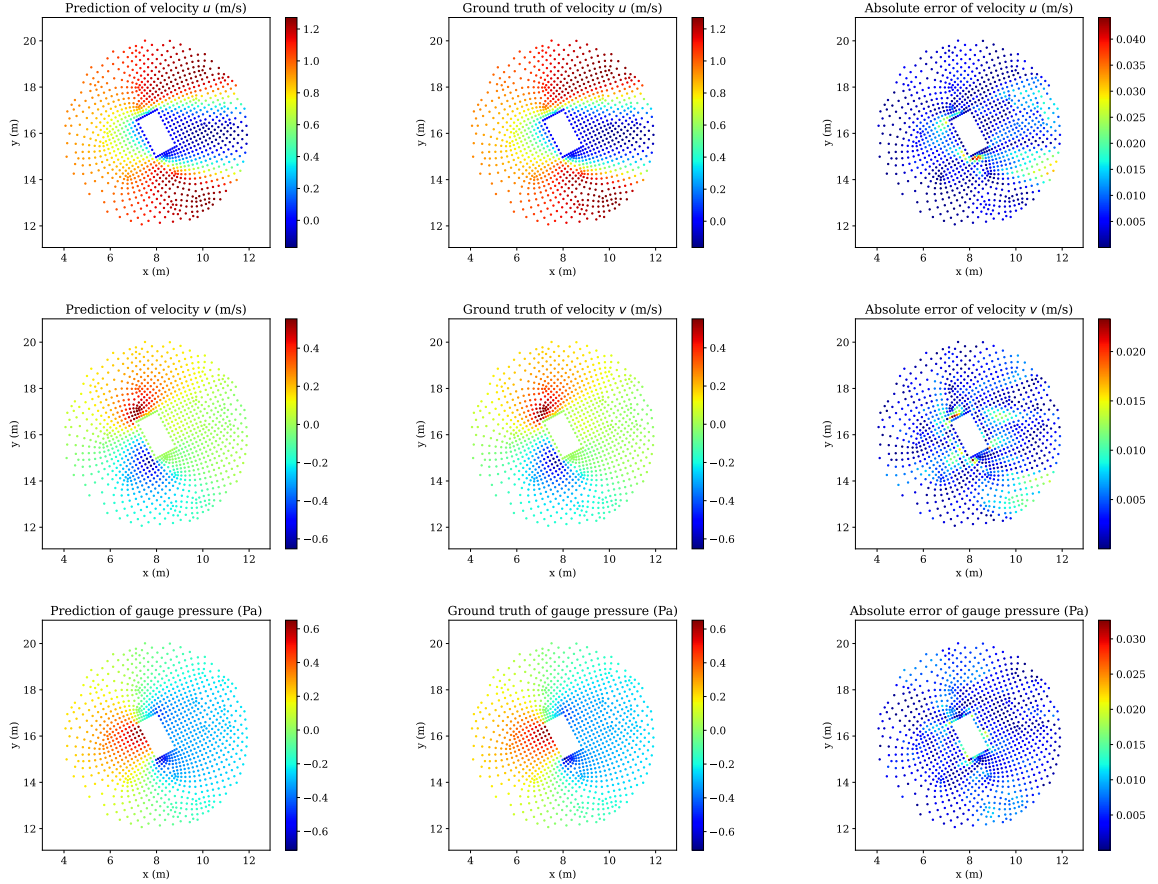


Figure 9: The fourth set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

degree of 2 to a degree of 3, the prediction accuracy decreases by 0.30160%, 0.47600%, and 1.25832% for the u , v , and p variables, respectively. On the other hand, the Jacobi polynomials of degrees 5 and 6 yield the lowest errors, as indicated in Table 2. Regarding the computational cost of training, higher degrees of Jacobi polynomials increase the number of trainable parameters (see Eq. (13)), thus extending the training time. Comparing the computational cost of using the higher-order Jacobi polynomials with the improvement in prediction accuracy, it appears that choosing even the polynomial of degree 2 is an optimized decision.

Next, we investigate the effect of the choice of α and β in the Jacobi polynomial (see Eqs. (14)–(17)) on the performance of KA-PointNet. To reach this goal, we set the degree of the Jacobi polynomial to 3 and $n_s = 1$. Specifically, we consider the following selections: $\alpha = \beta = 0$ (leading to the Legendre polynomial), $\alpha = \beta = -0.5$ (leading to the Chebyshev polynomial of the first kind), $\alpha = \beta = 0.5$ (leading to the Chebyshev polynomial of the second kind), $\alpha = \beta = 1$ (leading to the Gegenbauer polynomial), $2\alpha = \beta = 2$, and $\alpha = 2\beta = 2$. The outcomes of this investigation are presented in Table 3. According to Table 3, the Chebyshev polynomials (i.e., $\alpha = \beta = -0.5$ and $\alpha = \beta = 0.5$) lead to the highest accuracy among the options, considering the average relative pointwise error (L^2 norm) as the judgment criterion. More specifically, the Chebyshev polynomial of the first kind leads to more accurate predictions compared to the Chebyshev polynomial of the second kind. The lowest performance, in terms of prediction accuracy, occurs when $\alpha = 2\beta = 2$. Although the choice of α and β affects the accuracy of predictions, we observe that all the choices result in errors within an acceptable range. Considering all the choices, the average relative pointwise error (L^2 norm) of the u , v , and p variables over the test set containing 222 data with various geometries do not exceed 2.99%, 9.48%, and 6.40%, respectively.

Kolmogorov-Arnold PointNet

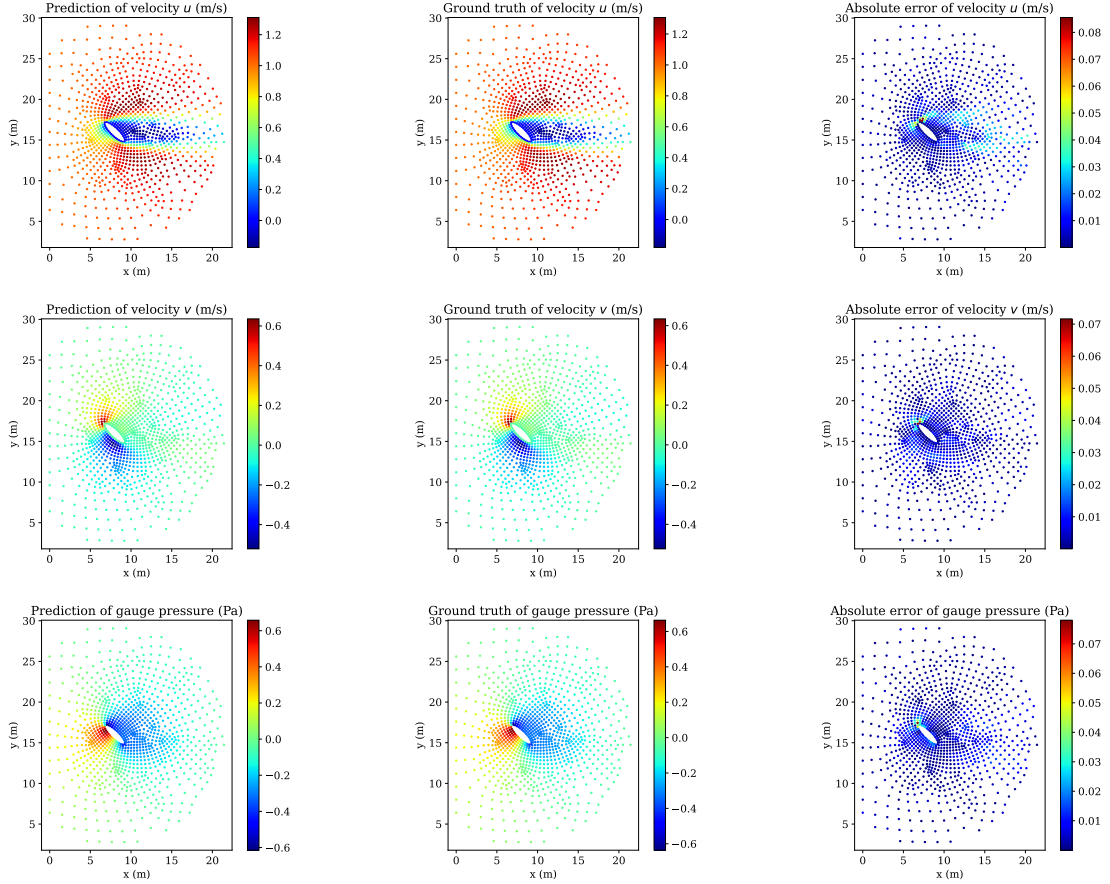


Figure 10: The fifth set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

To investigate the influence of the size of KA-PointNet, we tabulate the relative pointwise error (L^2 norm) of the predicted velocity and pressure fields as a function of the defined global scaling parameter (n_s) in Table 4. Using the Jacobi polynomial of degree 3 and setting $\alpha = \beta = 1$, we conduct machine learning experiments with different values of n_s . Based on the error analysis presented in Table 4, we observe that increasing n_s and thereby enlarging KA-PointNet enhances prediction accuracy as n_s is increased from 0.5 to 0.75 and from 0.75 to 1. This scenario changes when selecting n_s values greater than 1 (i.e., $n_s = 1.25$, $n_s = 1.5$, and $n_s = 2$). As shown in Table 4, there is no clear trend in the relative error for a specific field. For example, increasing n_s beyond 1 may lead to an increase in the error for the velocity prediction while reducing the error for the pressure prediction. Nonetheless, all errors for all fields are lower than those obtained with $n_s = 0.75$ and $n_s = 0.5$. Generally speaking, increasing the size of KA-PointNet while keeping the number of training data fixed raises the potential for overfitting. This explains why we observe a slight decrease in the accuracy for predicting some fields when n_s is increased beyond a certain threshold (e.g., $n_s = 1$ in the current experiment). According to Table 4, it is noticeable that increasing n_s leads to a rise in the number of trainable parameters and, consequently, the training time. For instance, the average training time per epoch for $n_s = 2$ is approximately 2.56 times greater than that for $n_s = 1$. Therefore, the design of KA-PointNet and the setting of hyperparameters such as n_s are critical considerations from both accuracy and computational cost perspectives. Additionally, the machine learning experiments discussed here demonstrate that despite the complexity of the KA-PointNet architecture, a desirable level of accuracy can be achieved by fine-tuning the single parameter, n_s .

Based on the design and architecture of KA-PointNet, the degree of Jacobi polynomials (n) and the global scaling parameter (n_s) are two important factors for controlling the size of the network. Considering Table 2 and Table 4

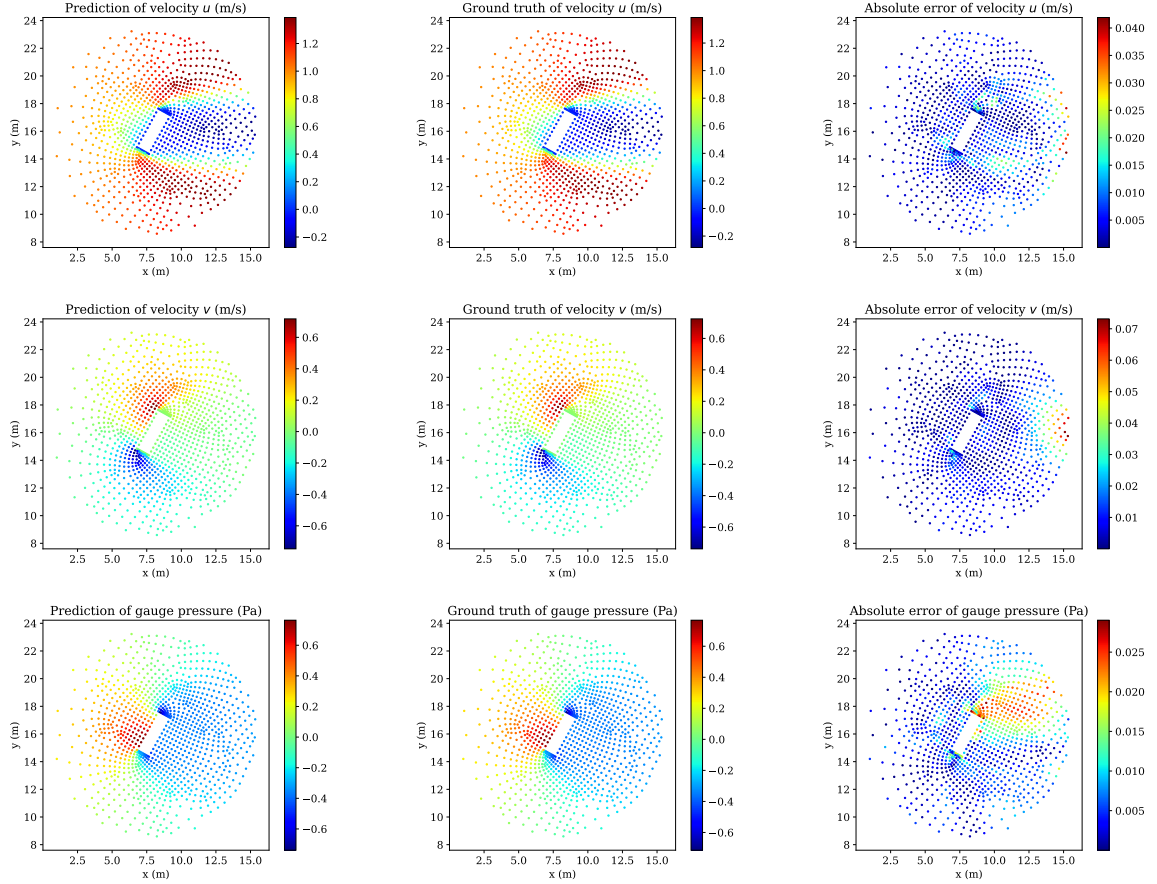


Figure 11: The sixth set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

simultaneously, for instance, the number of trainable parameters for KA-PointNet with the Jacobi polynomial degree of 5, $\alpha = \beta = 1$, and setting $n_s = 1$, is 5316224, with an average training time of 14.28542 seconds per epoch. On the other hand, consider KA-PointNet with the Jacobi polynomial degree of 3, $\alpha = \beta = 1$, and setting $n_s = 1.25$. This network has 5538080 trainable parameters, leading to an average training time of 8.85618 seconds per epoch. These two networks have approximately equal numbers of trainable parameters. Comparing the information in Table 4 and Table 2, the prediction accuracy for all the velocity and pressure fields is higher in the first configuration (i.e., higher-order Jacobi polynomials but a smaller value of n_s). Although the number of trainable parameters is approximately equal in both cases, the training time of the first configuration is approximately 1.614 times greater than that of the second configuration. From a computer science perspective, this means that given limited RAM, if there is no restriction on training time, the first configuration produces more accurate outcomes.

The last topic addressed in this subsection is the investigation of the performance of utilizing layer normalization in the architecture of KA-PointNet. One may refer to Ref. [110] for details of layer normalization and its differences with batch normalization from a computer science perspective. Note that all the results reported up to this point pertain to the architecture of KA-PointNet with batch normalization. As a machine learning experiment, we replace batch normalization with layer normalization in a KA-PointNet with a Jacobi polynomial of degree 3, setting $\alpha = \beta = 1$ and $n_s = 1$. The first consequence of using layer normalization is an increase in the number of trainable parameters from 3545728 to 8390656, requiring more RAM. Measuring the average relative pointwise error (L^2 norm) over the test set (222 data points) for the u , v , and p variables, they are respectively $1.32751\text{E}-2$, $4.77946\text{E}-2$, and $4.20826\text{E}-2$. The corresponding quantities for KA-PointNet with batch normalization are $1.73537\text{E}-2$, $5.75906\text{E}-2$, and $4.75586\text{E}-2$.

Kolmogorov-Arnold PointNet

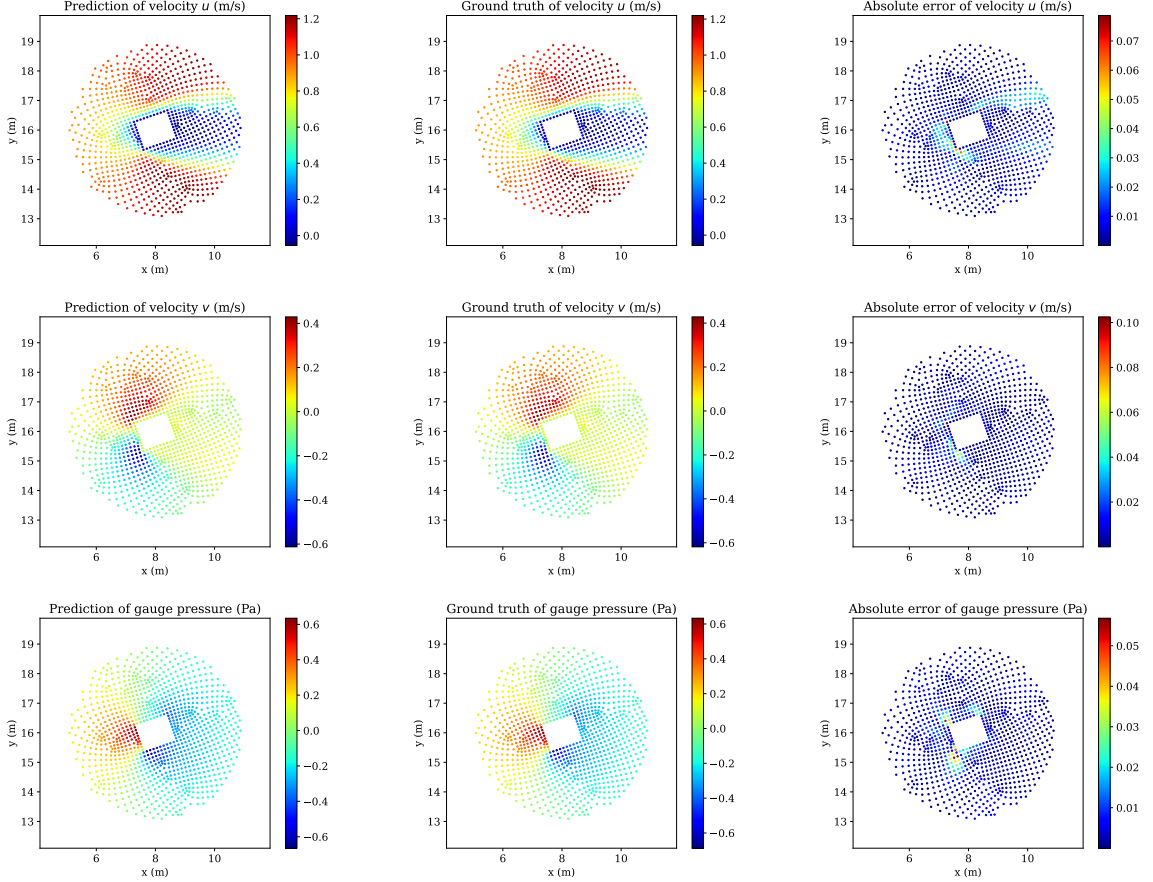


Figure 12: The seventh set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

Accordingly, we observe an improvement in the prediction accuracy, albeit with a relatively high cost for RAM usage. Depending on the user's criteria and available computational resources, the decision to use layer normalization may vary.

4.2. Comparison of PointNet with shared KANs and PointNet with shared MLPs

4.2.1. A brief overview of PointNet with shared MLPs

Before performing a comparison between PointNet with shared KANs (i.e., KA-PointNet) and PointNet with shared MLPs, we briefly review the structure of the latter, shown in Fig. 17. The goal is to make the differences clear so that the performance comparison becomes meaningful. As can be seen in Fig. 17, similar to the notation used for shared KANs, we use (A_1, A_2) and (A_1, A_2, A_3) for shared MLPs with two and three layers, respectively. The connection between the input vector \mathbf{r} of size d_{input} and the output vector \mathbf{s} of size d_{output} for one layer in an MLP can be formulated as follows:

$$\mathbf{s}_{d_{\text{output}} \times 1} = \sigma \left(\mathbf{W}_{d_{\text{output}} \times d_{\text{input}}} \mathbf{r}_{d_{\text{input}} \times 1} + \mathbf{b}_{d_{\text{output}} \times 1} \right), \quad (28)$$

where \mathbf{W} is the weight matrix and \mathbf{b} is the bias vector, containing trainable parameters. The nonlinear activation function is shown by σ , which operates elementwise. The number of trainable parameters in this layer is equal to $d_{\text{input}} \times d_{\text{input}} + d_{\text{output}}$. One may observe the difference in the structure of MLPs and KANs by comparing Eq. (11) and Eq. (28). The concept of shared MLPs is explained as follows. Let us apply the same example as we discussed in Sect.

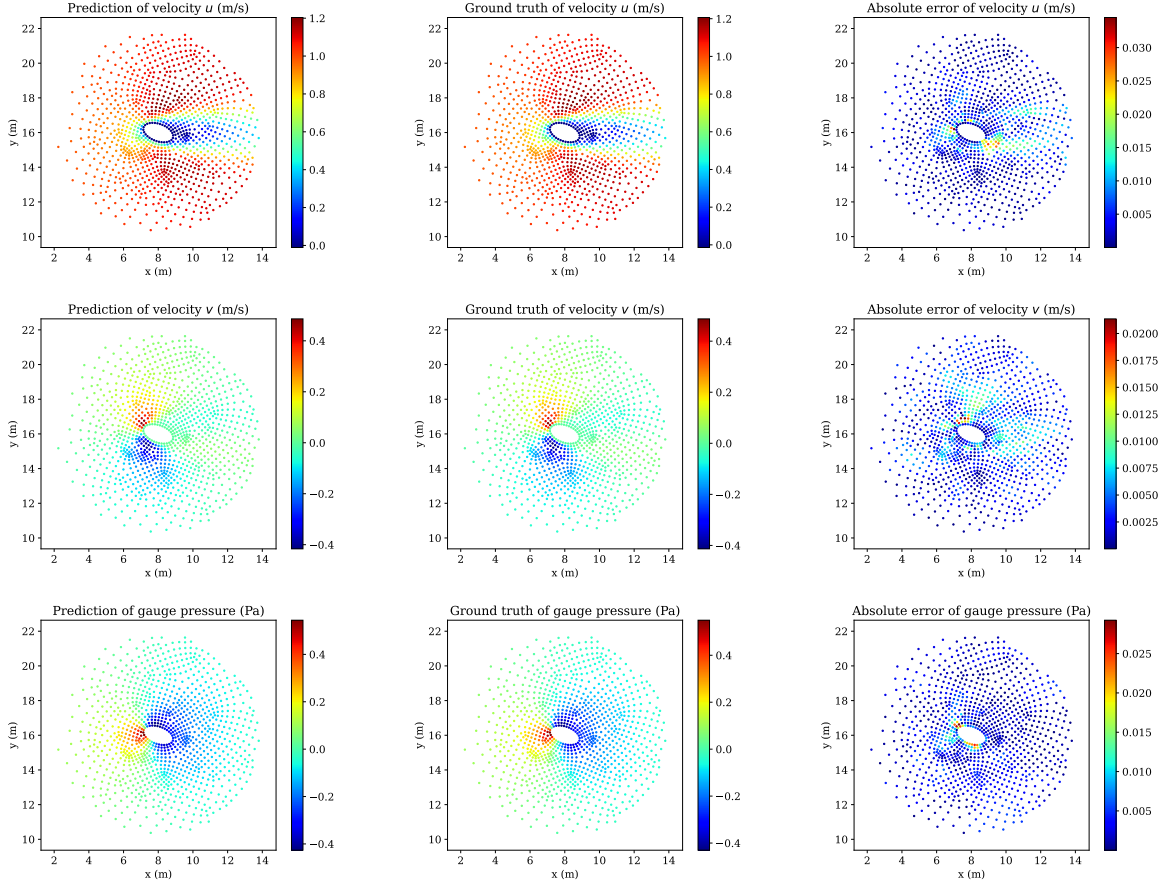


Figure 13: The eighth set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

3.3 for shared KANs. Taking $n_s = 1$ and focusing on the first layer of the first MLP in the first branch of PointNet (see Fig. 17), the resulting output vectors are:

$$\begin{aligned}
 \mathbf{s}_{64 \times 1}^{(1)} &= \sigma \left(\mathbf{W}_{64 \times 2} \begin{bmatrix} x'_1 \\ y'_1 \end{bmatrix} + \mathbf{b}_{64 \times 1} \right), \\
 \mathbf{s}_{64 \times 1}^{(2)} &= \sigma \left(\mathbf{W}_{64 \times 2} \begin{bmatrix} x'_2 \\ y'_2 \end{bmatrix} + \mathbf{b}_{64 \times 1} \right), \\
 &\vdots \\
 \mathbf{s}_{64 \times 1}^{(N)} &= \sigma \left(\mathbf{W}_{64 \times 2} \begin{bmatrix} x'_N \\ y'_N \end{bmatrix} + \mathbf{b}_{64 \times 1} \right),
 \end{aligned} \tag{29}$$

where \mathbf{W} and \mathbf{b} are respectively the shared weight matrix and shared bias vector applied to all $[x_j \ y_j]^{\text{tr}}$ for $1 \leq j \leq N$. One may compare Eqs. (26) and Eqs. (29). Further details about the implementation of shared MLP can be found in Ref. [111]. In the case of using PointNet with shared MLPs, our machine learning experiments demonstrate that the

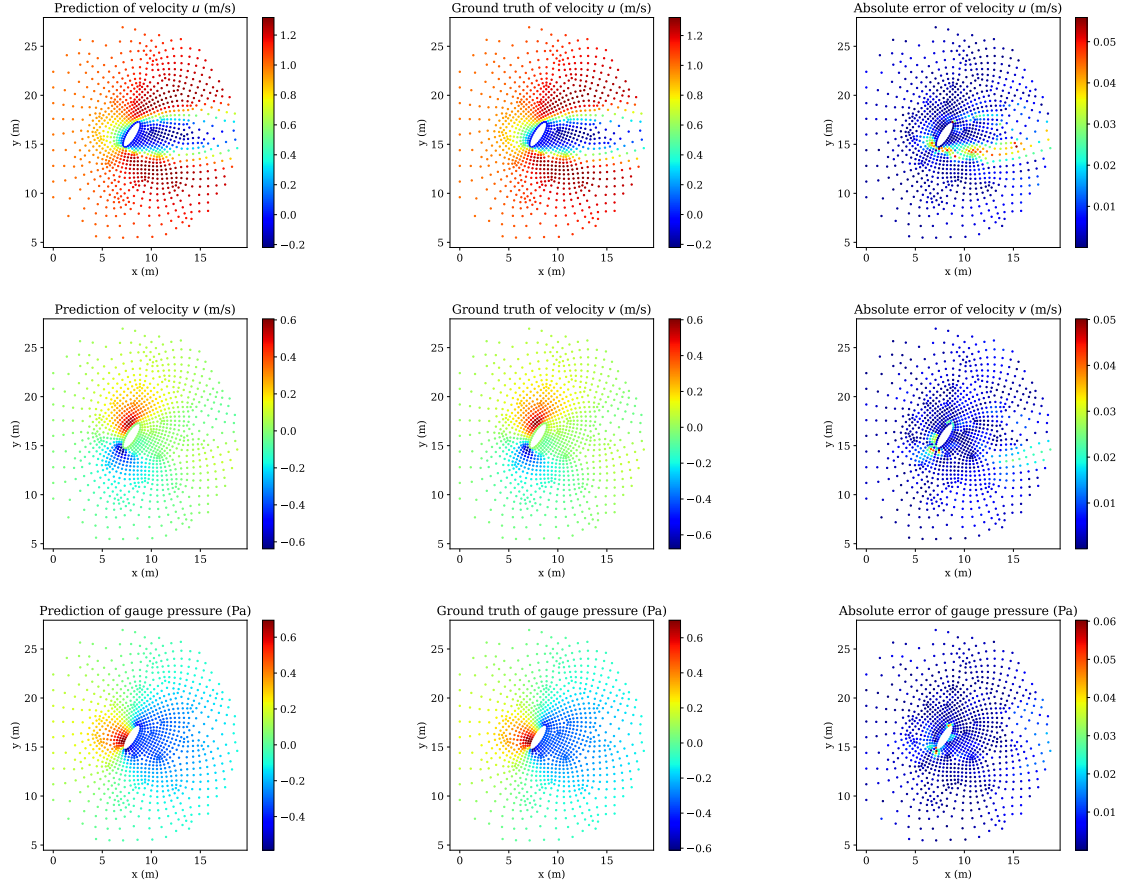


Figure 14: The ninth set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

accuracy of predictions is higher if we normalize the output of the network (i.e., u^* , v^* , and p^*) between $[0, 1]$, instead of $[-1, 1]$, using the following scaling formula:

$$\{\phi'\} = \frac{\{\phi\} - \min(\{\phi\})}{\max(\{\phi\}) - \min(\{\phi\})}. \quad (30)$$

Note that we still scale the spatial coordinates (i.e., x' and y') to the range of $[-1, 1]$ using Eq. (10). The activation function of the Rectified Linear Unit (ReLU), defined as follows:

$$\sigma(\gamma) = \max(0, \gamma), \quad (31)$$

is implemented after all the layers except the last layer. Similar to KA-PointNet, each layer is followed by a batch normalization [108], except the last layer. The activation function in the last layer is the sigmoid function defined as follows:

$$\sigma(\gamma) = \frac{1}{1 + e^{-\gamma}}, \quad (32)$$

which covers the range of $[0, 1]$. The loss function, batch size, learning rate, and the optimizer model with its associated hyperparameters are all set the same as those for KA-PointNet (see Sect. 3.4).

Kolmogorov-Arnold PointNet

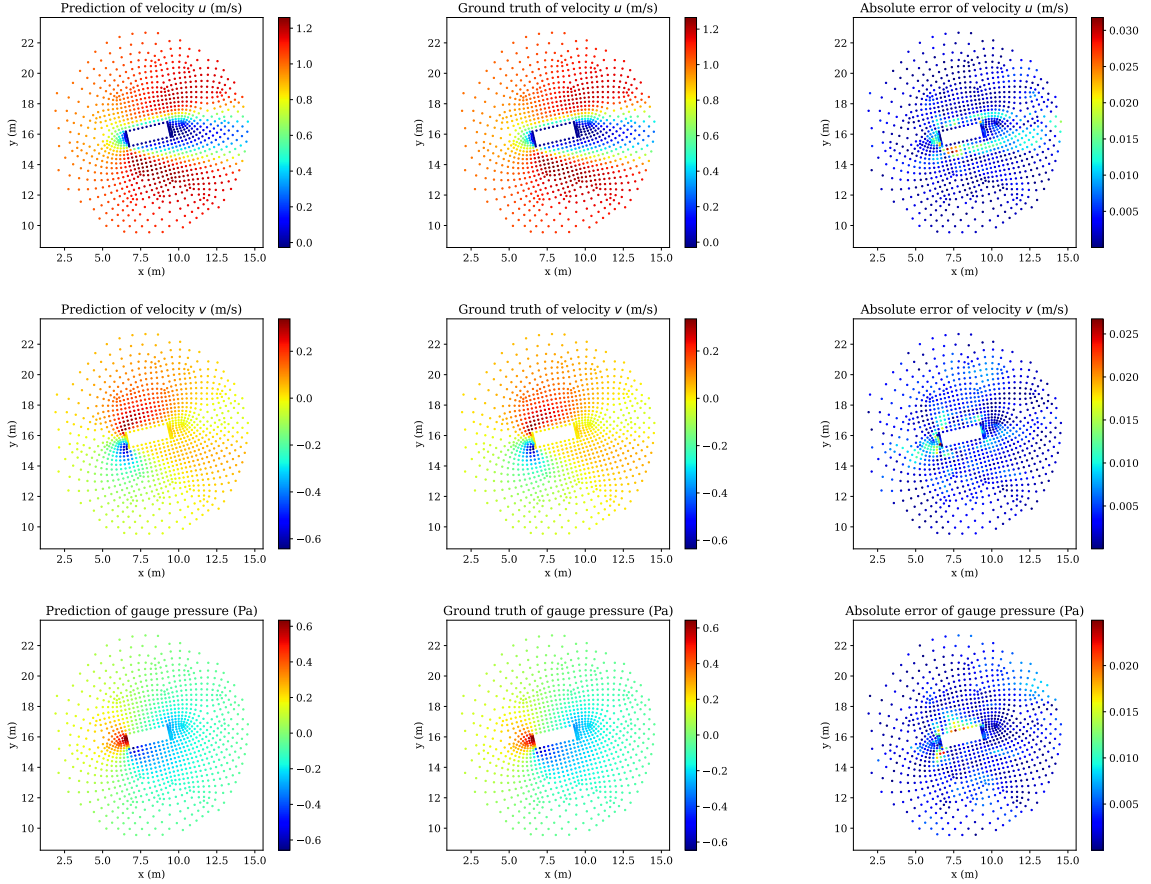


Figure 15: The tenth set of examples comparing the ground truth to the predictions of Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields from the test set. The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

4.2.2. Performance analysis

To ensure a fair comparison between PointNet with shared KANs (KA-PointNet) and PointNet with shared MLPs, we select hyperparameters for each network to achieve an approximately equal number of trainable parameters. Table 5 summarizes the results of this comparison, examining training time and error analysis. We conduct the comparison using three different setups. For PointNet with shared KANs, we set the degree of the Jacobi polynomial to 3 with $\alpha = \beta = 1$. Next, we choose n_s for each of these two networks. For instance, the second and third columns of Table 5 present the results for PointNet with shared KANs with $n_s = 0.5$ and 888128 trainable parameters, and PointNet with shared MLPs with $n_s = 1$ and 892355 trainable parameters, respectively. A similar scenario applies to the pairs in the fourth and fifth columns, as well as the pairs in the sixth and seventh columns of Table 5. In all three comparisons, we observe that the average relative pointwise error (L^2 norm) of the predicted velocity and pressure fields, over the test set, obtained by PointNet with shared KANs (i.e., KA-PointNet) is lower than those predicted by PointNet with shared MLPs. The pattern of errors is similar, with the v variable experiencing the highest level of errors and the u variable experiencing the lowest. This superior performance of PointNet with shared KANs (i.e., KA-PointNet) is particularly significant for the prediction of the v variable. For instance, the average relative pointwise error (L^2 norm) over 222 unseen geometries in the test set for PointNet with shared KANs with $n_s = 1$ (see the fourth column of Table 5) is approximately 5.76%, compared to 13.9% for PointNet with shared MLPs with $n_s = 2$ (see the fifth column of Table 5). A notable difference in performance is also observed for the p variable with this setup (4.76% vs. 10.6%). Interestingly, increasing the size of the PointNet with shared MLPs (i.e., choosing greater n_s) does not necessarily improve performance, while PointNet with shared KANs shows improvement. For example, comparing

Kolmogorov-Arnold PointNet

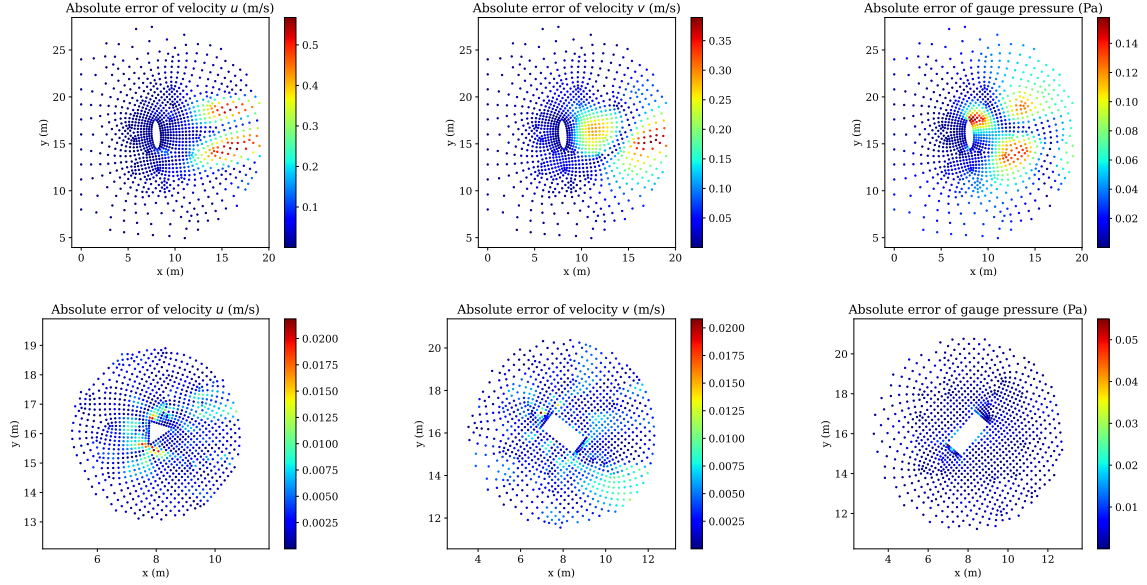


Figure 16: Distribution of absolute pointwise error for the prediction of the velocity and pressure fields by Kolmogorov-Arnold PointNet (i.e., KA-PointNet) for the velocity and pressure fields for the geometries from the test set when the relative pointwise error (L^2 norm) becomes maximum (first row) and minimum (second row). The Jacobi polynomial used has a degree of 5, with $\alpha = \beta = 1$. Here, $n_s = 1$ is set.

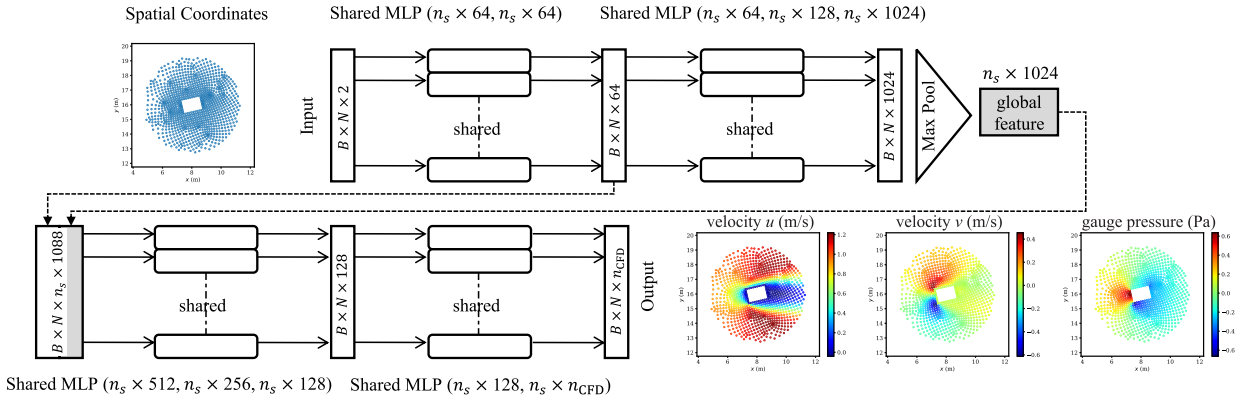


Figure 17: Architecture of PointNet with shared Multilayer Perceptron (MLP). Shared MLPs with the labels (A_1, A_2) and (A_1, A_2, A_3) are explained in the text. n_{CFD} denotes the number of CFD variables. N is the number of points in the point clouds. B represents the batch size. n_s is the global scaling parameter used to control the network size.

the error of the v variable listed in the fifth column ($n_s = 2$) and the seventh column ($n_s = 4$), the relative error of the v variable increases from 13.9% to 14.8%. In contrast, a similar comparison for PointNet with shared KANs shows that the relative error of the v variable decreases from 5.75% to 4.48% by increasing n_s from 1 to 2. Additionally, the average training time per epoch of KA-PointNet with shared KANs is approximately 6.7 times greater than that of PointNet with shared MLPs, as can be computed from the given training time in Table 5. The main reason for this is the recursive implementation of the Jacobi polynomial in PointNet with shared KANs. An improvement in this implementation from a software engineering perspective could potentially reduce the computational cost of shared KANs, making them more efficient compared to shared MLPs.

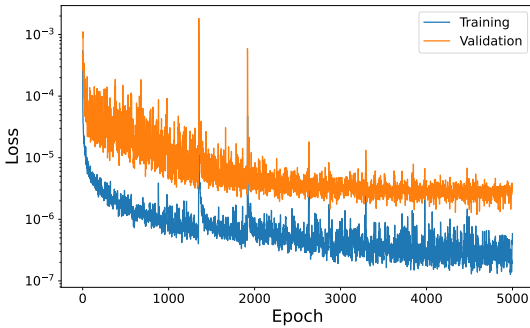
For a more in-depth comparison, let us focus on the last two columns of Table 5, where we have PointNet with shared KANs (using a Jacobi polynomial of degree 3, $\alpha = \beta = 1$, and $n_s = 2$), with PointNet with shared MLPs

Table 5

Comparison between the performance of PointNet with shared Kolmogorov-Arnold Networks (KANs), i.e. KA-PointNet, and PointNet with shared Multilayer Perceptrons (MLPs) for prediction of the velocity and pressure fields for the test set containing 222 unseen geometries. In the Jacobi polynomial of KA-PointNet, we set $\alpha = \beta = 1$. $\|\dots\|$ indicates the L^2 norm.

	KANs	MLPs	KANs	MLPs	KANs	MLPs
n_s	0.5	1	1	2	2	4
Polynomial degree	3	-	3	-	3	-
Number of trainable parameters	888128	892355	3545728	3554179	14169344	14186243
Training time per epoch (s)	2.83741	0.42597	6.52910	0.97970	16.69120	2.47669
Average $\ \tilde{u} - u\ /\ u\ $	2.63597E-2	5.01625E-2	1.73537E-2	3.83902E-2	1.69209E-2	3.80471E-2
Maximum $\ \tilde{u} - u\ /\ u\ $	1.30750E-1	1.57523E-1	1.40088E-1	1.51597E-1	1.50143E-1	1.47946E-1
Minimum $\ \tilde{u} - u\ /\ u\ $	1.30994E-2	2.03904E-2	7.22247E-3	1.24053E-2	1.08372E-2	1.40001E-2
Average $\ \tilde{v} - v\ /\ v\ $	1.00292E-1	1.61943E-1	5.75906E-2	1.39018E-1	4.48352E-2	1.47633E-1
Maximum $\ \tilde{v} - v\ /\ v\ $	4.19706E-1	5.07943E-1	4.42245E-1	4.85987E-1	4.85161E-1	4.87821E-1
Minimum $\ \tilde{v} - v\ /\ v\ $	5.58101E-2	6.56595E-2	2.05299E-2	5.60436E-2	2.49093E-2	5.42824E-2
Average $\ \tilde{p} - p\ /\ p\ $	9.86696E-2	1.54159E-1	4.75586E-2	1.05939E-1	4.25412E-2	7.22403E-2
Maximum $\ \tilde{p} - p\ /\ p\ $	2.88919E-1	4.03240E-1	1.55694E-1	2.36421E-1	1.63125E-1	1.70085E-1
Minimum $\ \tilde{p} - p\ /\ p\ $	4.67306E-2	4.04501E-2	1.75673E-2	3.77024E-2	1.74415E-2	2.84674E-2

PointNet with shared KANs



PointNet with shared MLPs

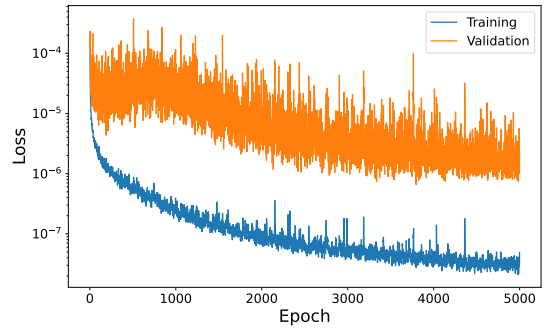


Figure 18: Comparison of the loss evolution for the training and validation sets between PointNet with shared KANs, i.e., KA-PointNet, ($n_s = 2$, Jacobi polynomial of degree 3, $\alpha = \beta = 1$, and 14169344 trainable parameters) and PointNet with shared MLPs ($n_s = 4$ and 14186243 trainable parameters)

($n_s = 4$). The loss evolution for the training set (1772 data) and the validation set (241 data) of these two networks is illustrated in Fig. 18. For a few epochs, we observe a tendency for divergence in KA-PointNet; however, it eventually stabilizes during the training process, a tendency not present in PointNet with shared MLPs. Furthermore, PointNet with shared MLPs exhibits a higher degree of bias (i.e., a greater difference between the training loss and the validation loss) compared to PointNet with shared KANs. This suggests a tendency for PointNet with shared MLPs to overfit the training data, as evidenced by tracking the training loss values in Fig. 18. Considering this case for comparison, the distribution and frequency of relative pointwise error (L^2 norm) of the velocity and pressure fields predicted by PointNet with shared KANs (i.e., KA-PointNet) and PointNet with shared MLPs are plotted in Fig. 19. We observe that the frequency of predictions with lower errors is significantly higher for PointNet with shared KANs compared to PointNet with shared MLPs. In both architectures, there are a few predictions that appear as outliers, experiencing

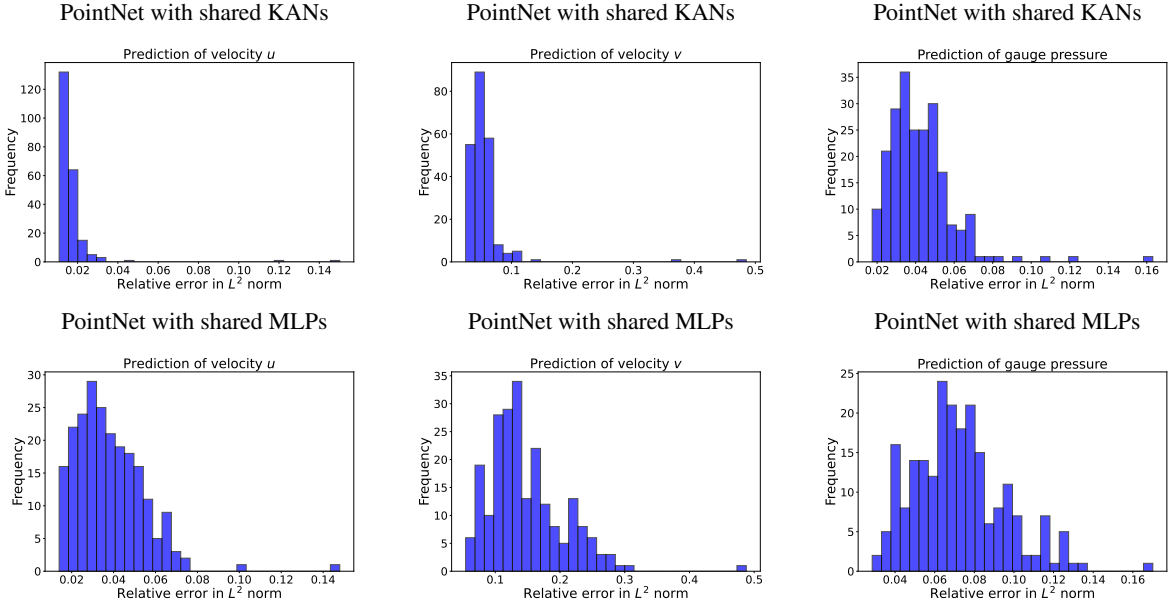


Figure 19: Comparison of histograms of the relative error in L^2 norm for the velocity and pressure fields predicted by PointNet with shared Kolmogorov-Arnold Networks (KANs), i.e., KA-PointNet, ($n_s = 2$, Jacobi polynomial of degree 3, $\alpha = \beta = 1$, and 14169344 trainable parameters) and PointNet with shared Multilayer Perceptrons (MLPs) ($n_s = 4$ and 14186243 trainable parameters).

extremely high errors. Overall, PointNet with shared KANs (i.e., KA-PointNet) provides more accurate results with higher computational expenses for training compared to PointNet with shared MLPs, despite having approximately the same number of trainable parameters.

5. Summary and future studies

In this work, we presented Kolmogorov-Arnold PointNet (KA-PointNet), a novel supervised deep learning framework for predicting fluid flow fields in irregular geometries. Instead of using regular shared Multilayer Perceptrons (MLPs), this framework incorporated shared Kolmogorov-Arnold Networks (KANs) within the segmentation branch of PointNet [71]. Jacobi polynomials were utilized to construct the shared KANs, and batch normalization [108] was applied after each shared KAN layer. A two-dimensional steady-state incompressible flow over a cylinder, with varying cross-sectional geometries, was considered as a test case. The Adam optimizer [91] was efficiently used to train KA-PointNet on the training set. Quantitative and visual error analyses were conducted to examine the capability of KA-PointNet. Specifically, the effects of the degree of Jacobi polynomials, specific types of Jacobi polynomials (e.g., Legendre polynomials, Chebyshev polynomials of the first and second kinds, and Gegenbauer polynomials), and the global size of KA-PointNet on the accuracy of the predicted velocity and pressure fields of the test set were assessed. According to one of our machine learning experiments, the average relative pointwise error (L^2 norm) of the x component of the velocity vector, the y component of the velocity vector, and the pressure field of the test set predicted by KA-PointNet were approximately 1.2%, 4.9%, and 3.4%, respectively, demonstrating the strong performance of KA-PointNet. Jacobi polynomials of degree 5 and degree 6 generally led to more accurate results but required higher computational training time compared to lower-order Jacobi polynomials. Among the specific cases of Jacobi polynomials, KA-PointNet with Chebyshev polynomials of the first and second kinds showed the lowest error in predicting the velocity and pressure fields of the test set. The single global scaling parameter defined in this study to control the size of KA-PointNet was shown to be an effective parameter for fine-tuning KA-PointNet, despite its complex architecture. A comparison between KA-PointNet with batch normalization and layer normalization demonstrated that, although the prediction accuracy with layer normalization is higher, it requires a significant increase in RAM, which may or may not be available depending on the user's resources. A comprehensive comparison was

conducted between PointNet with shared KANs (i.e., KA-PointNet) and PointNet with shared MLPs. PointNet with shared KANs (i.e., KA-PointNet) provided more accurate results than PointNet with shared MLPs, despite having approximately the same number of trainable parameters. However, this increased accuracy came at the expense of higher computational costs for training.

Because KANs are a relatively new deep learning paradigm, there are still many unexplored areas that can be investigated. A few possible directions include extending KA-PointNet to three-dimensional computational mechanics problems, investigating the capacity of KA-PointNet in solid mechanics, such as the prediction of stress and displacement fields in nonlinear elasticity and plasticity problems, and developing physics-informed KA-PointNet for solving inverse problems in scientific computing. Another direction is combining KANs with advanced versions of PointNet [71], such as PointNet++ [72] and KPConv [79], to improve the accuracy of predictions. Integrating KANs into Fourier neural operators [112, 113, 114, 115, 116] and large language models [117, 118, 119, 120, 121] could be potentially beneficial to the fields of computational physics and computational mathematics.

CRedit authorship contribution statement

Ali Kashefi: Conceptualization, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The author will make the code and data publicly available upon the acceptance of the manuscript.

References

- [1] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, M. Tegmark, Kan: Kolmogorov-arnold networks, arXiv preprint arXiv:2404.19756 (2024).
- [2] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Mathematics of control, signals and systems* 2 (1989) 303–314.
- [3] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural networks* 2 (1989) 359–366.
- [4] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016. <http://www.deeplearningbook.org>.
- [5] V. I. Arnold, On the representation of functions of several variables as a superposition of functions of a smaller number of variables, *Collected works: Representations of functions, celestial mechanics and KAM theory, 1957–1965* (2009) 25–46.
- [6] V. I. Arnold, On functions of three variables, *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965* (2009) 5–8.
- [7] A. K. Kolmogorov, On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition, *Doklady Akademii Nauk SSSR* 114 (1957) 369–373.
- [8] R. Hecht-Nielsen, Kolmogorov’s mapping neural network existence theorem, in: *Proceedings of the international conference on Neural Networks*, volume 3, IEEE press New York, NY, USA, 1987, pp. 11–14.
- [9] F. Girosi, T. Poggio, Representation properties of networks: Kolmogorov’s theorem is irrelevant, *Neural Computation* 1 (1989) 465–469.
- [10] J. Braun, M. Griebel, On a constructive proof of kolmogorov’s superposition theorem, *Constructive approximation* 30 (2009) 653–675.
- [11] A. Ismayilova, V. E. Ismailov, On the kolmogorov neural networks, *Neural Networks* 176 (2024) 106333.
- [12] N. Borri, D. Chetverikov, Y. Liu, A. Tsyvinski, One Factor to Bind the Cross-Section of Returns, Technical Report, National Bureau of Economic Research, 2024.
- [13] Y. Wang, J. Sun, J. Bai, C. Anitescu, M. S. Eshaghi, X. Zhuang, T. Rabczuk, Y. Liu, Kolmogorov arnold informed neural network: A physics-informed deep learning framework for solving pdes based on kolmogorov arnold networks, arXiv preprint arXiv:2406.11045 (2024).
- [14] K. Shukla, J. D. Toscano, Z. Wang, Z. Zou, G. E. Karniadakis, A comprehensive and fair comparison between mlp and kan representations for differential equations and operator networks, arXiv preprint arXiv:2406.02917 (2024).
- [15] A. A. Howard, B. Jacob, S. H. Murphy, A. Heinlein, P. Stinis, Finite basis kolmogorov-arnold networks: domain decomposition for data-driven and physics-informed problems, arXiv preprint arXiv:2406.19662 (2024).
- [16] J. D. Toscano, T. Käufer, M. Maxey, C. Cierpka, G. E. Karniadakis, Inferring turbulent velocity and temperature fields and their statistics from lagrangian velocity measurements using physics-informed kolmogorov-arnold networks, arXiv preprint arXiv:2407.15727 (2024).
- [17] Y. Wang, J. Sun, J. Bai, C. Anitescu, M. S. Eshaghi, X. Zhuang, T. Rabczuk, Y. Liu, Kolmogorov arnold informed neural network: A physics-informed deep learning framework for solving pdes based on kolmogorov arnold networks, arXiv preprint arXiv:2406.11045 (2024).
- [18] S. Rigas, M. Papachristou, T. Papadopoulos, F. Anagnostopoulos, G. Alexandridis, Adaptive training of grid-dependent physics-informed kolmogorov-arnold networks, arXiv preprint arXiv:2407.17611 (2024).

- [19] S. Patra, S. Panda, B. K. Parida, M. Arya, K. Jacobs, D. I. Bondar, A. Sen, Physics informed kolmogorov-arnold neural networks for dynamical analysis via efficient-kan and wav-kan, arXiv preprint arXiv:2407.18373 (2024).
- [20] D. W. Abueidda, P. Pantidis, M. E. Mobasher, Deepokan: Deep operator network based on kolmogorov arnold networks for mechanics problems, arXiv preprint arXiv:2405.19143 (2024).
- [21] B. C. Koenig, S. Kim, S. Deng, Kan-odes: Kolmogorov-arnold network ordinary differential equations for learning dynamical systems and hidden physics, arXiv preprint arXiv:2407.04192 (2024).
- [22] B. Azam, N. Akhtar, Suitability of kans for computer vision: A preliminary investigation, arXiv preprint arXiv:2406.09087 (2024).
- [23] M. Cheon, Kolmogorov-arnold network for satellite image classification in remote sensing, arXiv preprint arXiv:2406.00600 (2024).
- [24] S. T. Seydi, Exploring the potential of polynomial basis functions in kolmogorov-arnold networks: A comparative study of different groups of polynomials, arXiv preprint arXiv:2406.02583 (2024).
- [25] S. T. Seydi, Unveiling the power of wavelets: A wavelet-based kolmogorov-arnold network for hyperspectral image classification, arXiv preprint arXiv:2406.07869 (2024).
- [26] H.-T. Ta, Bsrbf-kan: A combination of b-splines and radial basic functions in kolmogorov-arnold networks, arXiv preprint arXiv:2406.11173 (2024).
- [27] M. Cheon, Demonstrating the efficacy of kolmogorov-arnold networks in vision tasks, arXiv preprint arXiv:2406.14916 (2024).
- [28] V. Lobanov, N. Firsov, E. Myasnikov, R. Khabibullin, A. Nikonov, Hyperkan: Kolmogorov-arnold networks make hyperspectral image classifiers smarter, arXiv preprint arXiv:2407.05278 (2024).
- [29] R. Yu, W. Yu, X. Wang, Kan or mlp: A fairer comparison, arXiv preprint arXiv:2407.16674 (2024).
- [30] V. D. Tran, T. X. H. Le, T. D. Tran, H. L. Pham, V. T. D. Le, T. H. Vu, V. T. Nguyen, Y. Nakashima, Exploring the limitations of kolmogorov-arnold networks in classification: Insights to software training and hardware implementation, 2024. arXiv:2407.17790.
- [31] M. G. Altarabichi, Rethinking the function of neurons in kans, arXiv preprint arXiv:2407.20667 (2024).
- [32] C. Li, X. Liu, W. Li, C. Wang, H. Liu, Y. Yuan, U-kan makes strong backbone for medical image segmentation and generation, arXiv preprint arXiv:2406.02918 (2024).
- [33] T. Tang, Y. Chen, H. Shu, 3d u-kan implementation for multi-modal mri brain tumor segmentation, arXiv preprint arXiv:2408.00273 (2024).
- [34] Y. Wang, X. Yu, Y. Gao, J. Sha, J. Wang, L. Gao, Y. Zhang, X. Rong, Spectralkan: Kolmogorov-arnold network for hyperspectral images change detection, arXiv preprint arXiv:2407.00949 (2024).
- [35] Z. Bozorgasl, H. Chen, Wav-kan: Wavelet kolmogorov-arnold networks, arXiv preprint arXiv:2405.12832 (2024).
- [36] C. J. Vaca-Rubio, L. Blanco, R. Pereira, M. Caus, Kolmogorov-arnold networks (kans) for time series analysis, arXiv preprint arXiv:2405.08790 (2024).
- [37] R. Genet, H. Inzirillo, Tkan: Temporal kolmogorov-arnold networks, arXiv preprint arXiv:2405.07344 (2024).
- [38] M. E. Samadi, Y. Müller, A. Schuppert, Smooth kolmogorov arnold networks enabling structural knowledge representation, arXiv preprint arXiv:2405.11318 (2024).
- [39] M. Liu, S. Bian, B. Zhou, P. Lukowicz, ikan: Global incremental learning with kan for human activity recognition across heterogeneous datasets, arXiv preprint arXiv:2406.01646 (2024).
- [40] Z. Li, Kolmogorov-arnold networks are radial basis function networks, arXiv preprint arXiv:2405.06721 (2024).
- [41] A. A. Aghaei, fkan: Fractional kolmogorov-arnold networks with trainable jacobi basis functions, arXiv preprint arXiv:2406.07456 (2024).
- [42] K. Xu, L. Chen, S. Wang, Kolmogorov-arnold networks for time series: Bridging predictive power and interpretability, arXiv preprint arXiv:2406.02496 (2024).
- [43] J. Xu, Z. Chen, J. Li, S. Yang, W. Wang, X. Hu, E. C.-H. Ngai, Fourierkan-gcf: Fourier kolmogorov-arnold network—an effective and efficient feature transformation for graph collaborative filtering, arXiv preprint arXiv:2406.01034 (2024).
- [44] Y. Peng, M. He, F. Hu, Z. Mao, X. Huang, J. Ding, Predictive modeling of flexible ehd pumps using kolmogorov-arnold networks, arXiv preprint arXiv:2405.07488 (2024).
- [45] R. Genet, H. Inzirillo, A temporal kolmogorov-arnold transformer for time series forecasting, arXiv preprint arXiv:2406.02486 (2024).
- [46] G. Nehma, M. Tiwari, Leveraging kans for enhanced deep koopman operator discovery, arXiv preprint arXiv:2406.02875 (2024).
- [47] L. F. Herbozo Contreras, J. Cui, L. Yu, Z. Huang, A. Nikpour, O. Kavehei, Kan-eeg: Towards replacing backbone-mlp for an effective seizure detection system, medRxiv (2024) 2024–06.
- [48] M. Liu, D. Geißler, D. Nshimiyimana, S. Bian, B. Zhou, P. Lukowicz, Initial investigation of kolmogorov-arnold networks (kans) as feature extractors for imu based human activity recognition, arXiv preprint arXiv:2406.11914 (2024).
- [49] E. Poeta, F. Giobergia, E. Pastor, T. Cerquitelli, E. Baralis, A benchmarking study of kolmogorov-arnold networks on tabular data, arXiv preprint arXiv:2406.14529 (2024).
- [50] A. Kundu, A. Sarkar, A. Sadhu, Kanqas: Kolmogorov arnold network for quantum architecture search, arXiv preprint arXiv:2406.17630 (2024).
- [51] A. A. Aghaei, rkan: Rational kolmogorov-arnold networks, 2024. URL: <https://arxiv.org/abs/2406.14495>. arXiv:2406.14495.
- [52] X. Li, Z. Feng, Y. Chen, W. Dai, Z. He, Y. Zhou, S. Jiao, Coeff-kans: A paradigm to address the electrolyte field with kans, arXiv preprint arXiv:2407.20265 (2024).
- [53] P. Pratyush, C. Carrier, S. Pokharel, H. D. Ismail, M. Chaudhari, D. B. KC, Calmposkan: Prediction of general phosphorylation sites in proteins via fusion of codon aware embeddings with amino acid aware embeddings and wavelet-based kolmogorov arnold network, bioRxiv (2024) 2024–07.
- [54] H. Liu, J. Lei, Z. Ren, From complexity to clarity: Kolmogorov-arnold networks in nuclear binding energy prediction, arXiv preprint arXiv:2407.20737 (2024).
- [55] A. D. Bodner, A. S. Tepsich, J. N. Spolski, S. Pourteau, Convolutional kolmogorov-arnold networks, arXiv preprint arXiv:2406.13155 (2024).
- [56] M. Kiamari, M. Kiamari, B. Krishnamachari, Gkan: Graph kolmogorov-arnold networks, arXiv preprint arXiv:2406.06470 (2024).

- [57] R. Bresson, G. Nikolentzos, G. Panagopoulos, M. Chatzianastasis, J. Pang, M. Vazirgiannis, Kagnns: Kolmogorov-arnold networks meet graph learning, arXiv preprint arXiv:2406.18380 (2024).
- [58] F. Zhang, X. Zhang, Graphkan: Enhancing feature extraction with graph kolmogorov arnold networks, arXiv preprint arXiv:2406.13597 (2024).
- [59] G. De Carlo, A. Mastropietro, A. Anagnostopoulos, Kolmogorov-arnold graph neural networks, arXiv preprint arXiv:2406.18354 (2024).
- [60] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, Springer, 2015, pp. 234–241.
- [61] X. Guo, W. Li, F. Iorio, Convolutional neural networks for steady flow approximation, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 481–490.
- [62] H. Gao, L. Sun, J.-X. Wang, Phygeonet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state pdes on irregular domain, *Journal of Computational Physics* 428 (2021) 110079.
- [63] J. Tompson, K. Schlachter, P. Sprechmann, K. Perlin, Accelerating eulerian fluid simulation with convolutional networks, in: International conference on machine learning, PMLR, 2017, pp. 3424–3433.
- [64] N. Thuerey, K. Weibenow, L. Prantl, X. Hu, Deep learning methods for reynolds-averaged navier–stokes simulations of airfoil flows, *AIAA Journal* 58 (2020) 25–36.
- [65] S. Bhatnagar, Y. Afshar, S. Pan, K. Duraisamy, S. Kaushik, Prediction of aerodynamic flow fields using convolutional neural networks, *Computational Mechanics* 64 (2019) 525–545.
- [66] K. Xu, W. Hu, J. Leskovec, S. Jegelka, How powerful are graph neural networks?, arXiv preprint arXiv:1810.00826 (2018).
- [67] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, P. W. Battaglia, Learning mesh-based simulation with graph networks, arXiv preprint arXiv:2010.03409 (2020).
- [68] F. D. A. Belbute-Peres, T. Economou, Z. Kolter, Combining differentiable pde solvers and graph neural networks for fluid flow prediction, in: international conference on machine learning, PMLR, 2020, pp. 2402–2411.
- [69] M. Maurizi, C. Gao, F. Berto, Predicting stress, strain and deformation fields in materials and structures with graph neural networks, *Scientific reports* 12 (2022) 21834.
- [70] H. Gao, M. J. Zahr, J.-X. Wang, Physics-informed graph neural galerkin networks: A unified framework for solving pde-governed forward and inverse problems, *Computer Methods in Applied Mechanics and Engineering* 390 (2022) 114502.
- [71] C. R. Qi, H. Su, K. Mo, L. J. Guibas, Pointnet: Deep learning on point sets for 3d classification and segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 652–660.
- [72] C. R. Qi, L. Yi, H. Su, L. J. Guibas, Pointnet++: Deep hierarchical feature learning on point sets in a metric space, *Advances in neural information processing systems* 30 (2017).
- [73] A. Kashefi, D. Rempe, L. J. Guibas, A point-cloud deep learning framework for prediction of fluid flow fields on irregular geometries, *Physics of Fluids* 33 (2021).
- [74] A. Kashefi, T. Mukerji, Physics-informed pointnet: A deep learning solver for steady-state incompressible flows and thermal fields on multiple sets of irregular geometries, *Journal of Computational Physics* 468 (2022) 111510.
- [75] A. Kashefi, L. J. Guibas, T. Mukerji, Physics-informed pointnet: On how many irregular geometries can it solve an inverse problem simultaneously? application to linear elasticity, *Journal of Machine Learning for Modeling and Computing* 4 (2023).
- [76] Y. Shen, W. Huang, Z.-g. Wang, D.-f. Xu, C.-Y. Liu, A deep learning framework for aerodynamic pressure prediction on general three-dimensional configurations, *Physics of Fluids* 35 (2023) 107111.
- [77] J. He, S. Koric, D. Abueidda, A. Najafi, I. Jasiuk, Geom-deeponet: A point-cloud-based deep operator network for field predictions on 3d parameterized geometries, *Computer Methods in Applied Mechanics and Engineering* 429 (2024) 117130.
- [78] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: A review of methods and applications, *AI Open* 1 (2020) 57–81.
- [79] H. Thomas, C. R. Qi, J.-E. Deschaut, B. Marcotegui, F. Goulette, L. J. Guibas, Kpconv: Flexible and deformable convolution for point clouds, in: Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 6411–6420.
- [80] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, M. Bennamoun, Deep learning for 3d point clouds: A survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43 (2021) 4338–4364.
- [81] C. R. Qi, W. Liu, C. Wu, H. Su, L. J. Guibas, Frustum pointnets for 3d object detection from rgb-d data, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [82] X. Zeng, A. Vahdat, F. Williams, Z. Gajic, O. Litany, S. Fidler, K. Kreis, Lion: Latent point diffusion models for 3d shape generation, in: Advances in Neural Information Processing Systems (NeurIPS), 2022.
- [83] H. Zhao, L. Jiang, J. Jia, P. H. Torr, V. Koltun, Point transformer, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 16259–16268.
- [84] X. Huang, T. Xie, Z. Wang, L. Chen, Q. Zhou, Z. Hu, A Transfer Learning-Based Multi-Fidelity Point-Cloud Neural Network Approach for Melt Pool Modeling in Additive Manufacturing, *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg* 8 (2021) 011104.
- [85] R. Zhi, Z. Li, F. Wen, L. Su, S. Wang, Research on performance predictions using single-hole film cooling based on PointNet, *Physics of Fluids* 35 (2023) 025108.
- [86] A. Kashefi, T. Mukerji, Point-cloud deep learning of porous media for permeability prediction, *Physics of Fluids* 33 (2021) 097109.
- [87] A. Kashefi, T. Mukerji, Prediction of fluid flow in porous media by sparse observations and physics-informed pointnet, *Neural Networks* 167 (2023) 80–91.
- [88] M. Raissi, P. Perdikaris, G. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational Physics* 378 (2019) 686–707.

- [89] J. Zhang, Rpn: Reconciled polynomial network towards unifying pgms, kernel svms, mlp and kan, 2024. URL: <https://arxiv.org/abs/2407.04819>. arXiv:2407.04819.
- [90] S. SS, K. AR, G. R, A. KP, Chebyshev polynomial-based kolmogorov-arnold networks: An efficient architecture for nonlinear function approximation, 2024. URL: <https://arxiv.org/abs/2405.07200>. arXiv:2405.07200.
- [91] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).
- [92] H. Ding, C. Shu, K. Yeo, D. Xu, Simulation of incompressible viscous flows past a circular cylinder by hybrid fd scheme and meshless least square-based finite difference method, *Computer Methods in Applied Mechanics and Engineering* 193 (2004) 727–744.
- [93] M. Behr, D. Hastreiter, S. Mittal, T. Tezduyar, Incompressible flow past a circular cylinder: dependence of the computed flow field on the location of the lateral boundaries, *Computer Methods in Applied Mechanics and Engineering* 123 (1995) 309–316.
- [94] A. Kashefi, Coarse grid projection methodology: A partial mesh refinement tool for incompressible flow simulations, *Bulletin of the Iranian Mathematical Society* 46 (2019) 177–181.
- [95] A. Kashefi, A coarse-grid incremental pressure projection method for accelerating low reynolds number incompressible flow simulations, *Iran Journal of Computer Science* 3 (2019) 13–23.
- [96] A. Kashefi, A coarse-grid projection method for accelerating incompressible mhd flow simulations, *Engineering with Computers* 38 (2021) 1539–1553.
- [97] C. Geuzaine, J.-F. Remacle, Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities, *International journal for numerical methods in engineering* 79 (2009) 1309–1331.
- [98] H. G. Weller, G. Tabor, H. Jasak, C. Fureby, A tensorial approach to computational continuum mechanics using object-oriented techniques, *Computers in physics* 12 (1998) 620–631.
- [99] L. Caretto, A. Gosman, S. Patankar, D. Spalding, Two calculation procedures for steady, three-dimensional flows with recirculation, in: *Proceedings of the third international conference on numerical methods in fluid mechanics*, Springer, 1973, pp. 60–68.
- [100] A. Kashefi, A. E. Staples, A finite-element coarse-grid projection method for incompressible flow simulations, *Advances in Computational Mathematics* 44 (2017) 1063–1090.
- [101] A. Kashefi, A coarse grid projection method for accelerating free and forced convection heat transfer computations, *Results in Mathematics* 75 (2020).
- [102] S. Sen, S. Mittal, G. Biswas, Flow past a square cylinder at low reynolds numbers, *International Journal for Numerical Methods in Fluids* 67 (2011) 1160–1174.
- [103] A. Kumar De, A. Dalal, Numerical simulation of unconfined flow past a triangular cylinder, *International journal for numerical methods in fluids* 52 (2006) 801–821.
- [104] W. Zhong, L. Deng, Z. Xiao, Flow past a rectangular cylinder close to a free surface, *Ocean Engineering* 186 (2019) 106118.
- [105] R. Mittal, S. Balachandar, Direct numerical simulation of flow past elliptic cylinders, *Journal of Computational Physics* 124 (1996) 351–367.
- [106] Z. Abedin, N. A. Khan, M. M. Rizia, M. Q. Islam, Simulation of wind flow over square, pentagonal and hexagonal cylinders in a staggered form, in: *AIP Conference Proceedings*, 1, AIP Publishing LLC, 2017, p. 020004.
- [107] G. Szegő, Orthogonal polynomials, 1939. URL: <http://dx.doi.org/10.1090/coll/023>. doi:10.1090/coll/023.
- [108] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: *International conference on machine learning*, pmlr, 2015, pp. 448–456.
- [109] M. Abramowitz, *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., USA, 1974.
- [110] J. L. Ba, J. R. Kiros, G. E. Hinton, Layer normalization, arXiv preprint arXiv:1607.06450 (2016).
- [111] M. Lin, Q. Chen, S. Yan, Network in network, 2014. URL: <https://arxiv.org/abs/1312.4400>. arXiv:1312.4400.
- [112] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar, Fourier neural operator for parametric partial differential equations, arXiv preprint arXiv:2010.08895 (2020).
- [113] A. Anandkumar, K. Azizzadenesheli, K. Bhattacharya, N. Kovachki, Z. Li, B. Liu, A. Stuart, Neural operator: Graph kernel network for partial differential equations, in: *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*, 2020.
- [114] G. Wen, Z. Li, K. Azizzadenesheli, A. Anandkumar, S. M. Benson, U-fno—an enhanced fourier neural operator-based deep-learning model for multiphase flow, *Advances in Water Resources* 163 (2022) 104180.
- [115] B. Bonev, T. Kurth, C. Hundt, J. Pathak, M. Baust, K. Kashinath, A. Anandkumar, Spherical fourier neural operators: Learning stable dynamics on the sphere, in: *International conference on machine learning*, PMLR, 2023, pp. 2806–2823.
- [116] A. Kashefi, T. Mukerji, A novel fourier neural operator framework for classification of multi-sized images: Application to three dimensional digital porous media, *Physics of Fluids* 36 (2024).
- [117] A. Lewkowycz, A. Andreassen, D. Dohan, E. Dyer, H. Michalewski, V. Ramasesh, A. Slone, C. Anil, I. Schlag, T. Gutman-Solo, et al., Solving quantitative reasoning problems with language models, *Advances in Neural Information Processing Systems* 35 (2022) 3843–3857.
- [118] S. Imani, L. Du, H. Shrivastava, Mathprompter: Mathematical reasoning using large language models, arXiv preprint arXiv:2303.05398 (2023).
- [119] S. Frieder, L. Pinchetti, R.-R. Griffiths, T. Salvatori, T. Lukasiewicz, P. Petersen, J. Berner, Mathematical capabilities of chatgpt, *Advances in neural information processing systems* 36 (2024).
- [120] A. Kashefi, A misleading gallery of fluid motion by generative artificial intelligence, *Journal of Machine Learning for Modeling and Computing* 5 (2024) 113–144.
- [121] A. Kashefi, T. Mukerji, Chatgpt for programming numerical methods, *Journal of Machine Learning for Modeling and Computing* 4 (2023).