FluidZero: Mastering Diverse Tasks in Fluid Systems through a Single Generative Model

Haodong Feng

 Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University

Haoren Zheng

 School of Mathematics and Statistics, Nanjing University of Science and Technology

Peiyan Hu

 Academy of Mathematics and Systems Science, Chinese Academy of Sciences

Hongyuan Liu

 Department of Electronic and Information Engineering, School of Engineering, Westlake University

Chenglei Yu

 Department of Artificial Intelligence, School of Engineering, Westlake University

Long Wei

 Department of Artificial Intelligence, School of Engineering, Westlake University

Ruiqi Feng

 Department of Artificial Intelligence, School of Engineering, Westlake University

Jinlong Duan

 Institute of Mechanics, Chinese Academy of Sciences

Dixia Fan

 fandixia@westlake.edu.cn

 Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University

Tailin Wu

 wutailin@westlake.edu.cn

 Department of Artificial Intelligence, School of Engineering, Westlake University

---

Research Article

DOI: https://doi.org/10.21203/rs.3.rs-6881567/v2

Additional Declarations: The authors declare no competing interests.

FluidZero: Mastering Diverse Tasks in Fluid Systems through a Single Generative Model

Haodong Feng[1†], Haoren Zheng[4†‡], Peiyan Hu[5‡], Hongyuan Liu[3], Chenglei Yu[2], Long Wei[2], Ruiqi Feng[2], Jinlong Duan[6‡], Dixia Fan[1,3*], Tailin Wu[2*]

[1]Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University, Hangzhou, Zhejiang, 310030, China.
[2]Department of Artificial Intelligence, School of Engineering, Westlake University, Hangzhou, Zhejiang, 310030, China.
[3]Department of Electronic and Information Engineering, School of Engineering, Westlake University, Hangzhou, 310030, Zhejiang, China.
[4]School of Mathematics and Statistics, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, China.
[5]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China.
[6]Institute of Mechanics, Chinese Academy of Sciences, Beijing, 100190, China.

*Corresponding author(s). E-mail(s): fandixia@westlake.edu.cn;
wutailin@westlake.edu.cn;
Contributing authors: fenghaodong@westlake.edu.cn;
zhenghaoren@westlake.edu.cn; hupeiyan@westlake.edu.cn;
liuhongyuan@westlake.edu.cn; yuchenglei@westlake.edu.cn;
weilong@westlake.edu.cn; fengruiqi@westlake.edu.cn;
duanjl@imech.ac.cn;
[†]These authors contributed equally to this work.
[‡]Work done as an intern at Westlake University.

Abstract

Fluid mechanics is a cornerstone of science and engineering, such as spacecraft, submersibles, and biomedicine, which are important to understand and optimize,

1

involving four key tasks throughout history: prediction, parameter identification, design, and control. Although each task has made significant strides individually, current approaches remain fragmented; existing models are limited to their specific domains and lack the capability to generalize across different tasks. To overcome this issue while utilizing the correlation between tasks to improve the performance on each task, we propose **FluidZero**, a unified deep generative model to tackle all these tasks using one single model. FluidZero is encouraged by the success of large foundation models in several domains, inspiring it possible to develop a unified fluid model capable of handling all these tasks. The key advantage of FluidZero is that it facilitates cross-modal and cross-task interactions on diverse data, enabling effective physical principle learning and significantly enhancing performance on all tasks across a wide range of fluid scenarios. We evaluate FluidZero across multiple datasets, including simulation data and real-world measured data obtained through Particle Image Velocimetry (PIV). Moreover, the designed foil is directly transferred to real-world experiments through 3D printing. Notably, FluidZero shows remarkable generalization capabilities in all scenarios, achieving superior performance even in out-of-distribution (OOD) situations and real-world applications. By integrating diverse fluid system tasks across varied scenarios into a unified model, FluidZero demonstrates the revolutionary impact of generative Artificial Intelligence (AI) approaches in fluid mechanics, opening new avenues for the adoption of integrated fluid system understanding and optimization throughout scientific and engineering domains.

1 Introduction

Fluid mechanics is a cornerstone of science and engineering, underpinning a vast array of critical applications across diverse domains. It enables space exploration through aerodynamic optimization of spacecraft like SpaceX's Dragon [1, 2], advances deep-sea research via the propulsion and stability design of submersibles such as the Deepsea Challenger [3, 4], and informs precision biomedical flow prediction and optimization for cardiovascular treatments [5–7]. These applications span a wide spectrum of tasks—from predicting unsteady flow dynamics to optimizing designs and controls—where understanding and manipulating fluid behavior is essential for ensuring reliability, efficiency, and safety. As such, fluid mechanics is not only a foundational discipline but also a cross-cutting enabler of progress in transportation, energy, medicine, and environmental systems.

Historically, fluid mechanics has been studied since the 17th century by pioneering scientists such as Newton, Bernoulli, Euler, et al. Building on several centuries of progress, in the 21st century, Artificial Intelligence (AI) is rapidly making inroads in fluid mechanics research with the development of AI, which is an unprecedented change [8]. Brunton et al. [8] reviews the AI methods to model fluid dynamic systems for prediction, identify hidden parameters or latent system dynamics for understanding, optimize flow for design, and control fluid behavior through action planning. It

shows that the above four tasks, prediction, identification, design, and control, are fundamentally important parts of fluid system research. Each brings together a large amount of work. Among them, many researchers focus on accurately simulating and predicting fluid system evolution [9–11], others investigate system identification to infer classifications or parameters of fluid systems from observational data [10, 12–18], some concentrate on optimizing the design of structures within fluid systems [19–22], while others dedicate their efforts to developing flow control strategies [23–29]. While each task has made significant strides individually, current approaches remain fragmented, requiring distinct architectures, training data, and optimization schemes for each scenario. Existing models are limited to their specific domains and lack the capability to generalize across different tasks. In light of this, we therefore raise a pivotal question in fluid systems.

Can a single AI model tackle all these tasks in fluid systems?

Addressing this question represents a fundamental shift in how we conceptualize and operationalize fluid systems. These four tasks are not independent as they form an interdependent workflow in both simulation and real-world applications. A unified model could dramatically reduce redundancy, lower repetitive development costs, and streamline the entire modeling-to-optimizing pipeline [30, 31]. Moreover, the shared physical principles underlying all four tasks suggest the possibility of cross-task knowledge transfer, potentially leading to improved generalization and performance [32]. For example, prediction capabilities are fundamental and can provide more information to effective design and control [32, 33].

This vision is increasingly supported by breakthroughs in large foundation models trained on vast and diverse datasets [31, 34–44]. In natural language processing, models like BERT [34], GPT [35], and DeepSeek [36] have demonstrated strong generalization across diverse linguistic tasks. In the vision, CLIP [45] shows impressive zero-shot transfer capabilities. More importantly for scientific domains, models such as the Nature Language Model (NLM) [39] exhibit cross-disciplinary abilities spanning proteins, chemistry, and materials. WeatherGFM [31] integrates several tasks within the prediction range, such as super-resolution, cross-modal inference, and temporal forecasting, into a single framework. Such a paradigm shift in AI development, using a single generalizable framework rather than developing small specialized models for each scenario, offers promising new perspectives for fluid system research. The core insight of these foundation models is that expressive architectures trained on diverse data develop generalizable capabilities that transfer across problems. For fluid systems, which naturally involve multiple modalities (e.g., fluid fields, boundary conditions, physical parameters, forces, and control actions), this shift presents a promising opportunity to move beyond task-specific tools and toward unified, general-purpose modeling.

Despite this promise, no prior work has successfully unified the four canonical tasks in fluid systems. Two major challenges stand in the way: (1) the nonlinear complexity of the Navier–Stokes equations (NSE), one of the most difficult unsolved problems in mathematical physics [46]; and (2) the heterogeneous modalities required for each task—ranging from high-dimensional fields for prediction, to physical parameters, to design geometries, to action sequences in control—making joint learning technically nontrivial.

To address these challenges, we propose **FluidZero**, a unified generative model that unifies all these tasks as a multimodal conditional probabilistic generation task, shown in Fig. 1. Rather than designing separate models for prediction, parameter identification, design, and control, FluidZero casts each as the inference of missing modalities, such as fluid fields, parameters, boundaries, or control actions, conditioned on known ones. Built on a diffusion model backbone, FluidZero learns to transform noise into structured outputs through a denoising process [47], effectively modeling the nonlinear, high-dimensional distributions inherent to fluid dynamics [33, 48]. In particular, FluidZero has three main features. (1) It employs modality-specific encoders to convert different types and dimensions of modalities into latent token representations, enabling effective cross-modal fusion. The design allows FluidZero to handle diverse types and numbers of multimodal data. (2) It adopts a diffusion process to learn probabilistic distributions rather than deterministic mappings. The probabilistic modeling approach enables the model to capture the inherent stochasticity of fluid dynamics and generate diverse yet physically reasonable solutions. (3) It randomly masks the unknown modalities to simulate various task situations in training, enabling the model to learn to complete unknown modalities. The training strategy enables cross-task knowledge transfer and multiplies task generalization, allowing the model to learn fundamental physical principles of fluid dynamics rather than direct mappings between data. Through comprehensive experiments spanning diverse simulations, real-world measurements, and multiple out-of-distribution (OOD) scenarios, we demonstrate that FluidZero achieves superior performance across all tasks. Notably, we demonstrate the practical applicability of FluidZero through the successful transfer of the designed foil to real-world scenarios and show remarkable generalization on real-world measurement data, even for OOD cases, which bridge the gap between numerical methods and practical engineering applications.

2 Results

In this section, we first introduce the datasets and environments (2.1) to provide the foundation for subsequent training and evaluations. We then systematically demonstrate the performance of a single FluidZero model across multiple important tasks through a series of comprehensive experiments, including prediction (2.2), parameter identification (2.3), design (2.4), and control (2.5). We further validate the flexibility of our unified model by examining various task combinations (2.6), demonstrating FluidZero's ability to integrate multiple tasks seamlessly. Finally, we examine OOD scenarios (2.7) to validate the generalization capability and robustness of FluidZero, demonstrating that it maintains superior performance even when faced with previously unseen data distributions.
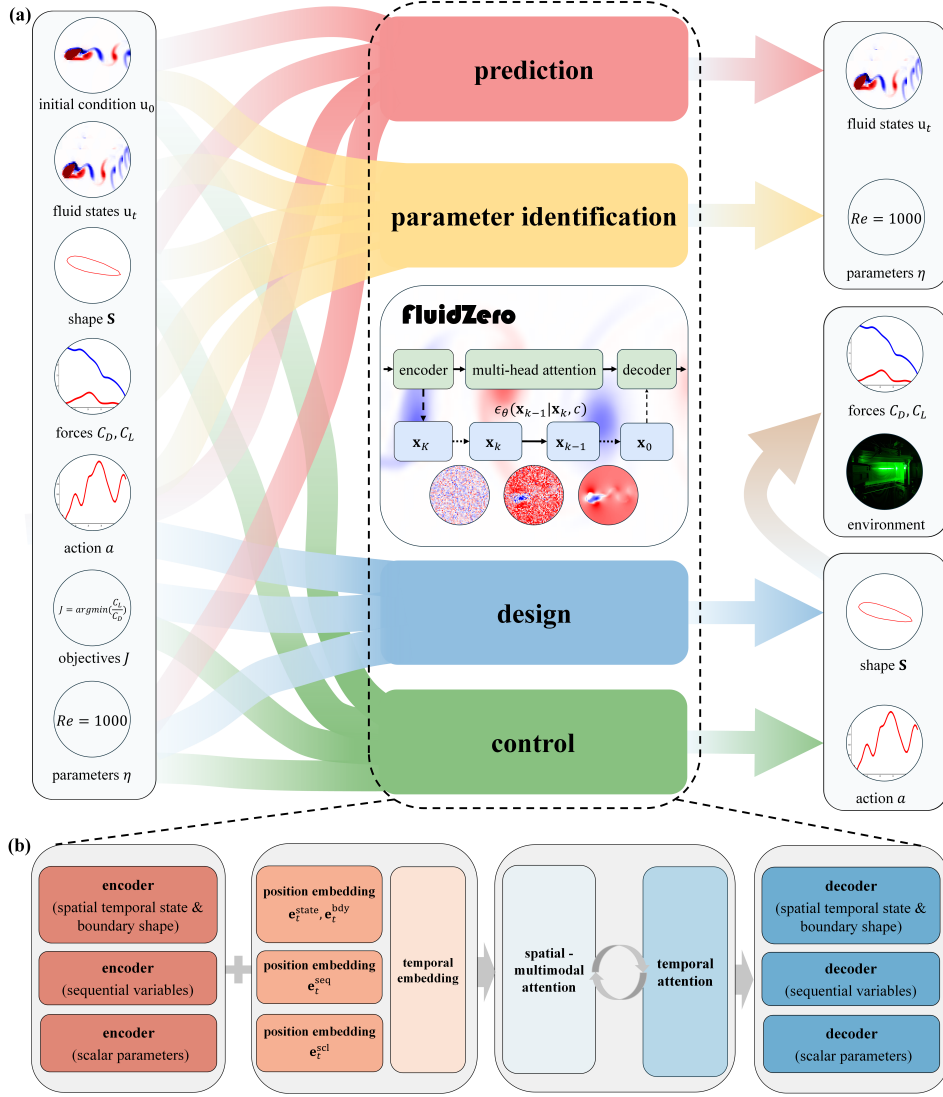
Fig. 1 Schematic illustration of the proposed FluidZero. The sub-figure **(a)** illustrates our proposed FluidZero that unifies multiple tasks in fluid systems, including prediction, parameter identification, design, and control, into a conditional probabilistic generation task. FluidZero, based on the diffusion model and transformer architecture, generates unknown modalities using known modalities as conditions. It starts from Gaussian noise and gradually denoises to generate solutions for various tasks. The sub-figure **(b)** illustrates the framework of FluidZero. Firstly, different modalities are encoded to the token level using modality-specific encoders, and then different positional embedding methods are utilized. The purpose is to map modalities of different dimensions to tokens with the same dimension, allowing for information distillation at the token level between different modalities. Then, tokens enter the transformer, which consists of two main components: spatial-multimodal attention and temporal attention. These parts are leveraged to capture spatial correlations and modal fusion, as well as learn temporal evolution features. After several iterations, the corresponding modalities are output through modality-specific decoders.
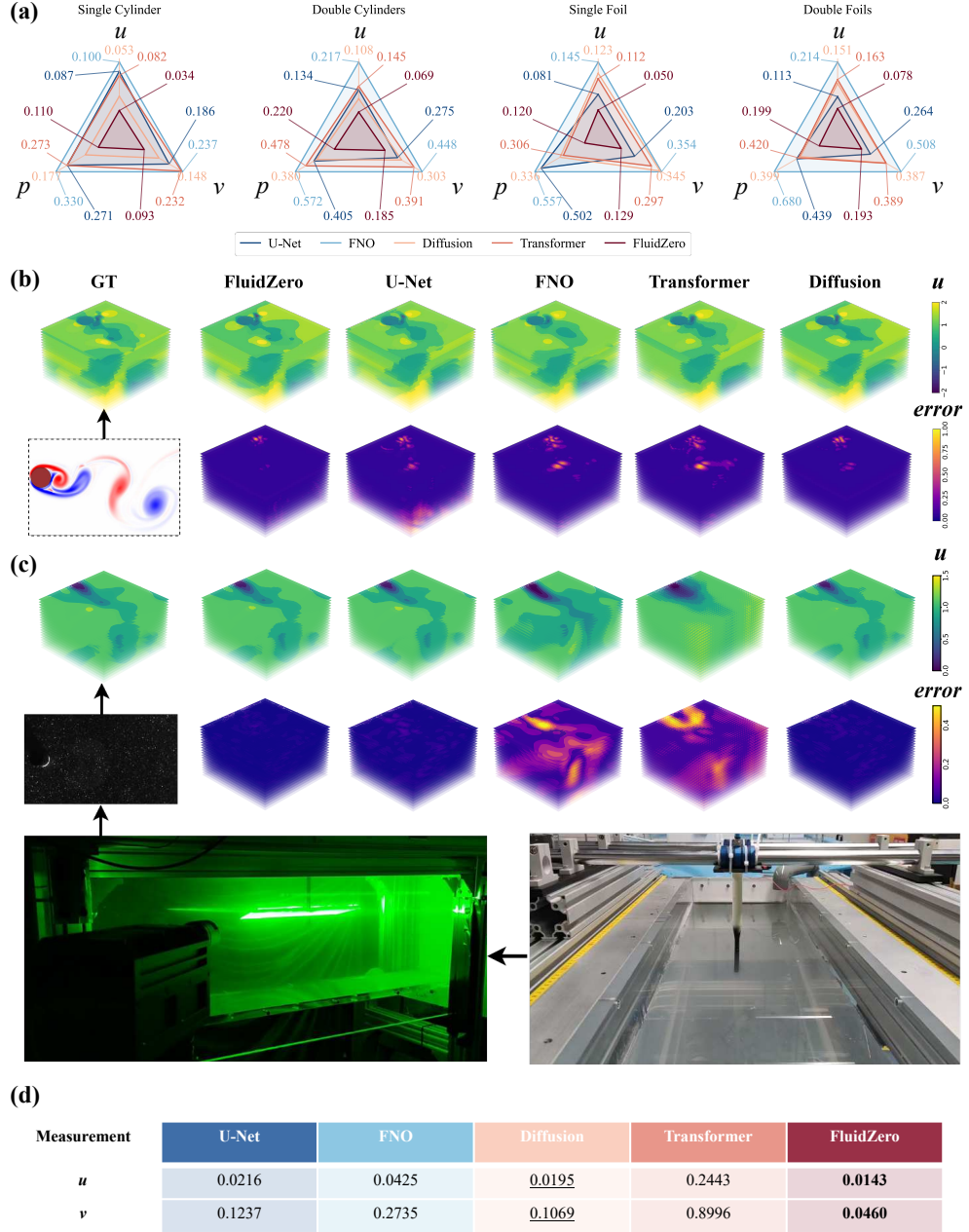
(a)

Single Cylinder
u
0.053 0.082
0.100 0.034
0.087
0.110 0.186
0.273 0.237
p 0.17 0.148 v
0.330 0.232
0.271 0.093

Double Cylinders
u
0.108 0.145
0.217 0.069
0.134
0.220 0.275
0.478 0.448
p 0.386 0.303 v
0.572 0.391
0.405 0.185

Single Foil
u
0.123 0.112
0.145 0.050
0.081
0.120 0.203
0.306 0.354
p 0.338 0.345 v
0.557 0.297
0.502 0.129

Double Foils
u
0.151 0.163
0.214 0.078
0.113
0.199 0.264
0.420 0.508
p 0.399 0.387 v
0.680 0.389
0.439 0.193

— U-Net — FNO — Diffusion — Transformer — FluidZero

(b) GT | FluidZero | U-Net | FNO | Transformer | Diffusion | u

error

(c) | | | | | | u

error

(d)

Measurement	U-Net	FNO	Diffusion	Transformer	FluidZero
u	0.0216	0.0425	<u>0.0195</u>	0.2443	**0.0143**
v	0.1237	0.2735	<u>0.1069</u>	0.8996	**0.0460**

Fig. 2 Results of the prediction task. (a) The average relative L_2 error between the predicted value and the ground truth, with each radar chart containing FluidZero and four baselines. Each radar chart contains results for a fluid scenario. Each scenario contains velocities in two directions (u, v) and pressure (p) fields. The fluid states $\{\mathbf{u}_t\}_{t=0}^T$ of the experimental data in **(d)** includes velocities fields in two directions (u, v). From the figure, it can be seen that FluidZero has the smallest innermost area, indicating the smallest error in all scenarios compared to all baselines. **(b,c)** The following visualizations contain the ground truth, predicted values, and error between them in 50 time steps (height represents the timeline). In the main text, we present the velocity states in the x-direction of a simulation sample and an experimental data sample. For more visualizations and tables, please refer to Supplementary Materials G. The figures demonstrate that the proposed method FluidZerooutperforms the baselines with significantly fewer errors. The videos of prediction on simulation data and measurement are available in the link 1 and link 2, respectively.

2.1 Datasets and Environments

To evaluate FluidZero's performance across diverse fluid systems, we utilize two types of datasets: simulation data and real-world experimental measurements.

For simulation, we employ Lily-pad [49, 50], an immersed boundary method-based CFD solver for the incompressible Navier-Stokes equation as Equation 1,

$$
\begin{cases}
\nabla \cdot \mathbf{u} = 0, & x \in \Omega,\, t \in [0, T], & \text{(continuity equation)} \\
\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \frac{1}{Re} \nabla^2 \mathbf{u} + \mathbf{f}, & & \text{(momentum equation)} \\
\mathbf{u}(x, 0) = \mathbf{u}_0(x), & x \in \Omega, & \text{(initial condition)} \\
B[\mathbf{u}](x, t) = 0, & x \in \partial\Omega,\, t \in [0, T], & \text{(boundary condition)}
\end{cases}
\tag{1}
$$

where $\mathbf{u}$ is the fluid vector field, t is time, ∇ is the gradient operator, p is pressure, Re is the Reynolds number (Re), $\mathbf{f}$ represents external body forces, $\mathbf{u}_0$ is the initial fluid field, $B[\mathbf{u}](x, t) = 0$ is the boundary conditions, Ω is the computational domain, $\partial\Omega$ denotes the boundary of the domain, and $[0, T]$ is the time interval of interest.

Lily-pad has been extensively validated in the prediction and control of machine learning studies [32, 51, 52]. To thoroughly assess FluidZero's generalization ability across various fluid systems, we generate four simulation datasets with the following configurations. **Single Cylinder:** flow around a single cylinder with randomly varying Re and cylinder diameters. **Single Foil:** flow around a foil with different Re, chord lengths, angles of attack, and thickness profiles. **Double Cylinders:** flow around two cylinders with different Re, cylinder diameters, and relative positions between the structures. **Double Foils:** Flow around a combined two foils with randomized parameters for both structures and their relative positioning. In all configurations, the simulation parameters (Re, geometric characteristics, actions, and relative positions of multiple structures) are randomly sampled within physically relevant ranges to ensure comprehensive coverage of possible fluid behaviors.

For experimental measurements, to validate FluidZero's capability to transfer to real-world conditions, we also collect experimental data using Particle Image Velocimetry (PIV) [53, 54]. The technique allows us to measure actual fluid velocity fields. The experimental setup was designed to correspond to the single cylinder scenario where the cylinder is fixed while the inflow velocities and Re are varying. The detail of experimental measurements is introduced in Section 4.5.

These configurations and scenarios are ubiquitous in both natural and engineering contexts. In nature, the cylinder configuration mirrors water flow around plant stems and rocks in ocean and river [55, 56]. The foil scenario represents bird and insect wings during flight [57, 58]. Multiple structure configurations simulate fish swimming in schools [59], birds flying in formation [60], or plant clusters responding to flow [55]. In engineering applications, these configurations are fundamental to wind turbine placement in wind farms [61], foil design [62], building aerodynamics in urban settings [63], and offshore structure stability [64].

2.2 Prediction

The prediction task involves forecasting future fluid states based on initial conditions, system parameters, and external forces. Accurate fluid system prediction is the cornerstone of all fluid mechanics applications, providing essential insights for engineering design, control systems, and scientific understanding [65]. We herein demonstrate the capability of the FluidZero for the prediction of fluid systems. Specifically, we treat the known modalities as condition information for the diffusion model as illustrated in Equation 2. By conditioning the generative process on these modalities, FluidZero learns to sample physically reasonable future states that are consistent with the initial conditions and governing physics. We conduct comprehensive experiments using FluidZero and baseline methods across all datasets, including both simulation and experimental measurement data. The baseline methods, experimental results, and analysis are described in the following paragraphs.

Task Formulation: The prediction task can be formulated as a mapping function:

$$f_{\text{pred}} : \{\mathbf{u}_0, C_L^0, C_D^0, \eta, \mathbf{S}, a_t\}_{t=0}^T \mapsto \{\mathbf{u}_t\}_{t=1}^T, \tag{2}$$

where the goal is to predict the future fluid states ($\mathbf{u}_t$) over a time horizon T, given the initial conditions ($\mathbf{u}_0 \in \{\mathbf{u}_t\}_{t=0}^T$), initial force coefficients (C_L^0, C_D^0), parameters (η), boundary shape ($\mathbf{S}$), and control actions (a_t) over a time horizon T.

To compare the performance of FluidZero with existing methods, we provide two baseline models trained on a separate dataset for individual prediction tasks, namely, the U-Net [66] and Fourier Neural Operator (FNO) [67]. There are two other baseline methods that complete the ablation studies on the training dataset and training manner. The diffusion model (Diffusion) trained on a separate dataset is used to compare the effectiveness of training with diverse datasets and training with a single dataset, while the Transformer is trained by using supervised learning instead of the diffusion model training manner. Please note that the model frameworks of these two baselines are completely consistent with FluidZero, providing a fair comparison. The introduction of the baseline methods is in Supplementary Material F.

Quantitative results and visualizations are presented in Figure 2 comparing FluidZero's predictions against baseline methods. The experimental results in Figure 2(a) demonstrate that FluidZero consistently outperforms specialized baseline methods across all datasets. The average relative error decreases by 37.3%, 37.1%, 45.0%, and 35.9% compared to the best baseline, respectively. Notably, for experimental data, FluidZero achieves the remarkable result that is reduced by 41.9% compared to the error of the best baseline, as shown in Figure 2(d). The results demonstrate the model's exceptional ability to generalize to noisy experimental measurements that inherently differ from clean simulation data. Moreover, the ablation studies (compared to Diffusion) indicate that training with diverse datasets yields significantly better results compared to training on a single dataset. We attribute this improvement to the model's exposure to varied data distributions during training (similar to how GPT models benefit from diverse textual data), which facilitates the learning of fundamental fluid dynamics principles rather than superficial dataset-specific
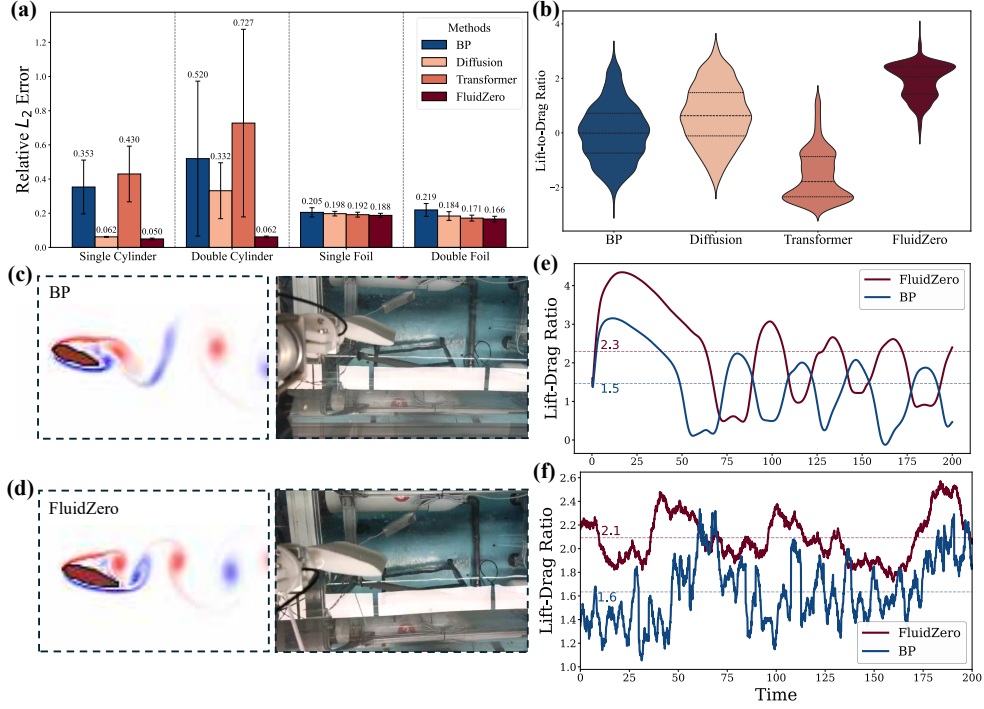
Fig. 3 Results of parameter identification and design tasks. (a) The bar chart contains the average relative L_2 error of parameter identification, including FluidZero and three baselines, and it can be seen that bars of FluidZero have the lowest error. **(b)** The violin plot of lift-to-drag ratio statistics across different methods for design, illustrating the statistical median values and variance for BP, Diffusion, Transformer, and FluidZero. Sub-figure **(c)** and **(d)** provide flow field visualizations of the design task at the final timestep alongside their corresponding experimental photographs for the best performing BP sample and FluidZero sample (they are located in the same Re), respectively. Sub-figure **(e)** displays the temporal evolution of lift-to-drag ratios for the above two samples in simulation, while sub-figure **(f)** shows the corresponding experimental measurement curves in the circulating water tank. For the other experimental measurement and their lift-to-drag ratios, please refer to Supplementary Material G.2. Results demonstrate that foil shapes designed by FluidZero consistently achieve higher lift-to-drag ratios not only in simulations but also when transferred to real-world testing environments, validating the sim-to-real transferability. The videos of design in the simulation and experimental measurement are available in the link.

patterns. Additionally, our ablation studies on training manners (compared to Transformer) reveal the significant advantages of diffusion models over supervised learning manners. While supervised learning directly maps inputs to deterministic outputs, diffusion models learn the underlying probabilistic distribution of the data, allowing them to capture the inherent stochasticity of fluid dynamics and generate diverse yet physically reasonable predictions.

9

2.3 Parameter Identification

As a key component of system identification, parameter identification tackles the inverse problem of recovering the governing physical parameters of the fluid from observed behavior [68, 69]. Precisely estimating quantities such as the Reynolds number from data is pivotal for understanding fluid systems, enabling reliable flow-pattern classification and rigorous validation of theoretical models [14, 16, 17, 70]. The task is typically challenging due to the complex, nonlinear relationship between fluid behavior and underlying physical parameters. We herein demonstrate the capability of FluidZero for identifying key physical parameters in fluid systems. Specifically, FluidZero takes as input a sequence of fluid states (velocity and pressure fields) and boundary conditions (structure geometry), and infers the Re as illustrated in Equation 3. We conduct comprehensive experiments across our four simulation datasets by using FluidZero and three baselines.

Task Formulation: The parameter identification task can be expressed as an inverse problem:

$$f_{\text{pi}} : \{\mathbf{u}_t, \mathbf{S}, a_t, C_L^t, C_D^t\}_{t=0}^T \mapsto \eta, \tag{3}$$

where the goal is to infer the paratmeter Re (η) from observed fluid states ($\mathbf{u}_t$), forces (C_L^t, C_D^t), control actions (a_t), and boundary shape ($\mathbf{S}$) over a time horizon T.

We employ three baselines: a backpropagation-based method (BP) that optimizes parameters by minimizing the difference between observed and predicted fluid states, the Diffusion trained on a single dataset, and the Transformer. The results across all datasets are presented in Figure 3(a). It demonstrates that FluidZero significantly outperforms baseline methods, achieving consistently lower parameter estimation error across all scenarios. Quantitatively, FluidZero achieves relative L_2 errors of 0.050, 0.062, 0.188, and 0.166 for four datasets, respectively, representing improvements of 85.8%, 88.1%, 2.1%, and 2.9% over the best baseline methods. Importantly, FluidZero demonstrates superior stability and robustness, as evidenced by its consistently small error bars (variance less than 0.0048). In contrast, baselines exhibit much larger variance, with BP showing high uncertainty (standard deviations up to 0.4533 for Double Cylinders) and Transformer displaying significant inconsistency (with error bars exceeding 0.5490 for Double Cylinders). Our ablation studies reveal that training with diverse datasets substantially improves parameter identification accuracy compared to single-dataset training, suggesting that exposure to varied flow regimes helps the model build a more comprehensive understanding of how physical parameters manifest in observable flow patterns.

2.4 Design

Foil design represents one of the most challenging and impactful applications of fluid mechanics in engineering [32]. The goal of optimizing designed boundaries to achieve specific performance objectives has traditionally required extensive computational resources and expert knowledge [71, 72]. We herein demonstrate the capability of FluidZero for efficiently designing foil shapes to maximize fluid dynamic performance. Specifically, our design task focuses on generating optimal foil boundaries that maximize the lift-to-drag ratio under given flow conditions as illustrated in Equation 4.

Compared to other tasks, the design problem presents a unique challenge as it operates with minimal initial information, FluidZero generates high-performance shape designs conditioned solely on the Re, without access to initial fluid states or existing boundary conditions. The sparse conditioning scenario tests the model's ability to leverage its learned understanding of fluid-structure interactions to directly propose optimized boundaries. In implementation, FluidZero employs the guided diffusion method [73] to steer the generative process toward boundaries with higher lift-to-drag ratios. By incorporating performance objectives into the sampling process, we enable FluidZero to balance between the exploitation of known high-performance boundaries and the exploration of novel design possibilities. We generate foils simultaneously using FluidZero and baselines with 50 Reynolds numbers as conditions. Although the performance of the best foil generated condition on one Re may also be best at other different Re, different Re validate the stochasticity of the model, thereby demonstrating its robustness. The baselines and analysis of results in both simulation and real-world transfer are described in the following paragraphs.

Task Formulation:

The design task can be formulated as an optimization problem:

$$\mathbf{S}^* = \mathrm{argmax}_{\mathbf{S}} \mathcal{J}(\mathbf{S}, \eta), \tag{4}$$

where the goal is to generate the optimal boundary shape ($\mathbf{S}^*$) that maximums the objective function ($\mathcal{J}$) under given parameters (η). For example, if the objective is to design the foil with the maximum lift-to-drag ratio, the objective function can be defined as $\mathcal{J} = \frac{C_L}{C_D}$ (maximizing the lift-to-drag ratio).

To evaluate the design capability of FluidZero, we compare the results of FluidZero against several baseline models. The primary baseline is a backpropagation-based method [74] (BP) that combines a pre-trained variational auto-encoder (VAE) [75] for foil representation with a U-Net [66] mapping boundaries to forces, allowing gradient-based optimization of foil shapes. Additional baselines for ablation studies include the Diffusion trained only on the single foil dataset and the Transformer trained via supervised learning instead of the diffusion framework.

The results comparing the lift-to-drag ratios of designed foils are presented in Figure 3(b), while Figure 3(c,d) visualizes both the generated foil boundaries and their corresponding flow fields. Statistical analysis of the lift-to-drag ratio distributions reveals that FluidZero achieves the most superior and consistent performance, with a median lift-to-drag ratio of approximately 2 and a tighter distribution. In contrast, baseline methods show significantly inferior performance: BP exhibits a wide distribution centered around 0, Diffusion shows moderate performance with a median around 0.5 but substantial scatter, and Transformer demonstrates the poorest performance with a median near -2 and limited range (-2.5 to 1.5). For the mean and max values, please refer to Table 13 in Supplementary Material G.3. The results demonstrate that FluidZero consistently generates foil designs with superior fluid dynamic performance compared to baselines across various Re conditions. A particularly noteworthy finding is that FluidZero generated designs with lift-to-drag ratios significantly higher than those found in the training dataset, where the average lift-to-drag ratios are close to zero (minimum and maximum thickness-to-chord ratios are 0.3 and 0.6,

11

respectively). It indicates that FluidZero effectively explores and discovers regions of the design space beyond those in the training data, demonstrating true optimization rather than simple imitation of existing designs. Furthermore, our ablation studies reveal that training with diverse datasets significantly enhances the model's design capabilities compared to training on a single dataset. To our knowledge, we are the first to demonstrate and quantify this beneficial effect of data diversity specifically for fluid system design tasks. The finding suggests that exposure to varied flow scenarios helps the model develop a more comprehensive understanding of how boundary modifications affect fluid behavior across different boundaries.

To validate the real-world applicability of FluidZero's designs, we 3D print the top-2 best-designed foils of both baseline and FluidZero and test them in a circulating water tank experiment at multiple Reynolds numbers[1]. The experiment setup is introduced in Section 4.5. These experimental tests confirm that designed foils of FluidZero transferred to experimental fluid environments have superiorly measured lift-to-drag ratios as shown in Figure 3(f). The successful transfer from computational design to physical implementation demonstrates the practical value of FluidZero for real-world engineering applications.

2.5 Control

Planning and control of fluid systems represents one of the most challenging yet promising frontiers in engineering, with applications ranging from aircraft foil control to marine vessel and platform stabilization [52, 76]. Compared to static design problems, control requires both the predictive understanding of fluid behavior and strategic decision-making. We herein present and analyze the results of FluidZero for planning the actions of structures in fluid environments to achieve specific performance objectives alongside baseline comparisons.

Task Formulation:

The control task can be expressed as a decision-making problem:

$$\{a_t^*\}_{t=0}^T = \text{argmax}_{\{a_t\}_{t=0}^T} \mathcal{J}(\{a_t\}_{t=0}^T, \mathbf{u}_0, \mathbf{S}, \eta), \tag{5}$$

where the goal is to generate control actions (a_t) that maximizes objective functions $(\mathcal{J})$ (e.g., drag minimization, lift maximization, or lift-to-drag ratio maximization) based on the initial condition $(\mathbf{u}_0)$, boundary shape $(\mathbf{S})$, and parameters η.

To evaluate the versatility of project, we implement and evaluate control strategies across all simulation environments, with a particular focus on varying control objectives. Specifically, we examine three distinct control objectives: drag minimization and lift maximization for both single cylinder and double cylinders configurations, and lift-to-drag ratio maximization for both single foil and double foils. The diversity of control objectives evaluates FluidZero's ability to learn general control principles rather than objective-specific strategies. FluidZero takes the initial fluid state, boundary condition, initial position, and Re as inputs, along with the given control objective. Then, FluidZero conditionally generates appropriate control actions, like angular velocity,

---

[1]To accommodate the requirements of numerical solver and experimental equipment, all numerical simulations and experimental measurements are conducted in liquid, considering hydrodynamic effects.
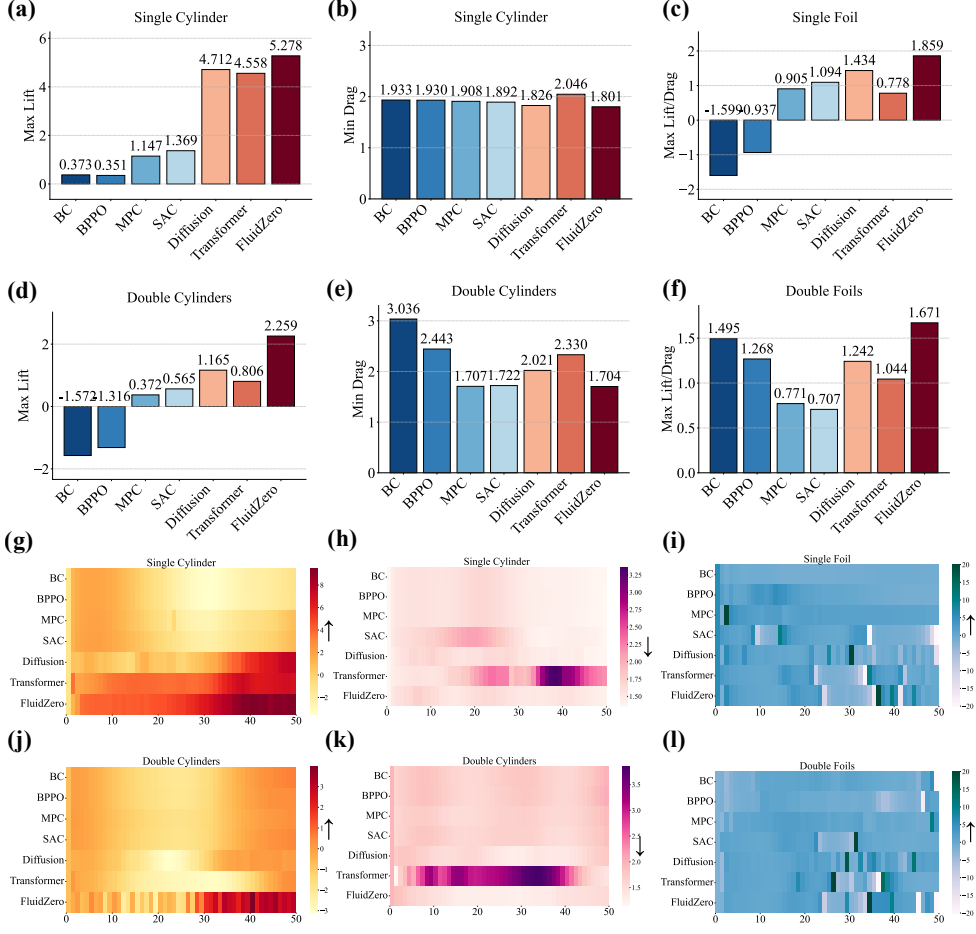
Fig. 4 Results of the control task. The figure presents a comprehensive comparison of control performance across multiple methods for different fluid dynamics scenarios. The visualization includes both quantitative bar charts **(a-f)** and heatmaps **(g-l)** for single cylinder **(a,b,g,h)**, double cylinders **(d,e,j,k)**, single foil **(c,i)**, double foils **(f,l)** configurations. The methods compared include BC, BPPO, MPC, SAC, diffusion model, and transformer, alongside our proposed FluidZero. The objectives in control include maximum lift **(a,d,g,j)**, minimum drag **(b,e,h,k)**, and maximum lift-to-drag ratios **(c,f,i,l)** across different flow configurations. The heatmaps reveal the temporal evolution of control effectiveness, while the bar charts quantify the average performance values. FluidZero consistently demonstrates superior performance across all scenarios. These results validate the effectiveness of our unified conditional generation approach for fluid system control tasks, outperforming both traditional control methods and other learning-based approaches. Moreover, they also demonstrate for the first time that training control strategies on diverse datasets can significantly improve performance compared to training on a single dataset in fluid systems. The videos of control results are available in the link.

to control the fluid system. We evaluate FluidZero against several existing baselines, including conventional Model Predictive Control (MPC) [77], Behavior Cloning (BC)

13

[78], and widely-used Reinforcement Learning algorithms, including Behavior Proximal Policy Optimization (BPPO) [79] and Soft Actor-Critic (SAC) [80]. For ablation studies, we also include two additional baselines: the Diffusion trained on individual datasets and the Transformer trained via supervised learning instead of the diffusion framework.

Quantitative control performances and visualizations for representative evaluation cases are presented in Figure 4. The results demonstrate that FluidZero achieves superior control performance compared to specialized algorithms across all scenarios and control objectives. For lift maximization, our controlled strategies achieve 48.43% higher lift coefficients in the double cylinders scenario, while for lift-to-drag ratio maximization tasks, FluidZero discovers control strategies that achieve the largest lift-to-drag ratio, which is 22.87% more than the best baseline method in the single foil scenario. What is particularly notable is FluidZero's ability to generalize across different control objectives without requiring objective-specific training or tuning.

2.6 Task Compositions

The superior performance of FluidZero across scenarios and tasks compared to all baselines—achieved with a single model—demonstrates its generalization to unify diverse tasks. This generalization capability is achieved through unifying tasks into a probabilistic generation task and training on diverse datasets for multiple tasks. However, beyond this fundamental advantage, we explore the model's flexibility by testing on various task compositions. These experiments, conducted on the single foil scenario, showcase the unique capabilities enabled by FluidZero. We investigate two task compositions.

Task Formulation: The task composition involves combining multiple tasks sequentially. For example:

Parameter identification and control:

$$f_{\text{comp1}} : \{\mathbf{u}_t, \mathbf{S}, C_L^t, C_D^t, a_t\}_{t=0}^T \mapsto \eta \overset{\mathcal{J}}{\mapsto} \{a_t^*\}_{t=0}^T, \tag{6}$$

where a_t is the action in the data sample, while a_t^* is the generated action given the objective J.

Parameter identification and design:

$$f_{\text{comp2}} : \{\mathbf{u}_t, \mathbf{S}, C_L^t, C_D^t, a_t\}_{t=0}^T \mapsto \eta \overset{\mathcal{J}}{\mapsto} \mathbf{S}^*, \tag{7}$$

where $\mathbf{S}$ is the boundary shape in the data sample, while $\mathbf{S}^*$ is the generated shape given the objective J.

First, we evaluate FluidZero's ability to perform sequential parameter identification and control in scenarios where Re is unknown. In this workflow, FluidZero first infers Re from fluid states, then uses this identified parameter as a condition for generating control actions. Results presented in Figure 5(a, left) demonstrate that FluidZero still achieves impressive lift-to-drag ratios while performance shows a modest decrease compared to scenarios with known parameters. The capability is particularly valuable for real-world applications where exact flow parameters may be unknown or uncertain,

14

allowing for adaptive control based on observed fluid behavior. Second, we evaluate the composition of parameter identification and design tasks under conditions of unknown Re. Similar to the above scenario, FluidZero first infers Re from fluid states and then generates optimized foil designs based on the identified parameter. As shown in Figure 5(a, right), the combined method maintains strong design performance, demonstrating FluidZero's robustness and flexibility.

These combined task experiments demonstrate the remarkable flexibility and robustness of FluidZero, enabling novel workflows that leverage the complementary capabilities of different fluid system tasks.

2.7 Evaluation in OOD Scenarios

In the utilization of the model, out-of-distribution (OOD) scenarios are often encountered, which differ significantly from the training distribution. Here, we illustrate the performance of FluidZero in various OOD scenarios, including the prediction of unseen fluid fields and control under substantially higher parameter (Re) configurations for extrapolation, showing that our model captures the underlying physical principles governing fluid dynamics, which enables superior performance.

Firstly, in training, the dataset only includes scenarios of cylinders and foils, but for unseen scenes such as the square, FluidZero still generalizes and achieves better results than baselines, which is reduced by 14.9% compared to the relative error of the best baseline, as shown in Figure 5(c,d). Secondly, we validate FluidZero's generalization capabilities to unseen flow parameters by extending control tests for single foil configurations to substantially higher Re as shown in Figure 5(b). While our training data encompassed Re in the range of 5,000-10,000, FluidZero successfully generalizes control strategies to fluid systems with Re between 10,000-100,000. This remarkable OOD generalization significantly expands the practical applicability of our approach, demonstrating its ability to transfer learned control principles to flow parameters well beyond those encountered during training.

We attribute the OOD generalization capability to FluidZero's cross-modal and cross-task learning paradigm, which enables the model to learn fundamental physical principles underlying fluid dynamics rather than memorizing dataset-specific patterns. The unified model facilitates this knowledge transfer by creating shared representations that capture the invariant physics across different tasks and boundaries. Through simultaneous training on diverse tasks across multiple boundary configurations, FluidZero develops an understanding of the universal relationships between fluid physics, boundary conditions, and system responses.

3 Discussion

This work introduces one single model that masters diverse tasks in fluid systems called FluidZero, unifying prediction, parameter identification, design, and control into a probabilistic generation task for complementing missing modalities of each task. We demonstrate through comprehensive experiments that FluidZero achieves better performance than all baseline methods in diverse scenarios. In addition to the Section 2 shows that FluidZero achieves consistent superior performance over baselines in all
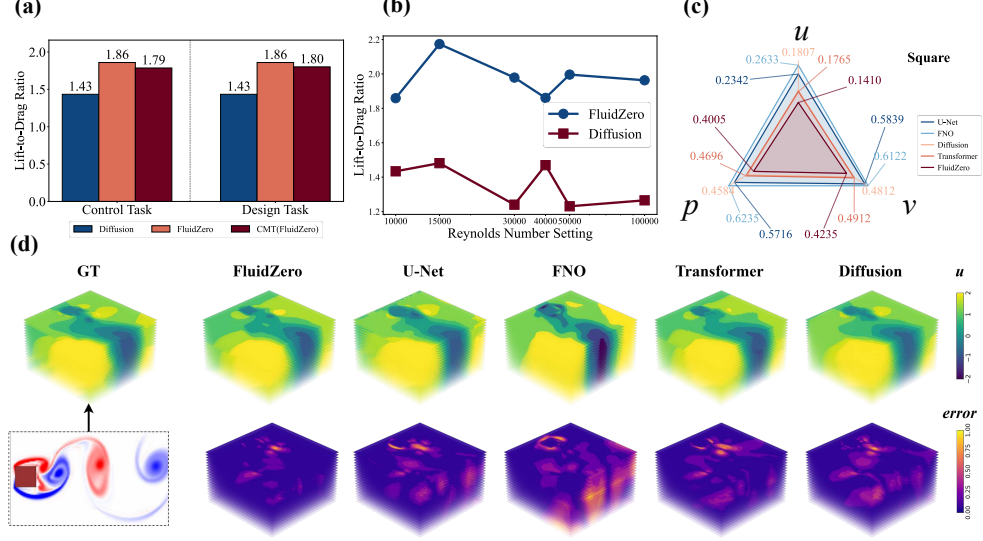
15

Fig. 5 The figure demonstrates FluidZero's generalization capabilities to OOD scenarios and task compositions. The bar chart **(a)** presents a comprehensive performance comparison of the best control baseline (Diffusion) and FluidZero on two compositions of multi-task (CMT): parameter identification and control (left group), and parameter identification and design (right group). The y-axis represents the average lift-to-drag ratios. The results show that when it comes to inferring parameters first, although FluidZero may have a slight decrease in performance, it is still better than the best baseline that knows all the information. The results highlight the critical advantage of our unified model: when physical parameters are unknown, FluidZero can leverage its parameter identification capabilities to first infer these parameters and then subsequently perform downstream tasks such as control or design, enabling effective performance even in scenarios with incomplete information. Sub-figure **(b)** compares lift-to-drag ratio performance across large Reynolds number range (10,000-100,000) for FluidZero versus the Diffusion (the best baseline in the control task), despite being trained only on Reynolds numbers between 5,000-10,000. FluidZero consistently maintains superior performance across all tested Reynolds numbers, with a significant performance gap compared to the baseline. The radar chart **(c)** is the relative L_2 error of the square configuration, including the velocities in two directions and pressure fields. The visualization **(d)** compares predictions of the velocity field in the x-direction for the square, a configuration entirely absent from the training. Predictions of FluidZero remarkably close to the ground truth. The results demonstrate FluidZero's strong capability to transfer learned fluid dynamics principles to both novel configurations and flow regimes well beyond those encountered during training, significantly enhancing its practical applicability.

tasks, various insights can be obtained from different tasks because each task has its own characteristics and challenges. (1) Training on diverse datasets significantly improves performance. Because training with a large amount of diverse data can capture shared principles, which may be commonly present in fluid systems, the insight is reflected in the results of all tasks, especially we attribute the success of OOD scenarios to this reason. (2) Through cross-task learning, the generalization of the model is improved, and the insight is illustrated in the results of design and control tasks compared to the task-specialized baselines. For example, the performance of the designed foil is better than that of the dataset and the generalization of control within OOD scenarios. We attribute it to the insight brought to the design and control by

16

learning other tasks, such as prediction. (3) The unified task model can flexibly combine different tasks to solve the problem of missing modalities, such as not knowing fluid parameters in design and control tasks. Such a model avoids developing many task-specialized models and reduces the cost of repetitive development.

Our proposed FluidZero will have a wide-ranging impact on the fluid mechanics community. The development of FluidZero marks a significant advancement toward unified understanding and optimization of fluid systems. By integrating diverse tasks within a unified probabilistic model, this approach eliminates the need for specialized models while improving performance through cross-task knowledge transfer. It streamlines research workflows and enables seamless integration of fluid mechanics tasks, demonstrating considerable potential across various applications. For example, FluidZero demonstrates the potential for the comprehensive spacecraft development workflow, from initial flow modeling around complex geometries, through aerodynamic design optimization of control surfaces, to flight control [81, 82]. Similarly, for underwater submersibles, the unified model will seamlessly integrate hydrodynamic modeling of turbulent wake flows, hull shape optimization for drag reduction, and control strategies for varying mission requirements [83–85]. Beyond the above applications, FluidZero will extend to critical biomedical domains where it can simultaneously predict blood flow patterns, identify patient-specific cardiovascular parameters, and optimize vessel geometries for surgical planning [86–88]. Moreover, FluidZero also demonstrates the remarkable potential for comprehensive structural analysis and design workflows in the domain of material mechanics. For instance, in semiconductor manufacturing, FluidZero will predict thermal stress distributions and defect formation in silicon wafers, identify material parameters such as elastic moduli and thermal expansion coefficients, and optimize chip packaging designs for thermal management and mechanical reliability [89–91].

Although FluidZero shows promise through extensive experiments, there are several exciting directions for future expansion by integrating complementary approaches. (1) The model can be significantly enhanced by incorporating much more diverse training data across different scenarios to develop a foundation model for fluid mechanics [43, 44]. FluidZero has demonstrated that diverse data is helpful for the generalization of tasks and OOD scenarios. Scaling up with significantly more data spanning broader scenarios beyond cylinders, foils, and simple geometries is expected to achieve better generalization at scale. It potentially creates a universal fluid mechanics foundation model capable of handling arbitrary configurations. (2) The control framework can be further advanced by integrating closed-loop feedback mechanisms, combining asynchronous denoising algorithms [92] to achieve efficient real-time feedback control that adapts to dynamic environmental changes and system uncertainties. (3) The unified model presents opportunities for integration with large language models (LLMs) to enable language-to-task capabilities, building upon existing works [93–96] that combines LLMs with prediction tasks. This integration will allow users to specify complex fluid mechanics tasks through natural language descriptions, automatically translating high-level objectives into specific modeling, identification, design, or control tasks within the FluidZero.

4 Methods

We herein introduce the method of the proposed FluidZero, including an overview, architecture, training, and sampling methods. After that, the experiment setup and evaluation metrics are briefly described. More details can be found in the Supplementary Material C, D, and E.

4.1 Overview of FluidZero

As we mentioned in the previous discussion, one of the important features of FluidZero is the provision of specific encoders for modalities with different dimensions, which are encoded into latent token representations. These tokens, along with position embeddings, are then input into the Transformer, which utilizes the multi-level attention mechanism to learn the patterns of modalities, spatial distribution, and temporal evaluation. The diagram of the architecture is shown in Figure 1(b). With the above multi-task compatible neural network architecture, the training method plays a critical role since FluidZero requires leveraging diverse training data to achieve strong generalization performance across different tasks and scenarios. The training employs a diffusion model framework using balanced sampling to ensure that the model has uniform exposure to features of diverse data. Moreover, we apply task-masked optimization and randomly mask modalities to simulate different task situations, which facilitates cross-task knowledge transfer, enabling the model to learn fundamental physical principles and generalize across multiple tasks. After training, FluidZero performs inference by conditionally sampling from Gaussian noise using task-specific known modalities to generate corresponding solutions.

4.2 Model Architecture for Multimodal Fusion

The motivation behind the model architecture is to enable unified representation and processing of spatial, temporal, and scalar modalities while handling variable numbers of physical structures within a single model. In general, the main component of the architecture is the Transformer [97] that captures the relationships between modalities, spatial and temporal, through attention mechanisms to integrate multi-modalities flexibly and model heterogeneous physical variable data for fluid systems as shown in Figure 1(b). Specifically, the model applies modality-specific encoders to transfer different modalities into latent token representations tailored to their modal characteristics. Then, they are processed by the Transformer architecture unifiedly. The encoders for each modality are further described below, including spatial-temporal fluid states, point clouds representing boundaries, sequential variables like actions and forces, and scalar parameters.

Spatial-temporal fluid state u are two-dimensional fluid fields containing multiple time steps, generally including velocity fields in two directions and a pressure field. The encoder divides the state at each time step into patches, after which each patch is embedded into a token:

$$\mathbf{z}_t^{\text{state}} := \mathcal{F}_{\text{patch}}(\mathbf{u}_t) + \mathbf{e}_t^{\text{state}}, \quad t \in \{1, \ldots, T\}, \tag{8}$$

18

where $\mathcal{F}_{\mathrm{patch}}$ denotes the patch embedding function and $\mathbf{e}_t^{\mathrm{state}}$ are positional embeddings for grid coordinates. By doing this, the fluid state is converted into tokenized representations suitable for the Transformer to learn while preserving its spatial locality.

The boundary shape S is generally represented as point clouds with the coordinates of each point. In order to uniformly represent structures with different shapes and quantities, we first convert the point cloud into the form of Signed Distance Function (SDF) [98] that provides an implicit representation of surfaces of shape, used to describe the shortest distance from a point in an image to a surface, and then divide the SDF image into patches and embed them as tokens:

$$\mathbf{z}_t^{\mathrm{bdy}} := \mathcal{F}_{\mathrm{patch}}(\mathbf{S}_t) + \mathbf{e}_t^{\mathrm{bdy}}, \quad t \in \{1, \ldots, T\}, \tag{9}$$

where $\mathbf{e}_t^{\mathrm{bdy}}$ are positional embeddings.

Sequential variables (C_D, C_L, a) are treated as time-series vectors with length T. The dimensions of these sequential variables depend on the number of structures in fluid systems. For instance, when there is a foil in a fluid system, there will be a set of sequential variables (C_D, C_L, a), and if there are two foils, there will be two sets of them. We apply the encoder with the Fourier transform to abstract the globally periodic features and then nonlinear layers to capture the key features. The above transformation is mathematically reflected in equation 10, which not only allows for the fusion of tokens from different modalities but also makes the model scalable as the Transformer can handle any number of tokens, which means that the number of structures in the fluid system can be expanded to more quantities.

$$\mathbf{z}_t^{\mathrm{seq}} := \mathrm{concat}\left(\left\{\mathcal{F}_{\mathrm{nl}}\left(\mathcal{F}_{\mathrm{FFT}}(s_t^i)\right)\right\}\right) + \mathbf{e}_t^{\mathrm{seq}},$$
$$t \in \{1, \ldots, T\}, \ s \in \{C_D, C_L, a\}, \ i \in \{1, \ldots, I_{\max}\}, \tag{10}$$

where $\mathcal{F}_{\mathrm{nl}}$ and $\mathcal{F}_{\mathrm{FFT}}$ denotes nonlinear layers and Fourier transform, concat refers to concatenation along the token number dimension, $\mathbf{e}_t^{\mathrm{seq}}$ are position embedding for sequence variables, $I_{\max}$ is the number of structures and s represent variables among C_D, C_L, and a.

Scalar parameters η in fluid systems are another modality, like Re that determines the similarity of fluid systems. We apply the encoder with the same architecture as the sequential variables to embed scalar parameters, but their tokens do not change over time and are consistent in a fluid system as shown in equation 11:

$$\mathbf{z}^{\mathrm{scl}} := \mathrm{concat}(\{\mathcal{F}_{\mathrm{nl}}(\mathcal{F}_{\mathrm{FFT}}(\eta^n))\}) + \mathbf{e}^{\mathrm{scl}}, \quad n \in \{1, \ldots, N\}, \tag{11}$$

where N is the number of parameters and $\mathbf{e}^{\mathrm{scl}}$ is the positional embedding.

After converting all modalities to tokens, within each time step t, spatial tokens ($\mathbf{z}^{\mathrm{state}}$ and $\mathbf{z}^{\mathrm{bdy}}$) and non-spatial tokens ($\mathbf{z}^{\mathrm{seq}}$ and $\mathbf{z}^{\mathrm{scl}}$) are concatenated into $\mathbf{Z}_t \in \mathbb{R}^{(N_{\mathrm{spat}} + N_{\mathrm{non\text{-}spat}}) \times d}$, where d is the hidden dimension of tokens. Then, the spatial-multimodal multi-head self-attention (MSA) [97] is calculated across all spatial tokens and tokens of different modalities for every time step $t \in \{1, ..., T\}$ to capture the spatial features and enable cross-modal fusion. After that, in order to maintain temporal

correlation and learn features of fluid system evolution, we further calculate temporal MSA within a horizon when completing spatial-multimodal attention and fusion of different modalities at each iteration. After Transformer processing with an iterative multi-level attention mechanism, tokens of fluid state and shape are decoded to reconstruct the original dimension by unpatchifying. Finally, tokens of sequential variables and scalar variables are processed through the nonlinear layer and converted back to the original space using the inverse Fourier transform.

4.3 Model Training for Estimating the Conditional Distribution for Different Tasks

Based on the above architecture, FluidZero is trained using the diffusion model framework, where the fundamental goal is to train a denoiser to reverse the noise adding process. Its advantages have been mentioned in Section 1. The diffusion model operates by gradually corrupting clean data with noise over a series of diffusion steps and then learning to recover the clean data from the noise [47]. Essentially, the diffusion model learns a conditional probabilistic distribution, where the conditions in FluidZero are known modalities in different tasks.

The forward diffusion process is defined by a Markov chain. Given a clean data sample $\mathbf{x}_0$, which is a set of all modalities of fluid systems, the Gaussian noise is gradually added to $\mathbf{x}_0$ with the increase of diffusion steps K. At step k, the noisy sample $\mathbf{x}_k$ is obtained from $\mathbf{x}_{k-1}$ as follows:

$$q(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \sqrt{1 - \beta_k}\mathbf{x}_{k-1}, \beta_k\mathbf{I}), \tag{12}$$

where $\beta_k \in (0, 1)$ is the variance schedule at step k, and $\mathbf{I}$ is the identity matrix. According to the derivation of Ho et al. [47], we can directly sample $\mathbf{x}_k$ from $\mathbf{x}_0$:

$$q(\mathbf{x}_k|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_k; \sqrt{\bar{\alpha}_k}\mathbf{x}_0, (1 - \bar{\alpha}_k)\mathbf{I}), \tag{13}$$

where $\bar{\alpha}_k = \prod_{i=1}^{k}(1 - \beta_k)$.

The reverse denoising process aims to learn a denoiser $\boldsymbol{\epsilon}_\theta(\mathbf{x}_k, k, \mathbf{c})$ that predicts the noise added at step k in the forward process, where θ represents the model parameters and $\mathbf{c}$ is the condition information.

In implementation, we train FluidZero on all diverse datasets, including single cylinder, single foil, double cylinders, double foils, and PIV data as introduced in Section 2.1. These datasets have different numbers of trajectories. To balance the contributions of different datasets during training, we adopt a custom balanced batch sampling strategy. The strategy first calculates the effective sample count of each dataset, then generates proportionate batches, and finally applies weighted random sampling. By doing so, it ensures that the model has uniform exposure to features of diverse scenarios, which improves generalization and prevents overfitting. Furthermore, we systematically mask different combinations of modalities to create incomplete input situations that reflect different task requirements. The method forces the model to learn underlying physical relationships rather than direct mappings between inputs

20

and labels, as the model must infer missing information based on available modalities and physical consistency. By training on various incomplete situations, the model develops robust inference capabilities and learns to capture invariant physical principles governing fluid dynamics, ultimately achieving better generalization across different tasks.

The training process is divided into pre-training and fine-tuning stages. We conduct pre-training on approximately 120,000 simulation data trajectories. In the pre-training stage, FluidZero is optimized to learn the dynamics and intrinsic physical principles of fluid systems. After pre-training, FluidZero can achieve tasks such as prediction, parameter identification, and control. Furthermore, FluidZero is fine-tuned by a few training epochs to improve generalization, including using a fixed foil simulation dataset to learn how to design a better foil and a real measured PIV dataset with noise to generalize to real-world applications. In both stages, the training objective is to minimize the denoising loss in the context of the diffusion model. The denoising loss is based on the difference between the predicted noise $\boldsymbol{\epsilon}_\theta(\mathbf{x}_k, k, \mathbf{c})$ and the true noise $\boldsymbol{\epsilon}_k$ at each step k. Specifically, the loss function is:

$$\mathcal{L}_{\text{denoise}} = \mathbb{E}_{q(\mathbf{x}_k|\mathbf{x}_0)} \left[\|\boldsymbol{\epsilon}_\theta(\mathbf{x}_k, k, \mathbf{c}) - \boldsymbol{\epsilon}_k\|^2 \right]. \tag{14}$$

This denoising loss is applied across all modalities to optimize the model to learn the reverse denoising process. The comprehensive pre-training and fine-tuning enable FluidZero to effectively learn and perform in diverse tasks and scenarios of fluid systems.

4.4 Sampling Solutions for Each Task from the Single Model

In the inference, we sample solutions for all tasks from the conditional probabilistic distribution using the single model FluidZero. To accelerate sampling, we use the Denoising Diffusion Implicit Models (DDIM) [99], which enhances sampling efficiency while preserving generation quality.

Specifically, we impose different conditions that are known modalities in the task to generate the corresponding solution by sampling from Gaussian noise. (1) The conditions for prediction are consistent with those utilized during the training. Modalities given for prediction are initial condition $\mathbf{u}_0$, initial forces C_L^0 and C_D^0, parameters η, action a: $\mathbf{c} = \{\mathbf{u}_0, C_L^0, C_D^0, \eta, \mathbf{S}, a_t\}_{t=0}^T$ as shown in Equation 2. (2) The conditions for parameter identification are $\mathbf{c} = \{\mathbf{u}_t, \mathbf{S}, a_t, C_L^t, C_D^t\}_{t=0}^T$ as shown in Equation 3. (3) For the design task, the given modalities are parameters η, which is less than others. Therefore, we also attach the additional guided strategy during the sampling process, which transforms the sampling process into an optimization process as: $\text{argmax}_{\mathbf{S}}[\mathcal{J}(\eta, \mathbf{S})]$. By doing it in design, FluidZero balances between the exploitation of known high-performance boundaries and the exploration of novel design possibilities. (4) For the control task, the given modalities are initial condition $\mathbf{u}_0$, initial boundary shape $\mathbf{S}$, parameters η as shown in Equation 5. The conditions of $\{C_L^t, C_D^t\}_{t=0}^T$ are adjusted according to the optimization objective. The details of the sampling can be found in Supplementary Material E. By tailoring conditioning variables and integrating guidance, the single model efficiently performs diverse tasks in fluid systems.

4.5 Experiment Setup for Measurements of Fluid Systems

We conduct laboratory experiments in the circulating water tank to validate FluidZero's ability to transfer to real-world scenarios as demonstrated in Section 2.2 and 2.4. The circulating water tank is an important equipment for conducting inflow experiments in fluid mechanics. Our equipment can design accurate inflow velocities, which reduce the disturbance caused by the backflow using flow straighteners and honeycombs, effectively ensuring uniform flow conditions for precise measurements.

For the prediction task, we collect fluid field data using PIV. As mentioned before, the flow around a cylinder is a classic problem in fluid systems with widely-used applications. We set Re between 350 and 1250 by changing the incoming flow velocity, which includes both fully laminar and transition states to turbulence. The flow field changes in this range are complex, posing a challenge for the model's OOD generalization. The PIV technique can be used to obtain the complete fluid state, and we set the observing window at the wake of the cylinder to better capture its wake characteristics. The measured fluid states exhibit characteristics analogous to simulation data, including overlapping Re regimes while extending to both lower and higher Reynolds values in numerous instances. These experimental datasets feature fixed cylinder configurations and lack pressure information while introducing observation noise, which provides an ideal benchmark for evaluating the model's generalization capabilities when confronted with realistic data.

For the design task, to verify the ability of FluidZero to transfer to real-world scenarios, we 3D print the designed foil and measure the lift-to-drag ratio using the NI DAQ-USB6218 with a ME-K3D40 three-axis force sensor in the circulating water tank. In laboratory experiments, we select the top-2 best foil shapes from FluidZero and the baseline BP, and measure the hydrodynamic performance at multiple Reynolds numbers. The shape and angle of attack of 3D printed foils are consistent with the generated boundary.

4.6 Evaluation Metrics

To rigorously evaluate the performance of the unified model across diverse tasks, we define distinct evaluation metrics tailored to the specific objectives of each task. These metrics are mathematically formalized as follows.

The evaluation metric for the prediction task is the relative L_2 error, which measures the relative error between the predicted values and ground truth values, defined as:

$$\mathcal{L}(\hat{\mathbf{u}}, \mathbf{u}) = \frac{\|\hat{\mathbf{u}} - \mathbf{u}\|^2}{\|\mathbf{u}\|^2}, \tag{15}$$

where $\mathbf{u}$ represents the ground truth value, $\hat{\mathbf{u}}$ is the predicted value. The metric ensures accurate evaluation of predictions by emphasizing relative magnitude discrepancies. The relative L_2 error is also employed for the parameter identification task. Moreover, the evaluation metrics for the design and control task depend on the specific objective functions, including minimizing drag, maximizing lift, and maximizing the lift-to-drag ratio. They use the mean values of drag mean(C_D), lift mean(C_L) and lift-to-drag ratio mean($\frac{C_L}{C_D}$) to evaluate, respectively.

5 Acknowledgements

6 Data Availability

All the used datasets in this study will be available upon publication of the manuscript.

7 Code Available

All the source codes to reproduce the results in this study will be available upon publication of the manuscript.

References

[1] Somasekhar, Y., *et al.*: A review on drag reduction and stability optimization techniques for small satellite launch vehicles. Acceleron Aerospace Journal **3**(2), 427–435 (2024)

[2] Hultgren, E.L., Collins, P., Lettieri, C.: Advances in modeling clusters of parachutes for spacex dragon spacecraft. In: 26th AIAA Aerodynamic Decelerator Systems Technology Conference, p. 2709 (2022)

[3] Cui, W., Hu, Y., Guo, W.: Chinese journey to the challenger deep: The development and first phase of sea trial of an 11,000-m rainbowfish arv. Marine Technology Society Journal **51**(3), 23–35 (2017)

[4] Fletcher, B., Bowen, A., Yoerger, D.R., Whitcomb, L.L.: Journey to the challenger deep: 50 years later with the nereus hybrid remotely operated vehicle. Marine Technology Society Journal **43**(5), 65–76 (2009)

[5] Borazjani, I.: Fluid–structure interaction, immersed boundary-finite element method simulations of bio-prosthetic heart valves. Computer Methods in Applied Mechanics and Engineering **257**, 103–116 (2013)

[6] Bressloff, N.W., Ragkousis, G., Curzen, N.: Design optimisation of coronary artery stent systems. Annals of biomedical engineering **44**, 357–367 (2016)

[7] Marsden, A.L.: Optimization in cardiovascular modeling. Annual review of fluid mechanics **46**(1), 519–546 (2014)

[8] Brunton, S.L., Noack, B.R., Koumoutsakos, P.: Machine learning for fluid mechanics. Annual review of fluid mechanics **52**(1), 477–508 (2020)

[9] Dissanayake, M.G., Phan-Thien, N.: Neural-network-based approximations for solving partial differential equations. communications in Numerical Methods in Engineering **10**(3), 195–201 (1994)

[10] Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics **378**, 686–707 (2019)

[11] Vinuesa, R., Brunton, S.L.: Enhancing computational fluid dynamics with machine learning. Nature Computational Science **2**(6), 358–366 (2022)

[12] González-García, R., Rico-Martìnez, R., Kevrekidis, I.G.: Identification of distributed parameter systems: A neural net based approach. Computers & chemical engineering **22**, 965–968 (1998)

[13] Mezić, I.: Analysis of fluid flows via spectral properties of the koopman operator. Annual review of fluid mechanics **45**(1), 357–378 (2013)

[14] Duraisamy, K., Iaccarino, G., Xiao, H.: Turbulence modeling in the age of data. Annual review of fluid mechanics **51**(1), 357–377 (2019)

[15] Xu, Z., Zheng, G.: Gradient-based optimization method for experimental modal parameter estimation with finite element model. AIAA Journal **62**(9), 3544–3558 (2024)

[16] Kontogiannis, A., Elgersma, S.V., Sederman, A.J., Juniper, M.P.: Bayesian inverse navier–stokes problems: joint flow field reconstruction and parameter learning. Inverse Problems **41**(1), 015008 (2024)

[17] Christopher, J.D., Wimer, N.T., Lapointe, C., Hayden, T.R., Grooms, I., Rieker, G.B., Hamlington, P.E.: Parameter estimation for complex thermal-fluid flows using approximate bayesian computation. Physical Review Fluids **3**(10), 104602 (2018)

[18] Liu, P., Chu, X., Cui, W., Zhao, L., Ge, Y.: Bayesian inference based parametric identification of vortex-excited force using on-site measured vibration data on a long-span bridge. Engineering Structures **266**, 114597 (2022)

[19] Ewoldt, R.H., Saengow, C.: Designing complex fluids. Annual Review of Fluid Mechanics **54**(1), 413–441 (2022)

[20] Giannakoglou, K., Papadimitriou, D., Kampolis, I.: Aerodynamic shape design using evolutionary algorithms and new gradient-assisted metamodels. Computer methods in applied mechanics and engineering **195**(44-47), 6312–6329 (2006)

[21] Van Rees, W.M., Gazzola, M., Koumoutsakos, P.: Optimal morphokinematics

for undulatory swimmers at intermediate reynolds numbers. Journal of Fluid Mechanics **775**, 178–188 (2015)

[22] Wei, Z., Dufour, E.R., Pelletier, C., Fua, P., Bauerheim, M.: Diffairfoil: An efficient novel airfoil sampler based on latent space diffusion model for aerodynamic shape optimization. In: AIAA AVIATION FORUM AND ASCEND 2024, p. 3755 (2024)

[23] Rowley, C.W., Williams, D.R.: Dynamics and control of high-reynolds-number flow over open cavities. Annu. Rev. Fluid Mech. **38**(1), 251–276 (2006)

[24] Rowley, C.W., Dawson, S.T.: Model reduction for flow analysis and control. Annual Review of Fluid Mechanics **49**(1), 387–417 (2017)

[25] Kim, J., Bewley, T.R.: A linear systems approach to flow control. Annu. Rev. Fluid Mech. **39**(1), 383–417 (2007)

[26] Greenblatt, D., Williams, D.R.: Flow control for unmanned air vehicles. Annual Review of Fluid Mechanics **54**(1), 383–412 (2022)

[27] Choi, H., Jeon, W.-P., Kim, J.: Control of flow over a bluff body. Annu. Rev. Fluid Mech. **40**(1), 113–139 (2008)

[28] Lumley, J., Blossey, P.: Control of turbulence. Annual Review of Fluid Mechanics **30**(1), 311–327 (1998)

[29] Fan, D., Yang, L., Wang, Z., Triantafyllou, M.S., Karniadakis, G.E.: Reinforcement learning for bluff body active flow control in experiments and simulations. Proceedings of the National Academy of Sciences **117**(42), 26091–26098 (2020)

[30] Yang, L., Liu, S., Meng, T., Osher, S.J.: In-context operator learning with data prompts for differential equation problems. Proceedings of the National Academy of Sciences **120**(39), 2310142120 (2023)

[31] Zhao, X., Zhou, Z., Zhang, W., Liu, Y., Chen, X., Gong, J., Chen, H., Fei, B., Chen, S., Ouyang, W., *et al.*: Weathergfm: Learning a weather generalist foundation model via in-context learning. In: International Conference on Learning Representations (2025)

[32] Wu, T., Maruyama, T., Wei, L., Zhang, T., Du, Y., Iaccarino, G., Leskovec, J.: Compositional generative inverse design. In: International Conference on Learning Representations (2024)

[33] Wei, L., Hu, P., Feng, R., Feng, H., Du, Y., Zhang, T., Wang, R., Wang, Y., Ma, Z.-M., Wu, T.: Diffphycon: A generative approach to control complex physical systems. In: Advances in Neural Information Processing Systems (2024)

[34] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep

bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (long and Short Papers), pp. 4171–4186 (2019)

[35] Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., et al.: Improving language understanding by generative pre-training. OpenAI (2018)

[36] Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., et al.: Deepseek-v3 technical report. arXiv preprint arXiv:2412.19437 (2024)

[37] Bai, J., Bai, S., Chu, Y., Cui, Z., Dang, K., Deng, X., Fan, Y., Ge, W., Han, Y., Huang, F., et al.: Qwen technical report. arXiv preprint arXiv:2309.16609 (2023)

[38] Yu, Q., Zhang, Z., Zhu, R., Yuan, Y., Zuo, X., Yue, Y., Fan, T., Liu, G., Liu, L., Liu, X., et al.: Dapo: An open-source llm reinforcement learning system at scale. arXiv preprint arXiv:2503.14476 (2025)

[39] Xia, Y., Jin, P., Xie, S., He, L., Cao, C., Luo, R., Liu, G., Wang, Y., Liu, Z., Chen, Y.-J., et al.: Naturelm: Deciphering the language of nature for scientific discovery. arXiv preprint arXiv:2502.07527 (2025)

[40] Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., et al.: Learning skillful medium-range global weather forecasting. Science **382**(6677), 1416–1421 (2023)

[41] Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T.R., El-Kadi, A., Masters, D., Ewalds, T., Stott, J., Mohamed, S., Battaglia, P., et al.: Probabilistic weather forecasting with machine learning. Nature **637**, 84–90 (2025)

[42] Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., Tian, Q.: Accurate medium-range global weather forecasting with 3d neural networks. Nature **619**, 533–538 (2023)

[43] Hao, Z., Su, C., Liu, S., Berner, J., Ying, C., Su, H., Anandkumar, A., Song, J., Zhu, J.: Dpot: Auto-regressive denoising operator transformer for large-scale pde pre-training. In: International Conference on Machine Learning (2024)

[44] Herde, M., Raonic, B., Rohner, T., Käppeli, R., Molinaro, R., Bézenac, E., Mishra, S.: Poseidon: Efficient foundation models for pdes. Advances in Neural Information Processing Systems **37**, 72525–72624 (2024)

[45] Radford, A., Wook, J., Aditya, H., Gabriel, R., Sandhini, G., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., et al.: Learning transferable visual models from natural language supervision. In: International Conference on Machine Learning (2021)

[46] Lemarié-Rieusset, P.G.: The navier-stokes problem in the 21st century. In: Chapman and Hall/CRC (2023)

[47] Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)

[48] Gilpin, W.: Generative learning for nonlinear dynamics. Nature Reviews Physics **6**(3), 194–206 (2024)

[49] Weymouth, G.D., Yue, D.K.-P.: Conservative volume-of-fluid method for free-surface simulations on cartesian-grids. Journal of Computational Physics **229**(8), 2853–2865 (2010)

[50] Weymouth, G.D., Yue, D.K.: Boundary data immersion method for cartesian-grid simulations of fluid-body interaction problems. Journal of Computational Physics **230**(16), 6233–6247 (2011)

[51] Feng, H., Wang, Y., Xiang, H., Jin, Z., Fan, D.: How to control hydrodynamic force on fluidic pinball via deep reinforcement learning. Physics of Fluids **35**(4) (2023)

[52] Wang, Z., Lin, R., Zhao, Z., Chen, X., Guo, P., Yang, N., Wang, Z., Fan, D.: Learn to flap: Foil non-parametric path planning via deep reinforcement learning. Journal of Fluid Mechanics **984**, 9 (2024)

[53] Owen, J.C., Bearman, P.W., Szewczyk, A.A.: Passive control of viv with drag reduction. Journal of Fluids and Structures **15**(3-4), 597–605 (2001)

[54] Guo, P., Kaiser, F., Rival, D.E.: Dynamic separation on an accelerating prolate spheroid. Journal of Fluid Mechanics **975**, 51 (2023)

[55] Vargas-Luna, A., Crosato, A., Calvani, G., Uijttewaal, W.S.: Representing plants as rigid cylinders in experiments and models. Advances in water resources **93**, 205–222 (2016)

[56] Tanino, Y.: Flow and solute transport in random cylinder arrays: A model for emergent aquatic plant canopies. PhD thesis, Massachusetts Institute of Technology (2008)

[57] Hunter, P.: The nature of flight: the molecules and mechanics of flight in animals. EMBO reports **8**(9), 811–813 (2007)

[58] Lentink, D., Biewener, A.A.: Nature-inspired flight—beyond the leap. Bioinspiration & biomimetics **5**(4), 040201 (2010)

[59] Weihs, D.: Hydromechanics of fish schooling. Nature **241**(5387), 290–291 (1973)

[60] Lissaman, P.B., Shollenberger, C.A.: Formation flight of birds. Science **168**(3934),

810   1003–1005 (1970)

[61] Emami, A., Noghreh, P.: New approach on optimization in placement of wind
      turbines within wind farm by genetic algorithms. Renewable Energy **35**(7), 1559–
      1564 (2010)

[62] Du, X., He, P., Martins, J.R.: Rapid airfoil design optimization via neural
      networks-based parameterization and surrogate modeling. Aerospace Science and
      Technology **113**, 106701 (2021)

[63] Zhao, Y., Li, R., Feng, L., Wu, Y., Niu, J., Gao, N.: Boundary layer wind tunnel
      tests of outdoor airflow field around urban buildings: A review of methods and
      status. Renewable and Sustainable Energy Reviews **167**, 112717 (2022)

[64] Kampitsis, A., Kapasakalis, K., Via-Estrem, L.: An integrated fea-cfd simulation
      of offshore wind turbines with vibration control systems. Engineering Structures
      254, 113859 (2022)

[65] Rao, C., Ren, P., Wang, Q., Buyukozturk, O., Sun, H., Liu, Y.: Encoding physics
      to learn reaction–diffusion processes. Nature Machine Intelligence **5**(7), 765–779
      (2023)

[66] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomed-
      ical 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, pp. 234–241 (2015). Springer

[67] Li, Z., Kovachki, N.B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., Anand-
      kumar, A., *et al.*: Fourier neural operator for parametric partial differential
      equations. In: International Conference on Learning Representations (2021)

[68] Ljung, L.: System Identification: Theory for the User, 2nd edn. Prentice-Hall,
      New Jersey (1999)

[69] Isermann, R., Münchhof, M.: Identification of Dynamic Systems: an Introduction
      with Applications vol. 85. Springer, Berlin (2011)

[70] Wu, T., Neiswanger, W., Zheng, H., Ermon, S., Leskovec, J.: Uncertainty quan-
      tification for forward and inverse problems of pdes via latent global evolution.
      In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, pp.
      320–328 (2024)

[71] Wang, L., Xu, J., Wang, Z., Zhang, B., Luo, Z., Yuan, J., Tan, A.C.: A novel cost-
      efficient deep learning framework for static fluid–structure interaction analysis
      of hydrofoil in tidal turbine morphing blade. Renewable Energy **208**, 367–384
      (2023)

28

[72] Safdar, M.M., Anand, A., Marepally, K., Lee, B., Baeder, J.D.: A surrogate modeling framework for airfoil design and optimization. In: AIAA Scitech 2024 Forum, p. 0458 (2024)

[73] Ho, J., Salimans, T.: Classifier-free diffusion guidance. In: NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications (2021)

[74] Allen, K., Lopez-Guevara, T., Stachenfeld, K.L., Sanchez Gonzalez, A., Battaglia, P., Hamrick, J.B., Pfaff, T.: Inverse design for fluid-structure interactions using graph network simulators. Advances in Neural Information Processing Systems **35**, 13759–13774 (2022)

[75] Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In: International Conference on Learning Representations (2014)

[76] Sun, Y., Yan, C., Xiang, X., Zhou, H., Tang, D., Zhu, Y.: Towards end-to-end formation control for robotic fish via deep reinforcement learning with non-expert imitation. Ocean engineering **271**, 113811 (2023)

[77] Darby, M.L., Nikolaou, M.: Mpc: Current practice and challenges. Control Engineering Practice **20**(4), 328–342 (2012)

[78] Pomerleau, D.A.: Alvinn: An autonomous land vehicle in a neural network. Advances in neural information processing systems **1** (1988)

[79] Zhuang, Z., Lei, K., Liu, J., Wang, D., Guo, Y.: Behavior proximal policy optimization. In: International Conference on Learning Representations (2023)

[80] Haarnoja, T., Zhou, A., Abbeel, P., Levine, S.: Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International Conference on Machine Learning (2018)

[81] Shyy, W., Papila, N., Vaidyanathan, R., Tucker, K.: Global design optimization for aerodynamics and rocket propulsion components. Progress in Aerospace Sciences **37**(1), 59–118 (2001)

[82] Pérez-Roca, S., Marzat, J., Piet-Lahanier, H., Langlois, N., Farago, F., Galeotta, M., Le Gonidec, S.: A survey of automatic control methods for liquid-propellant rocket engines. Progress in Aerospace Sciences **107**, 63–84 (2019)

[83] Hong, L., Wang, X., Zhang, D.-S.: Cfd-based hydrodynamic performance investigation of autonomous underwater vehicles: A survey. Ocean Engineering **305**, 117911 (2024)

[84] Liu, Y., Yu, Z., Zhang, L., Liu, T., Feng, D., Zhang, J.: A fine drag coefficient model for hull shape of underwater vehicles. Ocean engineering **236**, 109361 (2021)

29

[85] Caccia, M., Veruggio, G.: Guidance and control of a reconfigurable unmanned underwater vehicle. Control engineering practice **8**(1), 21–37 (2000)

[86] Zélicourt, D.A., Kurtcuoglu, V.: Patient-specific surgical planning, where do we stand? the example of the fontan procedure. Annals of biomedical engineering **44**, 174–186 (2016)

[87] Raissi, M., Yazdani, A., Karniadakis, G.E.: Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. Science **367**(6481), 1026–1030 (2020)

[88] Marconi, S., Negrello, E., Mauri, V., Pugliese, L., Peri, A., Argenti, F., Auricchio, F., Pietrabissa, A.: Toward the improvement of 3d-printed vessels' anatomical models for robotic surgery training. The International Journal of Artificial Organs **42**(10), 558–565 (2019)

[89] Buratti, Y., Le Gia, Q.T., Dick, J., Zhu, Y., Hameiri, Z.: Extracting bulk defect parameters in silicon wafers using machine learning models. npj Computational Materials **6**(1), 142 (2020)

[90] Sadiqbatcha, S., Zhang, J., Amrouch, H., Tan, S.X.-D.: Real-time full-chip thermal tracking: A post-silicon, machine learning perspective. IEEE Transactions on Computers **71**(6), 1411–1424 (2021)

[91] Son, K., Kim, M., Park, H., Lho, D., Son, K., Kim, K., Lee, S., Jeong, S., Park, S., Hong, S., *et al.*: Reinforcement-learning-based signal integrity optimization and analysis of a scalable 3-d x-point array structure. IEEE Transactions on Components, Packaging and Manufacturing Technology **12**(1), 100–110 (2021)

[92] Wei, L., Feng, H., Yang, Y., Feng, R., Hu, P., Zheng, X., Zhang, T., Fan, D., Wu, T.: Closed-loop diffusion control of complex physical systems. In: International Conference on Learning Representations (2025)

[93] Li, S., Marwah, T., Shen, J., Sun, W., Risteski, A., Yang, Y., Talwalkar, A.: Codepde: An inference framework for llm-driven pde solver generation. arXiv preprint arXiv:2505.08783 (2025)

[94] Wu, Q., Gao, C., Chen, T., Huang, Y., Zhang, Y., Wang, J., Li, J., Zhou, H., Zhang, S.: Pinnsagent: Automated pde surrogation with large language models. In: International Conference on Machine Learning (2025)

[95] Feng, J., Xu, R., Chu, X.: Openfoamgpt 2.0: end-to-end, trustworthy automation for computational fluid dynamics. arXiv preprint arXiv:2504.19338 (2025)

[96] Zhou, A., Li, Z., Schneier, M., Buchanan Jr, J.R., Farimani, A.B.: Text2pde: Latent diffusion models for accessible physics simulation. In: International Conference on Learning Representations (2025)

[97] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)

[98] Osher, S., Fedkiw, R., Piechor, K.: Level set methods and dynamic implicit surfaces. Appl. Mech. Rev. **57**(3), 15–15 (2004)

[99] Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: International Conference on Learning Representations (2021)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf