

# Learning to Move Objects with Fluid Streams in a Differentiable Simulation

Kārlis Freivalds, Laura Leja, Oskars Teikmanis

*Institute of Electronics and Computer Science (EDI), Riga, Latvia*

{karlis.freivalds, laura.leja, oskars.teikmanis}@edi.lv

**Abstract**—We introduce a method for contactless manipulation of objects in three-dimensional space using controlled fluid streams. To achieve this, we train a neural network controller in a differentiable simulation and evaluate it in a simulated environment consisting of an  $8 \times 8$  grid of vertical emitters. By carrying out various horizontal displacement tasks such as moving objects to specific positions while reacting to external perturbations, we demonstrate that a controller, trained with a limited number of iterations, can generalise to longer episodes and learn the complex dynamics of fluid-solid interactions. Importantly, our approach requires only the observation of the manipulated object’s state, paving the way for the development of physical systems that enable contactless manipulation of objects using air streams.

## I. INTRODUCTION

Real-time control of systems involving fluid and solid interactions presents inherent challenges due to the complex dynamics arising from their interactions and the high-dimensional state space. Designing a controller for such a system, for example, one that moves an object to a prescribed location with a fluid jet, is highly impractical through manual programming alone. However, by leveraging a suitable simulator in conjunction with modern deep learning techniques, it becomes feasible to train such a controller, which is the focus of this paper. We present an approach for controlling fluids using differentiable physics (DP) that only requires monitoring the state of the displaced solids, making it highly applicable to the design of physical fluid-based control systems.

Advances in DP [1] have demonstrated significant potential for addressing simulated inverse problems and complex control applications [2]–[8]. In the context of computational fluid-solid interactions, there are several examples of applied DP [2], [4], [6], [9], [10]. However, to our knowledge, there is no existing research of DP in control tasks in which the fluid itself serves as the acting agent.

We propose a scenario in which controlled fluid streams are employed to displace objects in three-dimensional space with a grid of emitters. Using differentiable simulations, we train a neural network that produces a fluid velocity for each emitter such that objects are moved to a desired location, as shown in Fig. 1. The system observes only the state (position, velocity) of the objects to carry out its control policy. Consequently, unlike prior approaches that rely on the fluid state (velocity and

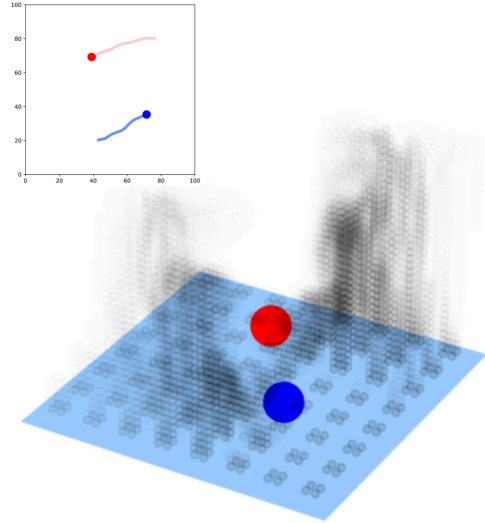


Fig. 1. **Controlled displacement of simulated objects.** The solid lines in the top left shows the trajectory of the objects as they are moved by the simulated fluid. The fluid is emitted from an  $8 \times 8$  grid, and is visualised as a grey plume. Animations can be found in the accompanied [video](#).

pressure fields) for decision-making [11], [12], our approach has the potential for a real-world implementation.

**Our key contributions** are the following: (1) present a sample-efficient DP-based learning method to train a neural network controller for moving objects around a horizontal surface with fixed vertical emitters, using only the state of the object as a control input, and (2) demonstrate versatile horizontal object displacement using fixed vertical emitters in a simulated environment.

The remainder of this paper is structured as follows: Section II describes the simulation environment and governing fluid dynamics equations. Section III offers insights into our controller design, and Section IV describes the training process. An evaluation of all experiments we performed with the controller is available in Section V. Lastly, Section VI puts the developed system into perspective, discussing its limitations and applicability in real-life counterparts.

### A. Related Work

Through the application of gradient-based optimisation, DP has been effectively applied to solve simulated inverse problems [2]–[4]. Hu et al. [2] leverage the differentiable programming language Taichi to conduct complex simulations

This research is funded by the Latvian Council of Science, project “Smart Materials, Photonics, Technologies and Engineering Ecosystem”, project No. VPP-EM-FOTONIKA-2022/1-0001.

involving multi-contact interactions between rigid and soft bodies, as well as computational fluid dynamics. Du et al. [4] present a dedicated differentiable simulator tailored for soft-body manipulation and showcase its use in system identification, trajectory planning and real-to-simulation transfer. Holl et al. [3] introduce PhiFlow, a differentiable partial differential equation (PDE) solver that can be used to handle control and planning tasks for high-dimensional problems, including Navier-Stokes equations for incompressible fluids. Among potential solvers, PhiFlow offers a good fit for our application.

DP has also proven effective in diverse control applications [5]–[7]. Ramos et al. [6] investigate a simulation involving a two-dimensional rigid object immersed in a fluid, and develop a DP-based controller for trajectory following amidst external perturbations. In a comparative analysis against a reference PID controller and a loop-shaping controller, the DP-based controller outperformed both counterparts, demonstrating lower average errors and reduced oscillations. The work of Ma et al. [9] demonstrates a differentiable pipeline for designing the geometry and controller of simulated swimmers. By applying gradient-based parameter optimization, their proposed method results in significantly faster convergence and better loss than gradient-free alternatives.

Reinforcement learning (RL) has been effectively applied in designing controllers for fluid-solid interactions [13], including drag reduction through active flow control [14], [15], efficient swarm-swimming strategies [16] and fluid-directed rigid body control [11], [12]. In earlier work by Ma et al. [11], the objective is to balance and displace 2D objects with simulated air flows. In this example, the controller uses the state of the object, the emitter and the fluid itself as inputs, and the loss is a function of the object’s deviation from the desired state. Although this approach proved successful in a simulated environment, the real-time measurement of the fluid state becomes highly impractical if it were to be applied on a physical device. Ren et al. [12] offer to improve upon Ma et al. through the application of an off-policy RL algorithm (Soft Actor-Critic). While this change does improve the sample-efficiency of the method for completing similar tasks, the overall approach still requires the state of the fluid as an input to the controller.

RL methods usually require a large number of samples, which are computationally expensive to obtain in 3D fluid simulations. In addition, they require careful parameter tuning to avoid unstable behaviour [12], and are challenging to scale to high dimensional parameter space. For instance, for their rigid body balancing scenario, Ma et al. [11] train a system with an action space of 3 (spout position, angle and the on/off state of the jet), while our setup requires 64 continuous parameters that control the emitted fluid velocities.

In this paper, we aim to demonstrate the effectiveness of DP-based learning of three-dimensional fluid-directed rigid body control by applying gradient-based optimisation for simulated fluid-solid interactions, and showcase the resulting controller for enabling complex manipulation tasks by using only vertical fluid emitters, *without* requiring any inputs relating to the state of the fluid during control.

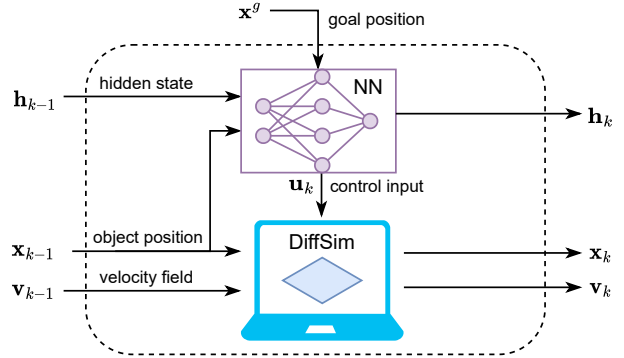


Fig. 2. **Differentiable control diagram.** Shown is one time step of the controlled simulation. The differentiable simulator (bottom) operates on the velocity field  $\mathbf{v}$  containing fluid velocity vectors for each spatial location at a time step  $k$ , the object’s 3D position  $\mathbf{x}$ , and the control signal  $\mathbf{u}$ , corresponding to the emitter velocity. The neural network controller (top) receives the object’s position from the simulator at the previous time step ( $k - 1$ ), its own hidden state  $\mathbf{h}$ , and the goal position of the controlled objects in order to update its hidden state and the control signal.

## II. SIMULATION

In this work we consider a scenario where objects are displaced with vertical fluid emitters arranged in a grid on a horizontal grid (also referred to as “table”), as shown in Fig. 1. The fluid velocities are controlled by a neural network controller, which takes the position and velocity of the objects as inputs, and outputs the required velocities for each emitter to guide the objects to their target locations.

The simulation domain is designed as a cuboid of size  $100 \times 100 \times 100$  units with all boundaries open, except for the bottom boundary, which functions as a solid surface. On this surface, emitters are arranged in an  $8 \times 8$  grid, each capable of emitting fluid at different velocities. Within this setup, one or more rigid objects are placed to demonstrate controlled solid-fluid interactions. Section V features examples with one and two spherical objects with a radius of 10 units, but the approach can handle any number of such objects.

The system behaviour is modelled using fluid dynamics within a bounded domain, and is governed by the incompressible Navier-Stokes (NS) equations, applicable for time and space instances  $x, t \in \Omega_f \times [0, t]$ :

$$\frac{\partial u}{\partial t} = -u \cdot \nabla u - \frac{\nabla p}{\rho} + \nu \nabla^2 u + g, \quad (1)$$

$$\nabla \cdot u = 0, \quad (2)$$

where the fluid velocity  $u(t, x)$  and pressure  $p(t, x)$  are spatiotemporal functions defined within the domain  $\Omega_f \subset \mathbb{R}^2$ . The term  $\frac{\partial u}{\partial t}$  indicates the rate of change of fluid velocity over time. The term  $-u \cdot \nabla u$  is the convective acceleration of the fluid. The pressure gradient force  $-\frac{\nabla p}{\rho}$ , arising from changes in pressure within the flow, is normalised by the fluid density  $\rho$ . The term  $\nu \nabla^2 u$  represents the viscous diffusion of the velocity, where  $\nu$  is the kinematic viscosity of the fluid, and  $\nabla^2 u$  denotes the diffusion of the velocity field due to viscosity. The term  $g$  represents external forces per unit mass acting on the fluid, such as gravity in our context.

The divergence constraint of the velocity field,  $\nabla \cdot u = 0$ , ensures the conservation of fluid mass and constant density within the flow. The negative sign in the pressure gradient term,  $-\frac{\nabla p}{\rho}$ , reflects a pressure drop in the direction of fluid flow, which is proportional to a given volumetric force  $f$ .

For our task of purposefully moving objects with fluid jets, it is essential to effectively manage fluid-solid interactions. This necessitates to account for the velocity and acceleration of the solids created by the fluid. To address this, we adapt the method from [6] which can be described with the following equations:

$$m \frac{\partial \mathbf{v}}{\partial t} = \mathbf{f} = - \oint_{\partial\Omega} p(x, t) \vec{n}(s) ds \quad (3)$$

$$I \frac{\partial \alpha}{\partial t} = \mathbf{s} = - \oint_{\partial\Omega} p(x, t) \vec{r}(s) \times \vec{n}(s) ds \quad (4)$$

where  $m$  is the body mass,  $\mathbf{v}$  is the velocity of the body,  $\frac{\partial \mathbf{v}}{\partial t}$  is the acceleration,  $I$  is the moment of inertia,  $\omega$  is the angular velocity, and  $\frac{\partial \omega}{\partial t}$  represents the angular acceleration. The forces ( $\mathbf{f}$ ) and torques ( $\mathbf{s}$ ) acting on the body are calculated by integrating the pressure  $p(x, t)$  over the surface of the body  $\partial\Omega$ , with  $\vec{n}(s)$  being the outward normal at the surface element  $ds$ , and  $\vec{r}(s)$  representing the position vector relative to the centre of mass.

In parallel, the velocity of the fluid at the boundary of the region  $\Omega$  equals the velocity of the immersed body, expressed as  $\mathbf{v} + \mathbf{r} \times \alpha$ . In NS simulations, Dirichlet boundary conditions are imposed on the velocity field. These conditions are specified as  $\mathbf{u}_\Omega = \frac{\partial \mathbf{v}}{\partial t} + \mathbf{r}_\Omega \frac{\partial \alpha}{\partial t}$  for the rigid body surface. Our approach involves discretizing the fluid region at each time step, where the discretization must accurately describe the boundary surface of the body.

Our simulation environment is implemented using the differentiable PDE solver and simulation toolkit PhiFlow [3]. It supports fluid simulations and one-way coupling with solids, i.e., solid objects influence the fluid but not vice-versa. To enable full coupling, we employ the implementation from Ramos et al. [6], adapted for 3D simulations. With this approach, we can calculate the force exerted by the fluid on solids. For the current implementation, we simulate only spherical objects, for which we implement Newtonian motion and collision handling using PhiFlow's integrated solver.

The domain is discretized at a  $32 \times 32 \times 32$  resolution, which is the maximum for training on one A100 GPU card. The time step is set to 0.5, representing the upper limit for maintaining training stability. The emitters are modelled as boundary conditions on the bottom plane of the domain. Each emitter functions as an inlet of a small area, producing fluid velocities as prescribed by the controller for each time step.

### III. NEURAL NETWORK CONTROLLER DESIGN

In a control scenario, the process depicted in Fig. 2 is applied for each time step. The simulation is initialised with a velocity field  $\mathbf{v}_0 = \mathbf{0}$ . For the hidden state vector  $\mathbf{h}_0$ , all elements are initialised with  $1/\sqrt{M_h}$ , a small nonzero value that prevents catastrophic errors during normalisation, which could occur if all values were zero inside the neural

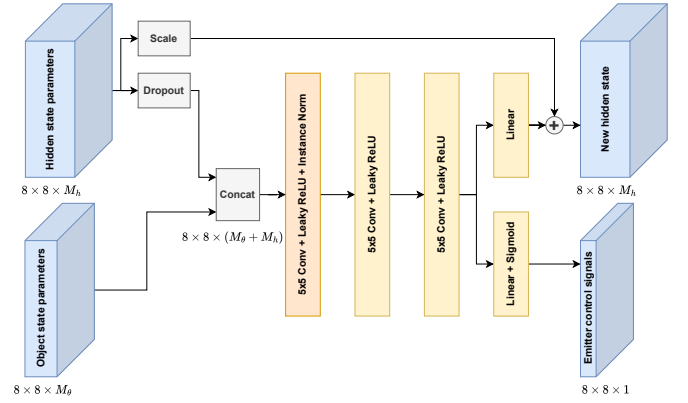


Fig. 3. **Neural network architecture.** We use a recurrent convolutional neural network (RCNN) that operates on an input grid of  $8 \times 8$  fluid emitters with a corresponding feature map of  $M_g$  elements, and a hidden state of size  $M_h$ . Dropout is applied to the hidden state to enhance generalisation to longer simulations. Two output heads performing linear projection are attached to the last layer, producing emitter velocities (range-limited via sigmoid activation) and an updated hidden state for the next time step by applying a scaled residual connection.

network depicted in Fig. 3. The initial state  $\mathbf{x}_0$  depends on the control task, but we usually set the initial velocity to zero and randomise the initial position within the boundaries of the environment. During training, this structure is unrolled through time, and the loss function is calculated depending on  $\mathbf{x}_N$ , where  $N$  is the number of time steps. Gradient backpropagation is performed to adjust the neural network controller's parameters to minimise the loss.

The neural network architecture is depicted in Fig. 3. It is designed as a recurrent convolutional network, receiving the object state representation and its own internal state as inputs at each time step. It produces an emitter control signal matrix and the next internal state tensor as outputs, and operates on spatial data with an  $8 \times 8$  shape, corresponding to the emitter grid. By design, convolutions are well suited for such spatial data processing. Recurrence is implemented as scaled residual connections similar to ResNet [17], facilitating stable training through many time steps.

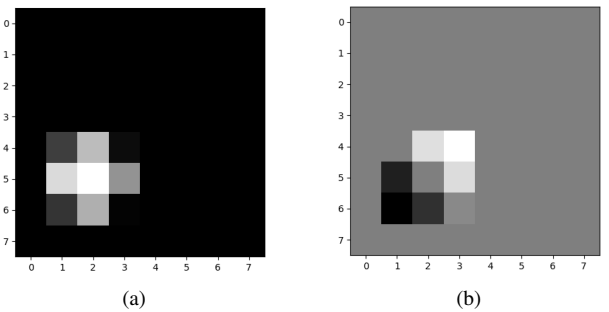


Fig. 4. **Object position and velocity representation.** (a) Object soft mask rendered onto an  $8 \times 8$  grid. A brighter shade indicates greater coverage of the object within a grid cell. (b) Velocity representation derived by subtracting two object masks displaced by velocity. Bright colours indicate positive values, dark colours represent negative values, and zero is depicted in grey.

To fit the convolutional structure, the object state is given to the neural network as an  $8 \times 8$  matrix with several feature

maps:

1. The object’s position in the  $xy$  plane is represented as an  $8 \times 8$  image onto which the object is rendered, see Fig. 4(a). To support gradient backpropagation, the object is rendered in PhiFlow using soft boundaries giving the volume fraction occupied by the object, similar to [18].
2. The object’s  $z$  coordinate is represented as the position mask scaled with the object’s  $z$  value.
3. To represent the  $x$  and  $y$  components of the object’s velocity, the difference of two object masks is calculated – one from the current time step and another from the subsequent time step, as illustrated in Fig. 4(b). The position mask for the next time step is estimated by adding the current velocity to the current object position.
4. For the  $z$  component of the velocity, the object mask obtained in step 1 scaled by the velocity’s  $z$  component is used.

In addition, the goal position  $\mathbf{x}^g$  is passed to the neural network as a distance field. Consequently, each spatial location  $(i, j)$  within the convolutional map is assigned the absolute value of the distance to the goal, calculated as  $|x^g - (i, j, 0)|$  for the  $x, y$  and  $z$  coordinates, resulting in 3 additional feature maps. This representation aids the neural network in discerning the direction towards the goal and remains effective even when the goal lies outside the domain, as illustrated in the scenario depicted in Fig. 7. In total, there are seven input feature maps per controlled object.

#### IV. TRAINING

To train the system, we initialise objects with randomised locations and velocities, conduct simulations, and use the object trajectory for loss calculation. The loss function primarily expresses the distance of the object to the desired target location  $\mathbf{x}^g$  with additional terms:

$$L = \|x_N - x^g\|^2 + \alpha \sum_{k=1}^N \|x_k - x^g\|^2 + \beta \sum_{k=1}^N \|\mathbf{u}_k\|^2. \quad (5)$$

The first term ensures that the final object position is close to the target position. The second term keeps all intermediate positions near the target, effectively maximising the object’s speed towards the target and maintaining its position once it arrives. We experimentally set  $\alpha = 1/N$  to achieve a good balance with the first and most important term. The third term aims to minimise the total energy expended by the emitters throughout the episode. We set a small weight  $\beta = 0.001$  for this term during training. The resulting reduction in fluid emissions based on this value is evaluated in Section V.

All steps of the simulation are differentiable, and we employ gradient descent with the Adam optimiser [19] using a learning rate of 0.0001. Based on experiments, the training is stable for values between 0.001 and 0.0001, resulting in minor differences in training speed. Training is performed for 2000 iterations, which takes about 14 hours on a single Nvidia A100 graphics processor.

Training requires a significant amount of GPU memory since all simulation steps must be preserved for backpropagation. With a  $32 \times 32 \times 32$  domain discretization, approximately

80 steps can be stored on an A100 GPU. This is too small to complete the control tasks we consider. To enable longer rollouts, we employ truncated backpropagation through time, wherein backpropagation is applied solely to the final 80 steps of the rollout, while the total simulation length varies randomly from 80 to 280 steps. We observed that 280 training steps are sufficient to train the tasks, and the trained controller effectively supports much longer simulations, up to 600 steps used in evaluation, as shown in Section V.

#### V. EVALUATION

We evaluate the proposed system in three different scenarios that showcase the capabilities of our controller to carry out versatile object displacement using controlled fluid flows:

1. keep the object in place,
2. move the object to a desired location,
3. independently control two objects.

Additionally, we investigate the system’s robustness in response to external perturbations. During these evaluations, simulations are conducted over 600 time steps, which is sufficient to accomplish the specified tasks.

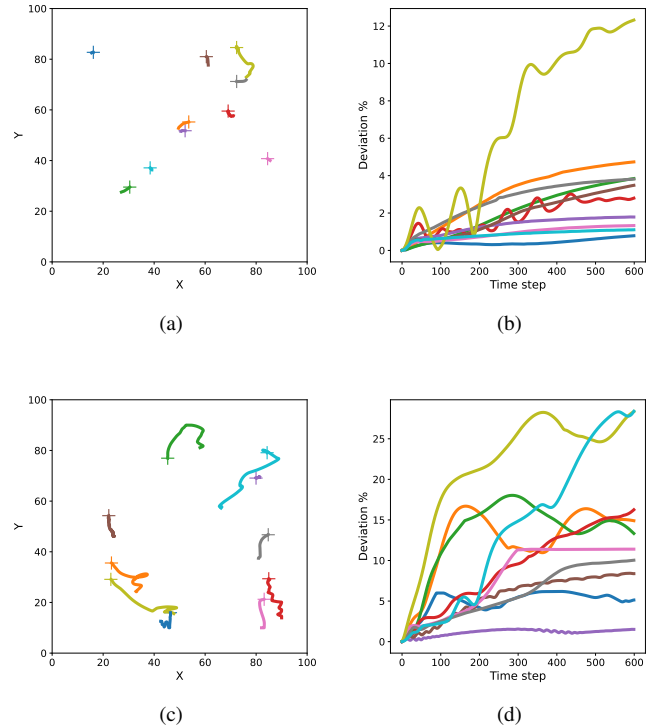


Fig. 5. Task: maintain the position of the object without side wind (top) and with side wind (bottom). Left side: trajectories of several objects in the  $xy$  plane. The target position is marked with a “+” sign. Right side: deviation from the target throughout the simulation time. The deviation is evaluated as the Euclidean distance to the desired location. Trajectory colours between left and right plots are set to match.

First, we consider the task of maintaining an object’s position in a predetermined location above the table. The object is initially placed at the desired location and the system has to learn to regulate the emitters to counteract the force of gravity, while also minimising horizontal displacement.

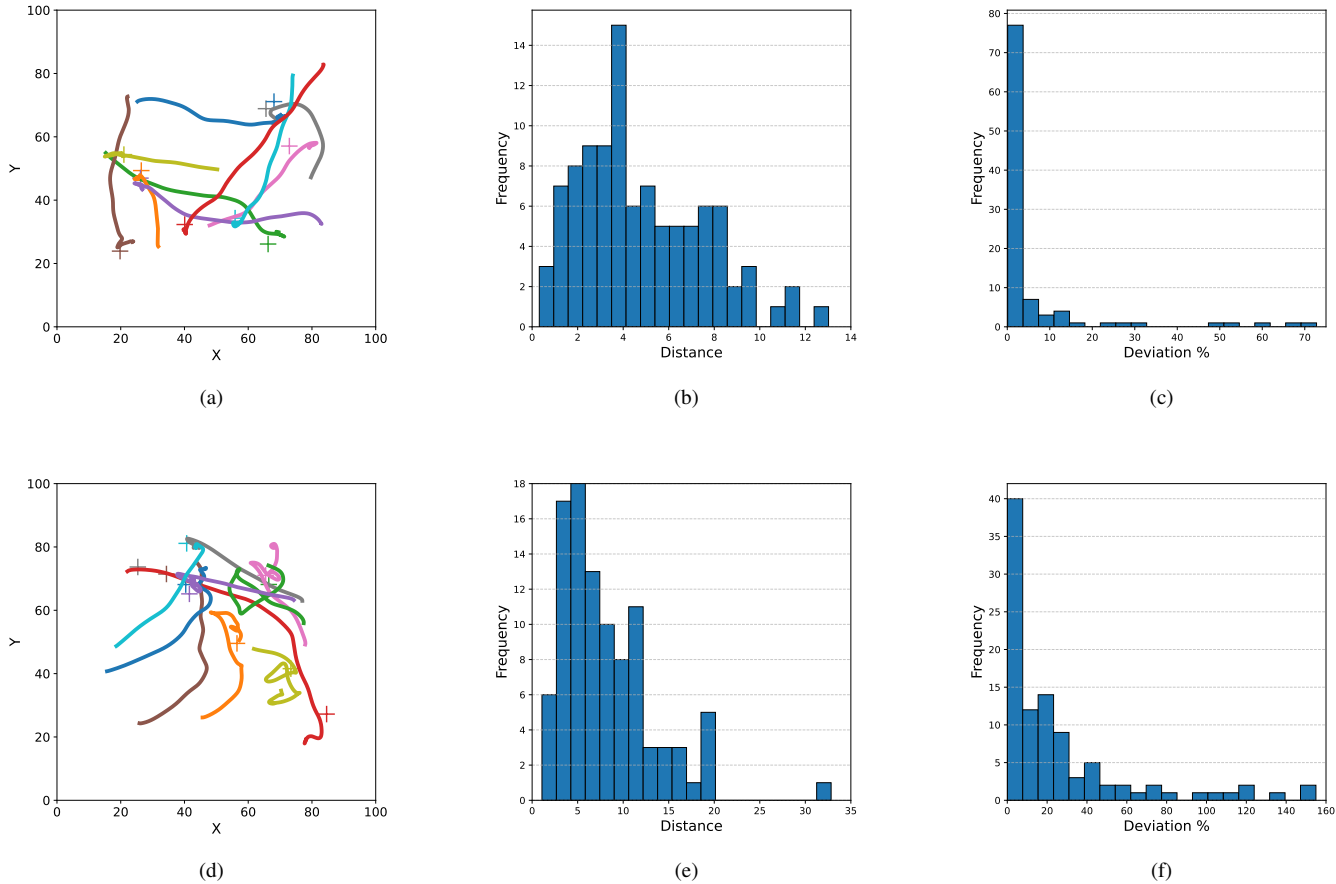


Fig. 6. Task: move an object to the desired location. (a) Sample trajectories of the moved objects. The goal locations are marked with a "+" sign. (b) A histogram displaying proximity of the objects' location to the target position, measuring the distance between these two points in the last time step, for 100 trajectories. (c) Percentage by which the actual path deviates from a straight line, for 100 trajectories, shown as a histogram. (d) Sample trajectories of moved objects that are disturbed by side wind. (e) and (f) Trajectory accuracy metrics under the influence of side wind, analogous to (b) and (c).

Sample trajectories projected onto the  $xy$  plane are depicted in Fig. 5(a). In Fig. 5(b), we measure the deviation from the target over time including the  $z$  coordinate. The objects stay close to their initial position, within a tolerance of 5 units for all trajectories but one, where the units are relative to the simulation domain where the table size is  $100 \times 100$  units. However, we also observe that the error tends to increase over time. This is due to the controller being trained on 280 time steps, and having to generalise for anything beyond that.

Next we analyse the system's robustness against external disturbances. To this end, we repeat the same scenario in the presence of side wind. A horizontal wind of random direction and magnitude (up to 20% of the emitter velocity) is applied during both training and testing. Its intensity is kept constant during a simulated episode, but it changes from one episode to the next. Thus, its magnitude cannot be stored inside the controller during training and has to be inferred from the object's movement during evaluation. The results are depicted in Fig. 5(c). Again, the controller succeeds in learning this task, though at a reduced performance compared to the previous experiment. This is an expected result, since the neural network has to acquire the wind strength by observing the ball motion before acting on it, and during this time the

ball has already deviated slightly from its optimal location.

The next task involves moving the object to a predetermined location. First, the object is randomly initialised on the table, and the task is to move it to another randomly specified location and hold it there until the end of the episode. A sample of the resulting trajectories is shown in Fig. 6(a). In order to determine the accuracy of these trajectories, we measure the distance of the moved object to its target at the end of the episode. We perform 100 simulations and the histogram of the measurements is shown in Fig. 6(b). In most cases the object reaches a 5 unit distance from the target and the median distance is 4.0. Additionally, in Fig. 6(c), we analyse the straightness of the trajectories, specifically comparing the length of the trajectory to the distance between the initial and final points. We see that the excess distance, shown as a percentage relative to the straight-line distance, is typically lower than 5% with the median being 1.3.

We repeat the same experiment with side wind. The results are depicted in Fig. 6(d) - 6(f). We see that the controller is again able to guide the object towards the target, although with less precision (median distance from the target is 6.7 in 100 experiments) and with more curved trajectories (median straightness deviation is 14.4).

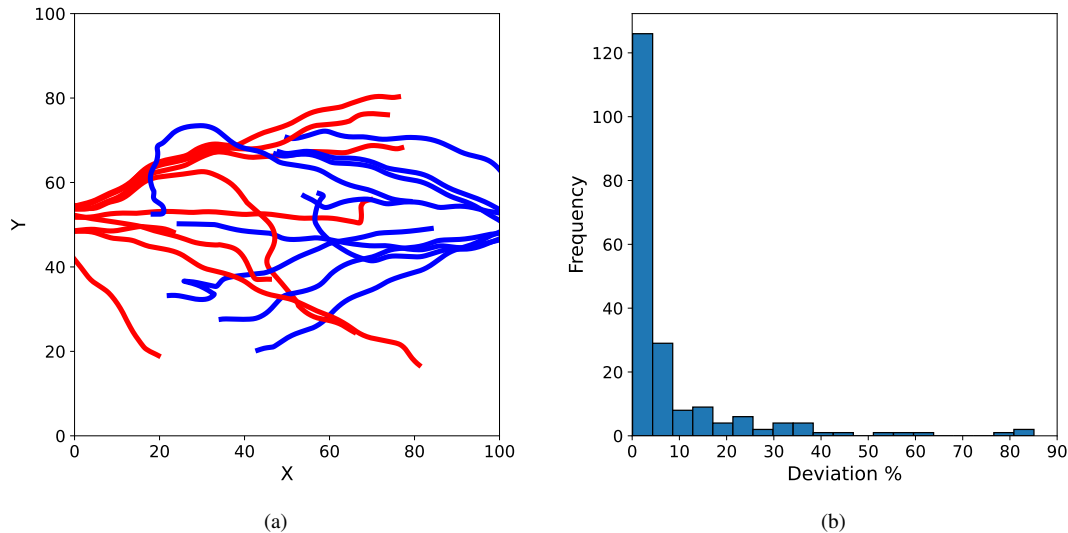


Fig. 7. Task: independently control two objects, where each object is required to arrive at a specified lateral position. (a) Sample trajectories of moved two objects, when the red object should reach the target point located at  $(-10, 50)$  on the left, and the blue object should reach the target point situated at  $(110, 50)$  on the right. (b) Percentage by which the actual path deviates from a straight line, for 100 trajectories.

Lastly, we investigate independent control of two objects. The task is to move the first object to the left of the table and the second one to the right simultaneously. The initial object locations are distributed randomly, meaning that their paths can collide. Results are shown in Fig. 7. We see in Fig. 7(a) that the controller has learnt to correctly move both objects to their respective targets. Additionally, when both objects come into close proximity, the emitters can separate them successfully and move them toward their respective targets, as seen in the accompanied video. In Fig. 7(b) we see the trajectory straightness depicted in the same way as in Fig. 6(c). Compared with the single-object case, the straightness is slightly lower, mainly resulting from curved trajectories to avoid collisions.

In order to avoid high energy consumption for real-life implementations of such devices, it is generally advantageous to minimise the fluid consumption of the emitters. To that end, we increase the constant  $\beta$  of the loss function in Eq. 5 to 0.02 and train the system for an additional 2000 iterations on top of the pre-trained model. Training it from scratch with the increased constant results in sub-optimal behaviour, with emitters turned off over entire episodes. The chosen value of 0.02 is near the maximum threshold, where the unwanted behaviour of emitters consistently shutting down does not occur. When tested on 100 simulations of the moving task, this adjustment results in an average reduction of emitted fluid volume from 0.32 to 0.05, while reducing the performance metrics only slightly. The average distance to the target increases from 3.32 to 4.74 but excess trajectory length increases from 5.28 to 7.78.

## VI. CONCLUSION

We have presented a method to train a neural network controller that can perform object displacement tasks using an array of fixed vertical fluid emitters. This approach has

demonstrated versatile control tasks across multiple simulated scenarios, including: (1) hovering an object in place, (2) moving an object perpendicular to the flow direction, and (3) separating two objects. Unlike previous work, our approach relies solely on observing the object’s state (its position and velocity) this way enabling to be implemented in real-world devices, where the object state can be readily monitored using computer vision techniques – a direction we aim to explore further.

Our controller has also proven robust against external disturbances, which is an essential feature for bridging the gap between simulations and real-world devices. It also achieves object control while minimising the energy consumption in the emitters without significant loss in performance. The controller, trained by leveraging a differentiable simulation, reaches good performance within just 2000 iterations and can handle much longer episodes than those used during training.

However, several challenges remain before this approach can be implemented in a real-life device. The most important one is the expected difference in controller behaviour between simulation and reality. Currently, our simulations use abstract units for fluid and solid properties without corresponding to actual physical quantities. The simulation’s resolution and episode lengths are also limited, which may not accurately reflect real-world dynamics. Moreover, we have only considered spherical objects with uniform rotational inertia across all axes.

Despite these limitations, which are largely technical and addressable, our work represents a meaningful step toward developing real-world systems for manipulating objects with fluid streams.

## REFERENCES

- [1] N. Thuerey, P. Holl, M. Mueller, P. Schnell, F. Trost, and K. Um, *Physics-based Deep Learning*. WWW, 2021. [Online]. Available: <https://physicsbaseddeeplearning.org>.
- [2] Y. Hu, L. Anderson, T.-M. Li, *et al.*, “DiffTaichi: Differentiable Programming for Physical Simulation,” in *Proceedings of the 7th International Conference on Learning Representations (ICLR)*, 2019.
- [3] P. Holl, V. Koltun, and N. Thuerey, “Learning to Control PDEs with Differentiable Physics,” in *Proceedings of the 8th International Conference on Learning Representations (ICLR)*, 2020.
- [4] T. Du, K. Wu, P. Ma, *et al.*, “Diffpd: Differentiable projective dynamics,” *ACM Trans. Graph.*, vol. 41, no. 2, Nov. 2021, ISSN: 0730-0301. DOI: [10.1145/3490168](https://doi.org/10.1145/3490168). [Online]. Available: <https://doi.org/10.1145/3490168>.
- [5] Y. Fang, J. Liu, M. Zhang, *et al.*, “Complex locomotion skill learning via differentiable physics,” *arXiv preprint arXiv:2206.02341*, 2022.
- [6] B. Ramos, F. Trost, and N. Thuerey, “Control of Two-way Coupled Fluid Systems with Differentiable Solvers,” in *ICLR 2022 Workshop on Generalizable Policy Learning in Physical World*, 2022.
- [7] O. Teikmanis, L. Leja, and K. Freivalds, “Applying a Differentiable Physics Simulation to Move Objects with Fluid Streams,” in *International Workshop on Embedded Digital Intelligence (IWEDI)*, 2023.
- [8] N. N. Tathawadkar, N. A. K. Doan, C. F. Silva, and N. Thuerey, “Incomplete to complete multiphysics forecasting: A hybrid approach for learning unknown phenomena,” *Data-Centric Engineering*, vol. 4, e27, 2023.
- [9] P. Ma, T. Du, J. Z. Zhang, *et al.*, “Diffaqua: A differentiable computational design pipeline for soft underwater swimmers with shape interpolation,” *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1–14, 2021.
- [10] Z. Li, Q. Xu, X. Ye, B. Ren, and L. Liu, “Difffr: Differentiable sph-based fluid-rigid coupling for rigid body control,” *ACM Transactions on Graphics (TOG)*, vol. 42, no. 6, pp. 1–17, 2023.
- [11] P. Ma, Y. Tian, Z. Pan, B. Ren, and D. Manocha, “Fluid directed rigid body control using deep reinforcement learning,” *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 1–11, 2018.
- [12] B. Ren, X. Ye, Z. Pan, and T. Zhang, “Versatile control of fluid-directed solid objects using multi-task reinforcement learning,” *ACM Transactions on Graphics*, vol. 42, no. 2, pp. 1–14, 2022.
- [13] P. Garnier, J. Viquerat, J. Rabault, A. Larcher, A. Kuhnle, and E. Hachem, “A review on deep reinforcement learning for fluid mechanics,” *Computers & Fluids*, vol. 225, p. 104973, 2021.
- [14] J. Rabault, M. Kuchta, A. Jensen, U. Réglade, and N. Cerardi, “Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control,” *Journal of fluid mechanics*, vol. 865, pp. 281–302, 2019.
- [15] D. Fan, L. Yang, Z. Wang, M. S. Triantafyllou, and G. E. Karniadakis, “Reinforcement learning for bluff body active flow control in experiments and simulations,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 42, pp. 26 091–26 098, 2020.
- [16] S. Verma, G. Novati, and P. Koumoutsakos, “Efficient collective swimming by harnessing vortices through deep reinforcement learning,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 23, pp. 5849–5854, 2018.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [18] T. Takahashi, J. Liang, Y.-L. Qiao, and M. C. Lin, “Differentiable fluids with solid coupling for learning and control,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, 2021, pp. 6138–6146.
- [19] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.