the flow dynamics. First, a zero-net mass-flux blowing/suction strategy is provided, comprised of two equal and opposite jets - one at the top and one at the bottom of the cylinder. The RL agent controls the mass flux of the upper jet, which automatically specifies the lower jet through the zero-net-mass constraint. In the default configuration, the jets are placed at 90° and 270° positions and cover 10° of the cylinder circumference each, though the HydroGym framework is designed to allow for flexible user-specified adjustments of the jet number, placement angles, and coverage areas. Second, a rotating surface actuator is implemented for which the RL agent directly controls the angular velocity of the cylinder surface. Both action spaces are normalized to the range $[-1, 1]$ to ensure consistent RL algorithm performance across different actuation modes.

The environment supports both two-dimensional and three-dimensional cylinder flow configurations across multiple Reynolds numbers ($Re = 100,\ 200,\ 1,000,\ 3,900$), enabling the investigation of flow control across different flow regimes from laminar vortex shedding to fully turbulent conditions. For three-dimensional cases at $Re = 1,000$ and $3,900$, the framework additionally supports multi-agent reinforcement learning through spanwise domain decomposition, where the cylinder is divided into 4 and 10 subcylinder segments respectively, each individually controlled by a dedicated agent. This multi-agent approach enables the investigation of cooperative control strategies and spanwise flow coordination. The framework provides flexibility in multi-agent coordination, allowing agents to share gradient information and experience buffer data if desired, though independent agent operation is also supported based on user requirements.

The state space representation consists of user-configurable point probes distributed throughout the computational domain, providing measurements of velocity components, pressure, density, and body forces at specified locations. By default, the state observations consist of the raw solver output, but different normalization strategies are implemented to normalize the state space in the range $[-1, 1]$ for consistent RL training performance. This flexible sensor configuration allows the investigation of minimal sensing strategies and optimal sensor placement for effective flow control.

Each training episode spans 200 control actions, corresponding to approximately 10 vortex shedding periods, providing sufficient time for the agent to observe and respond to flow dynamics. The temporal coupling between the CFD solver and RL agent employs Reynolds number-specific action frequencies: 200, 350, and 600 CFD timesteps per control action for 2D cases at $Re = 200,\ 1,000,\ 3,900$ respectively, and 400, 500, and 1,200 timesteps for corresponding 3D cases. These ratios are designed to match the characteristic timescales of vortex shedding dynamics while maintaining computational efficiency and control authority.

The environment framework provides a comprehensive testbed for investigating reinforcement learning-based active flow control strategies, supporting both fundamental research into control mechanisms and practical development of sensing and actuation approaches for cylinder wake management. All described configurations represent default settings that can be easily modified by users to accommodate specific research requirements, including custom reward functions, alternative actuation strategies, modified sensor placements, and adjusted temporal coupling parameters.

5.2 Fluidic pinball flow

Characteristic physics. The fluidic pinball serves as an exemplary test case for investigating complex wake dynamics and flow control strategies in transitional flow regimes. This benchmark configuration consists of three equal-diameter circular cylinders positioned at the vertices of an equilateral triangular arrangement, with one vertex directed toward the incoming flow [171]. Although geometrically simple, this arrangement demonstrates a diverse spectrum of nonlinear flow behaviors, including steady-state solutions, time-periodic oscillations, quasi-periodic motion, and fully chaotic dynamics across different operating parameters.

The fluidic pinball exhibits distinct flow regime transitions as the Reynolds number varies from 30 to 150, with critical thresholds that differ from classical single-cylinder wake behavior. Initial investigations by [172] identified specific transition points unique to this multi-cylinder configuration. For Reynolds numbers below $\mathrm{Re} < 18$, the flow field maintains global stability with symmetric recirculation zones forming downstream of the cylinder cluster. The transition from this steady state occurs via a supercritical Hopf bifurcation at

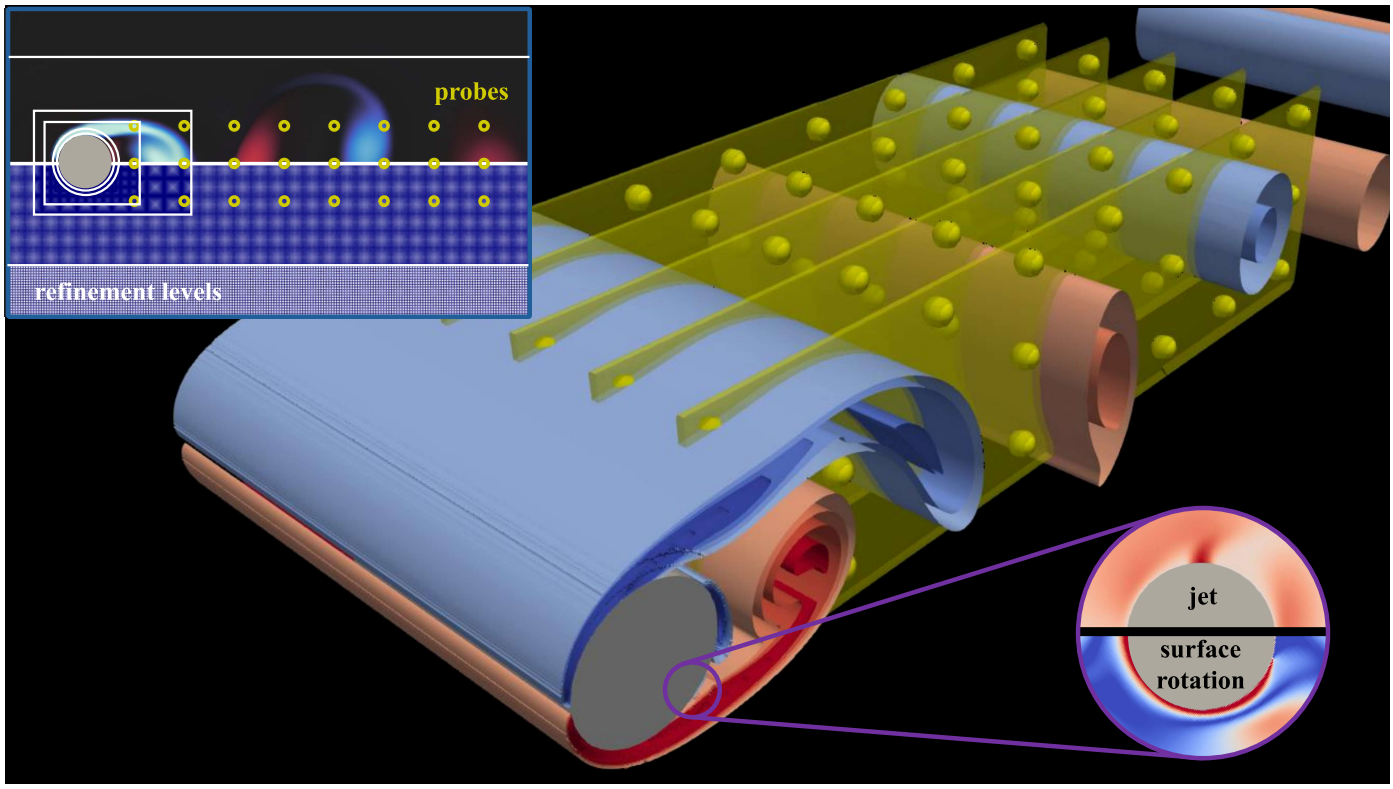

Figure SI 2: Computational domain setup, sensor probe distribution and actuation strategy for the circular cylinder flows: Three-dimensional view showing the probe distribution around the cylinder setup with vortical structures visualized through Q-criterion isosurfaces. The yellow dots in the zoom-in highlight sensor probe locations in a single spanwise plane. Active flow control is either performed using jet actuators positioned at the top and bottom of the cylinder or via surface rotation.

approximately $Re = 18$ [172]. This early onset of instability results from the enhanced wake interactions between the three cylinders. The Hopf bifurcation introduces time-periodic vortex shedding characterized by complex conjugate eigenvalues crossing the stability boundary in the linearized flow equations. The resulting flow exhibits von Kármán-type vortex shedding with dimensionless frequency (Strouhal number) of approximately 0.2. Importantly, this initial instability preserves the mirror-plane symmetry about the configuration centerline, maintaining symmetric shedding patterns until $Re \approx 68$.

At $Re \approx 68$, the fluidic pinball undergoes a second critical transition through a supercritical pitchfork bifurcation [173]. This transition fundamentally alters the wake dynamics by breaking the mirror symmetry preserved during the initial periodic regime. The physical mechanism involves the destabilization of the jet flow between the two downstream cylinders, which becomes susceptible to preferential deflection in either the upward or downward direction. This symmetry-breaking creates two mirror-symmetric stable solutions - one with upward jet deflection and another with downward deflection. The system exhibits bistable behavior, where initial conditions determine which asymmetric state the flow adopts [173]. This represents a qualitative change from the previous symmetric periodic shedding to asymmetric limit cycle dynamics.

Beyond $Re = 104$, the fluidic pinball system develops additional complexity through a secondary Neimark-Sacker bifurcation [174]. This mathematical bifurcation introduces quasi-periodic motion by creating torus-like attractors in the system's phase space. The primary vortex shedding frequency becomes modulated by secondary low-frequency oscillations, establishing multi-frequency dynamics characteristic of weakly chaotic flows. The transition to fully chaotic motion occurs above $Re \approx 130$, following what was identified by [174] as a Newhouse-Ruelle-Takens scenario. In this Re number regime, the quasi-periodic torus structure

becomes unstable and breaks down, leading to aperiodic motion with sensitive dependence on initial conditions - the hallmark of deterministic chaos.

Moreover, three-dimensional effects can introduce additional complexity beyond the two-dimensional analysis. Following the foundational work of [141] on cylinder wake transitions, the fluidic pinball exhibits spanwise instabilities originating from hyperbolic regions within the braid shear layers. These instabilities manifest as short-wavelength structures with characteristic spanwise dimensions comparable to the cylinder diameter. The three-dimensional modes create streamwise vortical structures that interact with the primary cross-stream vortex shedding. Multiple instability modes can coexist and compete, leading to complex scenarios including cooperative destabilization and vortex reconnection phenomena that cannot be captured by two-dimensional computational models.

To summarize, the fluidic pinball's value as a benchmark derives from the combination of geometric simplicity and dynamical richness, making it suitable for testing advanced control strategies while remaining computationally tractable. The independent rotation capability of each cylinder provides multiple actuation inputs, enabling implementation of various flow control mechanisms established in the literature. Classical flow control approaches applicable to this configuration include boundary layer manipulation through surface blowing (analogous to base bleeding techniques), high-amplitude periodic forcing for vortex synchronization, and circulation control via cylinder rotation utilizing Magnus force effects. Additionally, modern approaches such as opposition control and model-based feedback strategies can be evaluated.

Hence, the fluidic pinball presents several fundamental challenges for flow control design. The succession of bifurcations means that control strategies must adapt to qualitatively different flow physics as operating conditions change. Systems designed for the symmetric periodic regime may prove ineffective or counterproductive when applied to the asymmetric bistable regime. The multi-frequency dynamics create coupling between different temporal scales, where actuation at one frequency can influence the entire spectral content of the flow response. This cross-frequency coupling necessitates control architectures that account for broadband nonlinear interactions rather than targeting individual frequency components. Furthermore, the presence of multiple coexisting attractors (in the bistable regime) and hysteresis effects between flow states require control strategies capable of managing state transitions and maintaining desired operating points despite external disturbances.

Extension to three-dimensional control introduces requirements for spanwise-distributed actuation to address the various three-dimensional instability modes. The transition between different three-dimensional regimes creates scenarios where two-dimensional control approaches become inadequate, necessitating fundamentally different control frameworks. Control strategies optimized for one three-dimensional mode may lose effectiveness or become destabilizing when the flow transitions to alternative modal states. This highlights the importance of robust adaptive control approaches that maintain performance across regime boundaries.

Numerical setup and validation. The fluidic pinball configuration consists of three circular cylinders arranged in an equilateral triangle formation, creating a fundamentally different flow topology compared to the single cylinder case. Relative to the pinball center at $(x/D, y/D) = (10.0, 0.0)$ for the 2D simulations and $(8.0, 0.0)$ for the 3D counterpart, the cylinders are positioned with centers at $(x/D, y/D) = (-1.299, 0.0)$, $(0.0, 0.75)$, and $(0.0, -0.75)$, where the upstream cylinder interacts with two downstream cylinders to generate complex wake dynamics and vortex interactions.

The computational setup employs the same domain dimensions and boundary condition treatments as described for the single cylinder benchmark. However, the fluidic pinball operates at lower Reynolds numbers $Re \in [30, 100, 150]$, allowing for a uniform refinement strategy with $n = 12$ across all cases. This refinement level provides adequate resolution to capture the intricate flow features arising from the multi-cylinder interaction, including wake merging, vortex pairing, and the characteristic frequency modulation effects.

The triangular arrangement leads to inherently asymmetric flow patterns, where the downstream cylinders experience different flow conditions due to their positioning relative to the upstream cylinder's wake. This asymmetry manifests in the lift coefficient distributions and requires careful validation against established benchmarks to ensure accurate representation of the flow physics. Validation results are summarized in Table SI 4, demonstrating good agreement with the reference data of Wang et al. [175]. The simulations accurately predict the natural frequencies across all Reynolds numbers. The time-averaged drag coefficients show good agreement (within 2% deviation), while the asymmetric lift coefficients on the downstream cylinders are well-captured, confirming a proper resolution of the wake interaction dynamics. The 3D results maintain similar accuracy levels, validating the three-dimensional extension of the fluidic pinball configuration.

Re	Reference	Dim	f_0	$\overline{C_D}$	$\overline{C_{L,2}}$	$\overline{C_{L,3}}$
30	**Present study**	**2D**	**0.064**	**2.651**	**-0.361**	**0.361**
	Present study	**3D**	**0.061**	**2.597**	**-0.358**	**0.359**
	Wang et al.[175]	2D	0.065	2.656	-0.365	0.365
100	**Present study**	**2D**	**0.088**	**2.904**	**-0.079**	**0.110**
	Present study	**3D**	**0.087**	**2.795**	**-0.0107**	**0.119**
	Wang et al.[175]	2D	0.090	2.874	-0.0745	0.108
150	**Present study**	**2D**	**0.120**	**2.922**	**-0.0441**	**0.0330**
	Present study	**3D**	**0.113**	**2.949**	**-0.0467**	**0.0297**
	Wang et al.[175]	2D	0.118	2.960	-0.043	0.0319

Table SI 4: Validation of the fluidic pinball flow results at different Reynolds numbers. Dim: Dimensionality (2D/3D); f_0: Natural frequency; $\overline{C_D}$: Time-averaged drag coefficient; $\overline{C_{L,i}}$: Time-averaged lift coefficient for cylinder i

Environment setup. The fluidic pinball environment features a multi-cylinder flow control setup comprising three circular cylinders positioned in an equilateral triangular arrangement. The primary reward function r for the fluidic pinball targets total drag reduction across all three cylinders and is defined as $r = -|\sum_{i=1}^{3} C_{D,i}| - \omega|\sum_{i=1}^{3} C_{L,i}|$, where $C_{D,i}$ represents the normalized drag coefficient of cylinder i and $\omega = 1.0$ serves as the scaling parameter consistently applied across all Reynolds numbers to avoid control policies that introduce a significant lift generation. Users can readily adapt this reward structure to explore different objectives including thrust control or independent cylinder optimization.

The control mechanism employs surface rotation actuation, where the RL agent manages the angular velocity of each cylinder individually (see Fig. SI 3). This creates a three-dimensional continuous action space with components bounded within $[-1, 1]$, facilitating exploration of cooperative multi-cylinder control approaches. Although jet-based actuation is not part of the standard setup, the framework design accommodates user integration of blowing/suction mechanisms for extended research applications or comparative analysis.

The framework encompasses both two-dimensional and three-dimensional fluidic pinball flow scenarios across various Reynolds numbers ($Re = 30, 75, 100, 150$), covering flow regimes from steady symmetric configurations through periodic vortex shedding to bistable asymmetric behavior and quasi-periodic dynamics. This Reynolds number spectrum encompasses the distinctive bifurcation characteristics of the fluidic pinball, including the initial Hopf bifurcation at $Re \approx 18$, the symmetry-breaking pitchfork bifurcation at $Re \approx 68$, and progressive transitions toward chaotic flow patterns. The setup utilizes a single-agent configuration, where one agent manages all three cylinder rotations concurrently, promoting research into unified control methodologies for multi-body flow applications.

The observation space utilizes the same adaptable probe-based methodology as established in other environments, featuring user-definable point measurements of velocity components, pressure, density, and body forces positioned throughout the computational domain. Training episodes encompass 200 control actions, equivalent to roughly 10 vortex shedding cycles, offering an adequate temporal scope for agents to perceive and react to the intricate multi-frequency behavior inherent in the fluidic pinball. The CFD-RL coupling utilizes Reynolds number-dependent action intervals tailored to match evolving flow characteristics: 230, 190, 170, and 100 CFD timesteps per control action for 2D scenarios at $Re = 30, 75, 100, 150$ respectively, with corresponding 3D cases using 250, 235, 225, and 200 timesteps. These intervals adapt to the changing temporal scales and flow complexity encountered as the system progresses through distinct dynamical states, maintaining suitable control responsiveness while preserving computational tractability. All outlined

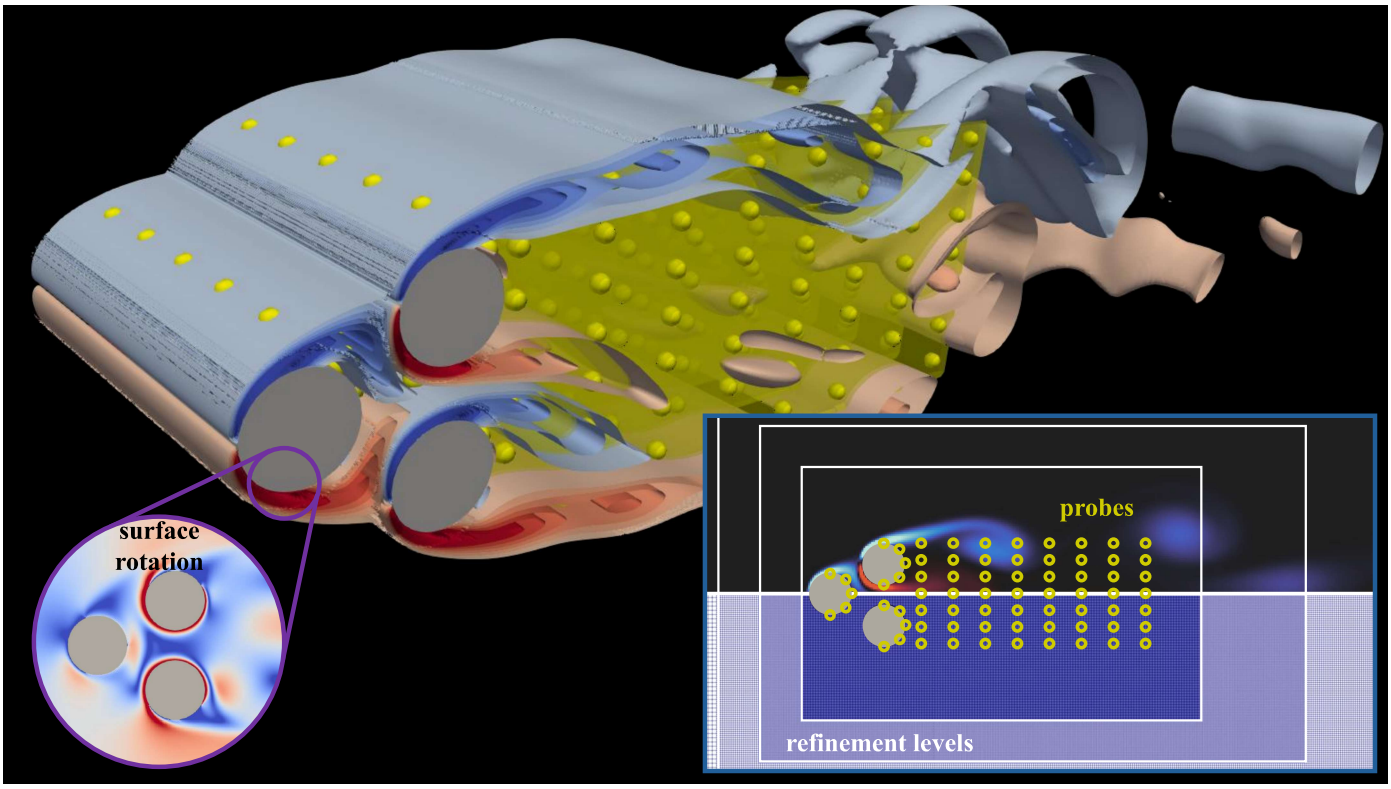

Figure SI 3: Computational domain setup, sensor probe distribution and actuation strategy for the fluidic pinball environment: Three-dimensional view showing the probe distribution around the fluidic pinball with vortical structures visualized through Q-criterion isosurfaces. The yellow dots in the zoom-in highlight sensor probe locations in a single spanwise plane. Active flow control is performed by individual surface rotation of each cylinder.

specifications serve as baseline configurations that users can readily adjust to meet particular research needs, encompassing modified reward structures, alternative control mechanisms, customized sensor layouts, and tailored temporal coupling settings.

5.3 Open cavity flow

Characteristic physics. The flow around open cavities in the Reynolds number range 4,000-7,500 based on the cavity height and the freestream velocity exhibits complex transitional behavior characterized by distinct regime evolutions and sophisticated three-dimensional instabilities that have profound implications for both fundamental aeroacoustics research and practical flow control applications. This transitional range encompasses the evolution from predominantly steady two-dimensional recirculation patterns to complex three-dimensional unsteady dynamics with strong acoustic coupling, presenting unique challenges for modeling, prediction, and control that distinguish it from both lower and higher Reynolds number regimes.

Within this Reynolds number range, the flow physics is governed by several critical transitions that fundamentally alter the cavity dynamics and acoustic characteristics. At Reynolds numbers around 4,000-5,000, the initially stable two-dimensional recirculation undergoes its first significant three-dimensional instability through centrifugal mechanisms arising from curved streamlines in the recirculation region [176, 177]. This transition manifests as weakly unsteady flow behavior with emerging spanwise variations that mark the onset of complex temporal dynamics. As the Reynolds number continues to increase, a second critical transition occurs around Re = 7,000-8,000, where strong temporal instabilities emerge through Hopf bifurcation characteristics, introducing multiple competing temporal modes and mode-switching phenomena