Particles 2025, Barcelona

OCEAN COLLEGE, ZHEJIANG UNIVERSITY

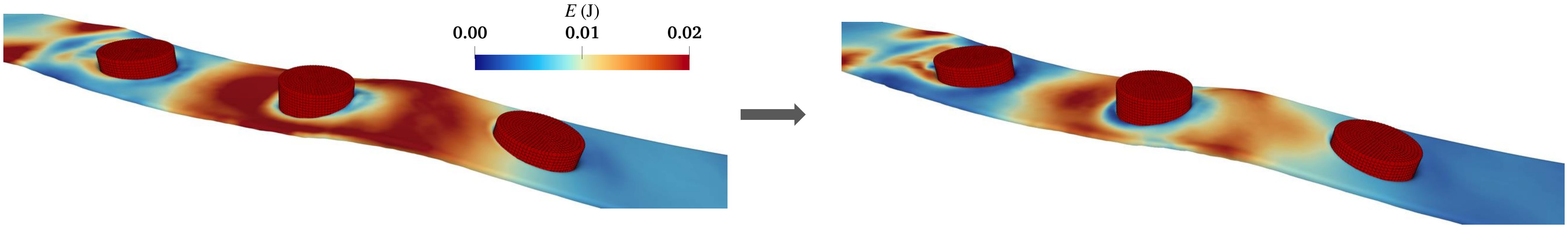
Coupling SPH with a Multi agent DRL framework for active flow control

Yi Zhan[1,2], Min Luo[1], Alejandro J. C. Crespo[2], Iván Martínez-Estévez[2], Abbas Khayyer[3]

[1] Ocean College, Zhejiang University [2] EPhysLab, University of Vigo

[3] Department of Civil and Earth Resources Engineering, Kyoto University

22nd October 2025



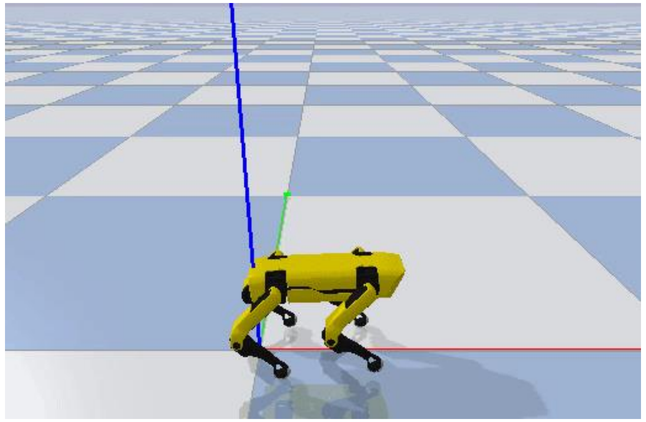
Contents

1.1 Motivation : Giving marine structures intelligence

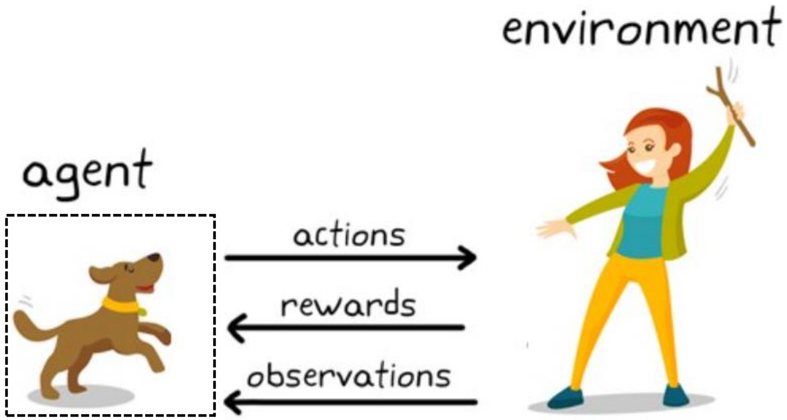
➢ Deep Reinforcement Learning (DRL)

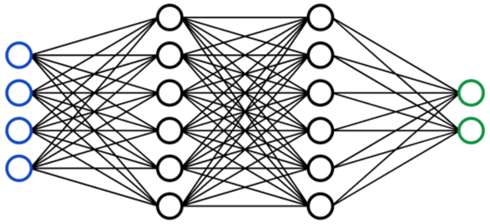
Embodied Intelligence

Quadruped Robots



AlphaGo

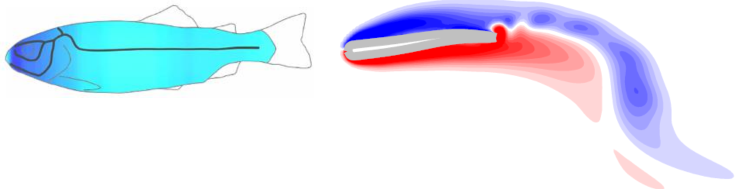




- Marine vehicles

Collision Avoidance ?

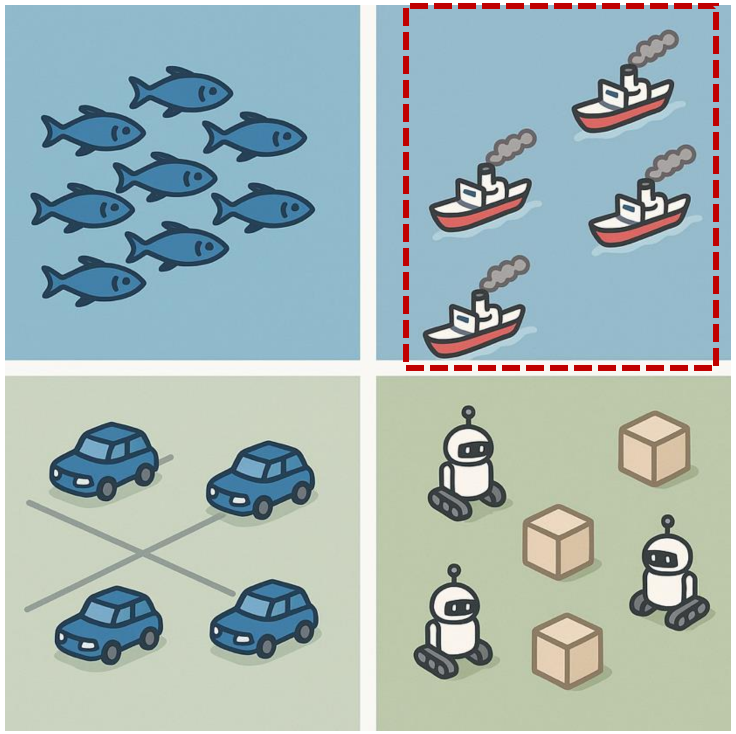
Efficient swimming ?



- Make decisions guided by neuron network
- Agent takes suitable action to maximize reward in a particular situation

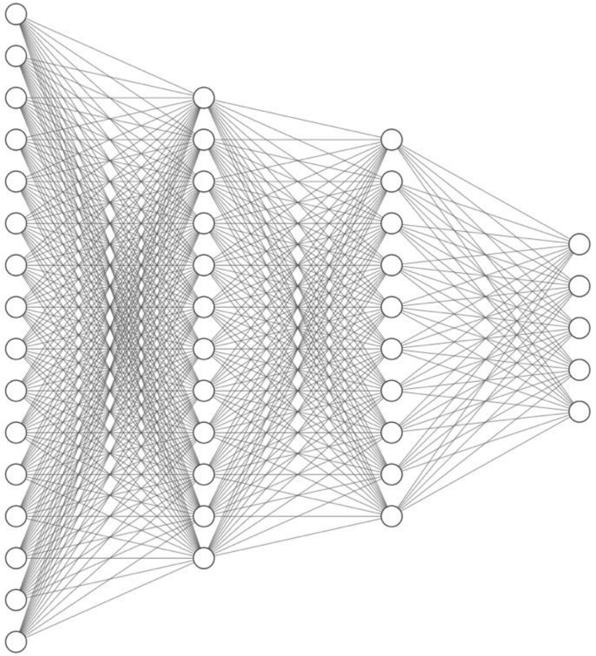
1.1 Motivation : Multi-agent decision-making

➢ Multi agents



● Multi-agent decision-making scenarios are common in engineering

➢ Single NN, multi output

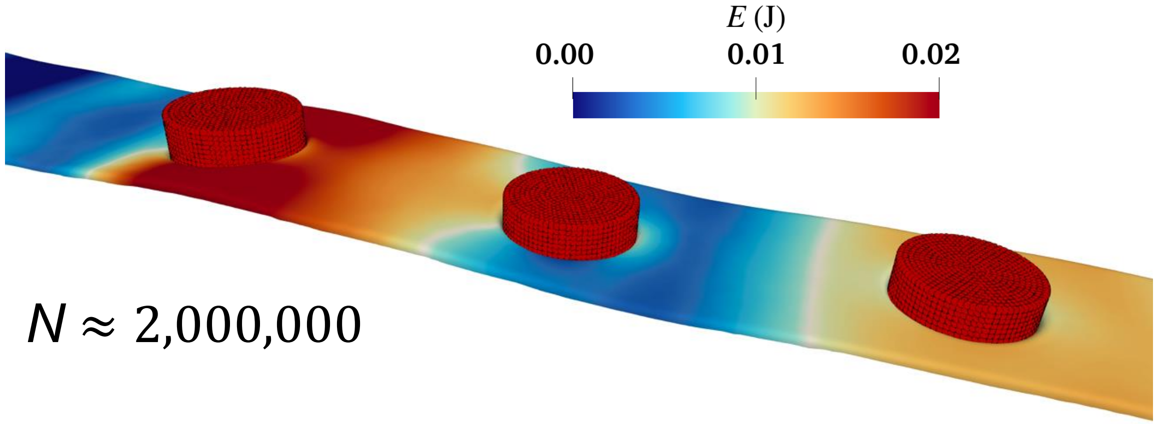


✗ More observations make single-agent input too large to train

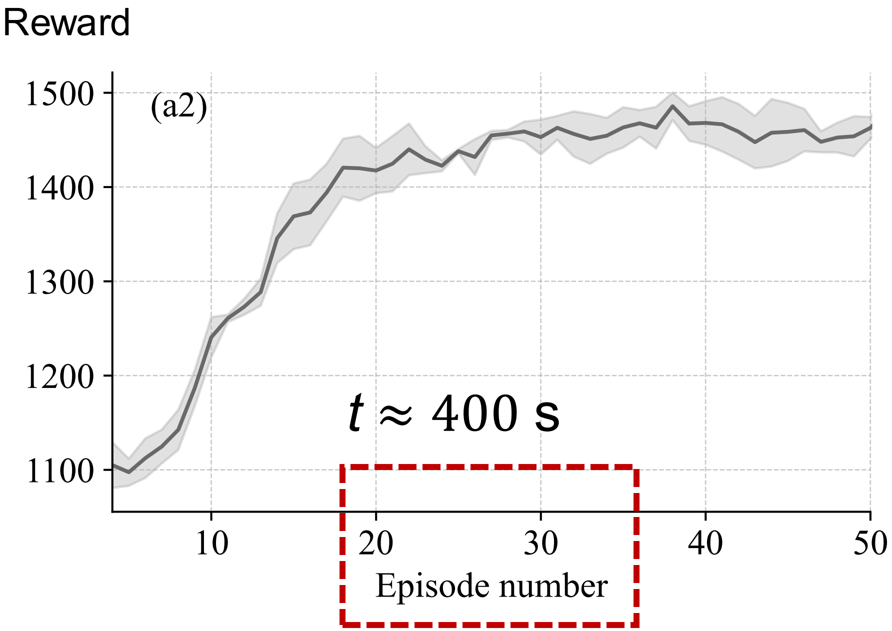
✗ Real-time communication to obtain all agents' observations is infeasible

Motivation : 3D simulation

➢ Increase energy capture of Wave Energy Converter



$N \approx 2{,}000{,}000$



$t \approx 400$ s

- The number of particles in 3D simulations typically reaches the order of millions
- At least 20 episodes are required for the training to reach stability

1.2 Research gap and solutions

- Active control in ocean engineering → Coupling SPH and DRL
- Multi-agent decision-making scenarios → Multi-agent reinforcement learning
- Long term 3D simulations → Parallelism in GPU to accelerate SPH and DRL

Contents

2.1 SPH solver

δ-SPH model in DualSPHysics

$$\frac{D\rho_a}{Dt} = -\rho_a \sum_b \boldsymbol{u}_{ab} \cdot \nabla_a W_{ab} V_b + D_a, \quad \frac{D\boldsymbol{u}_a}{Dt} = -\sum_b m_b \left(\frac{p_b + p_a}{\rho_b \rho_a} + \Pi_{ab} \right) \nabla_a W_{ab} + \boldsymbol{g}$$

DualSPHysics

Further enhancements corresponding to four aspects

i) Minimum energy dissipation ← Riemann stabilization term instead of artificial viscosity

ii) Consistent particle shifting at and in the vicinity of free surface ← OPS

iii) Enhanced resolution of the continuity equation ← VEM and VCS

iv) Effective cleaning of velocity-divergence errors ← Combination of VEM and HPDC

δR-SPH-OPS-VCS-VEM-HPDC in DualSPHysics+ [1]

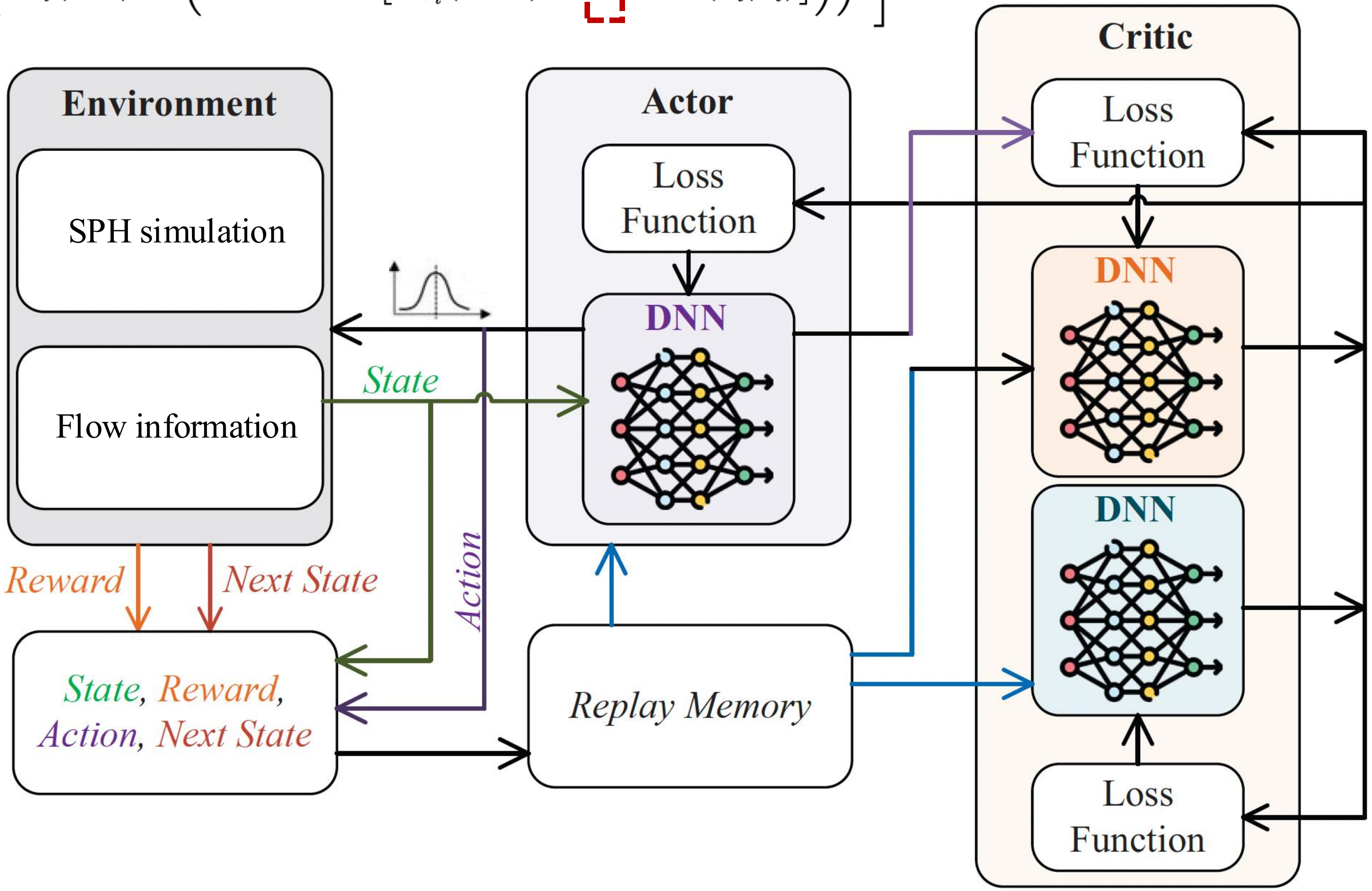
$$\frac{D\rho_a}{Dt} = -\rho_a \sum_b \boldsymbol{u}_{ab} \cdot \nabla_a W_{ab} V_b + D_a, \quad \frac{D\boldsymbol{u}_a}{Dt} = -2\sum_b m_b \left(\frac{p^*}{\rho_a \rho_b} \right) \nabla_a W_{ab} + \boldsymbol{g} + \boldsymbol{a}_a^{\mathrm{VEM}} - \nabla \psi_a, \quad p^* = \frac{1}{2} F(p_L, p_R) + \frac{1}{2} \phi \bar{\rho} (u_L - u_R), \quad \boldsymbol{r}_a' = \boldsymbol{r}_a + \delta \boldsymbol{r}_a^{\mathrm{OPS}} + \delta \boldsymbol{r}_a^{\mathrm{VCS}}$$

[1] Zhan, Yi, Min Luo, and Abbas Khayyer. "DualSPHysics+: An enhanced DualSPHysics with improvements in accuracy, energy conservation and resolution of the continuity equation." *Computer Physics Communications* (2024): 109389.

2.2 DRL Algorithm: Soft actor-critic

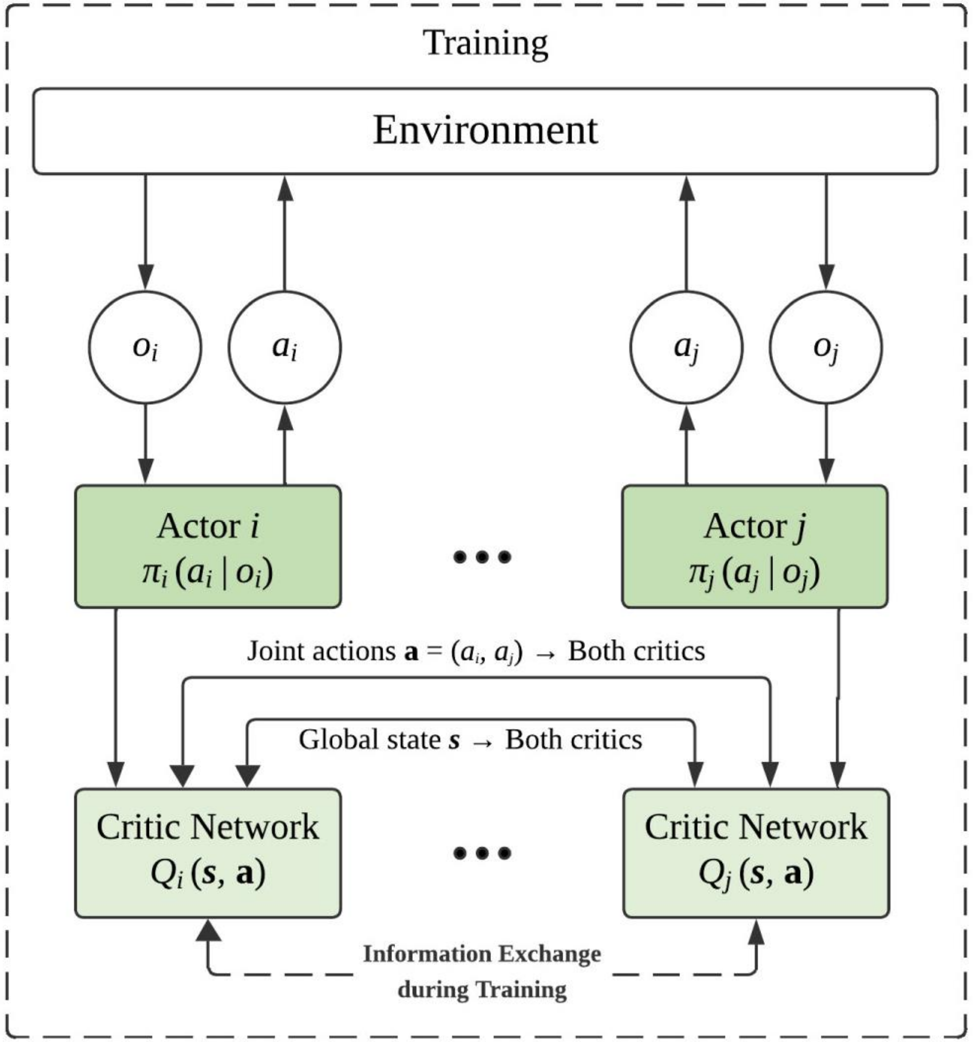
$$\boxed{J_Q(\theta_i)} = \mathbb{E}_{(s,a,r,s')\sim\mathcal{D}}\left[\left(Q_{\theta_i}(s,a) - \left(r + \gamma\,\mathbb{E}_{a'\sim\pi}\left[Q_{\bar{\theta}_i}(s',a') - \boxed{\alpha}\log\pi_i(a'_i|s'_i)\right]\right)\right)^2\right]$$

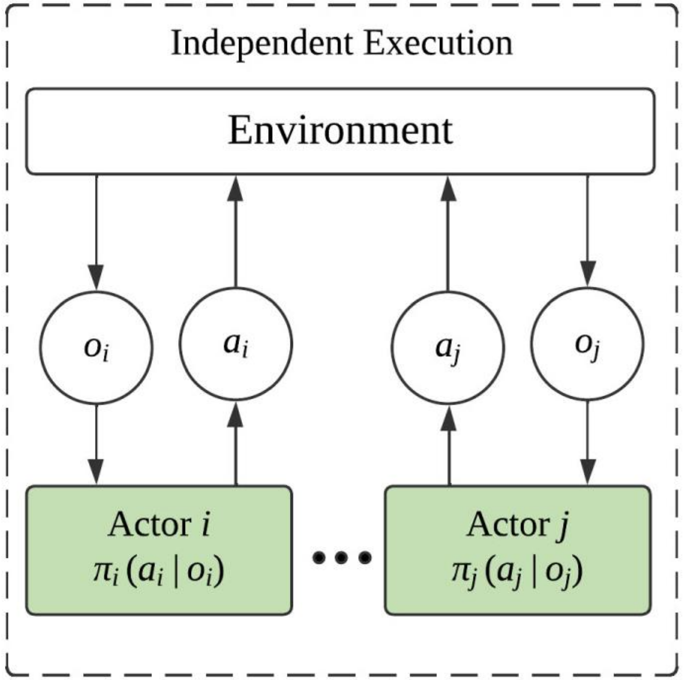
- **Actor–Critic Framework:** Actor → learns optimal policy; Critic → evaluates the action by Actor NN
- **Entropy-Regularized Learning:** Promotes exploration by maximizing both reward and entropy
- **Automatic Temperature Tuning:** Dynamically balances exploration and exploitation



[1] Haarnoja, Tuomas, et al. "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor." *International conference on machine learning*. Pmlr, 2018.
[2] Du, H., et al. "Enabling AI-generated content (AIGC) services in wireless edge networks. arXiv 2023." *arXiv preprint arXiv:2301.03220*.

2.2 MADRL: Centralized training decentralized execution

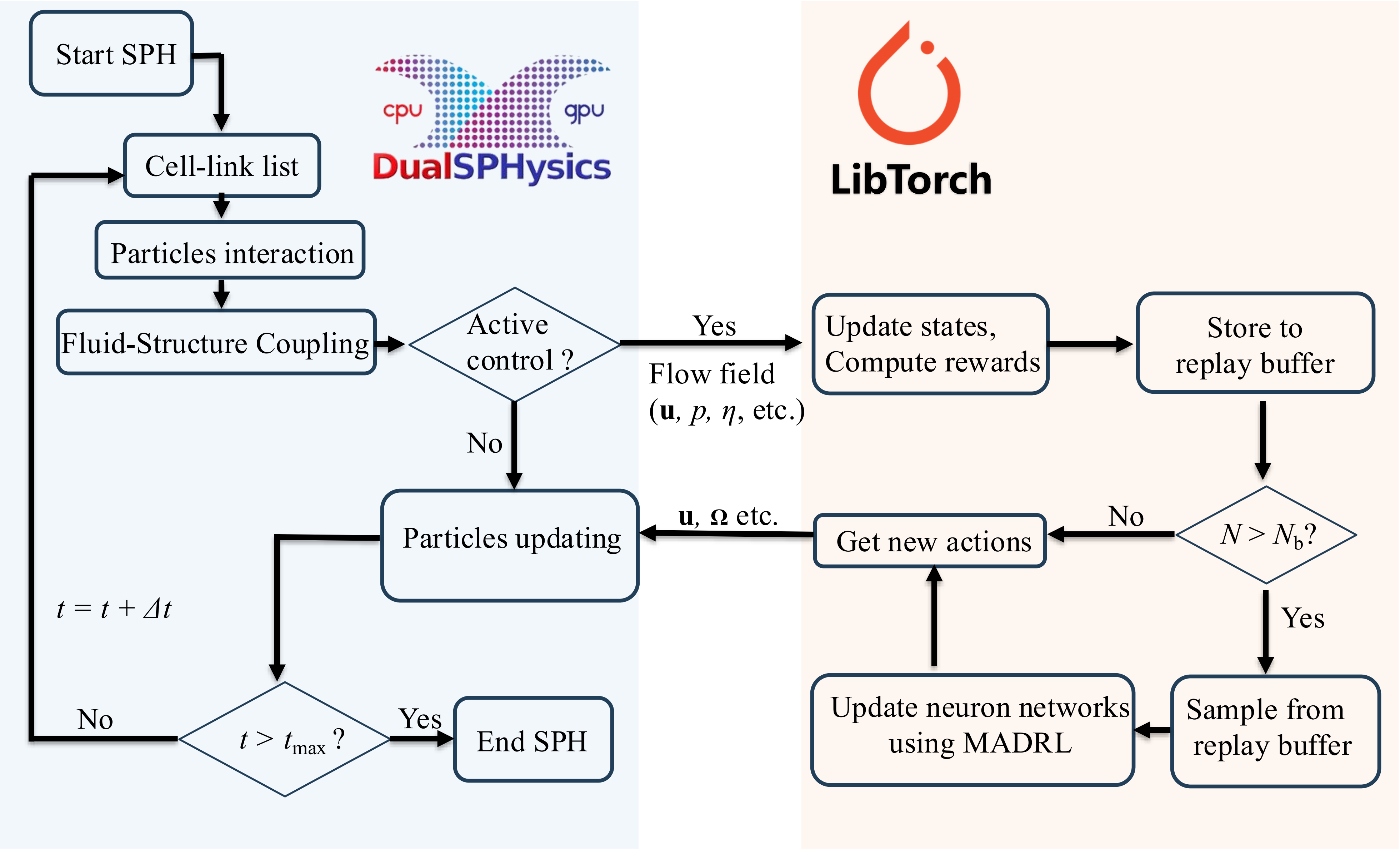




- **Centralized Training:** Critic network → Learn with global information for cooperative policy optimization
- **Decentralized Execution:** Actor network → Get actions independently using local observations

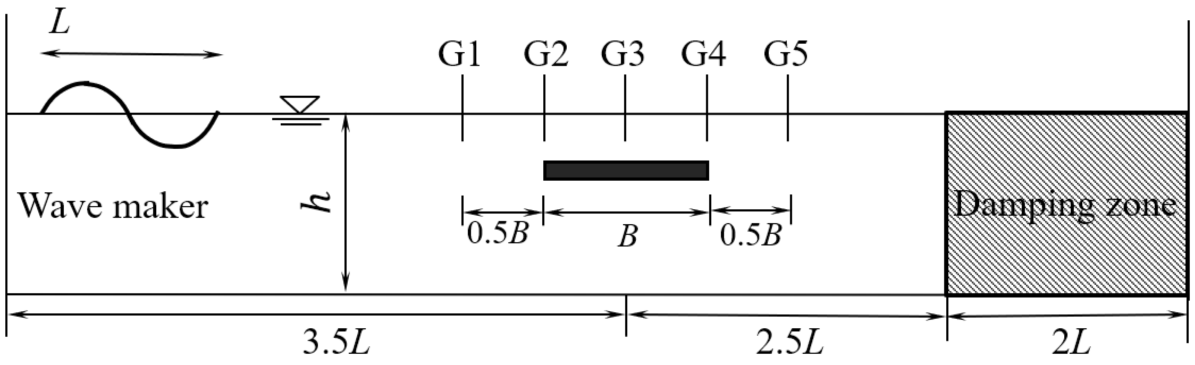
2.4 Coupling SPH with DRL

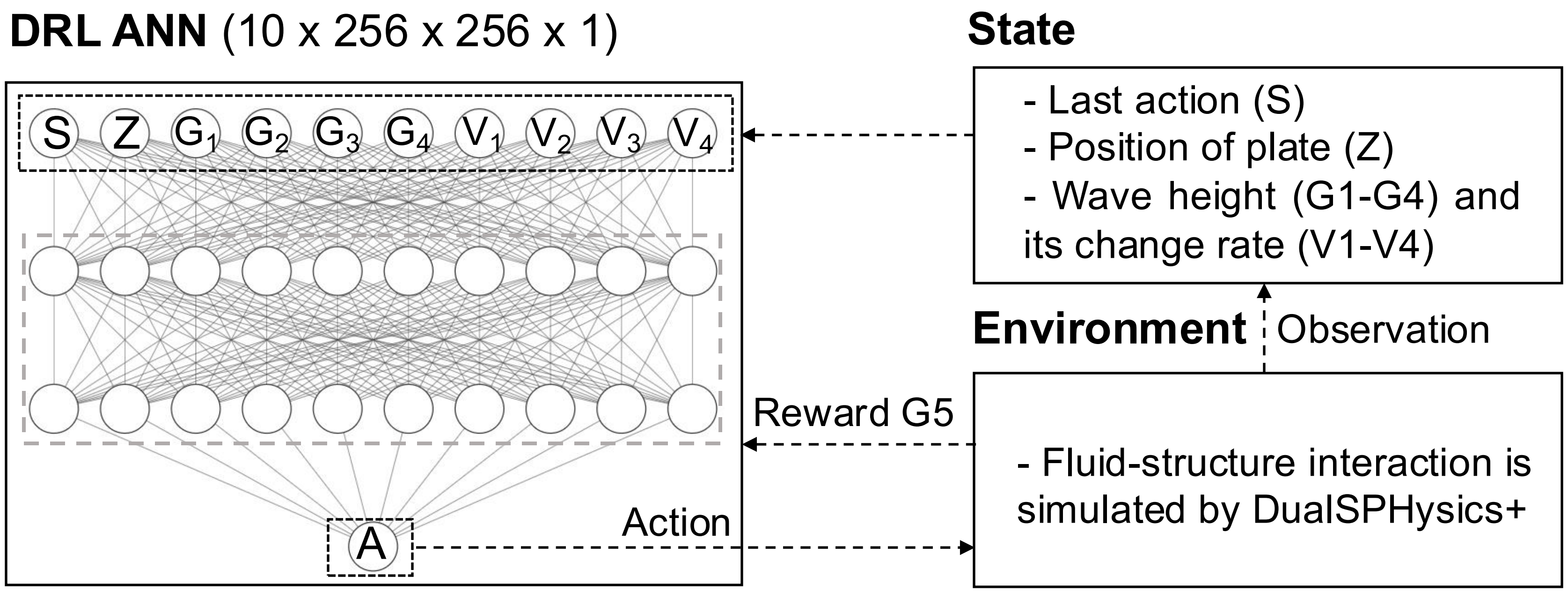
- Libtorch is linked to DualSPHysics+ as a dynamic library
- All codes are in C++ & CUDA and can be parallelized using GPU



Contents

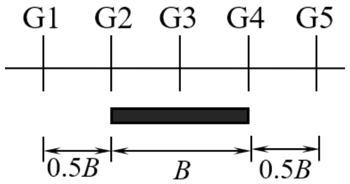
3.1 Active controlled plate breakwater : overview





- The SPH solver acts as environment during training
- The reinforcement learning model guides movement of plate

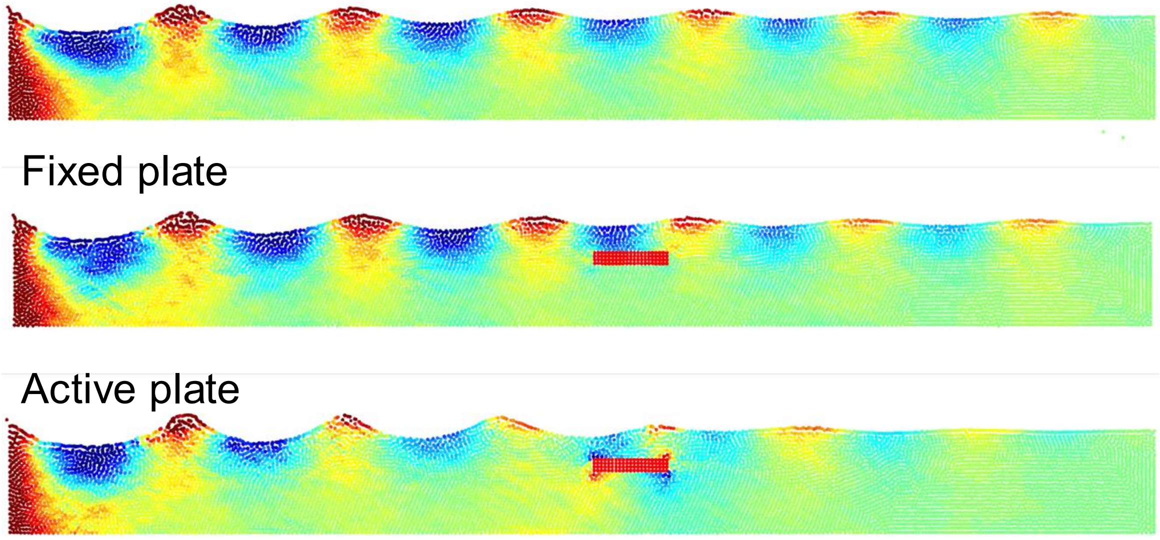
3.1 Active controlled plate breakwater : wave elevation

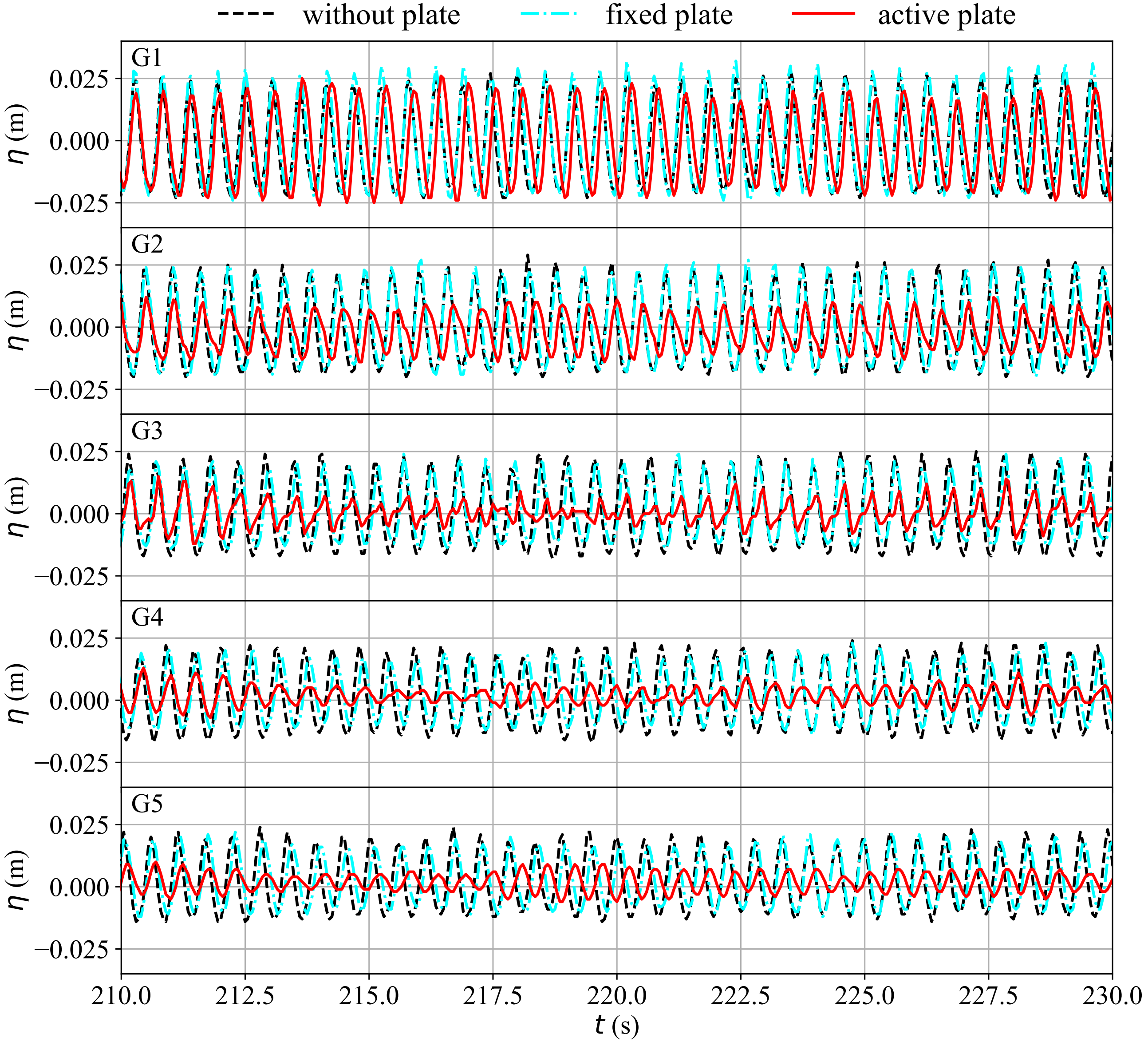


Without plate

Fixed plate

Active plate

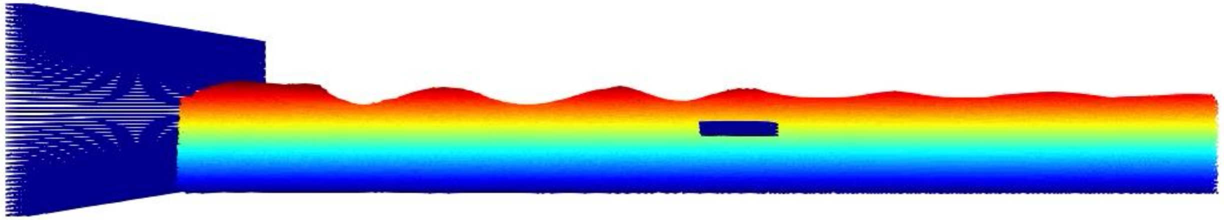


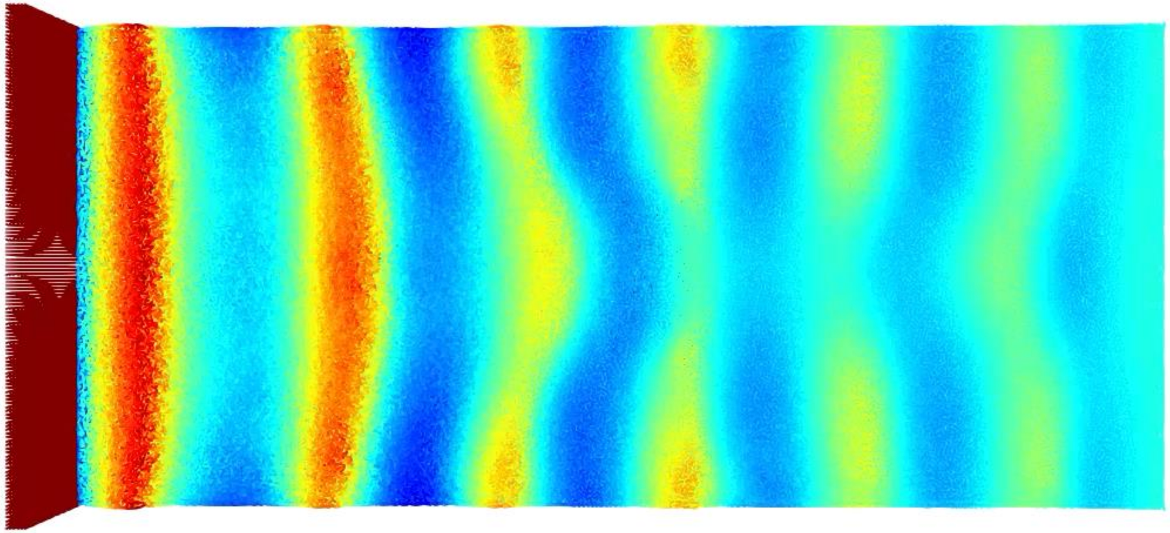


- Active control plate significantly reduces wave height
- Pulling the water body downward when the wave crest arrives

3.1 Active controlled plate breakwater : 3D simulation

3D simulation : Wave elevation field

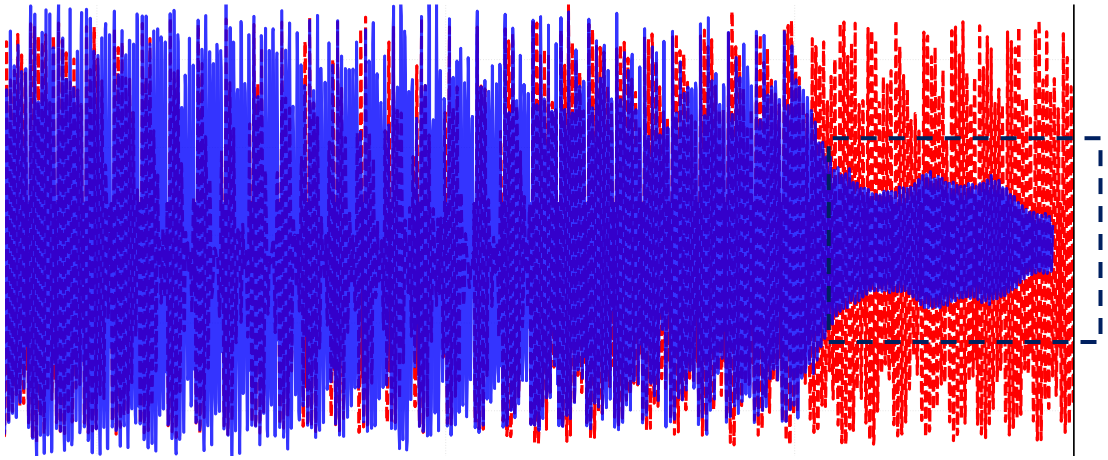




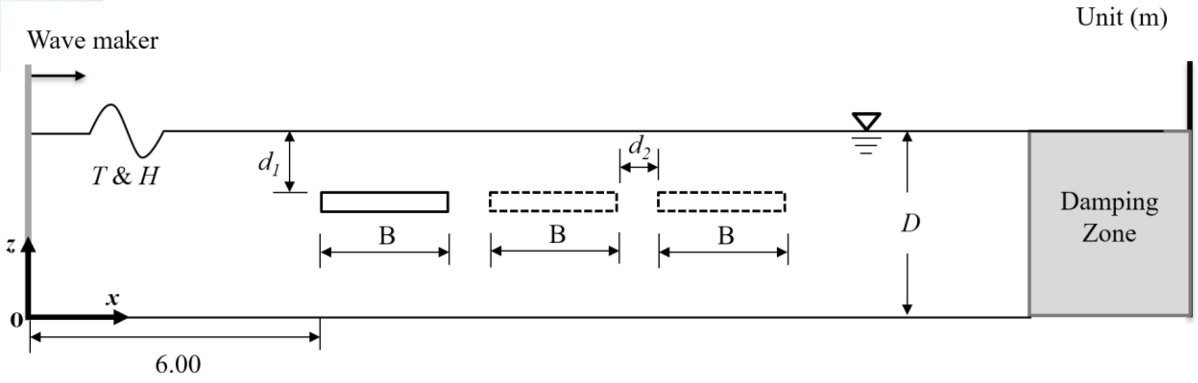
- Device: Nvidia RTX 4090 && AMD EPYC 9754
- N ≈ 2 mil. Simulation : 300 s

Total runtime 500⁺ h (cpu) → 7 h (gpu)

Wave elevation w/ and w/o active control



3.1 Active controlled plate breakwater : Multi agents



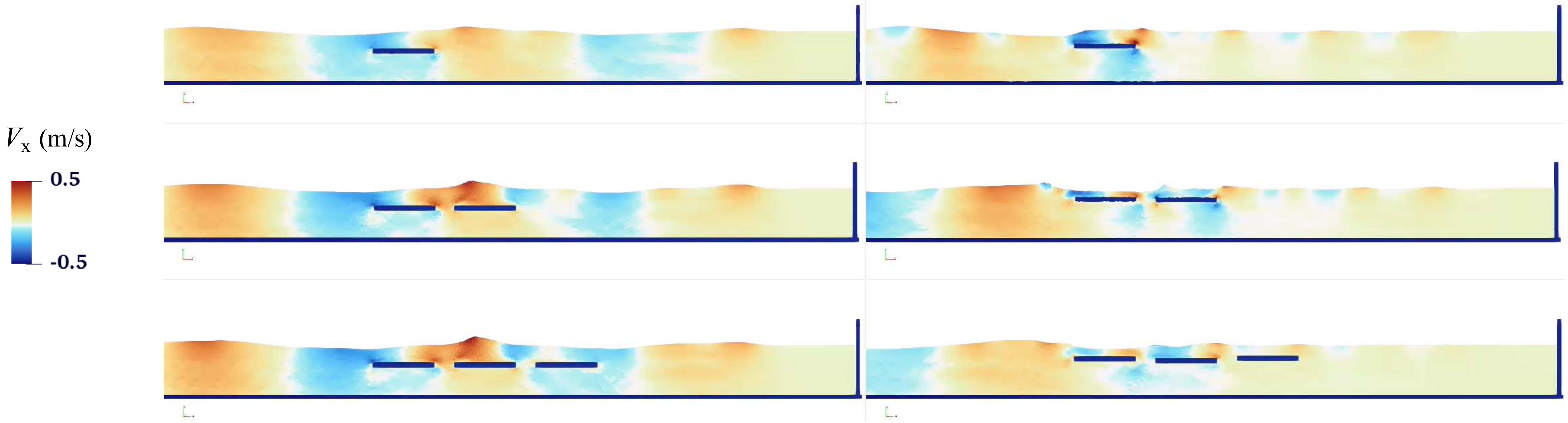
Relative motion between the plates

→ Wave reflection and interaction

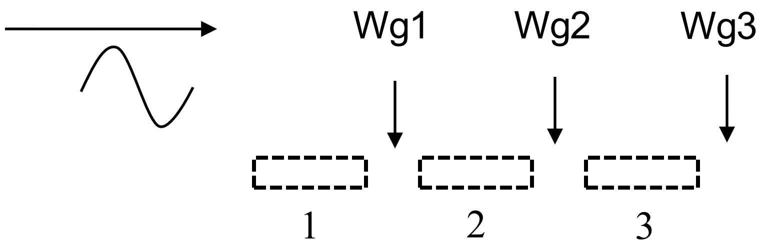
→ Dissipate wave energy

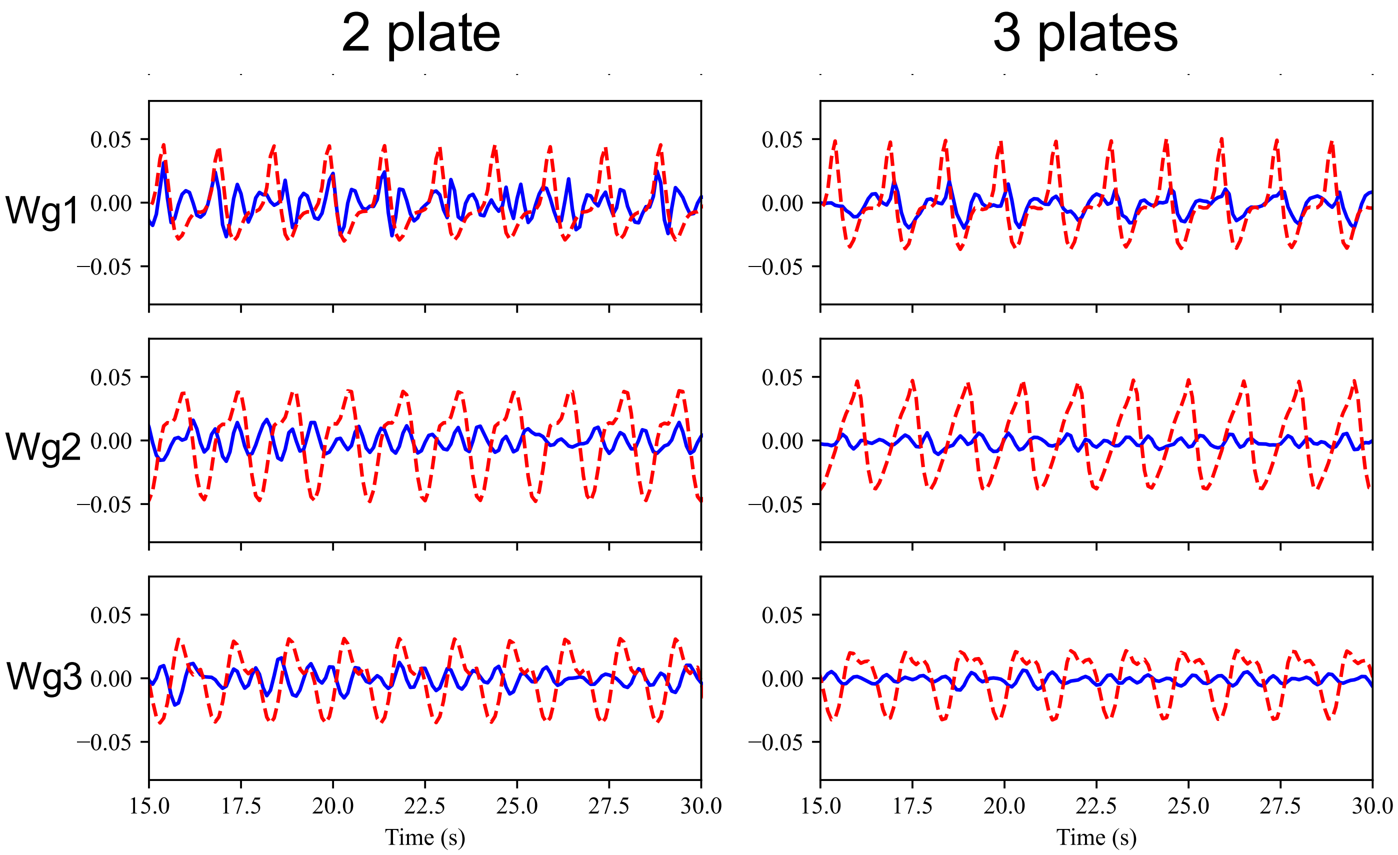
Fixed plate

Moving plate



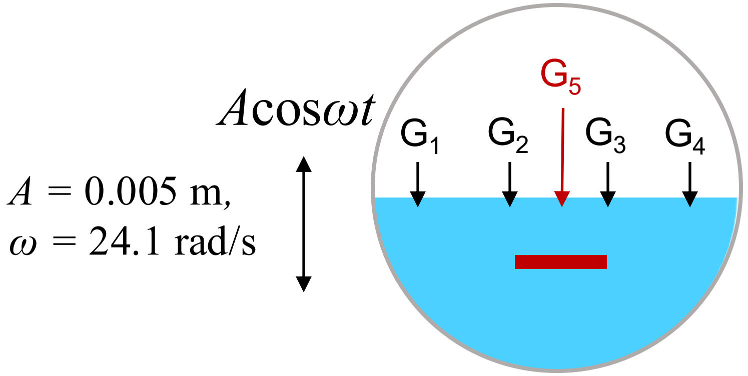
3.1 Active controlled plate breakwater : Multi agents





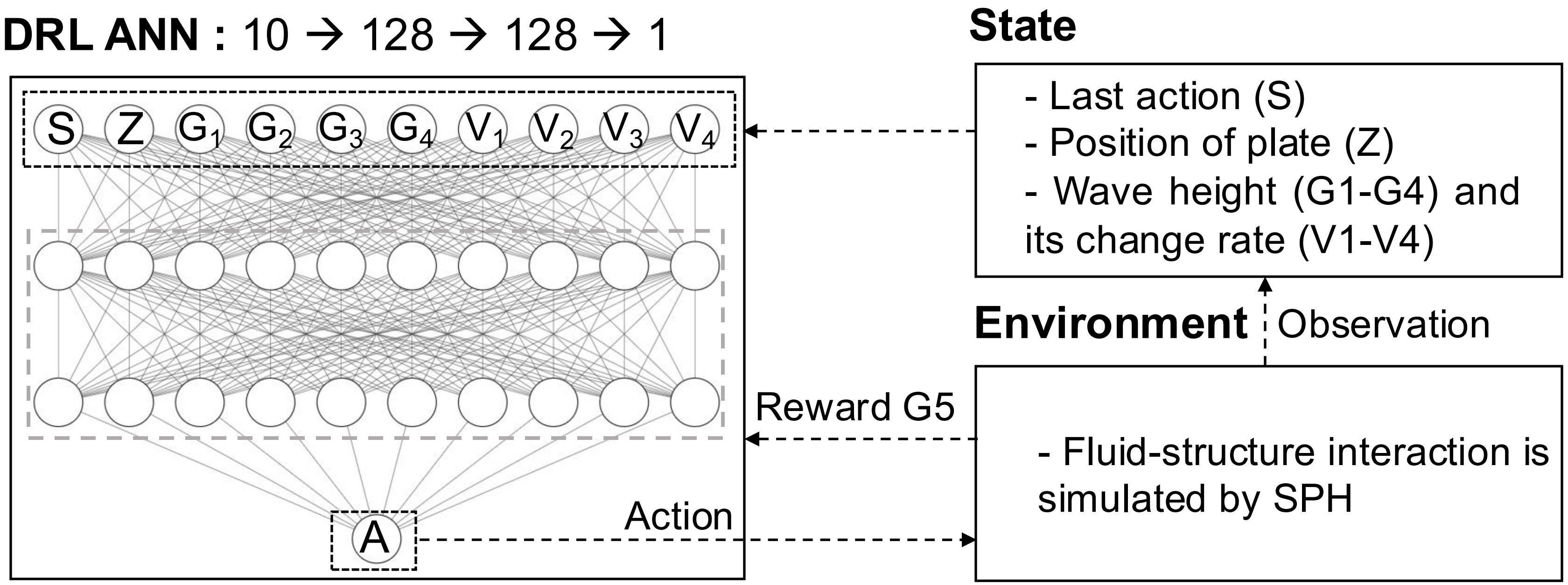
- Active control plate effectively reduces the wave height
- The wave attenuation effect increases with the number of plates

3.2 Sloshing flow with active controlled baffle : overview



- Training episodes : 90
- Evaluation episodes : 10
- Episode duration: 4 s

DRL ANN : 10 → 128 → 128 → 1



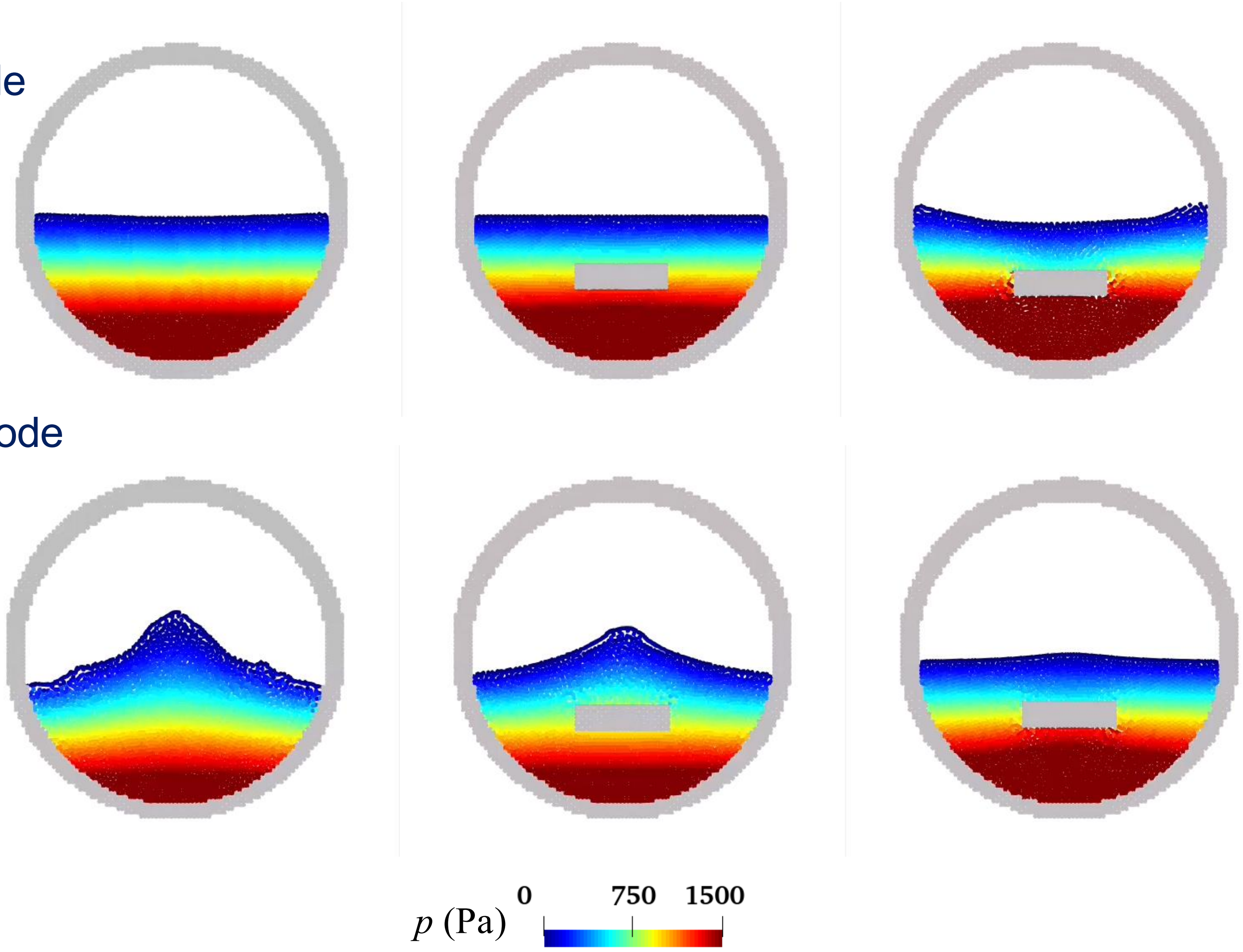
- SPH solver acts as environment during training
- DRL model guides movement of baffle to suppress sloshing

Sloshing flow with active controlled baffle : training process

Without plate | With fixed plate | With active control plate

Training episode

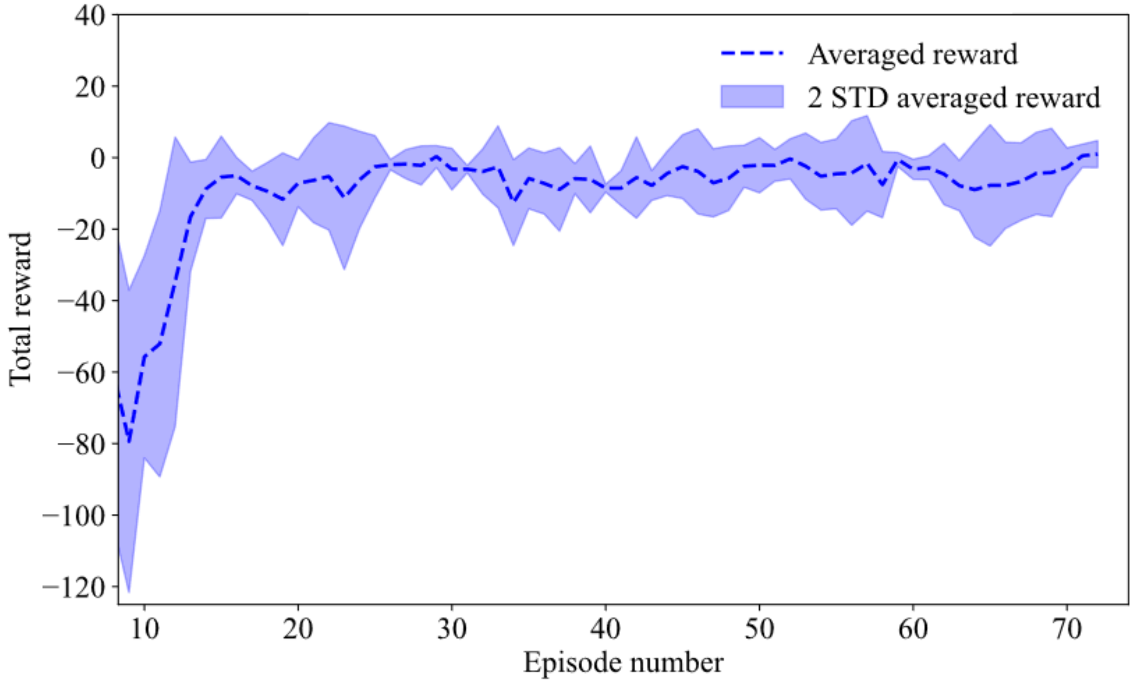
Evaluation episode

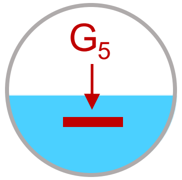


- Agent makes many attempts based on the neural network output
- DRL model effectively guides the agent's motion

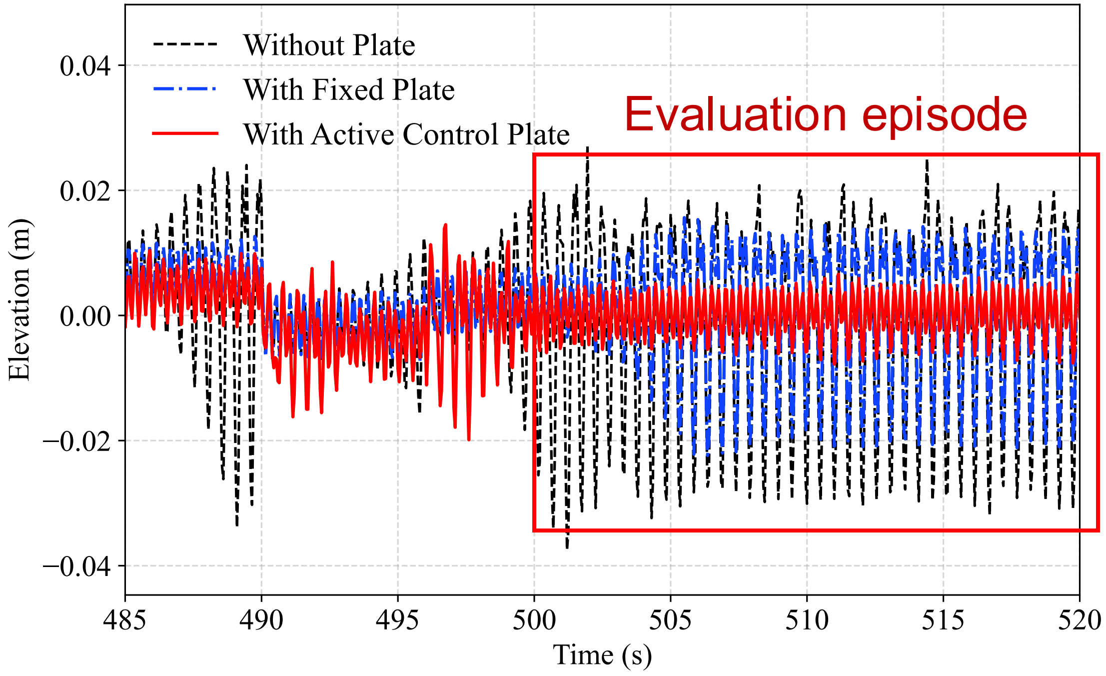
3.2 Sloshing flow with active control baffle : wave elevation

Total reward



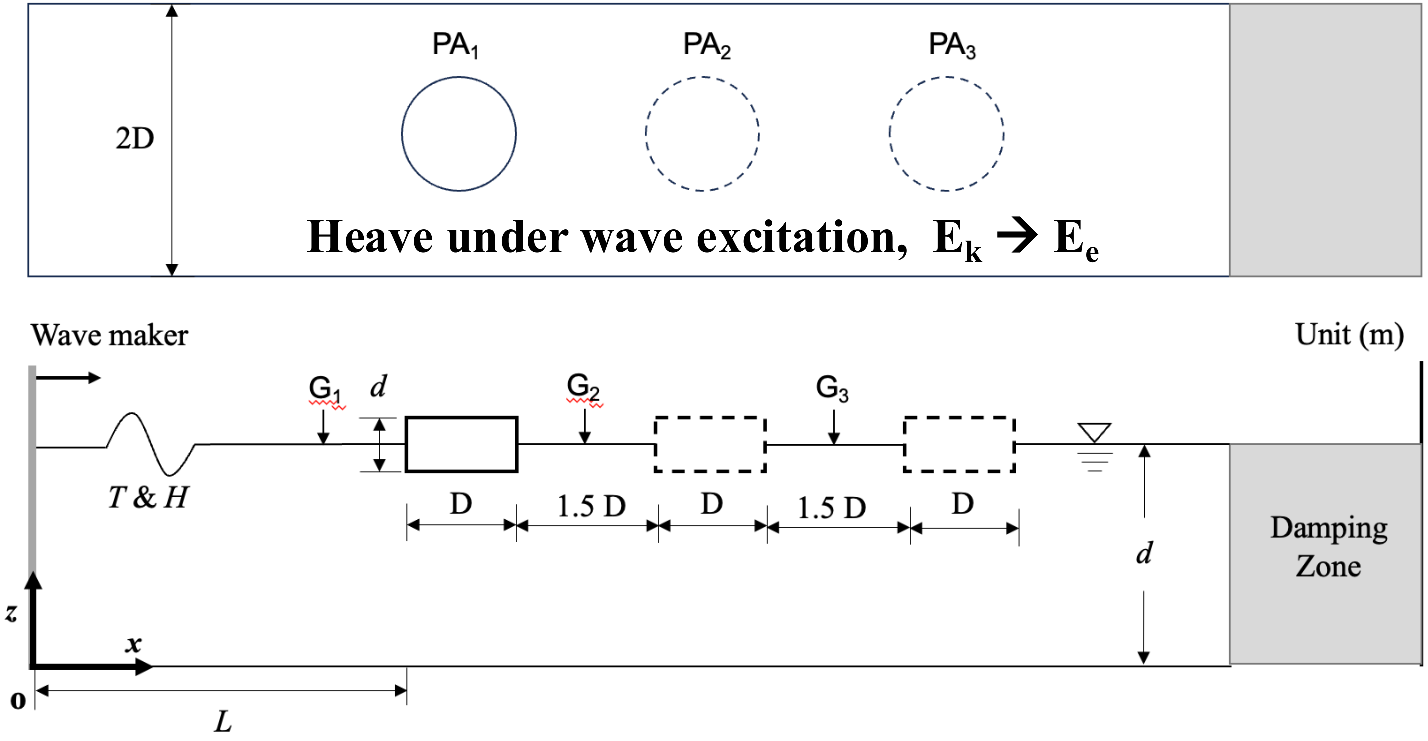


Elevation at G5



- The average reward keeps convergence and tuning slightly after 20 episodes
- 85% and 65% reduction with active control (vs no plate & fixed plate)

Point absorber wave energy converter : overview



$$M\frac{d\mathbf{V}}{dt} = \sum_j m_j \mathbf{f}_j + \mathbf{D}_t \quad \mathbf{D}_t = -k_p \mathbf{v}$$

Power output $\boxed{P_{abs} = k_p v_z^2(t)}$

negatively related

Output : o_i, $o_i \in [-1,1]$ $\quad k_{p,i} = k_{base} + o_i \Delta k_{max}$

Input : $G_{i,1} - G_{i,3}$, $d(G_{i,1} - G_{i,3})/dt$, $v_{z,i}$, z_i, a_i^{n-1}

Reward : $r_i = (1-\gamma_p)P_{out,i} + \gamma_p \frac{1}{N}\sum_{j=1}^{N} P_{out,j} \quad P_{out,i} = k_{p,i} v_i^2$

WEC Energy output

Encourage collaborating

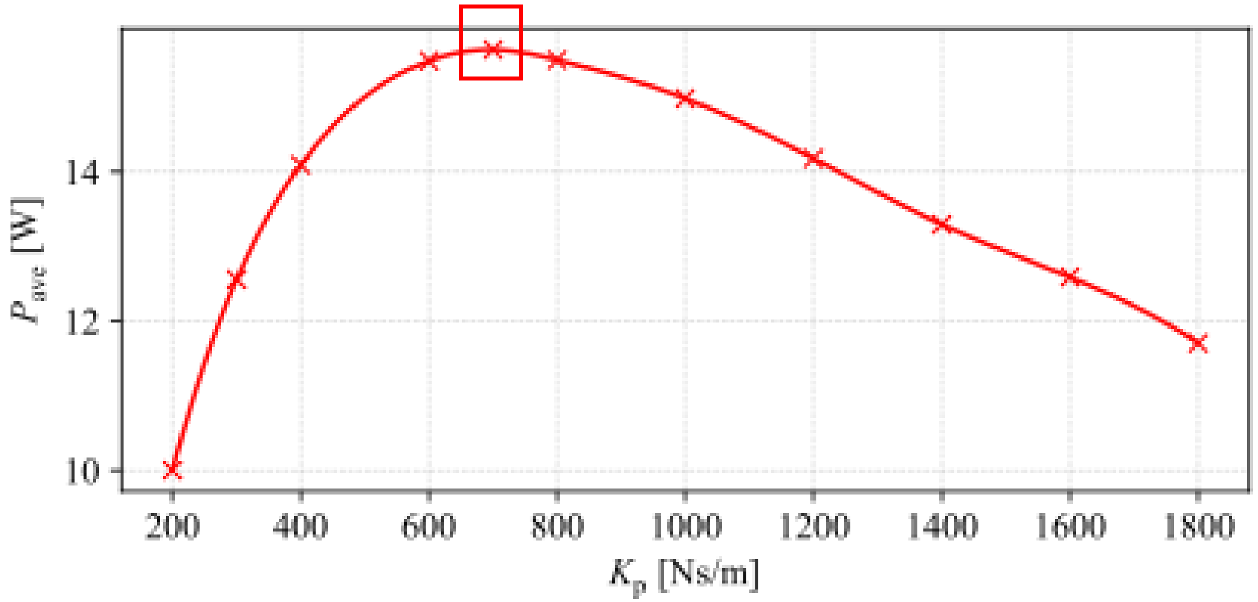
Make the damping parameter K_p adaptive to cope with different wave conditions → increase power output

➢ Determine damping coefficient k_{base}

T = 1.5 s, d = 0.5 m $P_{\text{ave}} = \frac{1}{T}\int_{t_0}^{t_0+T} P_{\text{abs}}(t)dt, \quad P_{\text{abs}} = k_{\text{p}} v_z^2(t)$

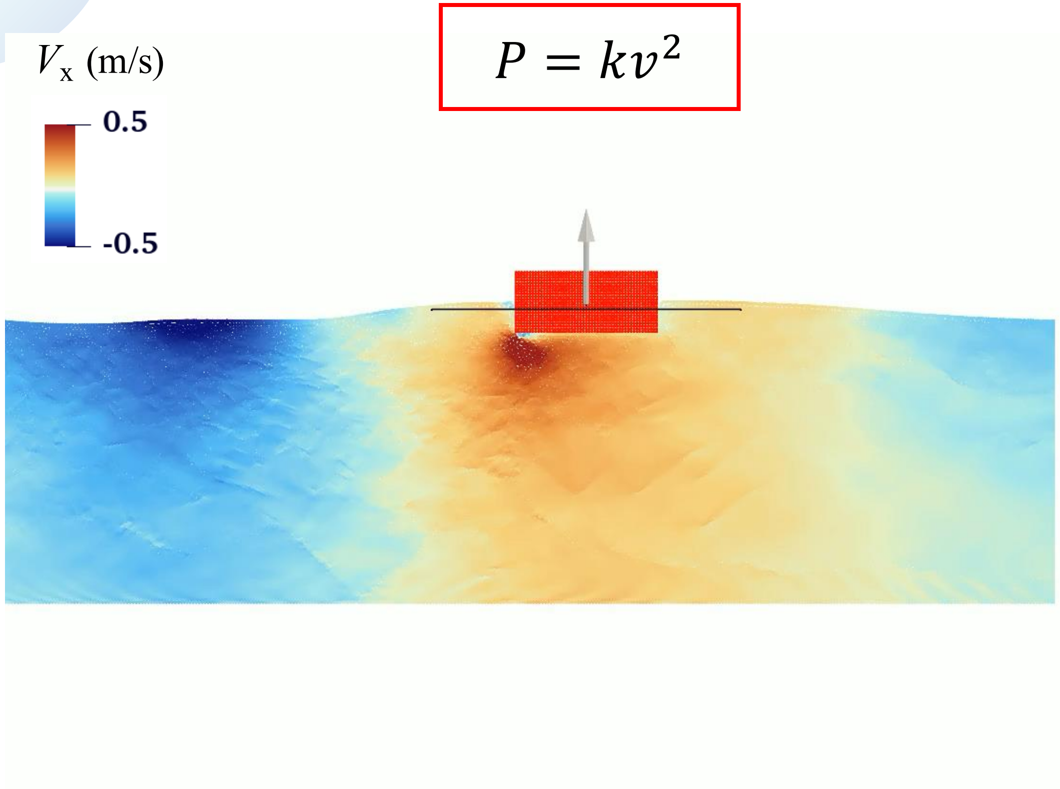
Output : o_i, $o_i \in [-1,1]$

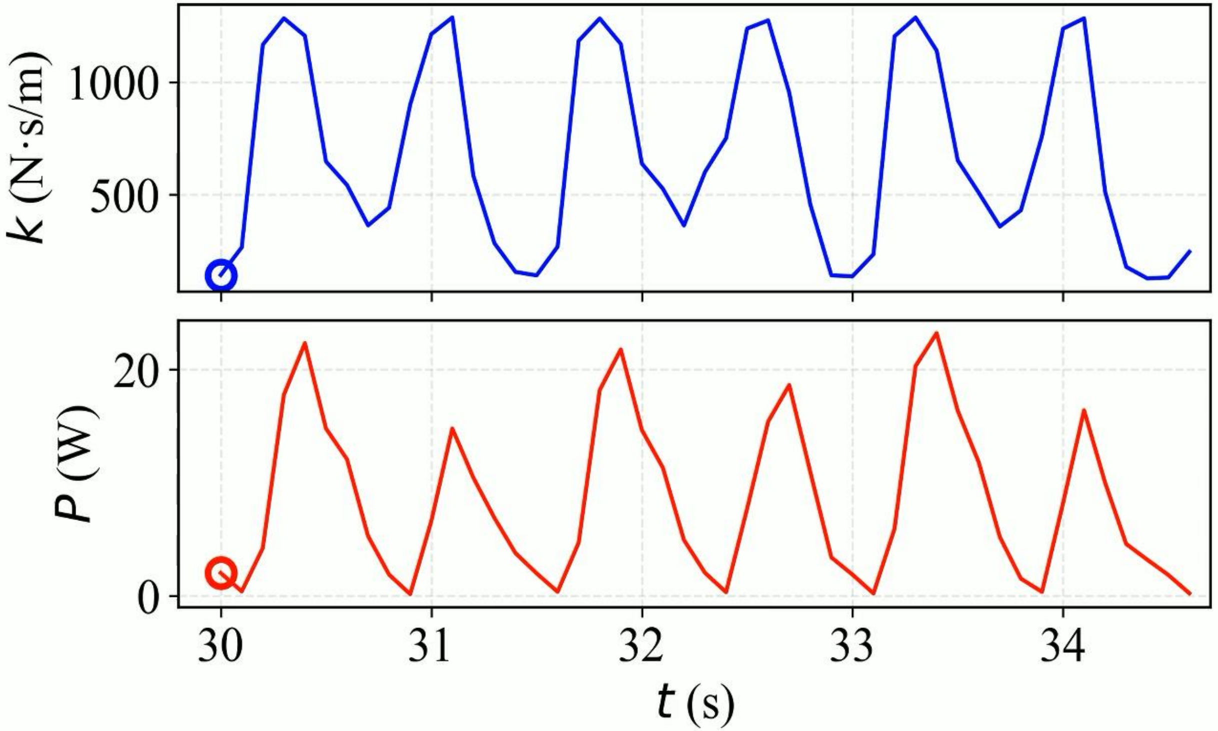
$$k_{\text{p},i} = \boxed{k_{\text{base}}} + o_i \Delta k_{\max}$$



- DRL optimization starts from an optimal initial value k_{base}

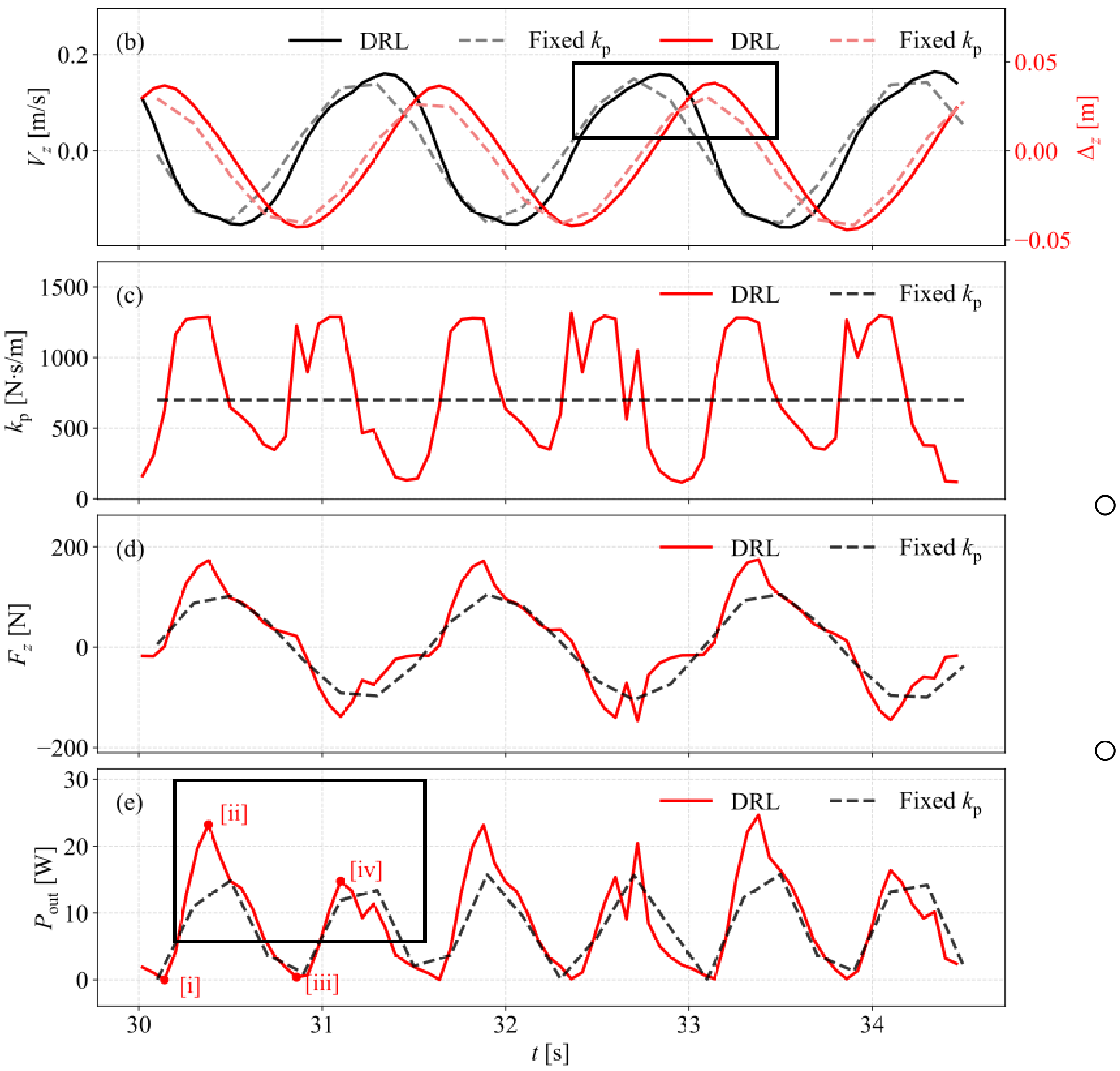
3.3 Point absorber wave energy converter : 2D regular wave

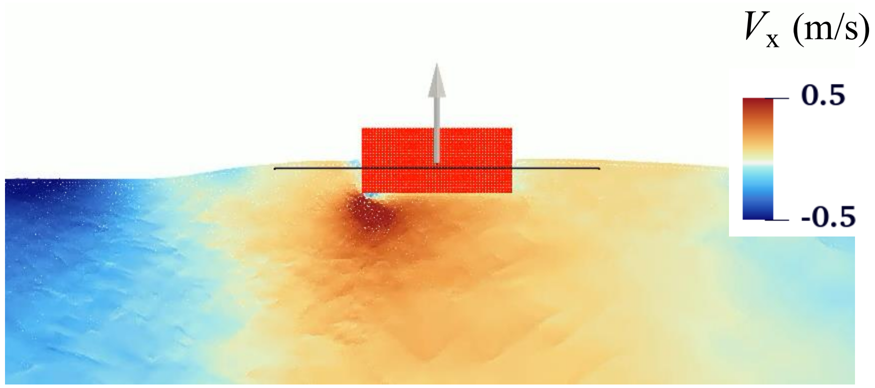




- k exhibits two peaks in one wave period, (1) near the wave crest (2) near the wave trough
- A higher energy output is observed when the wave passes the trough compared to the crest phase

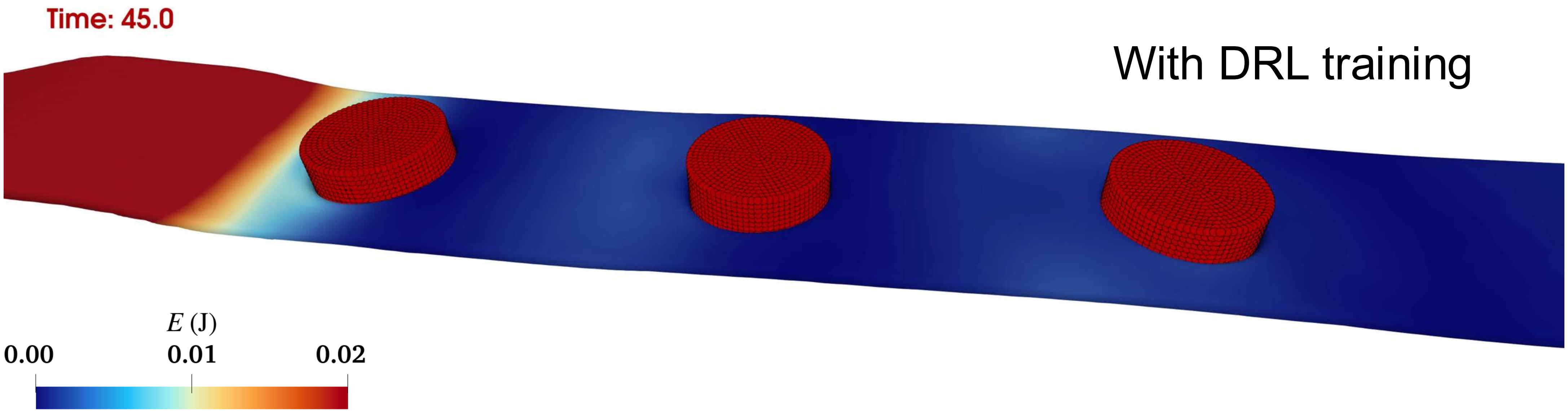
Point absorber wave energy converter : 2D regular wave

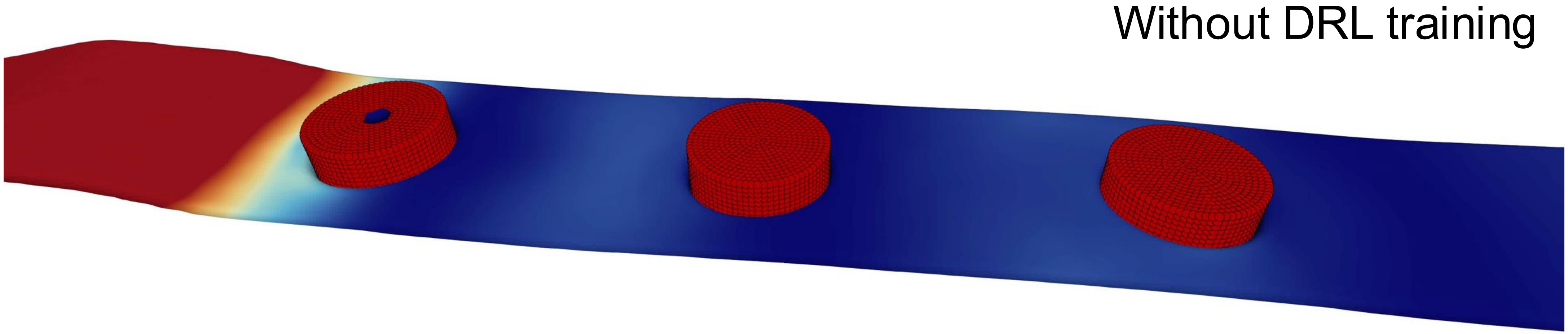




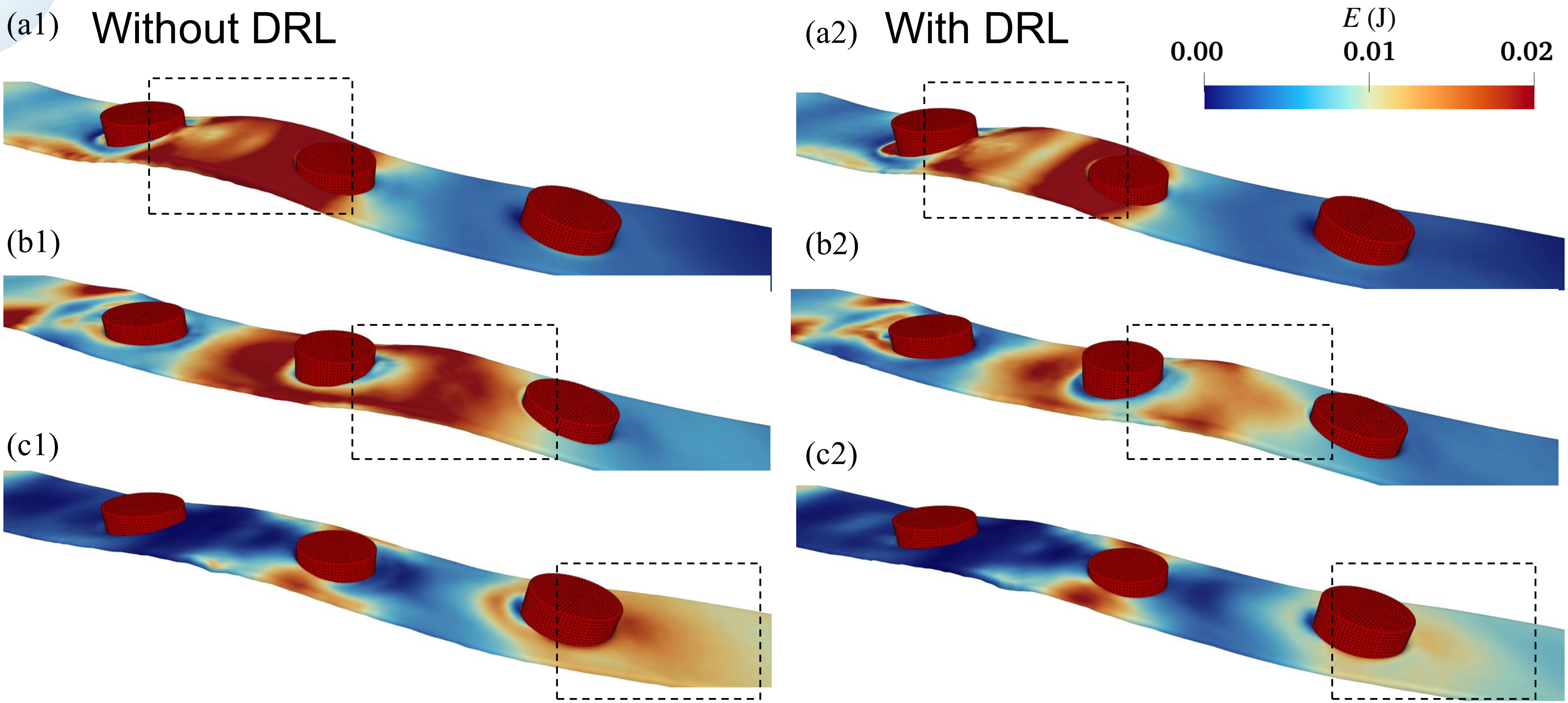
- DRL → larger velocity, displacement responses and damping force under wave excitation
- The total power increases by approximately 18% over three wave periods

3.3 Point absorber wave energy converter : 3D irregular wave





3.3 Point absorber wave energy converter : Energy distribution

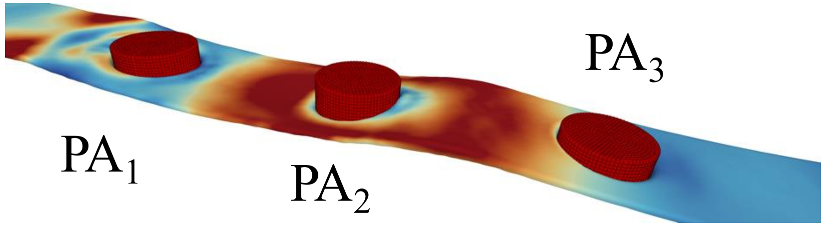


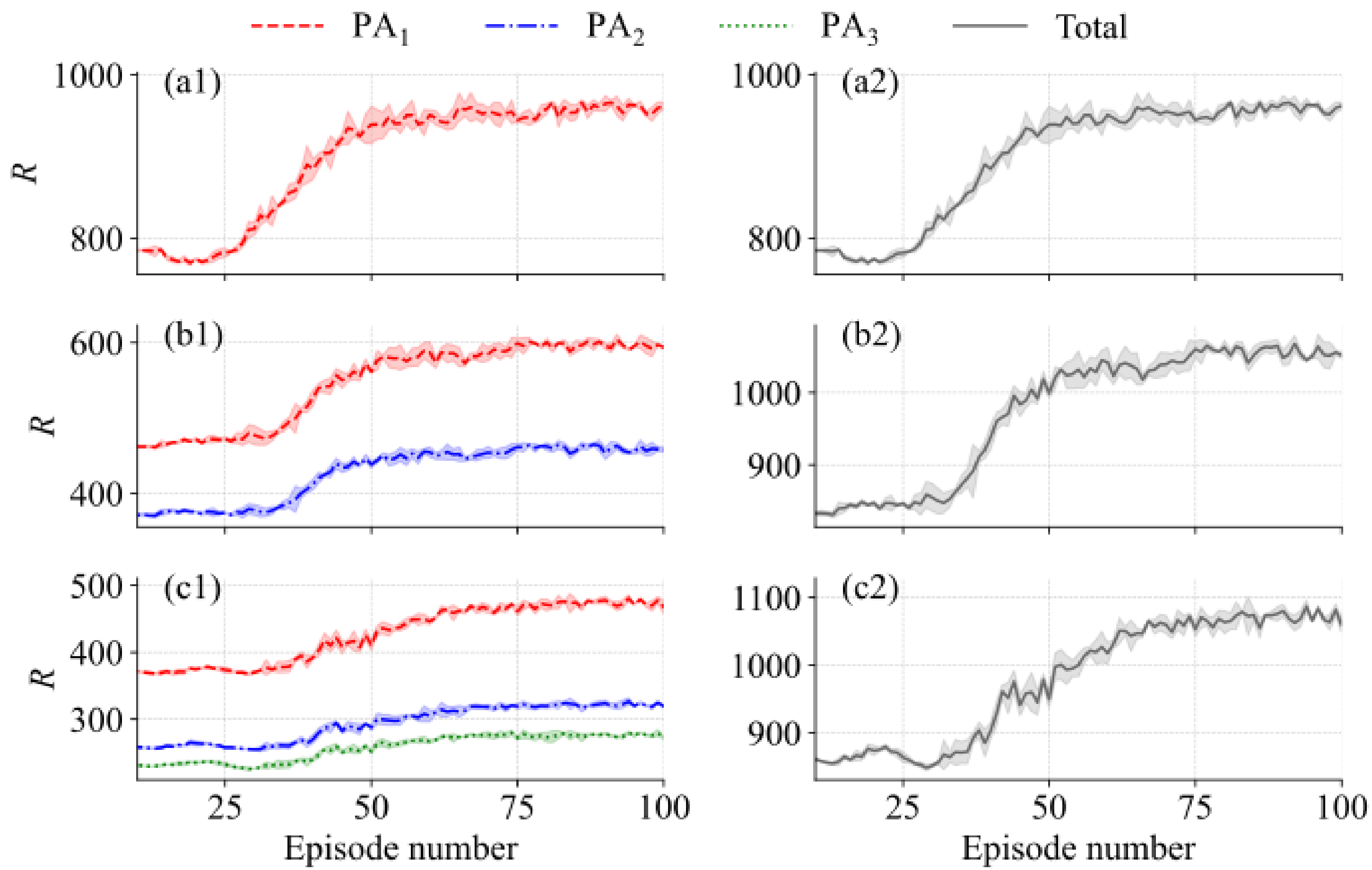
- Trained wave energy device improves absorption efficiency

Point absorber wave energy converter : 3D irregular wave

$$r_i = (1-\gamma_p)P_{\mathrm{out},i} + \gamma_p \frac{1}{N}\sum_{j=1}^{N} P_{\mathrm{out},j} \quad P_{\mathrm{out},i} = k_{\mathrm{p},i} v_i^2$$

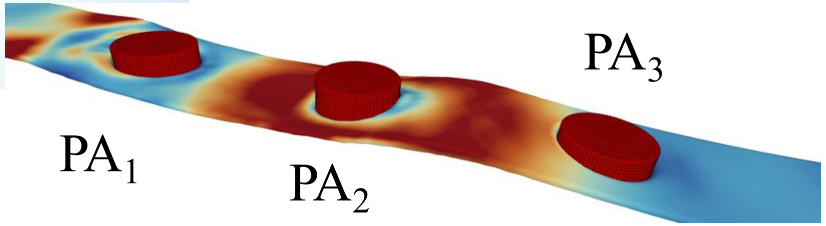
WEC Energy output   Encourage collaborating





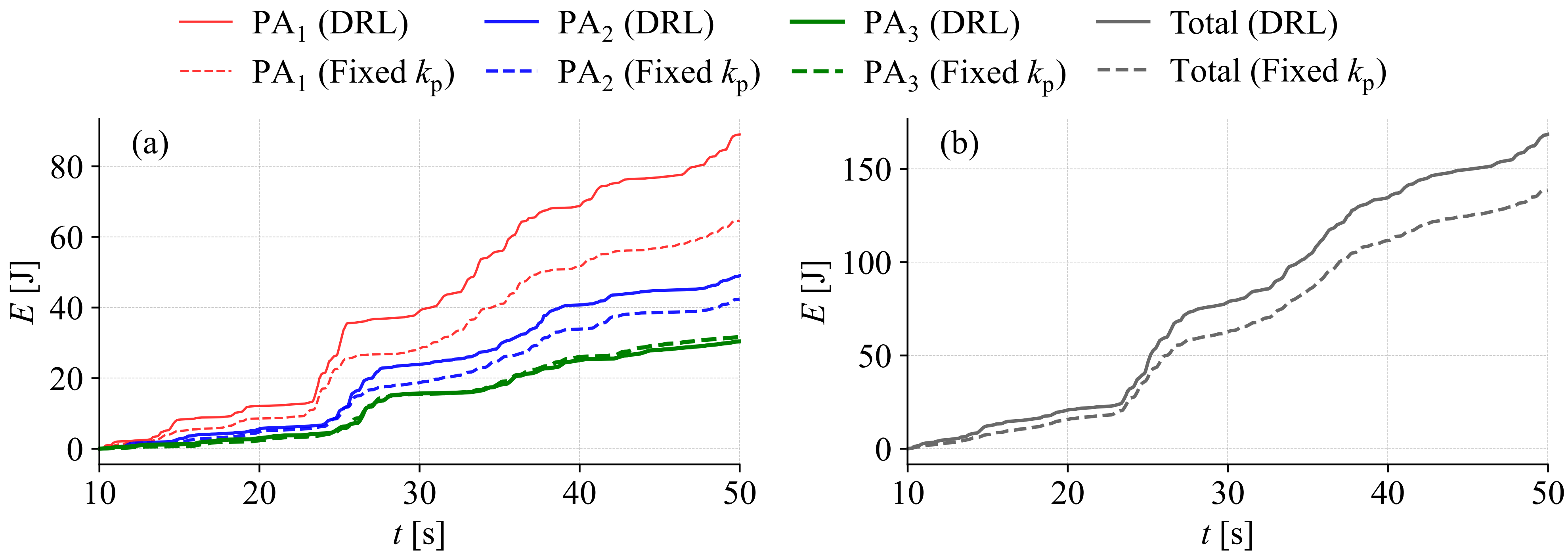
- The average reward keeps increasing and tuning slightly after 75 episodes
- Reward decreases from the first to the third PA

3.3 Point absorber wave energy converter : 3D irregular wave



$$r_i = (1-\gamma_p)P_{\mathrm{out},i} + \gamma_p \frac{1}{N}\sum_{j=1}^{N} P_{\mathrm{out},j} \quad P_{\mathrm{out},i} = k_{\mathrm{p},i} v_i^2$$

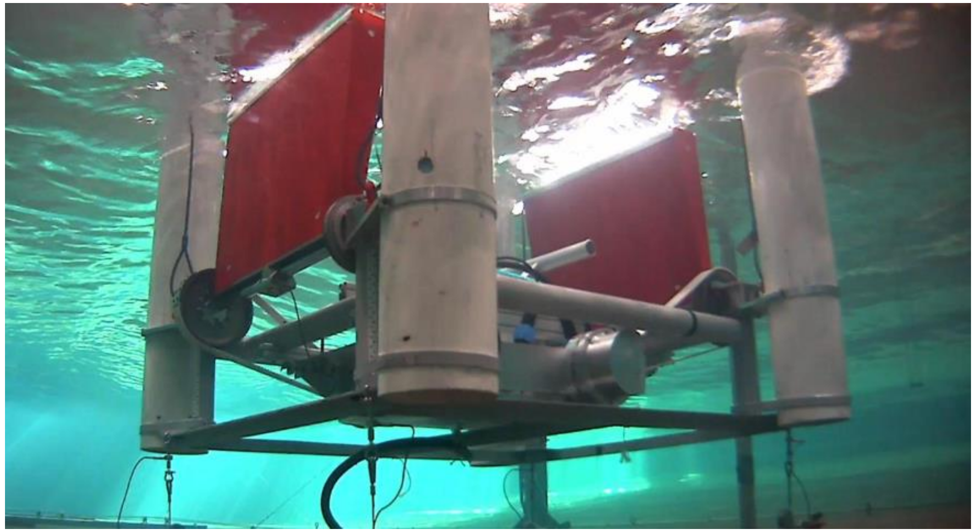
WEC Energy output   Encourage collaborating

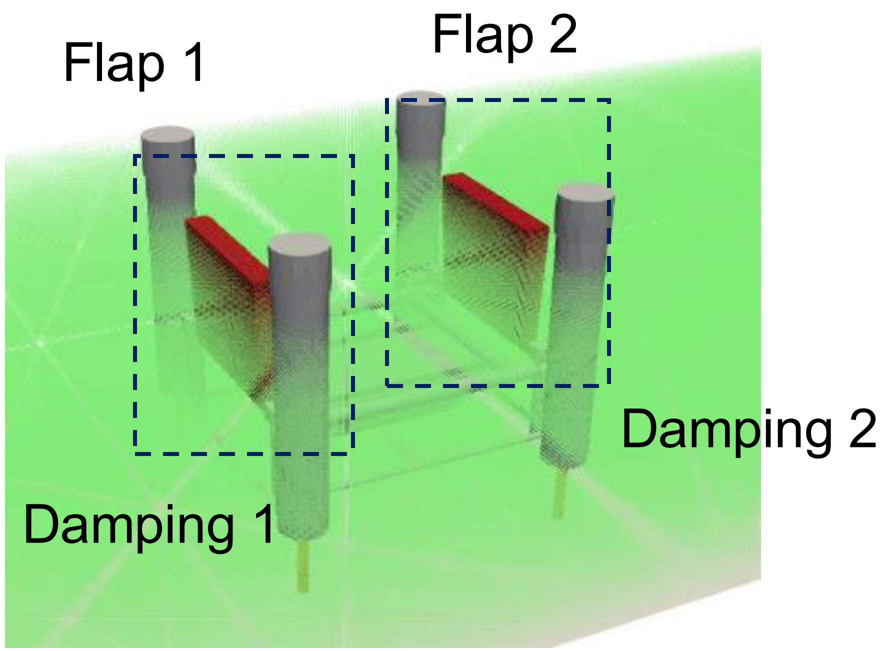


- Energy output increase for PA$_1$, PA$_2$, PA$_3$ and total system are 37.7%, 15.7%, −4.0% and 21.5%
- The slight decrease in the energy output of PA$_3$ reflects the cooperative effect among the agents

4 Summary: SPH-DRL coupling model

- SPH + DRL model → **Active control** in ocean engineering
- Multi Agent DRL → Achieve **cooperative optimization** among multiple agents
- GPU parallelism → **3D** practical engineering applications
- Future work: Improve the energy output (Mooring lines + Multibody rigid dynamics + Wave)





浙江大学 海洋学院
OCEAN COLLEGE, ZHEJIANG UNIVERSITY

Particles 2025, Barcelona

Environmental | Physics | Laboratory
EPhysLab

KYOTO UNIVERSITY

Thank you for your attention