ZHEJIANG UNIVERSITY

8th DualSPhysics Workshop
Ourense, Spain

KYOTO UNIVERSITY

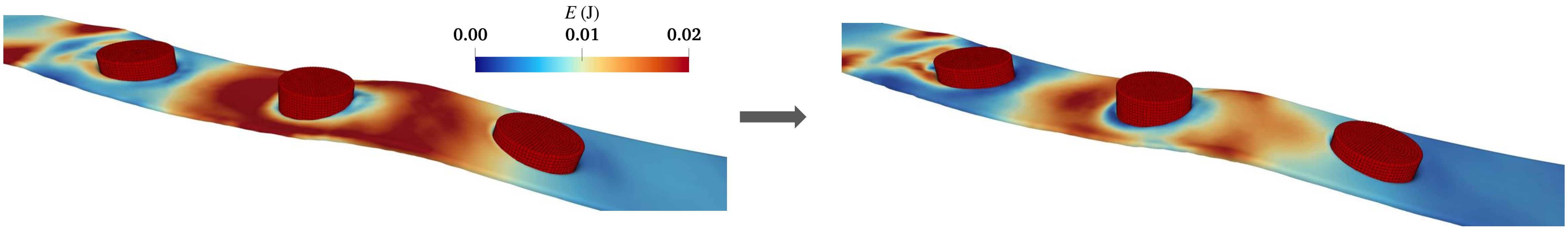
Coupling SPH with Multi-Agent DRL for Enhanced Energy Capture of WEC Arrays

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

[1] Ocean College, Zhejiang University [2] EPhysLab, CIM-UVIGO, Universidade de Vigo

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

29th January 2026

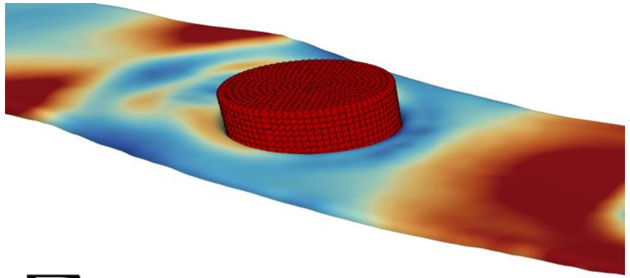


Contents

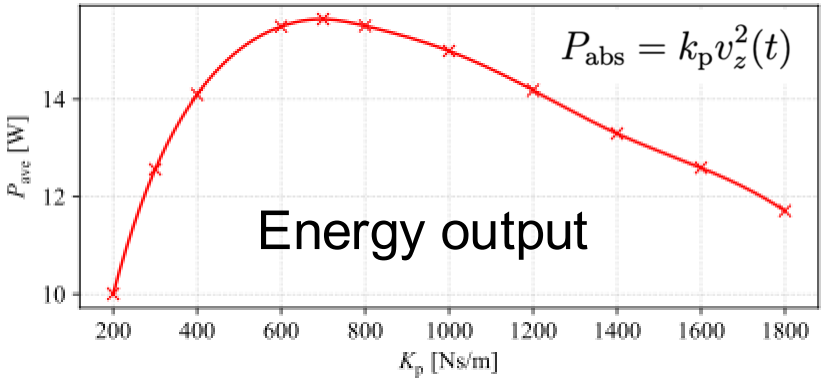
- Motivation
- SPH-MADRL coupling model
- Numerical validations

1.1 Motivation : How to improve energy absorption of WEC?

Resonance in regular wave

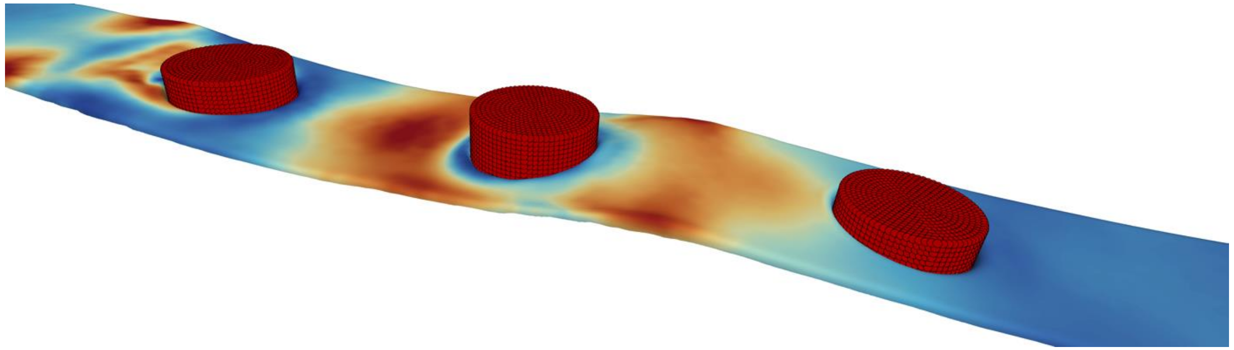


$$M\frac{d\mathbf{V}}{dt} = \sum_j m_j \mathbf{f}_j + \mathbf{D}_t \quad \mathbf{D}_t = -k_\mathrm{p}\mathbf{v}$$



Irregular / highly nonlinear waves

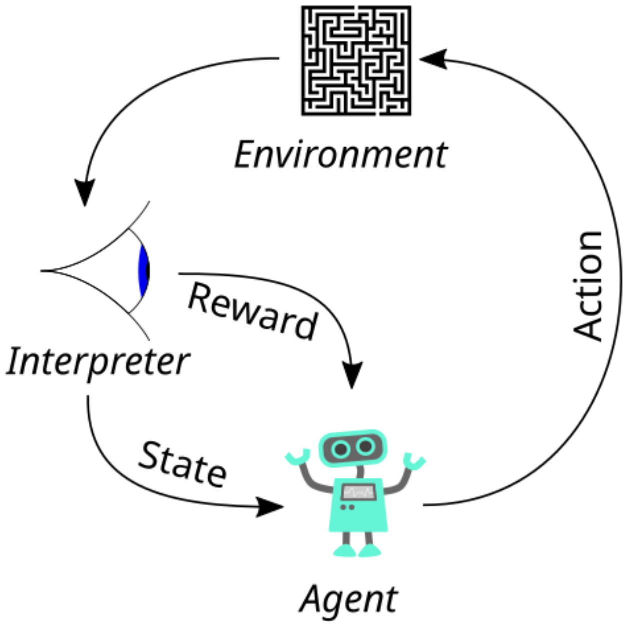
WEC Array?

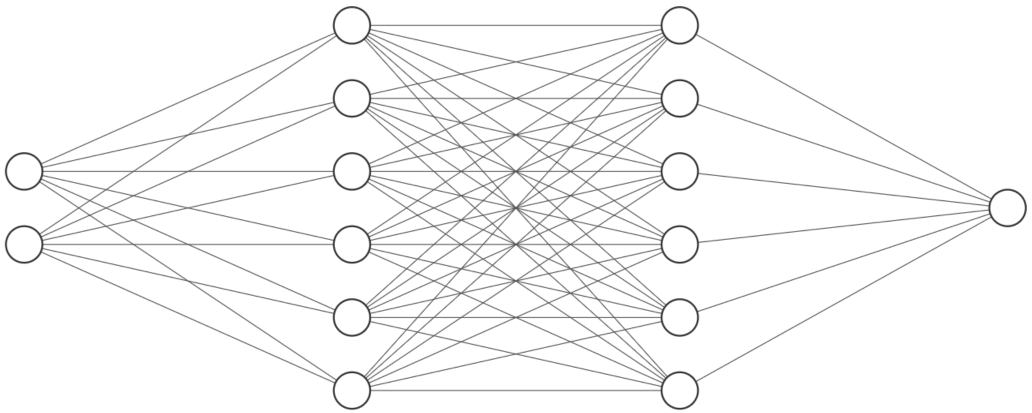


- We may find solutions from deep reinforce learning

1.1 Motivation : How to improve energy absorption of WEC?

➢ Concept of Deep Reinforcement Learning (DRL)





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

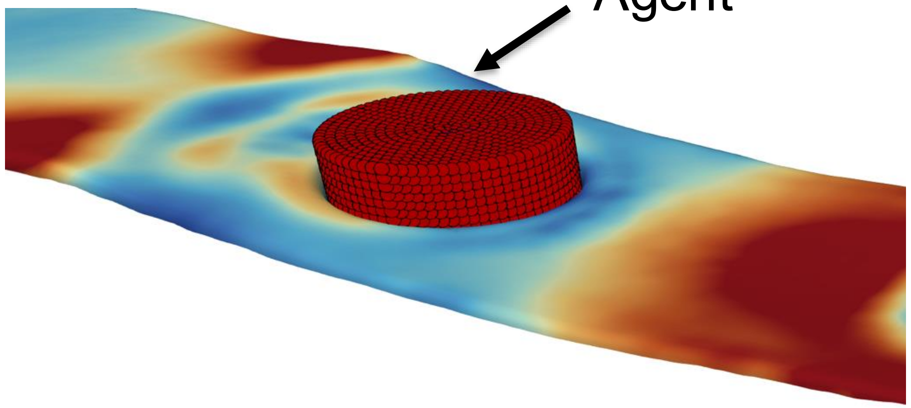
[1] wiki/Reinforcement_learning

1.1 Motivation : How to improve energy absorption of WEC?

➢ DRL for WEC

Environment : Ocean
State : Real time wave conditions

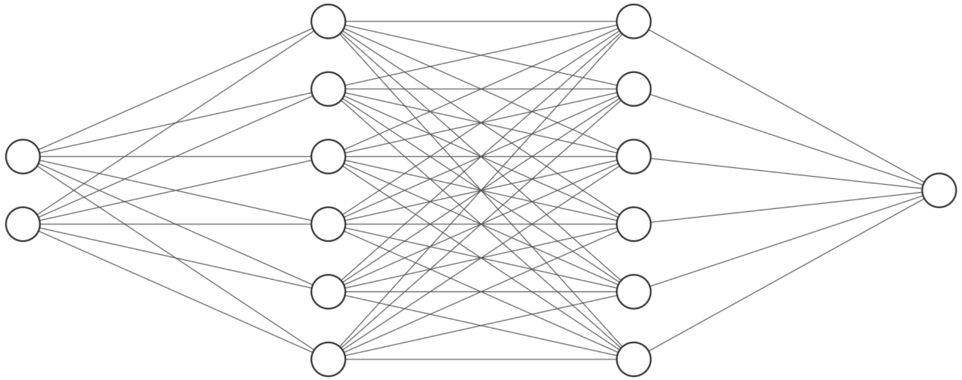
Agent



Reward : Energy output

Action : k_{p}

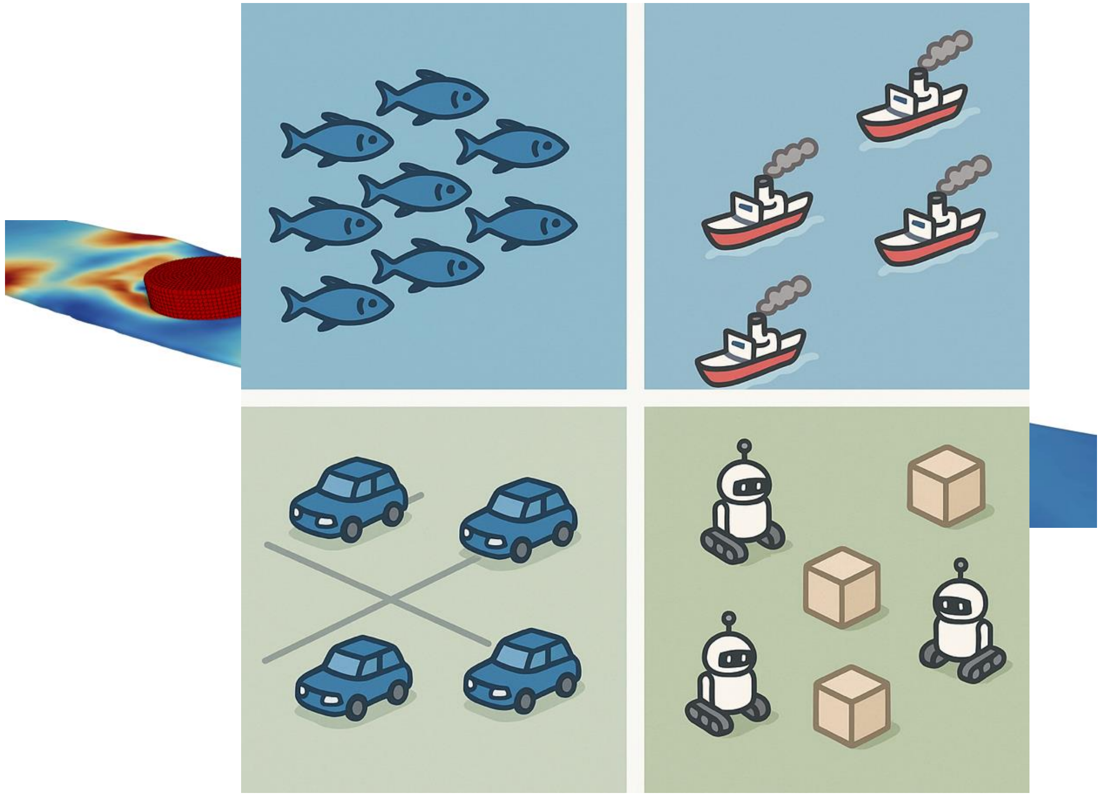
$$M\frac{d\mathbf{V}}{dt} = \sum_j m_j \mathbf{f}_j + \mathbf{D}_t \qquad \mathbf{D}_t = -k_{\mathrm{p}}\mathbf{v}$$



- SPH: provides environment; DRL: Train neuron network → Take actions → Maximize rewards

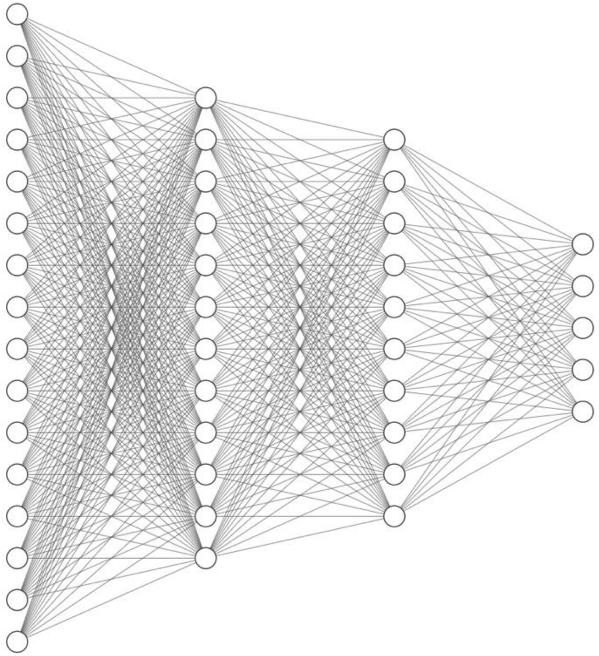
1.1 Motivation : Multi-agent decision-making

- Multi WECs / agents



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

- Single NN, multi outputs



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

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

1.2 Research gap and solutions

- Improve energy absorption of WEC → Coupling SPH with DRL
- Case with Multiple WECs → Multi-agent DRL

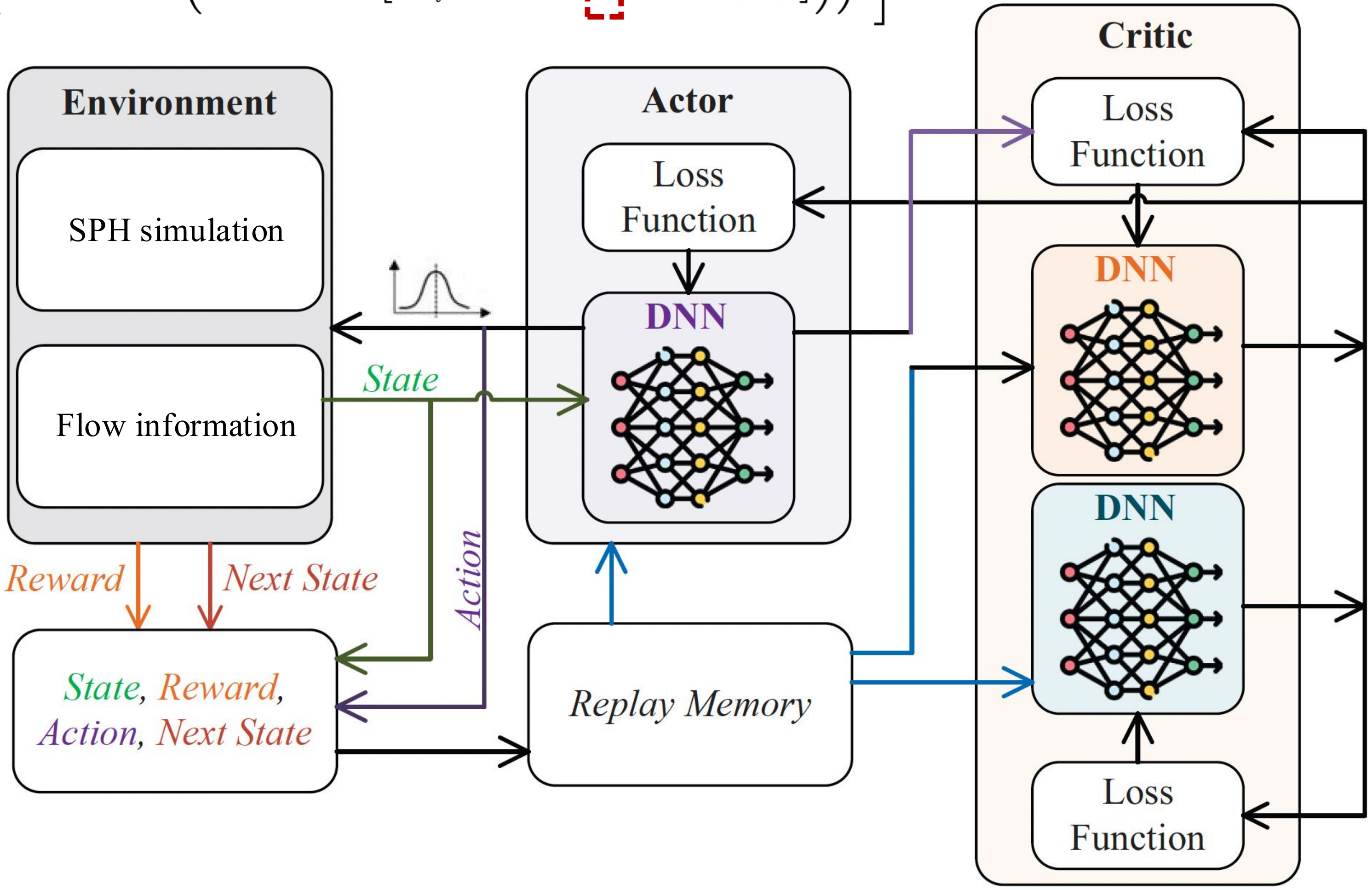
Contents

- Motivation
- SPH-MADRL coupling model
- Numerical validations

2.1 DRL Algorithm: Soft actor-critic

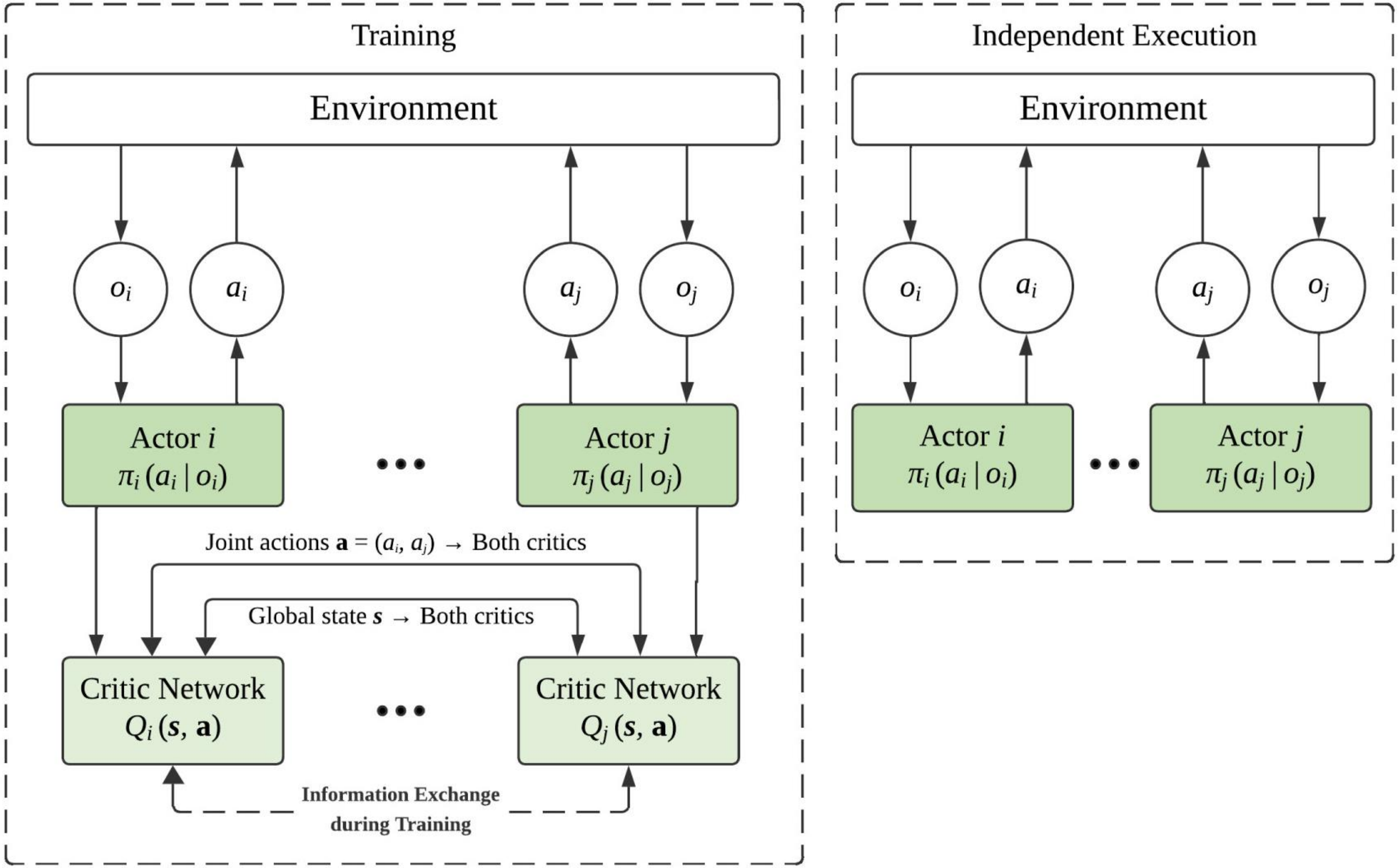
$$\boxed{J_Q(\theta_i)} = \mathbb{E}_{(s,a,r,s')\sim\mathcal{D}}\left[\left(Q_{\theta_i}(s,a) - \left(r + \gamma\,\mathbb{E}_{a'\sim\pi}\left[Q_{\bar{\theta}_i}(s',a') - \boxed{\alpha}\log\pi_i(a_i'|s_i')\right]\right)\right)^2\right]$$

- **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*.

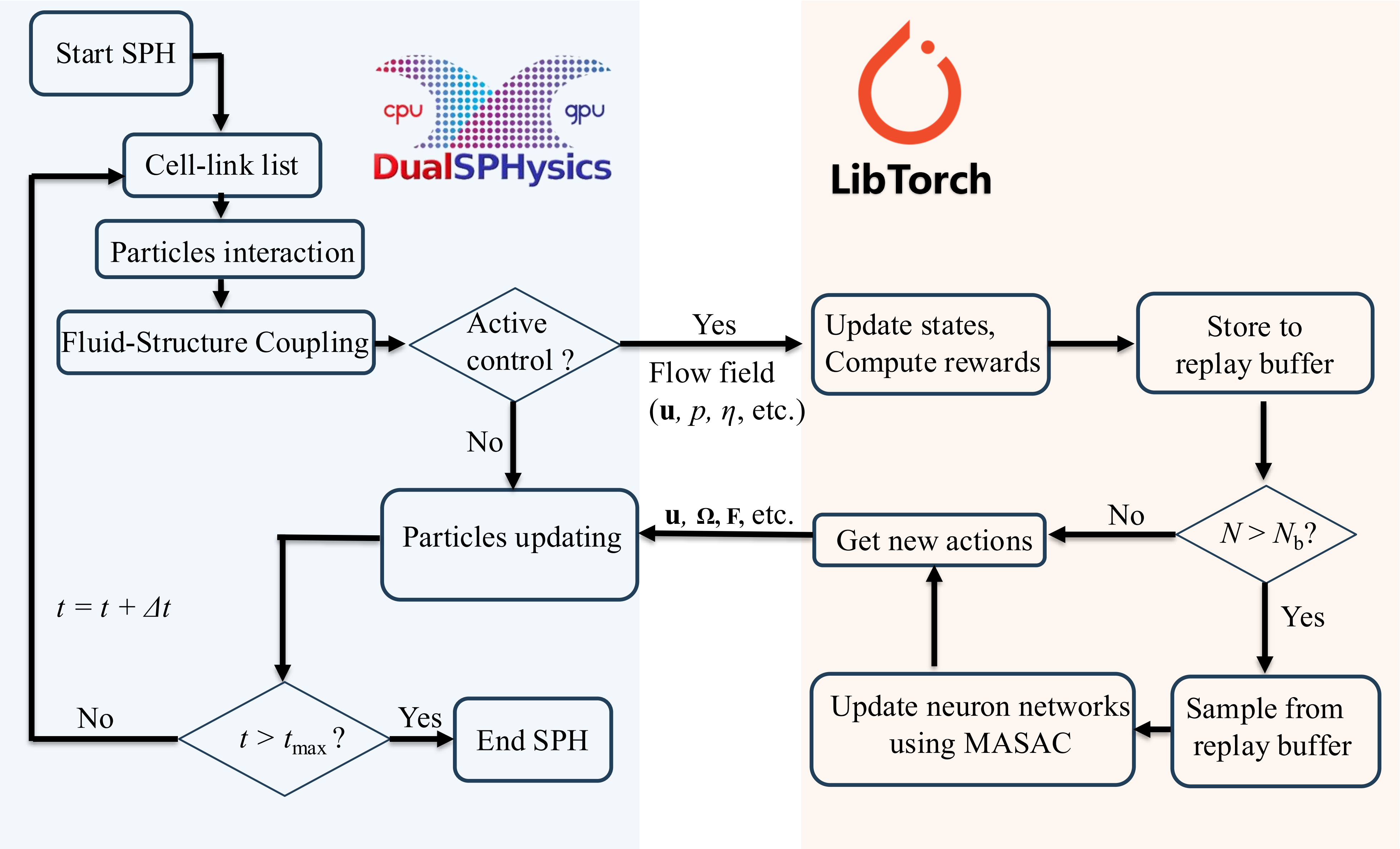
2.1 MADRL: Centralized training decentralized execution



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

2.2 Coupling SPH with MADRL

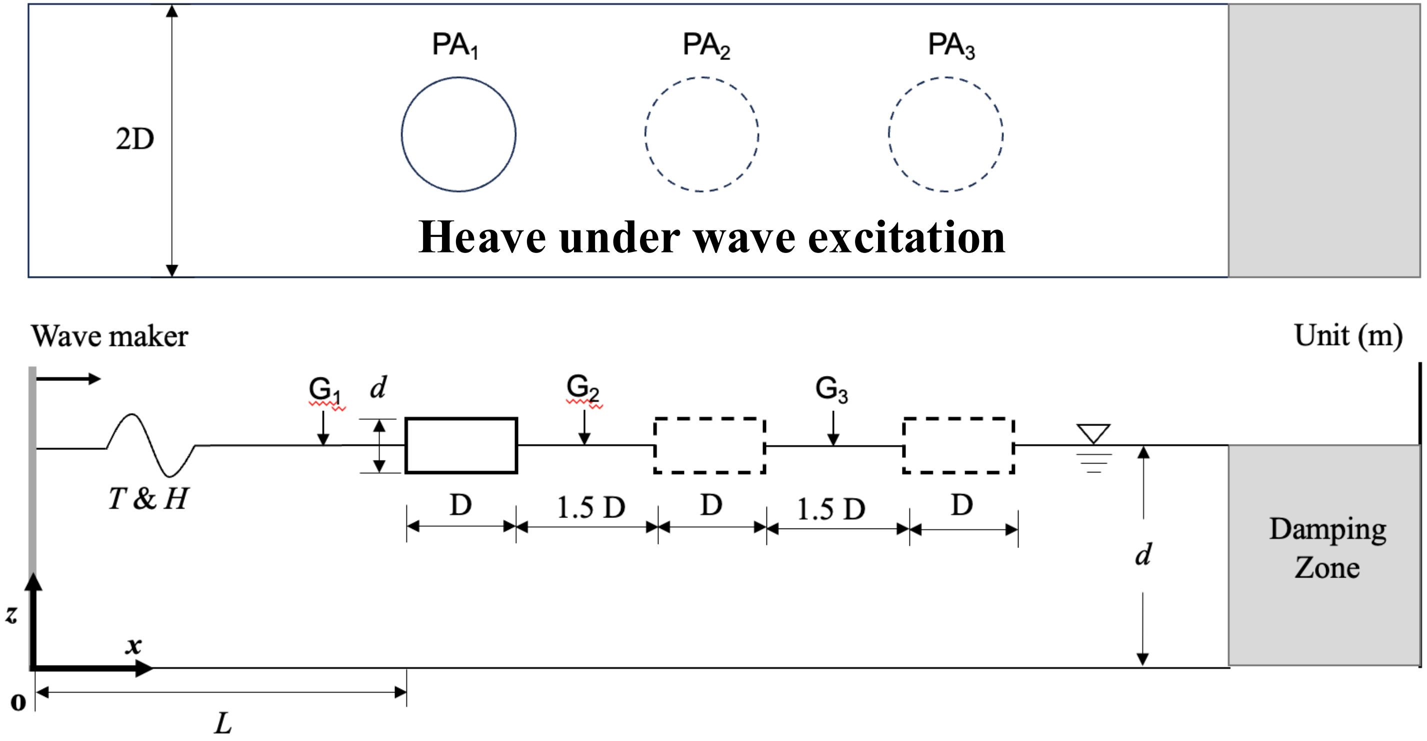
- Libtorch is linked to DualSPHysics as a dynamic library
- All codes are in C++ & CUDA and can be parallelized using GPU



Contents

- Motivation
- SPH-MADRL coupling model
- Numerical validations

3.1 Point absorber wave energy converter : overview



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

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

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

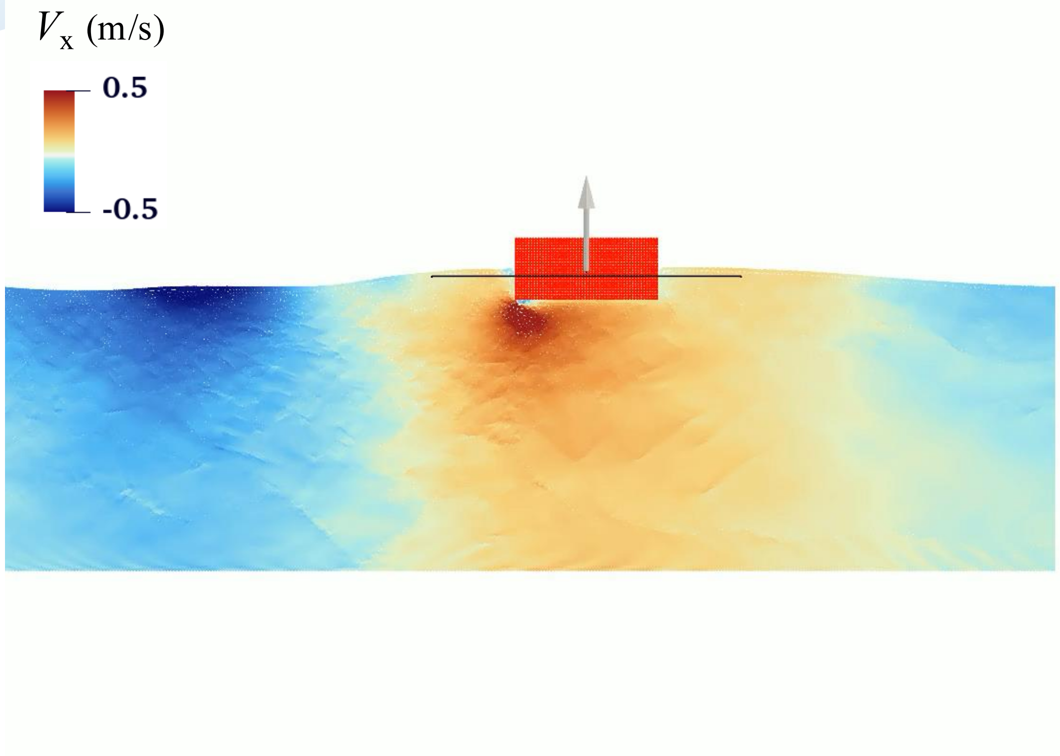
NN Input : $G_{i,1} - G_{i,3}$, $d(G_{i,1} - G_{i,3})/dt$, $v_{z,i}$, z_i, $\mathbf{a}_i^{n-1}$

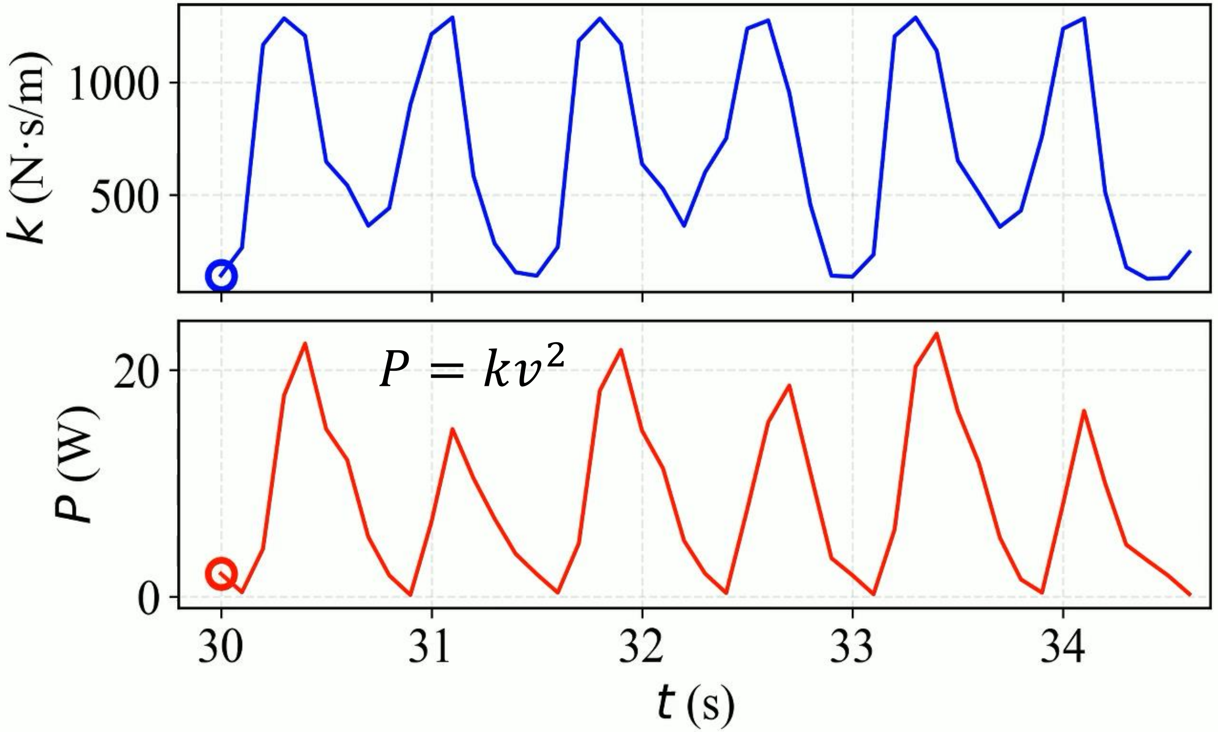
Reward : $r_i = (1-\gamma_p)P_{\mathrm{out},i} + \gamma_p \frac{1}{N}\sum_{j=1}^{N} P_{\mathrm{out},j} \quad P_{\mathrm{out},i} = k_{\mathrm{p},i} v_i^2$

WEC Energy output   Encourage collaborating

Make the PTO parameter k_{p} adaptive to cope with different incident wave → increase power output

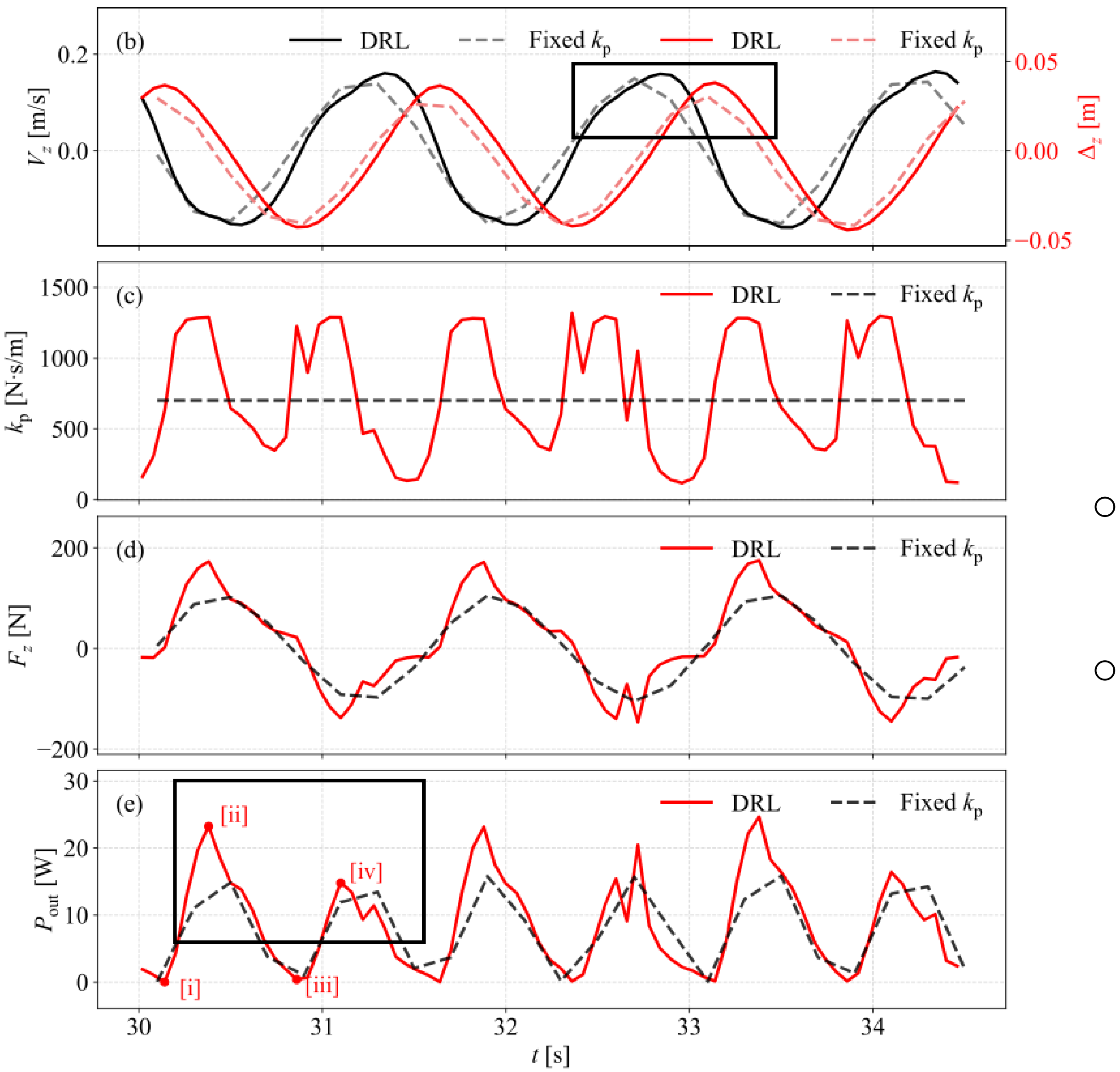
3.1 Point absorber wave energy converter : 2D regular wave

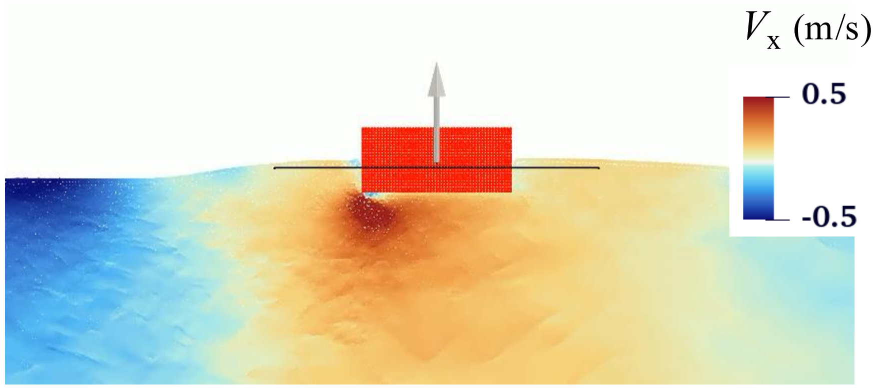




- *k* exhibits two peaks in one wave period, (1) near the wave crest (2) near the wave trough
- A higher energy output is observed when the wave passes the trough compared to the crest phase

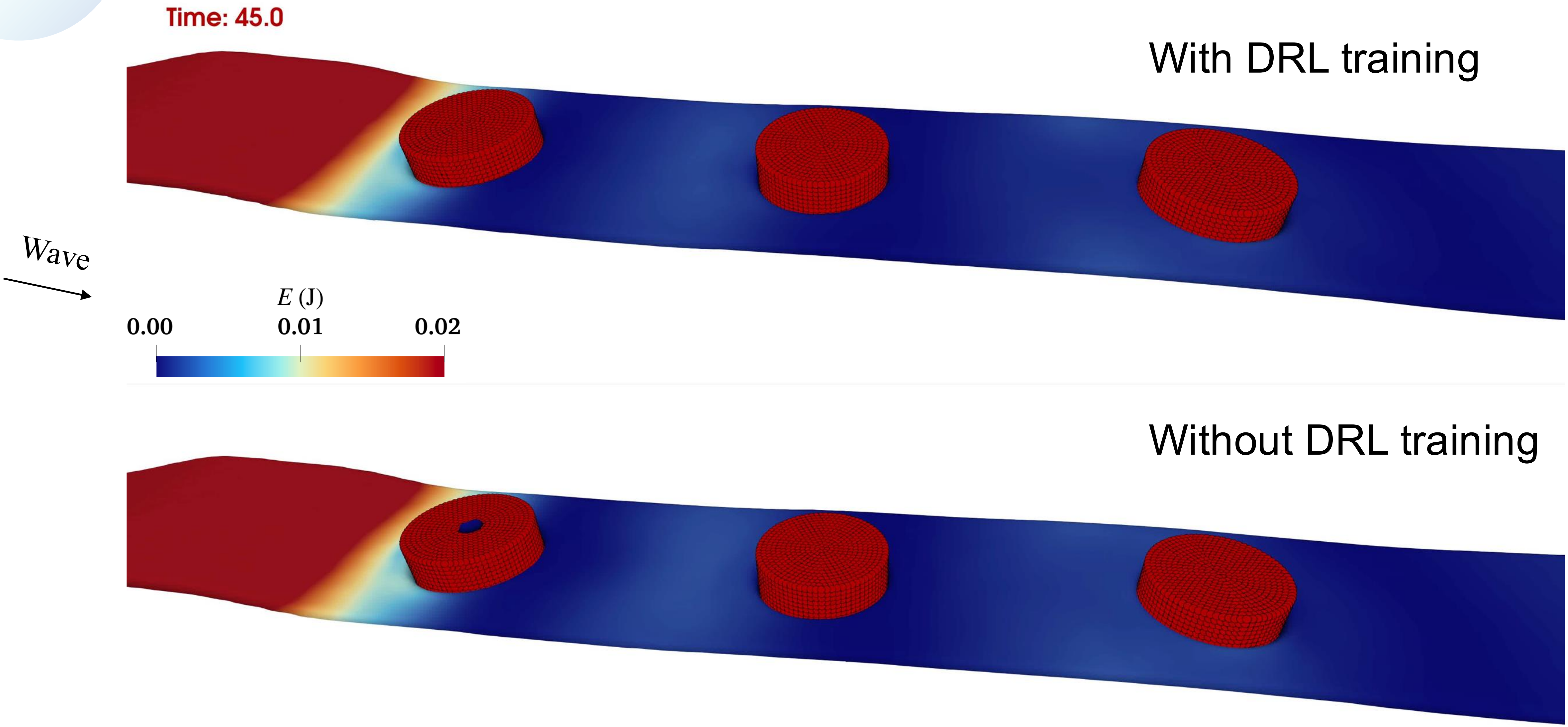
3.1 Point absorber wave energy converter : 2D regular wave



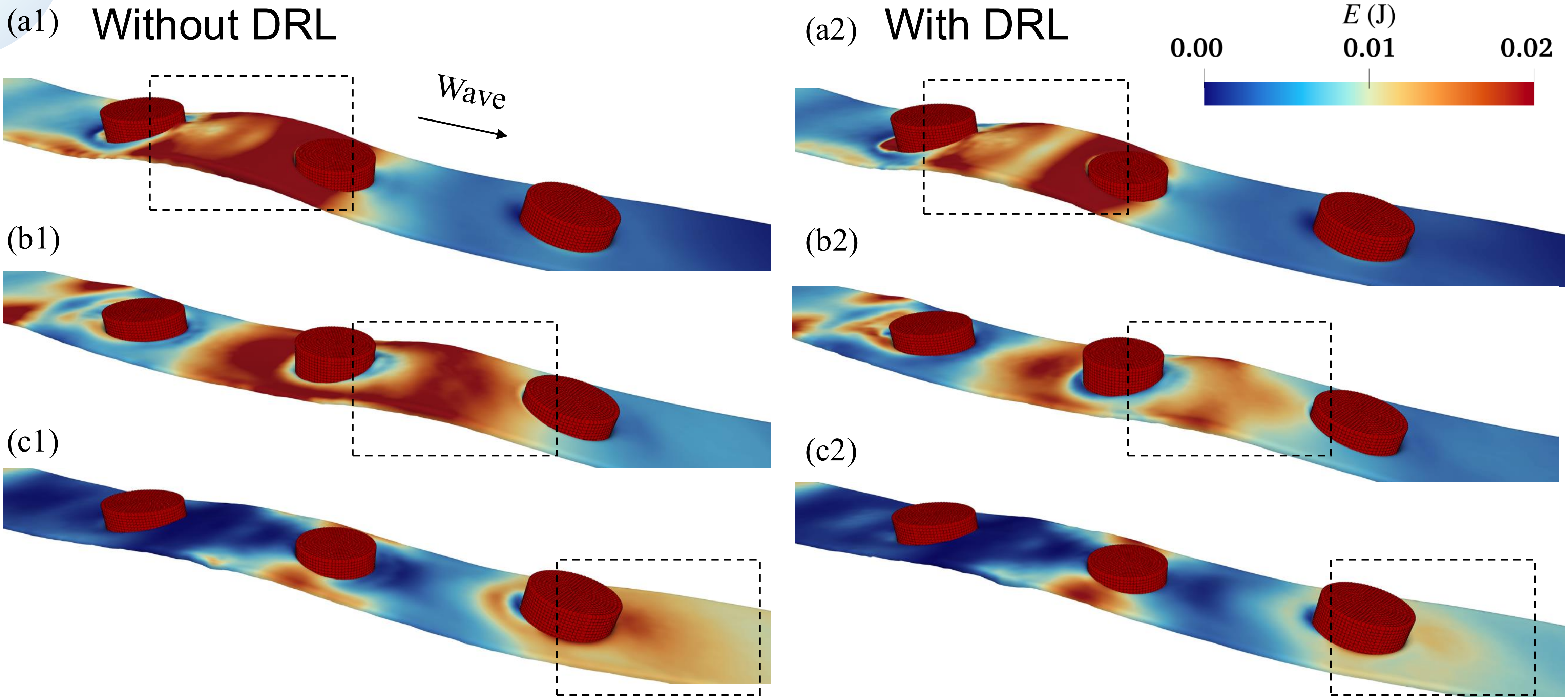


- DRL → larger velocity, displacement responses and damping force
- The total power increases by approximately 18% over three wave periods

3.2 Point absorber wave energy converter : 3D irregular wave

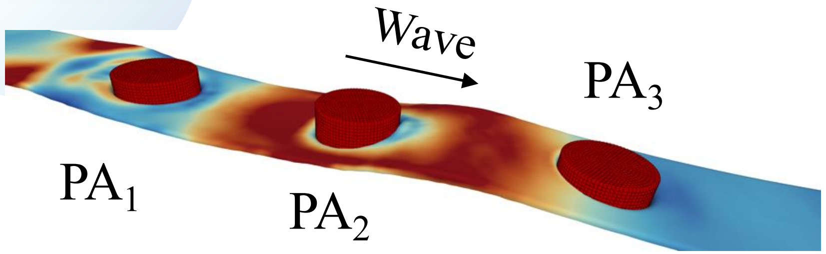


3.2 Point absorber wave energy converter : Energy distribution



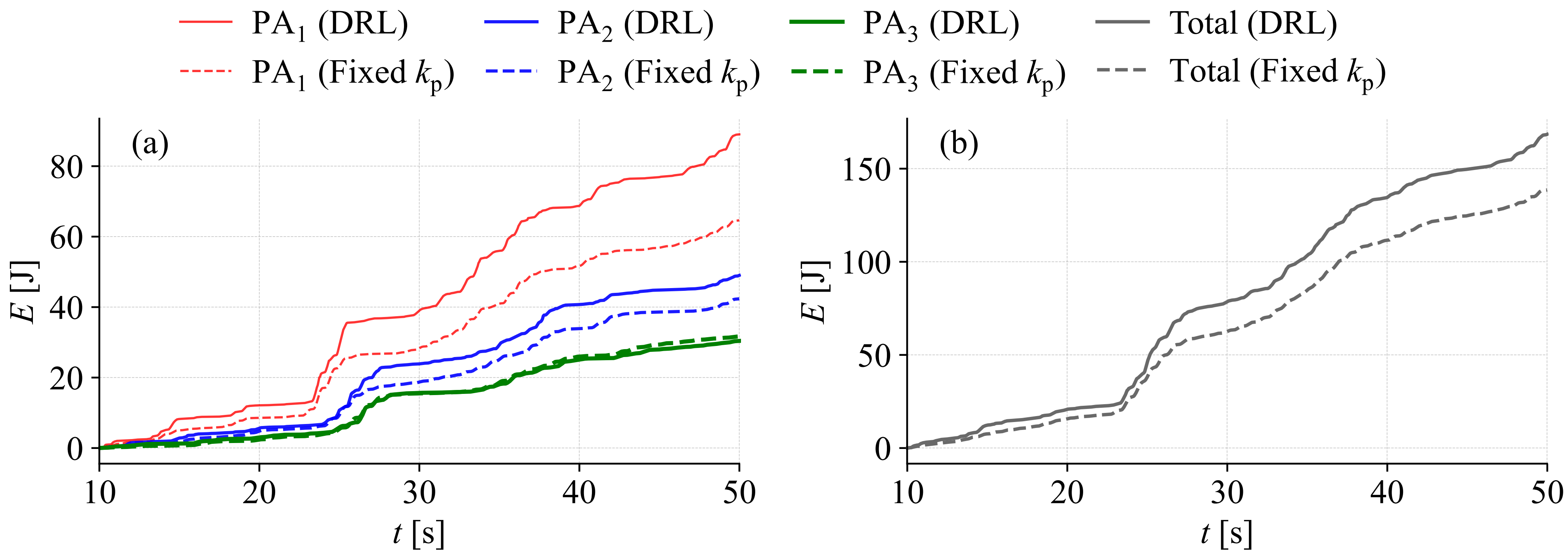
- Trained WEC improves absorption efficiency

3.2 Point absorber wave energy converter : 3D irregular wave



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

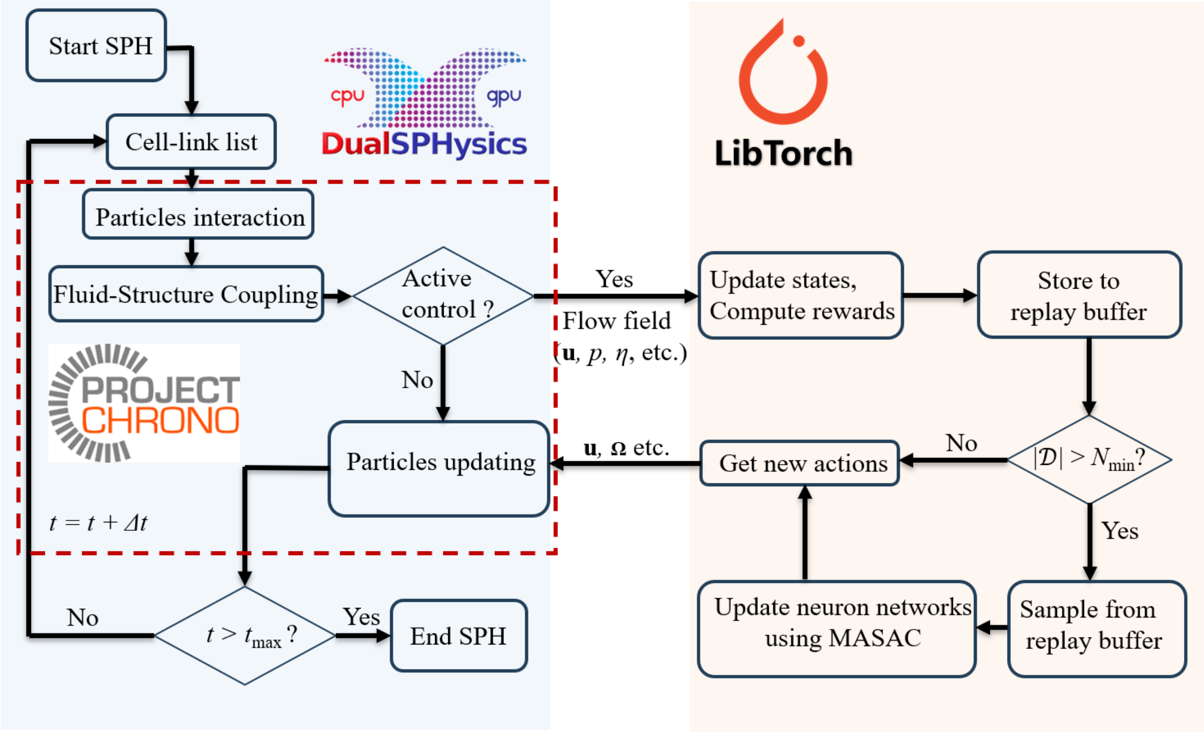
WEC Energy output   Encourage collaborating



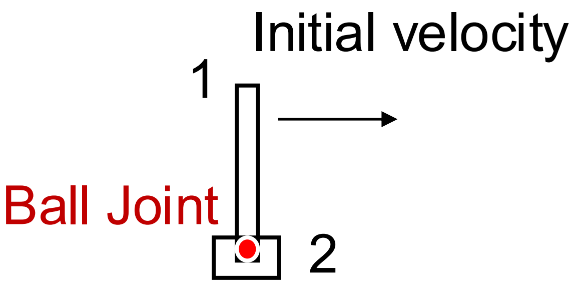
- Energy output increase for PA$_1$, PA$_2$, PA$_3$ and total system are 37.7%, 15.7%, −4.0% and 21.5%
- The slight decrease in the energy output of PA$_3$ reflects the cooperative effect among the agents

4 Summary: SPH-DRL coupling model

- SPH + DRL model → **Active control** in ocean engineering
- Multi Agent DRL → Achieve **cooperative optimization** among multiple WECs
- GPU parallelism → **3D** practical engineering applications
- Ongoing work: Multibody rigid dynamics



4 Ongoing work: Multibody rigid dynamics



Input: $\mathbf{v}_1, \mathbf{v}_2, \boldsymbol{\omega}_1, \mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\varphi}_2,$

Output: $\mathbf{v}_2$

Reward: $f\,(\mathbf{r}_1, \mathbf{v}_1, \boldsymbol{\omega}_1, \boldsymbol{\varphi}_2)$

- Before training

- During training

- After training

[CartPole in Pytorch Demo]

Thank you for your attention