Comparative experimental study on jet mixing enhancement using deep reinforcement learning and genetic programming

Zhi Wu[1], Jiayi Lin[1], Yangkun Zhang[1] and Yu Zhou[2]

[1] Center for Turbulence Control, Harbin Institute of Technology, Shenzhen, P.R. China
[2]College of Engineering, Eastern Institute of Technology, Ningbo, P.R. China

Keywords: Turbulent jet mixing, deep reinforcement learning, genetic programming

Abstract. Machine learning has experienced explosive development in the field of flow mechanics, particularly in turbulent modelling and flow control. Among these machine learning techniques, deep reinforcement learning (DRL) and genetic programming (GP) have attracted significant attention in flow control research due to their ability to handle highly nonlinear dynamic systems. However, few successful experimental applications have been reported, mainly due to the limitations of low dynamic response in sensors and actuators. This study aims to enhance the jet mixing using one pulsed radial minijet or two symmetrical pulsed minijets. Three hot-wire sensors, placed at 3*D* and 5*D* downstream of the jet exit (where *D* is the main jet outlet diameter), were used to measure the flow states. Proximal Policy Optimization (PPO) method and linear genetic programming (LGP) method were employed to optimize the feedback control law for the pulsed minijets. Experiments were conducted at Reynolds number of 8000, with the mass flow ratio of the minijets to the main jet fixed at 1.2%. The reward or cost function are defined based on the jet centerline velocity decay rate *K*. Experimental results show that both the PPO and LGP could run in real-time and optimize the control strategies. PPO is much faster than LGP at finding the optimal control law. The optimal control strategies identified by PPO and LGP show similar control performance under the designed condition, but the PPO proves to be more robust under off-design conditions. Additionally, the use of two symmetrical pulsed minijets does not improve the control performance for either control system.

Keywords: deep reinforcement learning, genetic programming, jet mixing

1. Introduction

The turbulent jet is one of the classic shear flows. Its control is of great importance in industries, such as dilution jets in combustors, fuel injection of combustion engines. Effective jet mixing enhances fuel-air homogeneity, which accelerates combustion reactions and reduces unburned hydrocarbon emissions. In supersonic flows, pulsed jets with optimal frequencies amplify shear vortex structures, enhancing mixing and non-premixed combustion efficiency by 24.5% through shock-wave interactions[1].

Jet mixing control can be categorized into passive control and active control. The former involves a change in geometry, thus does not need additional power input, such as tabs[2], non-circular nozzles[3] and chevron nozzles[4]. Although highly effective in designed conditions, passive techniques are difficult to be adjusted, making it less effective in varying environment. The latter requires additional power input but its control parameters could be adjusted easily. Examples are acoustic excitation[5], plasma actuators[6], synthetic jet[7], oscillating boundaries based on piezo-electric actuators[8] steady and unsteady control jets[9-11]. Active control techniques can be performed in an open-loop or closed-loop manner. With suitable sensors and control algorithms, closed-loop control has more potential to achieve better performance than open-loop control. Many closed-loop control schemes have been proposed and investigated, as discussed in references Brunton and Noack[12]; Choi et al.[13]; Collis et al.[14].

Recently, artificial intelligence (AI) techniques have been widely applied in flow control applications. Compared with the classical control method, AI technologies are featured by their exceptional capabilities in optimizing control laws and extracting nonlinear characteristics of complex systems. For instance, genetic programming (GP) and deep reinforcement learning (DRL) have achieved a great success in flow control applications, such as jet mixing enhancement control, flow separation control, and cylinder wake control[15]. In the field of jet mixing control, significant achievements have been made with the application of genetic programming (GP). Wu et al.[16] developed a control system based on GP to optimize the pulsed radial minijet, successfully enhancing the mixing rate of turbulent jet. Zhou et

al.[17] applied the control system to optimize 6 independent control laws of 6 radial minijets, resulting in a complex turbulent structures with periodic generation of mushroom-shaped morphologies. The deep reinforcement learning (DRL) algorithms have also achieved a great success in active flow control studies. A DRL control system for drag reduction of cylinder has reduced the drag by 8% and 30% for laminar flow[15] and weakly turbulent flow[18], respectively. Fan et al.[19] firstly employed TD3 (Twin Delayed Deep Deterministic Policy Gradient) to optimize the rotation velocity of two cylinders behind the main cylinder, which act as the actuators. They discovered a control strategy that is comparable to the optimal strategy found through lengthy systematically planned control experiments. It needs to be mentioned that, closed-loop feedback control systems based on DRL have not yet been widely applied in experiments. The primary challenge stems from the stringent requirements for high dynamic responsiveness from controllers, actuators, and sensors to achieve real-time closed-loop control.

To address the aforementioned issues, we developed a real-time closed-loop control system based on DRL algorithm, aiming to apply DRL at high Reynolds number with fast interaction frequency. The jet mixing enhancement performance under the actuation of one or two symmetrical pulsed minijets will be studied. The difference between DRL and LGP in finding the optimal feedback control law, as well as their robustness will be compared and discussed.

2. Experiments details

Figure 1(a) shows the jet experiment platform, which comprises a gas supply system, an axisymmetric main jet, and minijets. This experiment was conducted at Reynolds number $Re = \overline{U_j} D / \nu = 8000$, where $\overline{U_j}$ is the time-averaged velocity measured at the centerline of the main jet at the nozzle exit, ν is the kinematic viscosity of air, and $D = 20$ mm is the diameter of the main jet. The minijet is ejected through a 1 mm diameter orifice, located 17 mm upstream from the main jet exit. Hot-wire anemometers were employed as the primary sensors for measurements, with their arrangement depicted in Figure 1(b). Specifically, hot-wire 1 and hot-wire 2 are positioned at 3D downstream, symmetrically arranged from the centreline with a gap of 10mm. The hot-wire 3 is located along the centerline, at 5D downstream from the jet exit. As illustrated in Figure 1(c), one minijet (indicated by "1") or two symmetrical minijets (indicated by "1" and "2") were used for manipulating the main jet. Each minijet is connected to a solenoid valve to generate pulsed minijet with a maximum frequency of 500 Hz. The mass flow rate of each minijet is controlled by a mass flow controller with a range of 7 L/min. In this study, following Wu et al.[16], the mass flow rate ratio of each minijet to the main jet is fixed at 1.2%.

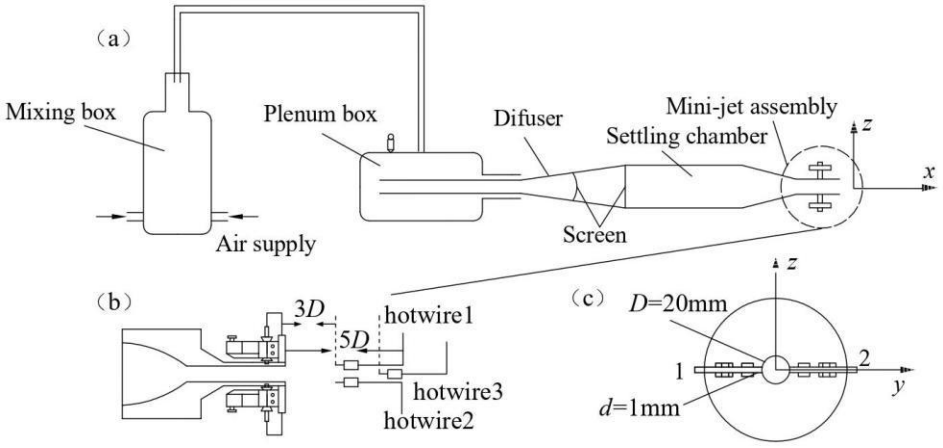

Figure 1 Schematic of experimental setup. (a) Main jet assembly; (b) minijets arrangement in x-z plane and (c) y-z plane.

3. Artificial Intelligent Control System

In this study, both DRL and LGP are deployed to optimize the feedback control law. The DRL is a machine learning paradigm where an agent learns optimal strategies through iterative interactions with

an environment to maximize cumulative rewards. Unlike supervised learning, DRL does not rely on labelled data but instead uses trial-and-error mechanisms guided by reward signals (positive/negative feedback) to refine decision-making policies. Key elements include states (environmental observations), actions (agent's responses), and rewards (performance evaluation). This framework is particularly suited for dynamic, real-time control systems, such as jet mixing enhancement by active control. Proximal Policy Optimization (PPO), a state-of-the-art DRL algorithm, addresses stability and efficiency challenges in policy updates. By introducing a clipped surrogate objective, PPO restricts policy deviations within a trust region, preventing destructive updates while maintaining sample efficiency. This approach enables multi-epoch minibatch updates, enhancing data utilization without compromising stability. In jet mixing optimization, PPO's ability to balance exploration (testing novel control actions) and exploitation (refining known strategies) makes it ideal for real-time feedback control, where rapid parameter adjustments are required to maximize mixing efficiency under turbulent conditions. Since the actuation is "ON/OFF" states of the pulsed minijet, thus discrete PPO is applied in the current study.

Figure 2(a) shows the schematic diagram of the DRL control system for jet mixing enhancement. The agent, which include the control policy, interacts with the experimental environment in a real-time control loop (indicated by the red arrows) by AD/DA converters with a sampling frequency of 1 kHz. The agent generates actions $\boldsymbol{a}$ based on the measured states $\boldsymbol{s}$ and policy. The $\boldsymbol{a}$ represents the "ON/OFF" state of the pulsed minijets, while $\boldsymbol{s} = [u_{3D1}, u'_{3D1}, u_{3D2}, u'_{3D2}, u_{5D}, u'_{5D}]$ denotes the instantaneous velocities (u) and fluctuating velocities (u') measured in the wake, where the subscripts 3D1, 3D2 and 5D correspond to the hotwire1, 2 and 3, respectively. The loop indicated by the green arrow in figure 2 indicates the policy updating and deployment process. Following Zhou et al.[17], the reward used to evaluate the performance of the control strategy is defined by the jet centerline velocity decay rate K, i.e., $r_t = K = 1 - u_{5D} / \overline{U_j}$. The control system runs as follows:

(1) The agent randomly generates a control law, which was applied to drive the pulsed minijets in a real-time control loop with a time step of 1 ms. The flow field states s_t, action single a_t, reward r_t as well as the next states s_{t+1} in each time step are saved in an experience replay memory.
(2) The agent updates the agent based on n experiences randomly sampled from the experience replay memory. The updated agent is compiled and deployed in LabVIEW environment, which interacts with the experimental environment by AD/DA converter.
(3) This whole process repeats iteratively until the optimal control strategy is identified.

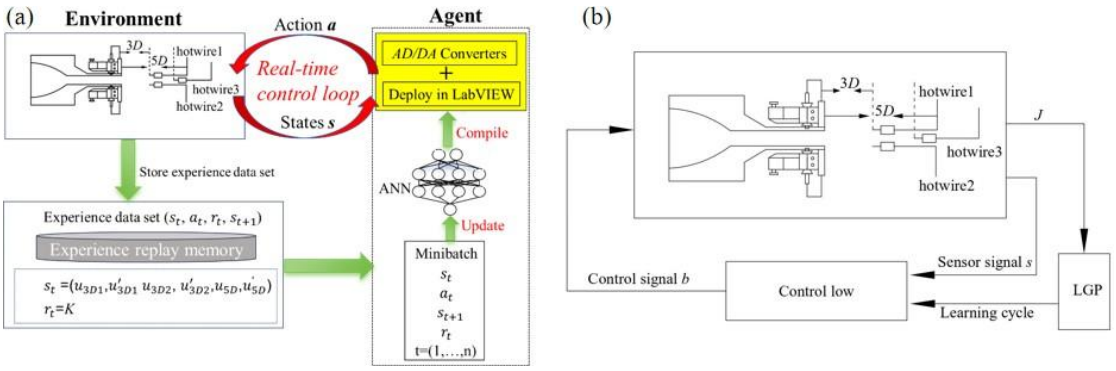

Figure 2 Schematic diagram of artificial intelligence algorithm. (a) DRL; (b) LGP.

The GP and DRL shared the same hardware facilities. LGP is a method based on the mechanism of natural selection, which optimizes control laws through population evolution. Its execution process begins with the generation of a large-scale initial population, where each individual represents a control law. Individuals in the population are then screened through fitness evaluation, with high-fitness individuals selected as parents. Subsequently, a new generation of population is created through elitism, mutation, and crossover. This process iterates continuously until convergence. Figure 2(b) shows the schematic diagram of the GP control system. The optimization framework follows methodologies established in Wu et al.[15] and Zhou et al.[20]. The J is defined as $J = U_{5D} / \overline{U_j}$ or 1-K. In this approach,

LGP serves as a regression-based solver to derive control laws, aiming to minimize the cost function *J*. The iterative process comprises four sequential phases: (i) initializing a population of 100 candidate control laws, (ii) quantifying the performance metrics of each law through *J*, (iii) verifying convergence of the minimum *J*, and (iv) generating a new generation of 100 control laws via evolutionary operations informed by predecessor fitness evaluations. This generational refinement cycle continues until convergence criteria are met.

4.Results and discussion

4.1 Jet mixing enhancement using one minijet

The Critic and Actor network is configured as two layers with 512 neurons and a single layer with 512 neurons, respectively. Figure 3(a) illustrates the learning process of DRL using one pulsed radial minijet, showing that the optimal control law was obtained at the 138th episode, with the maximum *K* of 0.44. Meanwhile, figure 3(b) shows the best individual in each generation in the learning process of LGP. In the first generation, the LGP achieved a relatively good control law with *K*=0.415. The maximum *K* of 0.44 was achieved at the 5th generation. By comparing the learning processes and control effects of the two algorithms, it is evident that both ultimately achieved the similar control performance (*K*=0.44). However, the DRL demonstrated a more significant advantage in searching speed.

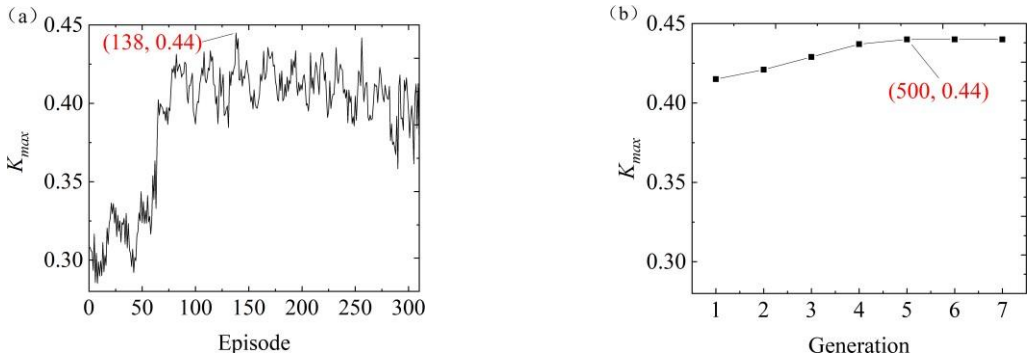

Figure 3 single minijet learning curve. (a) DRL; (b) LGP.

4.2 Jet mixing enhancement using two minijet

Figure 4(a) illustrates the learning process of DRL using two symmetrical pulsed minijets, showing that the algorithm reached the maximum $K = 0.43$ after 371 training iterations. Figure 4(b) shows the learning process of LGP for driving two minijets. The best *K* in each generation gradually increased, with a significant rise observed between the 3rd and 5th generations. Eventually, the *K* was converged to 0.43 at 5th generation. For the case of two symmetrical pulsed minijets, both DRL and LGP successfully identified the optimal control strategy, achieving a similar *K* of 0.43. However, DRL completed the optimization in 371 test iterations, while LGP required 500 tests. These results demonstrate that DRL exhibited a more significant advantage in search efficiency than LGP.

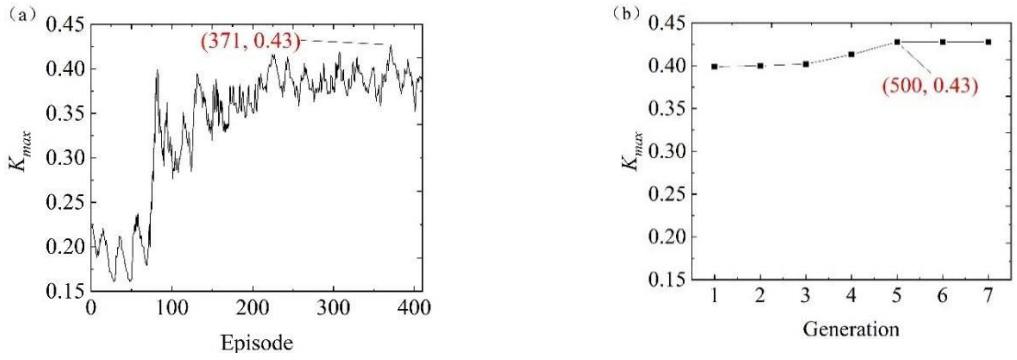

Figure 4 two symmetrical pulsed minijets learning curve. (a) DRL; (b) LGP.

Table 1 compared the performance of different control configurations. The DRL algorithm converged much faster than LGP in searching the optimal control law for both control cases using one or two minijets. The optimal control performance achieved by different control cases are similar, with $K \approx 0.43$-0.44. It indicates that the using of two symmetric pulsed minijets could not improve the control performance further as compared to the single pulsed minijet, but may increase the difficulty in searching the optimal control law.

Table 1 Different control modes

	Control Algorithm	Number of minijets	# of test before converged	K	Averaged duty cycle
Case 1	DRL	One minijet	138	0.44	15%
Case 2	DRL	Two minijets	371	0.43	6.5% and 13%
Case 3	LGP	One minijet	500	0.44	15%
Case 4	LGP	Two minijets	500	0.43	7% and 12%

4.3 Flow field change under different actuations

To illustrate the dynamic changes in the flow field, figure 5 presents the flow visualization results under the four control modes. The jet exhibits a symmetric pattern in the uncontrolled case (figure5a and 5d). In contrast, the jet column shows asymmetric pattern under different control modes (figure 5c, 5e, 5d, 5f), along with a significantly reduced potential core length. Additionally, the jet displays a distinct oscillatory motion.

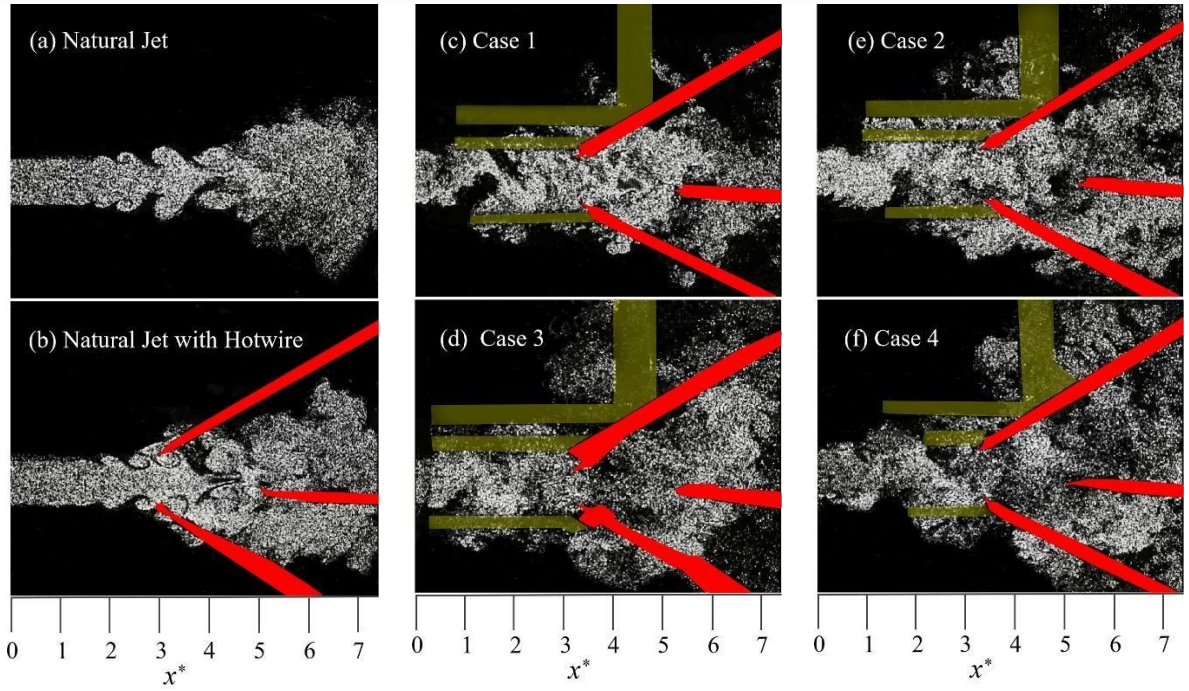

Figure 5 Flow diagram of different control modes. The red and yellow regions indicate the shadows or reflection of the hotwire supporting facilities.

The phase differences in the velocity signals measured by hotwire 1 and 2 at the $3D$ position under the four control modes were also calculated. It is found that the phase differences in Case 1 (DRL-one), Case 2 (DRL-two), Case 3 (LGP-one), and Case 4 (LGP-two) are 168°, 120°, 144°, and 120°, respectively. This result indicates that the jet undergoes significant oscillatory motion under the optimal actuation. This phenomenon is consistent with that reported in Wu et al. (2018), where the flapping motion in the jet column was found to be the main mechanism for jet mixing for single pulsed jet actuation. To further elucidate the mechanisms of different control modes, the duty cycles of actuation signals in the four control modes were also investigated (table 1). It can be found that the duty cycles of Case 1 and Case 3 are the same, the duty cycle of the Case 2 and Case 4 are similar. This indicates that the control law found by DRL and LGP are quite similar, thus the control mechanism is the same.

5. Conclusions

This study developed a real-time artificial intelligence control system utilizing deep reinforcement learning (DRL) for jet mixing enhancement, employing three hot-wire sensors in the wake and one or two pulsed minijet actuators near the jet exit. The principal findings can be summarized as follows:

(1) The DRL demonstrates significant speed advantages over LGP in searching the optimal control law in jet mixing control. The DRL achieved comparable control performance while significantly reducing iteration counts in both single- and dual-minijet configurations.

(2) Increasing the number of minijets does not improve control effectiveness, but rather increases optimization difficulty. The control performance of one minijet (K=0.44) and two minijets (K=0.43) is comparable, indicating that adding actuators does not significantly enhance jet mixing performance. Moreover, the number of iterations for DRL in the two minijets significantly increases, demonstrating that increasing the number of actuators negatively impacts the algorithm's search efficiency.

(3) The control mechanisms of different algorithms are fundamentally similar, both relying on jet oscillation to enhance mixing. Phase difference analysis of the flow field and flow visualization results show that the optimal control strategies of both DRL and LGP enhance jet mixing by inducing jet oscillation to shorten the potential core length.

6. Acknowledgments

Authors wishes to acknowledge supports from NSFC (12202123), Research Grants Council of the Shenzhen Government (GXWD20220811152110003), Aeronautical Science Foundation of China (2024M033077002), and Basic and Applied Basic Research Fund of Guangdong Province (No. 2020A1515110246)

7. References

[1] Zhao, M, Li, Q and Ye, T. Investigation of an optimal pulsed jet mixing and combustion in supersonic crossflow. *Combustion and Flame*, 2021. **227:** 186-201.
[2] Bradbury, L J S and Khadem, A H. The distortion of a jet by tabs. *Journal of Fluid Mechanics*, 1975. **70:** 801-813.
[3] Hussain, F and Husain, H S. Elliptic jets. Part 1. Characteristics of unexcited and excited jets. *Journal of Fluid Mechanics*, 1989. **208:** 257-320.
[4] Alkislar, M B, Krothapalli, A and Butler, G W. The effect of streamwise vortices on the aeroacoustics of a Mach 0.9 jet. *Journal of Fluid Mechanics*, 2007. **578:** 139-169.
[5] Zaman, K B M Q and Hussain, A K M F. Vortex pairing in a circular jet under controlled excitation. Part 1. General jet response. *Journal of Fluid Mechanics*, 1980. **101:** 449-491.
[6] Samimy, M, Kim, J H, Kastner, J, Adamovich, I and Utkin, Y. Active control of high-speed and high-Reynolds-number jets using plasma actuators. *Journal of Fluid Mechanics*, 2007. **578:** 305-330.
[7] Ho, C-M and Gutmark, E. Vortex induction and mass entrainment in a small-aspect-ratio elliptic jet. *Journal of Fluid Mechanics*, 1987. **179:** 383-405.
[8] Wiltse, J M and Glezer, A. Manipulation of free shear flows using piezoelectric actuators. *Journal of Fluid Mechanics*, 1993. **249:** 261-285.
[9] Davis, M R. Variable control of jet decay. *AIAA Journal*, 1982. **20:** 606-609.
[10] Yang, H and Zhou, Y. Axisymmetric jet manipulated using two unsteady minijets. *Journal of Fluid Mechanics*, 2016. **808:** 362-396.
[11] Zhou, Y, Du, C, Mi, J and Wang, X W. Turbulent Round Jet Control Using Two Steady Minijets. *AIAA Journal*, 2012. **50:** 736-740.
[12] Brunton, S L and Noack, B R. Closed-Loop Turbulence Control: Progress and Challenges. *Applied Mechanics Reviews*, 2015. **67:**
[13] Choi, H, Jeon, W-P and Kim, J. Control of Flow Over a Bluff Body. *Annual Review of Fluid Mechanics*, 2008. **40:** 113-139.
[14] Scott Collis, S, Joslin, R D, Seifert, A and Theofilis, V. Issues in active flow control: theory, control, simulation, and experiment. *Progress in Aerospace Sciences*, 2004. **40:** 237-289.
[15] Rabault, J, Kuchta, M, Jensen, A, Réglade, U and Cerardi, N. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics*, 2019. **865:** 281-302.
[16] Wu, Z, Fan, D, Zhou, Y, Li, R and Noack, B R. Jet mixing optimization using machine learning control. *Experiments in Fluids*, 2018. **59:**
[17] Zhou, Y, Fan, D, Zhang, B, Li, R and Noack, B R. Artificial intelligence control of a turbulent jet. *Journal of Fluid Mechanics*, 2020. **897:**
[18] Ren, F, Wang, C and Tang, H. Bluff body uses deep-reinforcement-learning trained active flow control to achieve hydrodynamic stealth. *Physics of Fluids*, 2021. **33:**
[19] Fan, D, Yang, L, Wang, Z, Triantafyllou, M S and Karniadakis, G E. Reinforcement learning for bluff body active flow control in experiments and simulations. *Proceedings of the National Academy of Sciences*, 2020. **117:** 26091-26098.