13th International Symposium on Turbulence and Shear Flow Phenomena (TSFP13)
Montreal, Canada, June 25-28, 2024

DEEP REINFORCEMENT LEARNING MODEL FOR REDUCING DRAG IN TURBULENT CHANNEL FLOW BASED ON VORTEX DEVELOPING AND DECAYING PROCESS

Taisei Sato
Department of Mechanical Systems Engineering
Tokyo University of Agriculture and Technology
2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
satot@st.go.tuat.ac.jp

Tomohiro Nimura
Department of Mechanical Systems Engineering
Tokyo University of Agriculture and Technology
2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
nimura@go.tuat.ac.jp

Akira Murata
Department of Mechanical Systems Engineering
Tokyo University of Agriculture and Technology
2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
murata@cc.tuat.ac.jp

Kaoru Iwamoto*
Department of Mechanical Systems Engineering
Tokyo University of Agriculture and Technology
2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
iwamotok@cc.tuat.ac.jp

ABSTRACT

A model which uses a convolutional neural network and a fully connected layer is proposed by applying the Soft Actor Critic (SAC) (Haarnoja et al., 2018), a deep reinforcement learning algorithm, to blowing and suction control. We performed reinforcement learning training to obtain further drag reduction rates in turbulent channel flow. A control law was adopted for the input for learning is the wall-normal velocity and the second invariant of the velocity gradient tensor on a detection plane parallel to the wall, and the output yields multiple coefficients that determine the blowing and suction flow on the wall. As training proceeded, the global coefficient of skin friction decreased, which means training proceeded properly. The control law obtained through the learning exhibits strong suction for regions where vortices move away from the wall, while it does weak blowing for regions where vortices move closer to the wall. The obtained control law is evaluated using the drag reduction rate, the net energy saving rate, and the gain. When the obtained control law was applied to turbulent channel flow, the drag reduction rate did not exceed that of the V-control (Choi et al., 1994). However, to uniformly adjust the velocities of blowing/suction returned by the obtained law, we achieved results that show the same drag reduction rate as the V-control with less input energy than the V-control. The gain of the obtained law is higher than the V-control, and higher control effects can be obtained with less input energy.

BACKGROUND AND OBJECTIVES

Controlling turbulence and reducing frictional drag are expected to solve industrial or environmental issues, such as designing highly efficient and safe machinery, ensuring high quality in manufacturing processes, etc. On the surface of transport equipment, including airplanes, streamwise vortices are created, which cause large turbulent frictional drag. There have been many studies on methods of turbulence control, which can be classified broadly into passive control and active control. In this study, we focus on feedback control, one of the active control methods, in which the control is varied according to the conditions of the turbulent field. As to the V-control, a typical example of feedback control, streamwise vortices are suppressed by imposing flow velocity of blowing and suction, which has an opposite phase of wall-normal velocity at the detection plane. Choi et al. achieved a drag reduction rate of 25 % at a bulk Reynolds number of 5000.

Recently, reinforcement learning, one of the machine learning methods, has been used to control turbulence. Reinforcement learning is a learning method in which the learner executes actions according to specific rules, aiming to discover a control law that maximizes long-term reward. It also has a feature that allows the learner to collect data and conduct the learning process in parallel. Since the invention of deep reinforcement learning, which uses neural networks, it has become possible to express control laws with a high degree of freedom. This has been applied to diverse fields, such as robot control and chess. As to turbulence control, Sonoda et al. (2023) applied deep reinforcement learning to blowing and suction control for fully developed turbulent channel flow and obtained a drag reduction rate of 31 % at a bulk Reynolds number of 5000. However, they determined control law based on the limited information at the detection plane, and further use of physical quantities is required. According to the control law they obtained, it physically means that strong blowing acts to a high-speed streak, whereas strong suction to a low-speed streak requires much energy input.

This study aims to obtain further drag reduction rates by focusing not only on the flow velocity at the detection plane but also on vortex structures near the wall. This paper reports a new control law and drag reduction effect on the blowing and suction control in fully developed turbulent channel flow.

CALCULATION METHOD

In this study, the flow field of turbulent channel flow was calculated by direct numerical simulation (DNS). The continuity and Navier-Stokes equations are the governing equations. The friction Reynolds number is set to $Re_\tau = u_\tau \delta / \nu = 110$, where u_τ is the friction velocity based on mean wall shear stress, δ channel half width, and ν the kinematic viscosity. The computational domain is shown in Figure 1. Here, x, y, and z denote the streamwise, spanwise, and wall-normal directions, respectively. The computational domain in each direction is (L_x, L_y, L_z) = ($1.25\pi\delta$, 2δ, $0.5\pi\delta$). The number of computational cells is (N_x, N_y, N_z) = (48, 108, 48), and grid resolution is (Δx^+, Δy^+, Δz^+) = (9.00, 0.55-4.80, 3.60). Here, $()^+$ denotes the value normalized by u_τ and ν.

The flow velocity of blowing and suction on the walls is determined by the control law f, which reinforcement learning decides using physical quantities at a virtual sensor surface

* corresponding author

called the "detection plane" located at $y_d^+ = 15$. Schematic of the control method is shown in Figure 2. The control law f is expressed as follows:

$$\varphi(x,0,z) = f(v'(x,y_d,z), II(x,y_d,z)) \quad (1)$$

where φ is the flow velocity of blowing and suction on the walls, v' is the wall-normal velocity at the detection plane, and II is the second invariant of the velocity gradient tensor at the detection plane. The control law determines the relationship between physical quantities at the detection plane and the flow velocities of blowing and suction on the walls, which are at the same location with respect to x and z. Thus, velocities of blowing and suction at each grid point on the walls is determined only from the physical quantities at the detection plane directly above it.

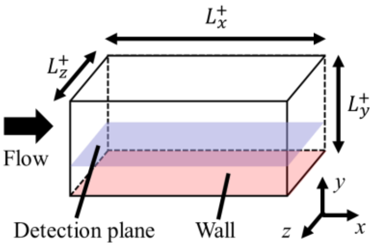

Fig. 1 Calculation condition in the present study.

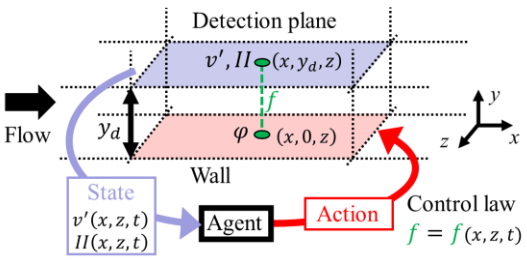

Fig. 2 Visualization of blowing and suction φ at the walls from physical quantities at the detection plane located at $y_d^+ = 15$.

DEEP REINFORCEMENT LEARNING

Learning Model

Deep reinforcement learning is a machine learning method in which a learner called an "agent" aims to maximize the total reward r within a predefined learning period in an "environment" that changes according to specific rules. Specifically, by performing multiple trials called "episodes," the agent discovers the strategy that maximizes the sum of the rewards r obtained in each episode. Figure 3 shows the schematic of a model in which SAC is applied to blowing and suction control in a fully developed turbulent channel flow. In this model, the state s is defined as two physical quantities, $s1$ and $s2$, as shown in Fig. 3. One is the field of the wall-normal velocity v' at the detection plane, and the other is the field of the second invariant of the velocity gradient tensor II at the detection plane. The action a is defined as the seven coefficients ω_1, ω_2…, ω_7, which constitute the control law f. In this paper, the control law f is defined as follows:

$$\varphi = \tanh\left(\tanh\left([s1 \quad s2]\begin{bmatrix}\omega_1 & \omega_2\\ \omega_3 & \omega_4\end{bmatrix}\right)\begin{bmatrix}\omega_5\\ \omega_6\end{bmatrix} + [\omega_7]\right) \quad (2)$$

which is inspired by the two fully connected layers (FC). The

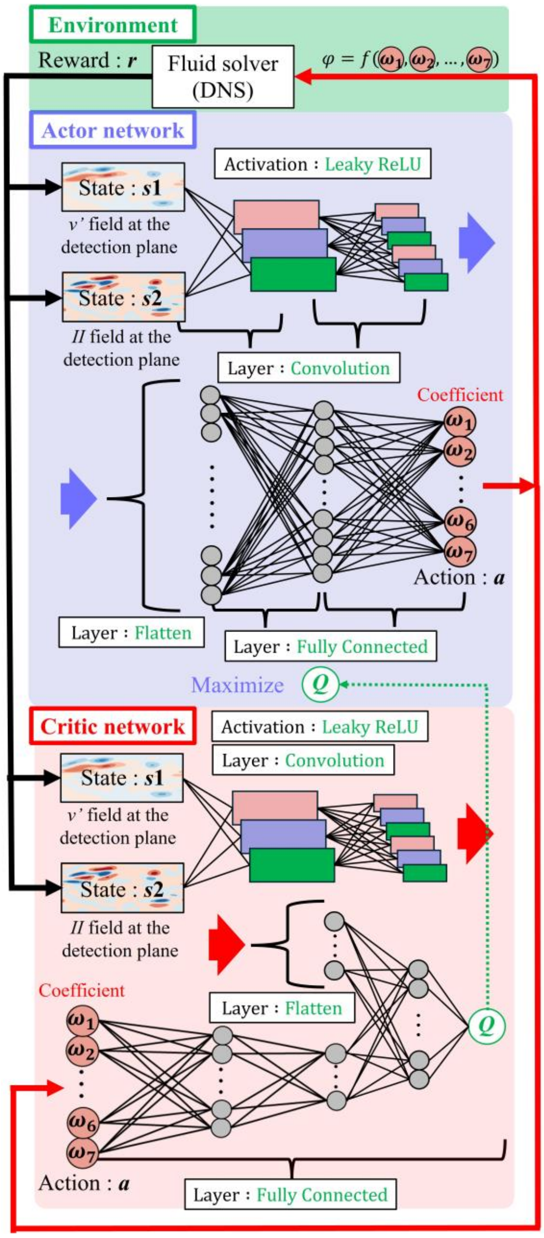

Fig.3 Schematic of SAC model. Environment corresponds to the green region, Actor network to the blue region and Critic network to the red region. The state s calculated by fluid solver (DNS) is input to the Actor network, which outputs action a, then the control f will be determined, and state s will be calculated again. To obtain a higher reward r, Critic network calculated the Q-value based on the state s and action a, and updates Actor network to obtain a higher reward r.

hyperbolic tangent function is used as activation function, which is expected to fix flow velocity of blowing and suction between –1 to 1 and prevent divergence of DNS. The agent consists of two neural networks called Actor and Critic. Actor corresponds to the policy which determines the action a based on state s, which is located on the blue region in Fig.3. Specifically, it consists of both a convolutional neural network (CNN) and an FC and learns to maximize the Q-value, which indicates the expected value of the reward r obtained in each episode. The reward is calculated as follows:

$$r = -C_f - \Delta \quad , \tag{3}$$

where Δ is the flow rate adjusted when setting the net frow rate of blowing and suction to zero. Critic corresponds to the evaluator which estimates an accurate *Q*-value based on state *s* and action *a*, which is located on the red region in Fig.3. Specifically, it consists of both a CNN and FC, and learns to minimize the error between the reward *r* actually obtained and *Q*-value. The environment corresponds to the control law *f* and the DNS, which is located on the green region in Fig.3. From the above, Actor network and Critic network are mutually learned so that the control law *f* is determined based on a long-term perspective.

Learning Conditions

In this model, 400 episodes of DNS, each with period $\Delta T^+ = 550$, were conducted to train the Actor network and Critic network. The Adam method was used for updating each network, with a learning rate of 0.001 for Actor network, 0.002 for Critic network. The coefficients of the control law *f* were determined for each $\Delta t^+ = 0.55$, and the flow velocity of blowing and suction at each position on the wall were updated using equation (2). Since the field of the detection plane changes from moment to moment, the coefficients of the control law *f* also change accordingly. Note that the control law *f* shown in equation (2) can change every $\Delta t^+ = 0.55$. While updating the flow velocity of the blowing and suction, a linearly interpolated value was given.

RESULTS AND DISCUSSION

Results of deep reinforcement learning

Figure 4 shows the learning process of reinforcement learning. The horizontal axis is the time progressed in DNS, and the vertical axis is C_f, corresponding to the reward *r* of reinforcement learning. The color contour indicates the number of episodes, with blue representing the early stage of learning and red representing the end of learning. The result of the V-control is shown as a black line for comparison. Figure 4 shows that the C_f decreased as the episode proceeded overall. It shows that although learning failed and the C_f took a large value in some episodes, the C_f decreased around episode 300, shown in orange, which shows less value than the V-control. The adjustment of flow rate Δ also decreased with each episode, confirming that the control input determined by the model was imposed on the wall as we intended. Since the reward was the sum of $-C_f$ and $-\Delta$, the reinforcement learning objective, maximizing the sum of the rewards *r* obtained in each episode, was achieved.

Figure 5 shows a visualization of the control law determined by equation (2). Figure 5(a) shows the control law *f* with the episode 300, where C_f decreased the most. Due to space limitations, although only the control law for $t^+ = 100$ is shown in Fig. 5(a), the control law remained almost the same during the episode 300. Figure 5(b) shows the control law of the V-control for comparison. The horizontal axis is the wall-normal velocity *v*' at the detection plane $y_d^+ = 15$, and the vertical axis is the second invariant of the velocity gradient tensor *II* at the same plane. The color contour indicates the flow velocity of blowing and suction on the walls. Figure 5 shows that the obtained control law is almost the same as V-control

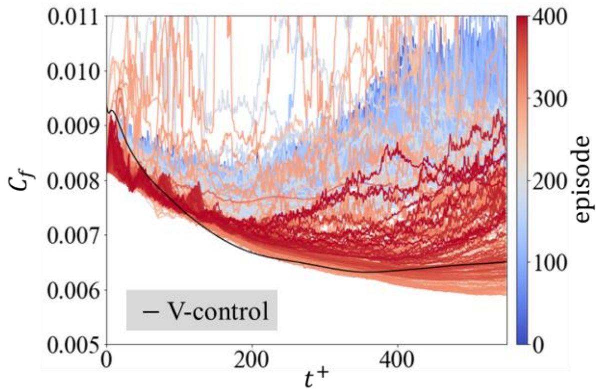

Fig.4 Learning process of C_f at each episode. Color contour indicates episode.

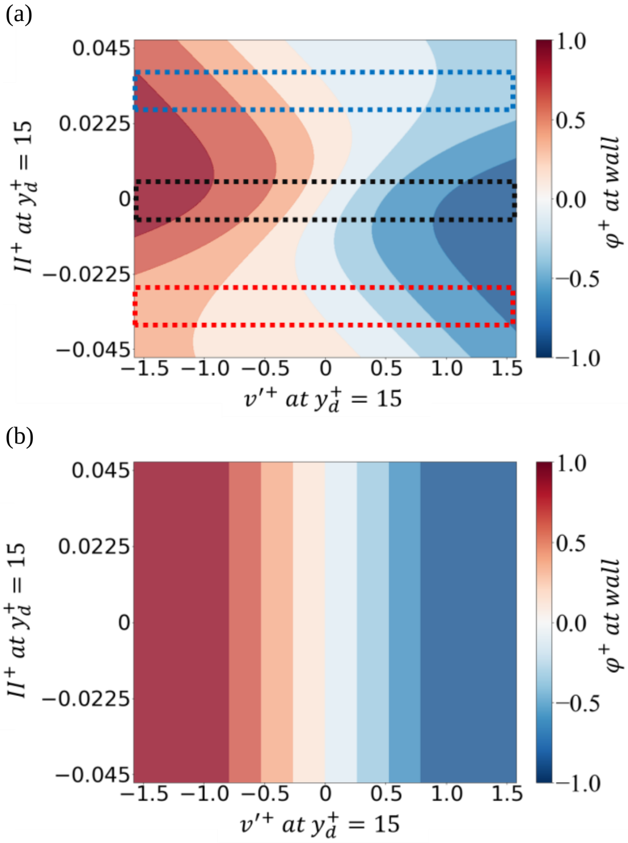

Fig. 5 Control law which determines the relationship between physical quantities at the detection plane and φ at wall: (a) Control law with the episode in which C_f decreased the most at $t^+ = 100$; (b) Control law of V-control.

when *II* at the detection plane is 0, indicated by the black dotted box. It means that for regions where no vortex exists, the same control as V-control is applied. Figure 5(a) shows that the obtained control law indicates strong suction when *II* at the detection plane is –0.03 and *v'* at the detection plane is positive; on the other hand, it indicates weak blowing when *II* at the detection plane is –0.03 and *v'* at the detection plane is negative. This means that the obtained law indicates strong suction for regions where vortices move away from the wall; on the other hand, for regions where vortices move closer to the wall, the obtained law indicates weak blowing. As to the region indicated by the blue dotted box in Fig. 5, where *II* at the detection plane is 0.03, it could be said the opposite of what the red dotted box indicates, but this is thought to be because the net flow rate is adjusted to zero.

Figure 6(a-c) shows the input and output of SAC models in the x–z planes at t^+ = 100. Figure 6(a, b) show the v' and II field at the detection plane (y_d^+= 15), respectively. Figure 6(c) shows the velocity field of blowing and suction φ at the wall (y = 0) calculated by the obtained control law shown in Fig. 5(a). Figure 6(d) shows the difference of φ between the obtained control law and V-control, which is calculated as follows:

$$\varphi(x,z)_{\text{Obtained law}} - \varphi(x,z)_{\text{V control}} \quad , \qquad (4)$$

which means that when the value of equation (4) is positive, φ, determined by the obtained control law, is larger than that of the V-control, and when negative, it is smaller than that of the V-control. In Fig.6(a) and (b), the red dotted regions show that v' at the detection plane is positive whereas II at the detection plane is negative, where Fig. 6(c) indicates strong suction. Conversely, the blue dotted regions show that both v' at the detection plane and II at the detection plane are negative in Fig.6(a) and (b), where Fig.6 (c) indicates weak blowing. We could confirm that for regions where vortices move away from the wall, strong suction is imposed on the walls, whereas for regions where vortices move closer to the wall, weak blowing is imposed on the walls. The obtained control law shown in Fig. 5(a) is physically consistent with the flow velocities of these blowing and suctioning on the walls. In Fig. 6(a) and (b), the black dotted regions show that v' at the detection plane is positive, and II at the detection plane is slightly positive, where Fig. 6(c) indicates strong suction as the obtained law shown in Fig.5(a) indicates. As to Fig. 6(d), the red dotted regions and the blue dotted regions, both of which have a negative value of II at the detection plane, indicate a negative value, which means φ at the wall determined by the obtained control law is smaller than that of the V-control for regions where vortex exists.

Evaluation of the obtained control law

In this section, the control law obtained through reinforcement learning is confirmed. Since the SAC model optimizes Actor network which determines the relationship between physical quantities at the detection plane and φ at the walls, we performed DNS using Actor network with the episode 300 when C_f most decreased. At this time, we fixed its weight, which means we performed DNS incorporating the obtained control. The initial field is the same as those used in reinforcement learning. The calculation time is set to t^+ > 50000, in which statistics will be converged and the other computational conditions are the same as those used in learning.

To take into account the energy needed in applying a control law, we investigate not only the drag reduction rate R_D but also the net energy saving rate S and the gain G, which Kasagi et al. (2009) suggested. These are defined as follows:

$$R = \frac{W_{p,\text{Dean}} - W_p}{W_{p,\text{Dean}}} \qquad (6)$$

$$S = \frac{W_{p,\text{Dean}} - (W_p + W_a)}{W_{p,\text{Dean}}} \qquad (7)$$

$$G = \frac{W_{p,\text{Dean}} - W_p}{W_a} \qquad (8)$$

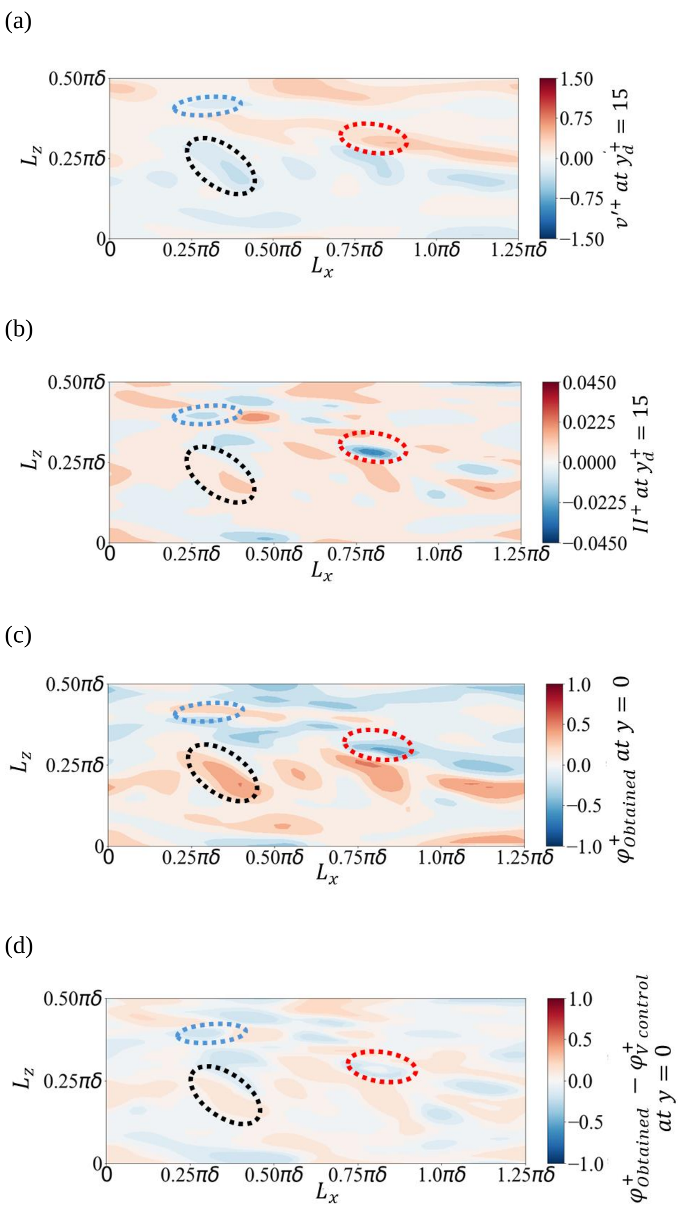

Figure 6. Input of SAC model at t^+ = 100: (a) field of v' at the detection plane; (b) field of II at the detection plane. Output of SAC model at t^+ = 100: (c) field of φ at the lower wall. (d) the difference of φ between the obtained law and V-control at the lower wall.

Here, $W_{p,\,Dean}$ and W_a are the pumping work of fluid flow and the input energy by imposing blowing and suction on the wall respectively. $W_{p,\,Dean}$ and W_a are defined as follows:

$$W_p = -\frac{\overline{\partial P}}{\partial x} u_b \qquad (9)$$

$$W_a = \frac{1}{L_x L_z}\int_0^{L_x}\int_0^{L_z} (p_{wall}\varphi + \varphi^3)\,dz\,dx \quad , \qquad (10)$$

where $-\partial P / \partial x$, u_b, and p_{wall} is the mean pressure gradient, the bulk mean velocity, and the pressure on the wall, respectively. $W_{p,\,\text{Dean}}$ is estimated by using following equation reported by Dean (1978):

$$C_{f,\,\mathrm{Dean}} = 0.073Re_{bulk}^{-1/4} \tag{11}$$

Figure 7 shows the results of the obtained control law and the V-control in terms of R_D, S, and G. The horizontal axis is the input energy on the walls shown in equation (10). The vertical axis is R_D, S, and G in the order of Fig. 7(a), Fig. 7(b), and Fig. 7(c). The horizontal and vertical axes are time-averaged values after $t^+ = 5000$ when the statistics are stable. In all figures, the red dot indicates the result of the obtained control law, and the blue dot indicates the result of the V-control for comparison. The red diamonds, squares, downward triangles, and upward triangles indicate the results in which the quantity of blowing and suction on the walls decided by the obtained control law are multiplied by 0.8, 0.6, 0.4, and 0.2, respectively. The quantity of blowing and suction was adjusted by uniformly multiplying the velocities determined from the obtained control law by a constant. In addition, the results of the V-control are also shown in the same manner, with the blue diamonds, squares, downward triangles, and upward triangles indicate the results in which the coefficient ω of the V-control shown in equation (4) is set to 0.8, 0.6, 0.4, and 0.2, respectively. Figure 7(a) shows that the R_D of the obtained control law, indicated by the red dot, is lower than that of the V-control indicated by the blue dot, although it requires more input energy. Therefore, the obtained control law is inefficient comparing to the V-control. This is thought to be the obtained control law is only optimized for the flow field during the reinforcement learning, because the time progressed in DNS for reinforcement learning is only 1 / 100 of that for evaluating the obtained control law. The blue dashed line in Fig. 7(a) shows the approximation line with respect to the results of the V-control. For the results of the V-control shown in the figure, there is an approximately proportional relationship between the input energy and R_D. Regarding the obtained control law, the result shown by the dots before adjustment requires more input energy than the V-control. However, in the results shown by red diamonds, squares, downward triangles, and upward triangles obtained by adjusting the quantity of blowing/suction, the input energy is smaller than that of the V-control. In other words, this adjustment does not directly change the input energy. In particular, as a black dotted line indicates, the red diamond, in which the obtained control law is multiplied by 0.8, has the same R_D compared to the blue diamond, in which the coefficient ω of the V-control is set to 0.8, despite the input energy being smaller. Also, the red diamond has almost the same input energy compared to the blue square, which is set to 0.6. Fig. 7(b) shows that the relationship between the input energy and S is almost the same as the one between the input energy and R_D shown in Fig. 7(a). This is because the following relationship holds for R_D, S, and G:

$$S = R_D\left(\frac{G-1}{G}\right) \tag{12}$$

Since G takes a value larger than 100 in all cases, as shown in Fig. 7(c), R_D and S have almost the same value. Figure 7(c) shows that the gain G of the obtained control law is smaller than that of the V-control. On the other hand, as the black dotted line shows, G of the obtained control law which is multiplied by each coefficient are higher than that of the V-control, which means the higher control effect is obtained with a small amount of the input energy.

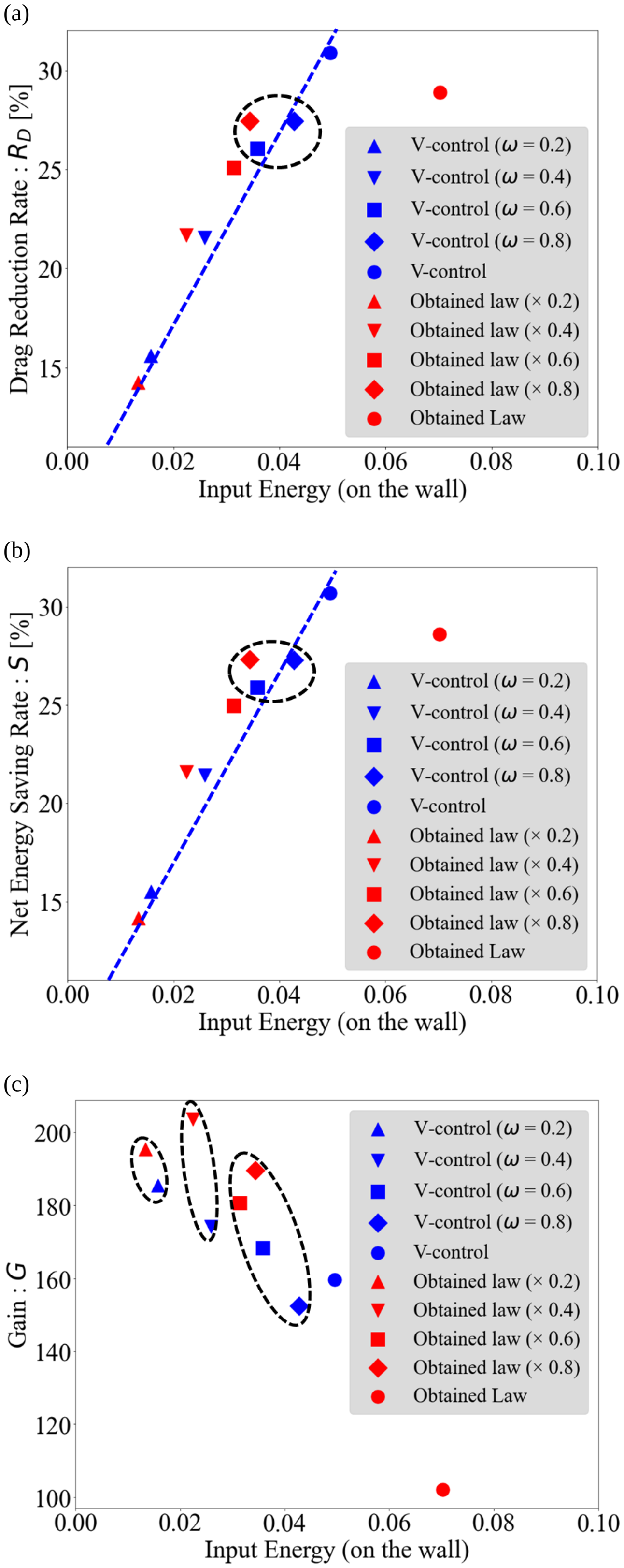

Fig. 7 The results of the obtained control law and the V-control: (a) Drag Reduction Rate, R_D; (b) Net Energy Saving Rate, S; (c) Gain, G

CONCLUSION

Training in reinforcement learning was conducted, and the control law for blowing and suction was derived using wall-normal velocity and the second invariant of the velocity gradient tensor field at the detection plane. The control law exhibits strong suction for regions where vortices move away from the wall; on the other hand, it does weak blowing for regions where vortices move closer to the wall. The control law obtained by reinforcement learning was applied to turbulent channel flow and evaluated using drag reduction rate R_D, net energy saving rate S, and gain G in comparison with the V-control. R_D and S of the obtained control law are lower than that of the V-control despite requiring more input energy on the wall, which indicates the obtained control law is inefficient compared to the V-control. This is thought to be the obtained control law is only optimized for the flow field during the reinforcement learning. When the quantity of blowing and suction on the walls decided by the obtained control law is multiplied by 0.8, the result shows the opposite; it needs less input energy despite R_D and S being the same as that of the V-control under an identical condition. Since G is larger than the V-control, high control effects can be obtained with less input energy.

REFERENCES

Haarnoja, T., et al., 2018, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor", *International conference on machine learning*. PMLR, pp. 1861-1870.

Choi, H., Moin, P., and Kim, J., 1994, "Active turbulence control for drag reduction in wall-bounded flows", *Journal of Fluid Mechanics* Vol. 262, pp. 75-110.

Sonoda, Takahiro, et al., 2023, "Reinforcement learning of control strategies for reducing skin friction drag in a fully developed turbulent channel flow", *Journal of Fluid Mechanics* Vol. 960, A30.

Kasagi, N., Hasegawa, Y, and Fukagata, K., 2009, "Toward cost-effective control of wall turbulence for skin friction drag reduction',' *In Advances in Turbulence XII: Proceedings of the 12th EUROMECH European Turbulence Conference, Marburg, Germany*, pp. 189-200.

Dean, Roger Bruce., 1978, "Reynolds number dependence of skin friction and other bulk flow variables in two-dimensional rectangular duct flow." *Journal of Fluid Engineering* 100, pp. 215–223.