REINFORCEMENT LEARNING FOR REDUCTION OF SKIN FRICTION DRAG IN A FULLY DEVELOPED TURBULENT CHANNEL FLOW

Takahiro Sonoda
Department of Mechanical Engineering
The University of Tokyo
Komaba 4-6-1, Meguro-ku, Tokyo, Japan
taka1030@iis.u-tokyo.ac.jp

Zhuchen Liu
Department of Mechanical Engineering
The University of Tokyo
Komaba 4-6-1, Meguro-ku, Tokyo, Japan
zhuchen@iis.u-tokyo.ac.jp

Toshitaka Itoh
Institute of Industrial Science
The University of Tokyo
Komaba 4-6-1, Meguro-ku, Tokyo, Japan
titoh@iis.u-tokyo.ac.jp

Yosuke Hasegawa
Institute of Industrial Science
The University of Tokyo
Komaba 4-6-1, Meguro-ku, Tokyo, Japan
ysk@iis.u-tokyo.ac.jp

ABSTRACT

Recently, reinforcement learning has attracted much attention in flow control because of its unique advantage for obtaining a long-term optimal policy. So far, it has been applied only to low-dimensional flows with simple actuations of limited degrees of freedom. In this study, we apply reinforcement learning to wall turbulence control in order to determine the complex spatio-temporal distribution of wall blowing and suction for reducing skin friction drag based on sensing information at a certain distance from the wall. It is demonstrated that reinforcement learning successfully finds an effective control policy showing non-linear relationship between the sensing information and the control input, and achieves better drag reduction performance than that obtained by the well-known opposition control.

BACKGROUND

While turbulence has the advantages of significantly enhancing heat and mass transfer, it also increases friction drag and noise. Therefore, smart control of turbulence is a key technology in engineering so as to promote its advantages with mitigating its drawbacks. Since turbulence is an extremely complex physical phenomenon with strong non-linearity and multi-scale nature, however, it is quite challenging to develop an effective control law to achieve various objectives encountered in engineering flows (Brunton et al, 2015).

With the rapid development of neural networks in recent years, reinforcement learning, which maximizes the long-term reward sum based on rewards obtained from short-term trials, has been attracting attention as a novel and effective approach to find new control strategies. Recently, promising results of reinforcement learning have been reported in a wide range of fields such as robot control (Kober et al, 2013) and Go (Silver et al, 2016), and its applications to flow control have also been attempted. For example, Rabault et al. (2019) applied reinforcement learning to flow control around a cylinder by optimizing wall blowing and suction on the surface of a cylinder. Fan et al. (2020) also considered a similar configuration with rotating small cylinders. Although large drag reduction rates have been confirmed in both the studies, their flow configurations are still simple two-dimensional flows and only control inputs with limited degrees of freedom are considered. Therefore, it remains unclear whether reinforcement learning is effective for the control of complex turbulent flows with high-dimensional control inputs.

The objective of this study is to establish a framework for optimizing a control policy for turbulent flow. Specifically, we attempt to optimize the spatio-temporal distribution of wall blowing and suction in a fully developed turbulent channel flow using reinforcement learning. The number of degrees of freedom of the control input corresponds to the total number of computational grid points on the wall, and therefore ranges from several thousands to several tens of thousands. Also, the control input is allowed to be changed in time. By applying the current framework, it will be shown that the reinforcement learning successfully finds an effective control policy which achieves a higher drag reduction rate than that obtained by the opposition control (Choi et al, 1994).

NUMERICAL SETUP

We consider a fully developed turbulent channel flow as shown in Figure 1. Wall blowing and suction with zero-net-mass-flux are given as a control input in order to reduce the skin friction drag. Direct numerical simulation (DNS) is conducted under a constant flow rate condition with the bulk Reynolds number of $Re_b = \frac{2U_b h}{\nu} = 4646.72$, where U_b is the bulk mean velocity of the flow, h is the channel half-width, and ν is the kinematic viscosity of the fluid. The considered bulk Reynolds number corresponds to the friction Reynolds number of $Re_\tau = \frac{u_\tau h}{\nu} = 180$ in the uncontrolled flow, where u_τ is the friction velocity.

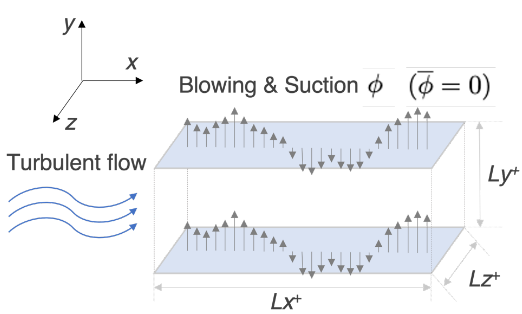


Figure 1. Flow configuration and coordinate system in the present study.

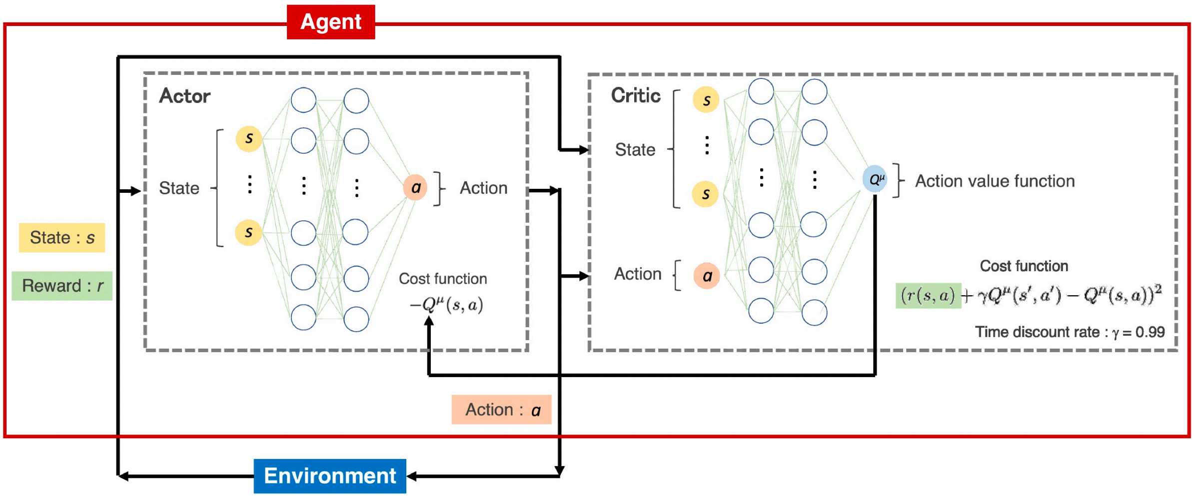


Figure 2. Schematic diagram of the DDPG algorithm employed in the present study

In order to reduce the computational cost for the training the neural networks employed in the present reinforcement learning, we consider the minimal channel (Jimenez and Moin, 1991), which is the smallest domain size for maintaining turbulence. The resultant control policy (control law) is then assessed in a larger domain to verify its control performance. Hereafter, the latter larger domain is referred to a full channel. The domain sizes of the minimum and full channels employed in the present study are $(L_x, L_y, L_z) = (2.67, 2.0, 0.8)$ and $(2.5\pi, 2.0, \pi)$, respectively, where L_x, L_y and L_z are the domain sizes normalized by the channel half height h in the x, y and z directions.

The governing equations for an incompressible fluid, i.e., the continuity and Navier-Stokes equations are solved by a pseudo-spectral method where Fourier transform is applied in the x and z directions, while Chebyshev polynomials are used in the wall-normal direction y. The number of modes used in the minimum and full channels are $(N_x, N_y, N_z) = (16, 65, 16)$ and $(64, 65, 64)$, respectively. The 3/2 rule is used to remove the aliasing errors, so that non-linear terms are evaluated on 1.5 times finer physical mesh.

For time advancement, a fractional step method is used to decouple the pressure term from the Navier-Stokes equation. The second-order Adams-Bashforth and Euler implicit methods are used for the convection and viscous terms, respectively. The time step is set as $\Delta t^+ = 0.06$ and 0.03 for the minimal and full channels. Here, the superscript of + indicates a quantity in the wall unit for the uncontrolled flow. These values ensure that the Courant number is less than unity even with the presence of the control input, i.e., wall blowing and suction considered in the present study.

REINFORCEMENT LEARNING

In reinforcement learning, the agent (learner) receives a state s from the environment (control target) and outputs an action a based on a policy $\mu(a|s)$. In the present study, a is a control input, i.e., wall blowing and suction. Its spatial distribution is determined by the state s, which is the instantaneous streamwise and wall-normal velocity fluctuations u' and v' at $y_d^+ = 15$. This sensing plane height is determined by following the opposition control (Choi et al, 1994), whose maximum drag reduction rate of 25% is obtained at $y_d^+ = 15$ for $Re_\tau = 180$ (Hammond et al, 1998). It should be noted that, in the previous studies, it is commonly assumed that the relationship between the sensing information and the control input is linear, and only the proportional constant is optimized so as to maximize the drag reduction rate. In the present study, however, the control input can be a complex non-linear function of the sensing data, since they are connected through neural network as explained below.

Our aim is to find the effective control policy which connects the state s and the control input a. For this purpose, we use the Deep Deterministic Policy Gradient (DDPG) algorithm (Lillcrap et al, 2015), in which an action a is obtained from a deterministic policy. DDPG consists of two neural networks called actor and critic, where actor is responsible for deciding actions and critic is responsible for evaluating actor. The schematic of the entire networks is shown in Figure 2.

By applying the action (or control input) a, the instantaneous state s is changed to the next state s', and the agent receives a reward r, which is a short-term (instantaneous) friction coefficient C_f obtained from DNS:

$$r = -C_f = -\frac{\overline{\tau_w}}{\frac{1}{2}\rho U_b^2}. \qquad (1)$$

where $\overline{\boldsymbol{\tau_w}}$ is the spatial average of the wall shear stress and $\boldsymbol{\rho}$ is the density of the fluid. We put a negative sign in Eq. (1), so that maximizing reward corresponds to minimizing the drag. By repeating the above interactions with the environment, the agent learns the optimal policy $\mu^*(a|s)$ so as to maximize the long-term drag reduction effects.

Critic estimates the long-term drag reduction based on the instantaneous reward r obtained by applying the control input determined by actor. It is achieved by minimizing the following squared residual of the Bellman equation:

$$L_{Critic} = \{r(s,a) + \gamma Q^{\mu}(s', a') - Q^{\mu}(s,a)\}^2, \qquad (2)$$

where $Q^{\mu}(s,a)$ is the expected total reward when a certain action a is taken under a certain state s. Meanwhile, actor network is trained so as to maximize the expected total reward $Q^{\mu}(s,a)$. During the training, the two networks, i.e., actor and critic, are trained alternatively, so that both networks will be optimized after a number of trials.

In this study, the network is trained by repeating DNS with one episode of $t^+ = 600$, where the superscript + denotes a dimensionless quantity in wall units of the uncontrolled flow.

Each episode starts with the same initial field, and the control input determined by the policy is applied to obtain its short-term reward. In this study, the agent obtains the state of the flow field every $\Delta t^+_{action} = 0.6$, and the control input between the updates is determined by linear interpolation. At the same time, the network is trained every $\Delta t^+_{training} = 0.6$ based on the short-term reward. The network has 1 hidden layer consisting of 8 nodes. ReLU is used as the activation function, while that of the output layer is tanh. These hyper-parameters are found to be optimal in a preliminary survey. The training is repeated until the control policy converges.

RESULTS

Figure 3 shows the time average of C_f in the final period of each episode, i.e., $500 \leq t^+ \leq 600$. It can be seen that C_f gradually decreases with increasing the episode number. It should also be noted that the present C_f is lower than that of the opposition control shown by the red line in Figure 3. This indicates that an effective control strategy is found from the present reinforcement learning (RL).

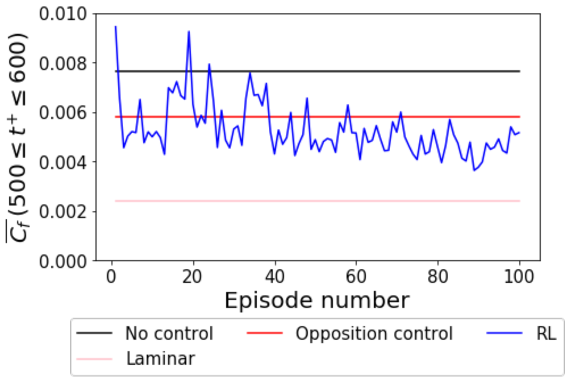
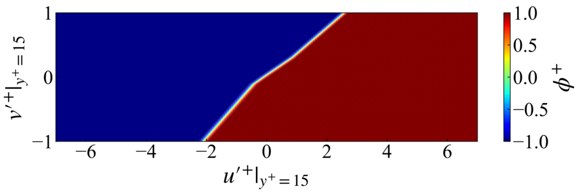


Figure 3. C_f obtained by the present reinforcement learning (RL) as a function of episode number

(a)



(b)

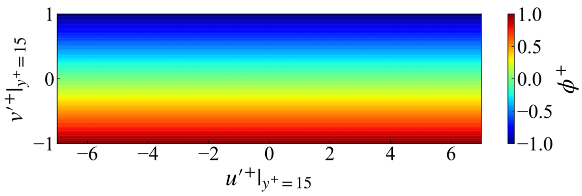
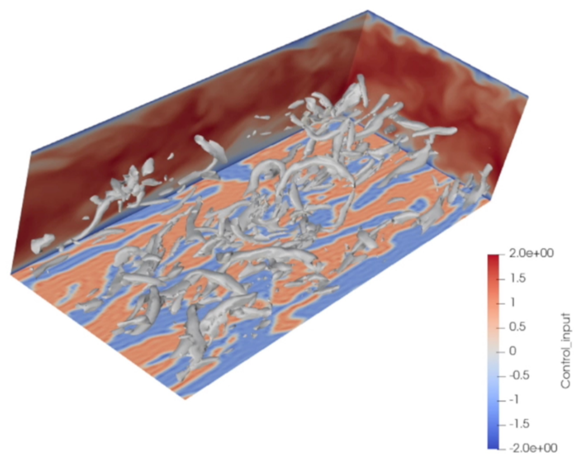


Figure 4. Control policies obtained by the present reinforcement learning (a) and the opposition control (b)

Figure 4 (a) shows the best control policy obtained in episode 89. Specifically, the control input ϕ at the bottom wall is plotted as a function of the state, i.e., u' and v' on the sensing plane at $y^+_d = 15$. It is found that the obtained control policy shows rapid change from blowing (blue) to suction (red). It is in contrast to the control policy of the opposition control shown in Figure (b), where the control input linearly depends on only the wall-normal velocity fluctuation at the sensing plane. It should be emphasized that the neural network employed for actor allows to find such a non-linear relationship between the sensing information and the actuation as shown Figure 4 (a). More specifically, the best policy applies blowing ($\varphi > 0$) when the high-speed fluid approaches the wall ($v' < 0, u' > 0$), while suction ($\varphi < 0$) for the low-speed fluid moves away from the wall ($v' > 0, u' < 0$). When this policy is applied to a full-size channel, about 31% drag reduction is achieved. Considering that the drag reduction rate of the opposition control (Choi et al, 1994) is about 23% in the present flow configuration, it can be concluded that a more effective control policy than the opposition control can be obtained by the present reinforcement learning.

(a)



(b)

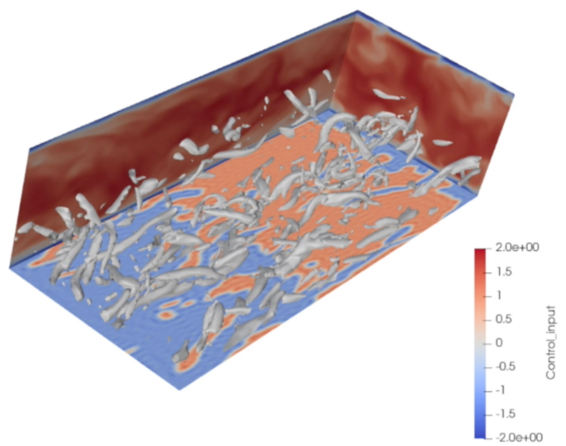


Figure 5. Instantaneous visualizations of the velocity field and the control input applied at the bottom wall at (a) $t^+ = 0.6$ and (b) $t^+ = 20.4$ after applying the best control policy obtained in the present reinforcement learning. Write contours show iso-surfaces of the second invariant of the deformation tensor. Red to blue colors on the bottom wall indicate wall blowing and suction, respectively.

The instantaneous flow fields at $t^+ = 0.6$ and 20.4 after the onset of the control with the best policy obtained in the present reinforcement learning are shown in Figure 5 (a) and (b), respectively. Just after applying the control, at $t^+ = 0.6$, it can be seen that the control input has an elongated structure in the streamwise direction. This corresponds to the near-wall coherent structures. Interestingly, with time passes, the control input transits to a coherent wave-like input as shown in Figure 5 (b), which is almost uniform in the spanwise direction and its streamwise wavelength is equal to the streamwise domain size.

In order to extract a coherent component from the control input, we take the spatial average of the instantaneous control input in the spanwise direction. The resultant spanwise-averaged control input $\widetilde{\phi^+}$ is shown as a function of the streamwise coordinate and the time in Figure 6. It can be seen that the wall blowing and suction switches to the other at a high frequency, while its wave nodes slowly move to upstream. Since the present control policy shown in Figure 4 (a) rapidly switches from strong wall blowing to suction depending on the state *u*' and *v*' at the detection plane, it induces a strong perturbation at the detection plane. This in turn affects the control input in the next time step. Such a feedback between the control input and the flow state at the detection plane should yields the wave-like coherent control input shown in Figure 6.

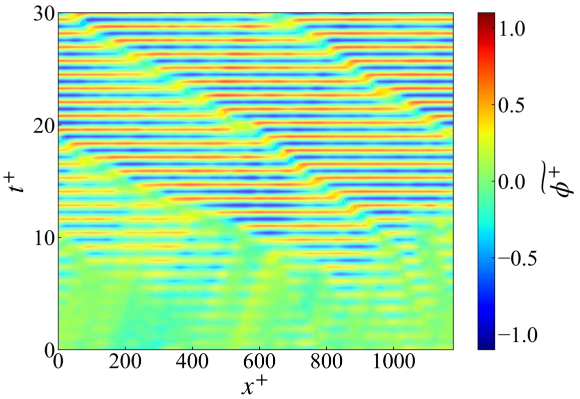


Figure 6. Spanwise-averaged control input as a function of the streamwise coordinate *x* and the time *t*.

It has been reported that drag reduction can be achieved by applying a traveling wave-like wall blowing and suction (for example, Ming et al., 2006, Lieu et al., 2010). Therefore, it is of interest to clarify whether the current drag reduction effects are caused by the coherent control input. For this purpose, we conducted additional computation where only the coherent control input shown in Figure 6, which is uniform in the spanwise direction, is applied. It should be noted that, in this case, the applied control should be regarded as a predetermined control rather than a feedback control, since the control input no longer depends on the instantaneous flow state. It is found that applying the coherent control input only does not lead to drag reduction effects. This indicates that the present control obtained by the reinforcement learning is essentially a feedback control.

CONCLUSIONS

Reinforcement learning was applied to a turbulent channel flow in order to find effective control policies for reducing skin friction drag for the first time. The obtained control policy based on the instantaneous streamwise and wall-normal velocity fluctuations at a sensing plane of $y_d^+ = 15$ achieved 31% drag reduction rate, which is higher than that obtained by the opposition control. The obtained contro policy is characterized by a sharp change from wall blowing and suction, and finding such a complex and non-linear control strategy became possible by a systematic learning framework leveraged by neural networks. The present results demonstrate the effectiveness of reinforcement learning for developing new strategies for turbulence control.

REFERENCES

Brunton, S. L. and Noack, B. R., 2015, "Closed-loop turbulence control: Progress and challenges.", *Applied Mechanics Reviews*, 67.5.

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

Fan, D., Yang, L., Wang, Z., Triantafyllou, M. S. and Kamiadakis, G. E., 2019, "Reinforcement learning for bluff body active flow control in experiments and simulations.", *Proceedings of the National Academy of Sciences*, 117.42, 26091-26098.

Hammond, E. P., Bewley, T. R., & Moin, P., 1998, “Observed mechanisms for turbulence attenuation and enhancement in opposition-controlled wall-bounded flows.”, *Physics of Fluids*, 10.9, 2421-2423.

Jiménez, J. and Moin, P., 1991, "The minimal flow unit in near-wall turbulence." *Journal of Fluid Mechanics*, 225, 213-240.

Kober, J., Bagnell, J. A. and Peters, J., 2013, "Reinforcement learning in robotics: A survey.", *The International Journal of Robotics Research*, 32.11, 1238-1274.

Lieu, B. K., Marref, R. and Jovanovic, M. R., 2010, “Controlling the onset of turbulence by streamwise travelling waves. part 2. direct numerical simulation” *Journal of Fluid Mechanics*, 663, 100–119.

Lillicrap, T. P., Hunt, L. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D.,2015, "Continuous control with deep reinforcement learning." a*rXiv preprint arXiv:1509.02971*.

Min, T., Kang, S. M., Speyer, J. L. and Kim, J., 2006, "Sustained sub-laminar drag in a fully developed channel flow." *Journal of Fluid Mechanics*, 558, 309-318.

Rabault, J., Kuchta, M., Jensen, A., Reglade, U. and Cerardi, N., 2019, "Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control.", *Journal of Fluid Mechanics,* 865, 281-302.

Silver, D, Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D.,2016, "Mastering the game of Go with deep neural networks and tree search.", *Nature*, 529.7587, 484-489.