DISCOVERING DRAG REDUCTION STRATEGIES IN WALL-BOUNDED TURBULENT FLOWS USING DEEP REINFORCEMENT LEARNING

Luca Guastoni
FLOW, Engineering Mechanics,
Swedish e-Science Research Centre (SeRC)
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
guastoni@mech.kth.se

Jean Rabault
Norwegian Metereological Institute
IT Department
Postboks 43, Blindern, 0313 Oslo, Norway
jean.rblt@gmail.com

Philipp Schlatter
FLOW, Engineering Mechanics,
Swedish e-Science Research Centre (SeRC)
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
pschlatt@mech.kth.se

Ricardo Vinuesa
FLOW, Engineering Mechanics,
Swedish e-Science Research Centre (SeRC)
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
rvinuesa@mech.kth.se

Hossein Azizpour
School Elect. Eng. and Comp. Sci.
Swedish e-Science Research Centre (SeRC)
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
azizpour@kth.se

ABSTRACT

The control of turbulent fluid flows represents a problem in several engineering applications. The chaotic, high-dimensional, non-linear nature of turbulence hinders the possibility to design robust and effective control strategies. In this work, we apply deep reinforcement learning to a three-dimensional turbulent open-channel flow, a canonical flow example that is often used as a study case in turbulence, aiming to reduce the friction drag in the flow. By casting the fluid-dynamics problem as a multi-agent reinforcement-learning environment and by training the agents using a location-invariant deep deterministic policy gradient algorithm, we are able to obtain a control strategy that achieves a remarkable 30% drag reduction, improving over previously known strategies by about 10 percentage points.

1 INTRODUCTION

Turbulent flows have been extensively studied in different fields, as they appear in a large number of engineering applications. Experimental studies are prohibitively complicated and costly, so computational fluid dynamics (CFD) has been an important investigation tool for scientific discovery in the last decades. Recently, machine learning models have been used to enhance several aspects of CFD, as summarized by Vinuesa & Brunton (2022). Added to these applications, the possibility to control turbulent flows has been investigated in a number of research works (Duriez et al., 2017; Pino et al., 2022). For this, deep learning tools such as convolutional neural networks (CNNs) have been employed: either directly, for the computation of the control (Park & Choi, 2020), or indirectly with *non-intrusive sensing*, helping to provide the flow measurements that are necessary for closed-loop control systems (Guastoni et al., 2021). The objective of these control studies is to mitigate the losses associated with wall-bounded turbulence, for example reducing the viscous drag on wings and the associated fuel consumption Spalart & McLean (2011). One possible approach to achieve such a reduction is reactive flow control, where a certain actuation is applied, typically to suppress the near-wall turbulent structures. Since turbulence is a stochastic, non-linear phenomenon with high dimensionality, it is difficult to design effective control strategies. Traditional well-established techniques such as opposition control (Choi et al., 1994) are typically based on ad-hoc hand-engineered control laws and they only provide a limited drag reduction.

In this regard, reinforcement learning represents a promising framework to discover novel turbulent flow control strategies because of its end-to-end formulation: all the features of the flow are included in the environment and as long as the simulation is physically-accurate, the agent can experience all consequences of the actions performed on the environment. Starting from the control of the wake behind a cylinder by Rabault et al. (2019) and until very recently, most application focused on two-dimensional environments at relatively low Reynolds numbers. The Reynolds number provides an intuition of the flow characteristics and how turbulent the flow is. It is a parameter that can significantly change the policy chosen by the reinforcement learning agent, as shown in the cylinder case by Varela et al. (2022). When even higher Reynolds numbers are considered, three-dimensional environments become necessary to simulate all the turbulence features and mechanisms. Recent works by Sonoda et al. (2022) and Zeng et al. (2022) have applied deep reinforcement learning (DRL) to three-dimensional flow cases. The former targets a Poisueille flow in a channel, while a Couette flow is controlled in the latter, using two streamwise parallel slots. Other notable applications of reinforcement learning in fluid mechanics include maximization of the efficiency of agents swimming in a turbulent flow (Biferale et al., 2019; Verma et al., 2018) and turbulence modelling (Novati et al., 2021; Kurz et al., 2023). In previous works a single DRL algorithm is typically tested for a specific task. By contrast, in this work we provide an environment that can be coupled with existing frameworks like Stable baselines (Raffin et al., 2021) to establish a new DRL benchmark problem.

In the present work, we use DRL algorithms to discover a policy to reduce the friction drag in a wall-bounded fluid flow, by imposing a wall-normal velocity distribution at the wall. Following the approach of Choi et al. (1994), the control is computed based on the flow field measurements at a given wall-normal distance. Our contributions can be summarized as follows:

- We introduce and publicly release a new reinforcement learning environment aiming to reduce the friction drag in a three-dimensional, fully-turbulent fluid flow. This represents a significant jump in complexity with respect to the two-dimensional flow cases used as benchmarks in the literature.
- We reproduce the results of the traditional state-of-the-art control strategy reported in the literature. Then, we use the deep deterministic policy gradient (DDPG, Lillicrap et al., 2015) algorithm to train a DRL agent on the proposed environment. We show that DDPG leads to the discovery of an efficient drag-reduction policy that performs significantly better than the strategies previously available in the literature.

2 METHODOLOGY

A summary of our work is shown in figure 1. In the left part, we show the numerical domain in which an open channel flow is simulated. We indicate the streamwise, wall-normal and spanwise directions with (x, y, z), respectively and their corresponding velocity components by (u, v, w). Two domain sizes are considered: the first one is a minimal channel (Jiménez & Moin, 1991), which represents the minimal flow unit in which a self-sustained turbulent flow is possible. A larger channel is also considered, with a different domain Ω and resolution $N_x \times N_y \times N_z$. Further details are available in appendix A. The Reynolds number used in this study and in the fluid-dynamics literature for wall-bounded flows is the friction Reynolds number, defined as $Re_\tau = u_\tau h/\nu$, where the friction velocity $u_\tau = \sqrt{\tau_w/\rho}$ (based on the the wall-shear stress τ_w and the fluid density ρ). ν is kinematic viscosity of the fluid. We consider $Re_\tau = 180$, to compare the drag reduction with the corresponding results in previous active flow control works (Choi et al., 1994).

In order to control the flow and reduce the overall drag, a wall-normal velocity is applied at the wall. For the action space, a naive design would be to let a single agent learn a policy that decides the wall-normal velocity in all the points. While this allows for the maximum control on all the simulated time- and length-scales in the flow, the resulting action space becomes prohibitively large, *i.e.* $\mathcal{A} \in \mathbb{R}^{N_x \times N_z}$, making it difficult to explore it in an efficient and comprehensive way.

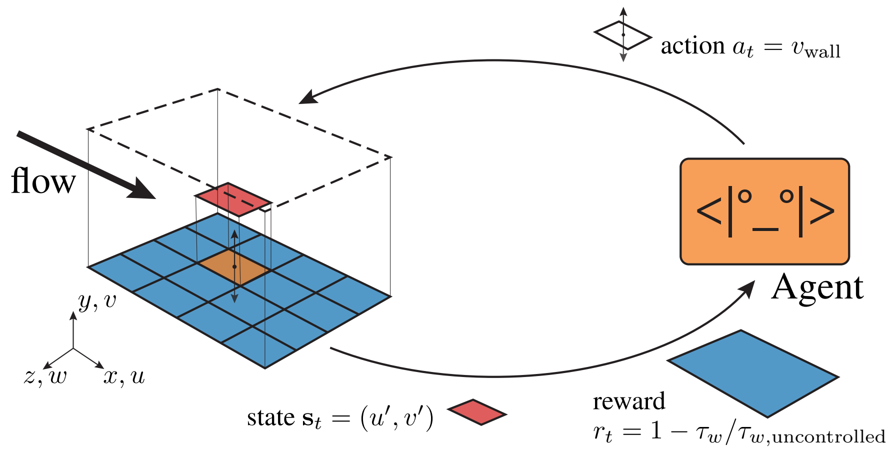

Figure 1: Overview of our multi-agent DRL approach to drag reduction. The simulation domain is shown on the left. The agents are organized in a grid $N_{\text{CTRLx}} \times N_{\text{CTRLz}}$. Each agents observes the velocity fluctuations in the streamwise (u') and wall-normal (v') direction. The reward is the percentage variation of the wall-shear stress τ_w. Based on the state, each agent acts by imposing a wall-normal velocity v at the wall.

Therefore, hinging on the inherent translational invariance of the physical system with respect to streamwise and spanwise directions, we opted for a multi-agent reinforcement-learning (MARL) approach where there are as many agents as action locations all of whom share, and thus contribute to learning, a single policy. We consider a grid of $N_{\text{CTRLx}} \times N_{\text{CTRLz}} = N_x \times N_z$ agents in the streamwise and spanwise directions, respectively. Taking advantage of the invariance by translation of the system can have a decisive impact on the success of the learning process, as highlighted by Belus et al. (2019). This approach makes the learning more sample efficient by not only significantly reducing the action space (down to a single scalar) but also effectively increasing the number of learning samples proportional to the number of agents. The observable state of each agent is limited to its local neighborhood, in fact every agent receives as observation the velocity fluctuations in the streamwise and wall-normal directions at a given distance from the wall, above the area where it is selecting the actuation. We measure the sampling height using the *inner-scaled* coordinate $y^+ = yu_\tau/\nu = y/\ell^* \in [0, Re_\tau]$, where ℓ^* is also termed viscous length. Based on these observations, the policy of the agent determines the intensity of the actuation within a prescribed range. Following the previous studies on opposition control (Hammond et al., 1998), we consider a sampling plane at $y^+ = 15$. The policy is then updated based on the reward, defined as the percentage variation of the wall-shear stress with respect to the uncontrolled flow. The reward is computed on the entire wall and the resulting value is provided to all MARL agents.

We perform the learning process using deep deterministic policy gradient (DDPG, Lillicrap et al., 2015) algorithm in the minimal channel. Once the learning is performed, the policy can be applied without modifications to a larger domain, with a more complete representation of all the physical features that characterize turbulent flows. Such policy transfer is possible thanks to the local nature of the control.

3 Experiments

In order to evaluate the performance of the learnt policy, we consider a well-established control baseline from the literature, namely *opposition control* (Choi et al., 1994). Furthermore, it is possible to compute an upper-bound for the drag reduction, by considering the case in which the flows relaminarize (Fukagata et al., 2009). At $Re_\tau = 180$, the maximum drag reduction is 73.9%. This value depends on the Reynolds number, the higher the Re, the higher is the potential drag reduction.The control strategies are tested in the larger channel with different initial conditions in order to verify their robustness in different settings. The drag reduction comparison between opposition control and the DRL policy learnt in the minimal channel is shown in the left panel of figure 2. The DRL

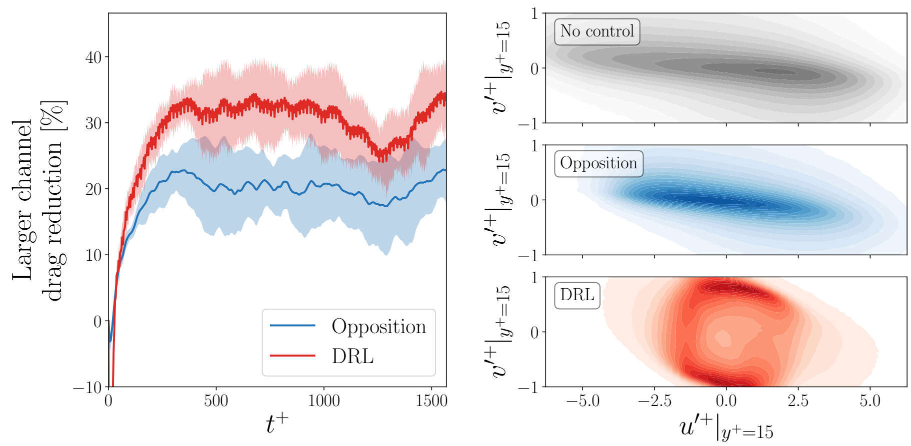

Figure 2: (Left) Drag reduction obtained in the larger channel with respect to the uncontrolled case, when using the DRL policy learnt in the minimal channel versus using opposition control. The result is averaged over six different initial conditions. The shaded area represents the variance of the drag reduction with the different initial conditions. (Right) Distribution of the inner-scaled velocity-fluctuation components after the initial transient ($t^+ > 500$) in the streamwise (u) and wall-normal (v) directions, for the uncontrolled case (top), with opposition control (middle) and when using DRL (bottom).

policy consistently provides a higher average drag reduction than the one obtained with opposition control. After the initial transient ($t^+ > 500$) the DDPG policy provides 30% drag reduction, whereas opposition control is limited to 20%. Further analysis are reported in appendix B. Figure 2 (right) shows the effect of the control on the velocity-fluctuations distribution. This joint probability density function is used in quadrant analysis (Wallace et al., 1972) to identify the different physical mechanisms characterizing the flow. Opposition control reduces the ranges of the fluctuations, both in streamwise and wall-normal direction. However, only a small variation of the overall distribution shape can be observed. This means that opposition control does not alter the physical mechanism in turbulence, rather it reduces their intensity. The fluctuations distribution after the application of the DRL policy is fundamentally different, with a reduction of the streamwise fluctuations and an increase of the wall-normal fluctuations. The DRL policy has a significant impact on the flow physics, modifying the distribution of the events among the four quadrants and the general flow configuration.

4 Conclusion

In this paper, we have described the coupling of a high-performance fluid dynamics solver and a (deep) multi-agent reinforcement-learning framework, aiming to discover novel drag-reduction strategies. By casting the fluid dynamics simulation as a multi-agent reinforcement-learning problem, we were able to discover a local and translational-invariant policy that provides a sensible improvement over existing control policies such as opposition control. We have also successfully transferred the learnt policy to a larger domain, where the training would be computationally more expensive. The drag reduction achieved in the larger channel is 30%, with around 10% improvement over opposition control, providing a new state-of-the-art reference for turbulent flow control channel flows. The significant improvement of this study over long-established designed control strategies demonstrates the potential of data-driven control methods for applications involving turbulent flows. Furthermore, the gap from the our achieved results to the theoretical upper bound of the relaminarized flow makes the released environment an exciting test-bed for novel reinforcement learning algorithms. Finally, despite the encouraging results, this represents only a first step towards the application of DRL policies in more realistic wall-bounded flows. For instance, the current policy does not take into account the cost of actuation, which could be included in the reward function to obtain more energy-efficient policies, but this extension is left for future work.

ACKNOWLEDGMENTS

The code supporting the findings in this paper including the environment implementation is available in the repository `https://github.com/KTH-FlowAI/MARL-drag-reduction-in-wall-bounded-flows`. The authors acknowledge the Swedish National Infrastructure for Computing (SNIC) for providing the computational resources by PDC, used to carry out the numerical simulations. This work is supported by the founding provided by the ERC grant no. "2021-CoG-101043998, DEEPCONTROL", the Swedish e-Science Research Centre (SeRC) and the Knut and Alice Wallenberg (KAW) Foundation.

REFERENCES

Vincent Belus, Jean Rabault, Jonathan Viquerat, Zhizhao Che, Elie Hachem, and Ulysse Reglade. Exploiting locality and translational invariance to design effective deep reinforcement learning control of the 1-dimensional unstable falling liquid film. *AIP Advances*, 9(12):125014, 2019.

Luca Biferale, Fabio Bonaccorso, Michele Buzzicotti, Patricio Clark Di Leoni, and Kristian Gustavsson. Zermelo's problem: Optimal point-to-point navigation in 2d turbulent flows using reinforcement learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(10):103138, 2019. doi: 10.1063/1.5120370.

Haecheon Choi, Parviz Moin, and John Kim. Active turbulence control for drag reduction in wall-bounded flows. *Journal of Fluid Mechanics*, 262:75–110, 1994.

Thomas Duriez, Steven L. Brunton, and Bernd R. Noack. *Machine Learning Control – Taming Nonlinear Dynamics and Turbulence*. Springer, 2017.

Irmgard Flugge-Lotz. *Discontinuous automatic control*, volume 2166. Princeton University Press, Princeton, NJ, 2015.

Koji Fukagata, Kazuyasu Sugiyama, and Nobuhide Kasagi. On the lower bound of net driving power in controlled duct flows. *Physica D: Nonlinear Phenomena*, 238(13):1082–1086, 2009. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2009.03.008. URL `https://www.sciencedirect.com/science/article/pii/S0167278909000840`.

Luca Guastoni, Alejandro Güemes, Andrea Ianiro, Stefano Discetti, Philipp Schlatter, Hossein Azizpour, and Ricardo Vinuesa. Convolutional-network models to predict wall-bounded turbulence from wall quantities. *Journal of Fluid Mechanics*, 928:A27, 2021.

Edward P. Hammond, Thomas R. Bewley, and Parviz Moin. Observed mechanisms for turbulence attenuation and enhancement in opposition-controlled wall-bounded flows. *Physics of Fluids*, 10 (9):2421–2423, 1998. doi: 10.1063/1.869759.

Javier Jiménez and Parviz Moin. The minimal flow unit in near-wall turbulence. *Journal of Fluid Mechanics*, 225:213–240, 1991. doi: 10.1017/S0022112091002033.

John Kim, Parviz Moin, and Robert Moser. Turbulence statistics in fully developed channel flow at low reynolds number. *Journal of Fluid Mechanics*, 177:133–166, 1987. doi: 10.1017/S0022112087000892.

Marius Kurz, Philipp Offenhäuser, and Andrea Beck. Deep reinforcement learning for turbulence modeling in large eddy simulations. *International Journal of Heat and Fluid Flow*, 99:109094, 2023. ISSN 0142-727X. doi: https://doi.org/10.1016/j.ijheatfluidflow.2022.109094. URL `https://www.sciencedirect.com/science/article/pii/S0142727X2200162X`.

Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning, 2015.

Guido Novati, Hugues Lascombes de Laroussilhe, and Petros Koumoutsakos. Automating turbulence modeling by multi-agent reinforcement learning. *Nature Machine Intelligence*, 3:87–96, 2021.

Jonghwan Park and Haecheon Choi. Machine-learning-based feedback control for drag reduction in a turbulent channel flow. *Journal of Fluid Mechanics*, 904:A24, 2020. doi: 10.1017/jfm.2020.690.

Fabio Pino, Lorenzo Schena, Jean Rabault, and Miguel A. Mendez. Comparative analysis of machine learning methods for active flow control, 2022.

Jean Rabault, Miroslav Kuchta, Atle Jensen, Ulysse Réglade, and Nicolas Cerardi. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics*, 865:281–302, 2019.

Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268):1–8, 2021. URL `http://jmlr.org/papers/v22/20-1364.html`.

Takahiro Sonoda, Zhuchen Liu, Toshitaka Itoh, and Yosuke Hasegawa. Reinforcement learning of control strategies for reducing skin friction drag in a fully developed channel flow. *arXiv preprint*, arXiv:2206.15355, 2022.

Philippe R Spalart and J Douglas McLean. Drag reduction: enticing turbulence, and then an industry. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1940):1556–1569, 2011.

Pau Varela, Pol Suárez, Francisco Alcántara-Ávila, Arnau Miró, Jean Rabault, Bernat Font, Luis Miguel García-Cuevas, Oriol Lehmkuhl, and Ricardo Vinuesa. Deep reinforcement learning for flow control exploits different physics for increasing Reynolds number regimes. *Actuators*, 11 (12), 2022. ISSN 2076-0825. doi: 10.3390/act11120359. URL `https://www.mdpi.com/2076-0825/11/12/359`.

Siddhartha Verma, Guido Novati, and Petros Koumoutsakos. Efficient collective swimming by harnessing vortices through deep reinforcement learning. *Proceedings of the National Academy of Sciences*, 115(23):5849–5854, 2018. doi: 10.1073/pnas.1800923115. URL `https://www.pnas.org/doi/abs/10.1073/pnas.1800923115`.

Ricardo Vinuesa and Steve Brunton. Enhancing computational fluid dynamics with machine learning. *Nature Computational Science*, 2:358–366, 2022.

James M. Wallace, Helmut Eckelmann, and Robert S. Brodkey. The wall region in turbulent shear flow. *Journal of Fluid Mechanics*, 54(1):39–48, 1972. doi: 10.1017/S0022112072000515.

Kevin Zeng, Alec J. Linot, and Michael D. Graham. Data-driven control of spatiotemporal chaos with reduced-order neural ODE-based models and reinforcement learning. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478(2267):20220297, 2022. doi: 10.1098/rspa.2022.0297.

A Training setup and implementation details

In this appendix we describe the solver and reinforcement learning setup used to obtain the results presented in the paper.

Our numerical solver is a pseudo-spectral solver, similar to the one used in Kim et al. (1987), with Chebyshev polynomials in the wall-normal direction. Periodic boundary conditions are imposed in the streamwise and spanwise directions, while a no-slip condition and a symmetry conditions are imposed at the two wall-normal boundaries. The numerical schemes used for time-advancement are a third-order Runge-Kutta method for the nonlinear terms and a second-order Crank-Nicolson algorithm for the linear ones. The timestep used in the simulation is $\Delta t^+ = 0.6$ (where the time units are scaled with $t^* = \nu/u_\tau^2$). In the minimal channel, the simulation domain has size $\Omega = L_x \times L_y \times L_z = 2.67h \times h \times 0.8h$ (where h is the open-channel height) and resolution $N_x \times N_y \times N_z = 16 \times 65 \times 16$. The larger channel has size $\Omega = L_x \times L_y \times L_z = 2\pi h \times h \times \pi h$ and resolution $N_x \times N_y \times N_z = 64 \times 65 \times 64$. The flow dynamics in the minimal channel is simpler while

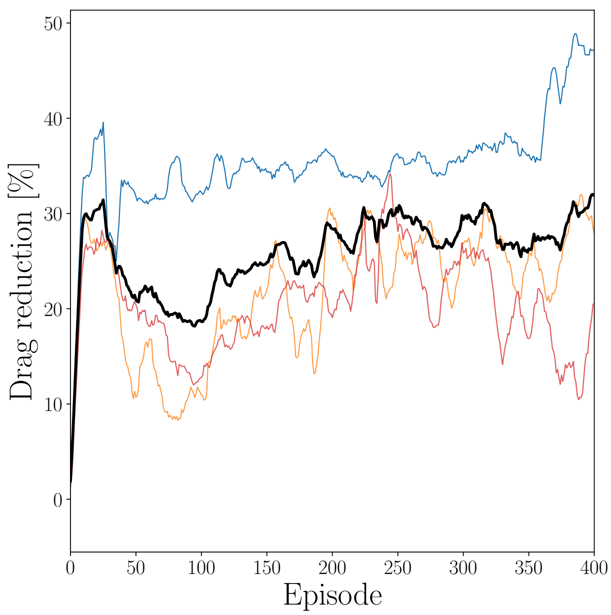

Figure 3: Running mean of the drag reduction with respect to the reference uncontrolled case during agent learning in the minimal channel. Coloured lines represent individual runs, the black line shows the average over the different runs.

maintaining a good agreement in the low-order statistics with larger domains and experimental data. It is also computationally cheaper, allowing for faster agent training.

Deep deterministic policy gradient is the reinforcement learning algorithm used to train the agents. The policy is updated every $\Delta t^+ = 180$, with 64 mini-batch gradient updates. The (state, action, reward) tuples are sampled from a replay buffer of size 5,000,000. The agents are trained in the minimal channel for 400 episodes, each one consisting of 3000 interactions. The overall temporal length of each episode is then $T^+ \approx 1500$.

Examples of the agent training runs in the minimal channel are shown in figure 3. The agents are able to reduce the drag in the open channel after 10-15 episodes, however the performance can change significantly during training, without converging to a specific policy. We do not show it in the figure, but it has to be noted that not all the training runs provide a successful policy. In these cases, the drag reduction achieved is close to zero.

The code is running entirely on CPUs. The training and evaluation were performed on recent Intel Xeon processors. The minimal channel solver was compiled to run on 4 cores, while the larger channel solver was run on 16 cores. The evaluation on a single episode takes about 4 core-hours (*i.e.* 15 minutes in wall-clock time). A detailed study on the scaling performance of the solver when coupled with the DRL framework is left for future work.

B Policy analysis and comparison

The difference in the actuation applied at the wall by opposition control and by the DRL policy can be qualitatively assessed in figure 4. Opposition control has a smoother actuation profile since the control is proportional to wall-normal velocity at the sensing plane, which the control aims to reduce. On the other hand, the DRL policy frequently switches between the minimum and maximum value in the actuation range, similarly to a *2-step* controller (Flugge-Lotz, 2015).

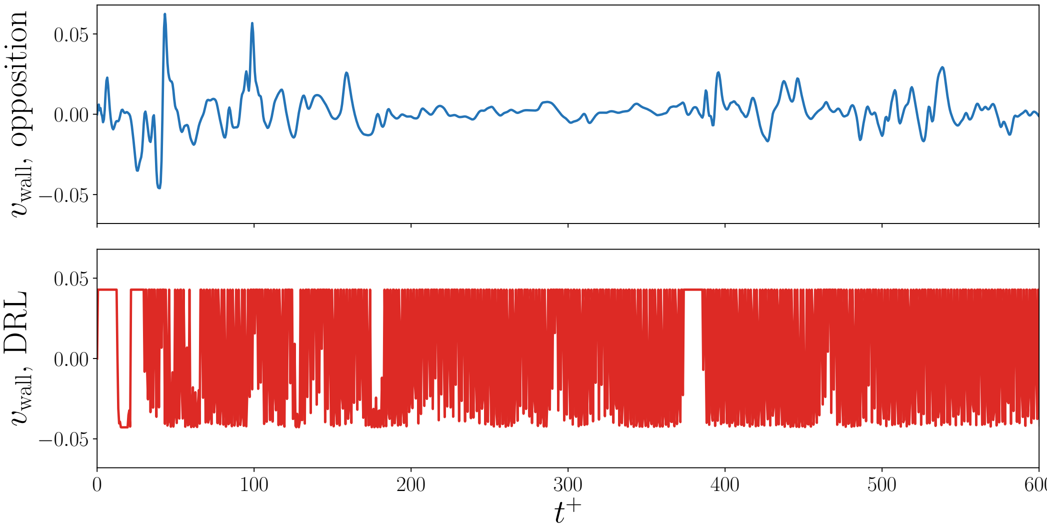

Figure 4: History of the wall-normal velocity applied at the wall by a single agent (up to $t^+ = 600$) using opposition control (top) and the DRL policy (bottom) in the larger channel, when using the policy trained in the minimal channel.

In the work by Sonoda et al. (2022), 27% drag reduction is obtained for a channel flow confined between two solid walls, with the possibility to apply a wall-normal velocity distribution on both surfaces. Since the control is mostly targeting the near-wall turbulent structures, their approach yields similar results compared to the one described in this work. The difference can be justified by the different domain boundary conditions.