14th International Symposium on Turbulence and Shear Flow Phenomena (TSFP14)
Heidelberg, Germany, July 28–31, 2026

# PHYSICS-GUIDED SURROGATE TRAINING FOR ZERO-SHOT REINFORCEMENT LEARNING CONTROL OF TURBULENT WINGS

**Pol Suárez**
FLOW, Engineering Mechanics,
KTH Royal Institute of Technology
114 28 Stockholm, Sweden.
polsm@kth.se

**Yuning Wang**
Department of Aerospace Engineering, University of Michigan
Ann Arbor, MI 48109, USA.
yuningw@umich.edu

**Mathis Bode**
Jülich Supercomputing Centre,
Forschungszentrum Jülich GmbH
52425 Jülich, Germany.
m.bode@fz-juelich.de

**Ricardo Vinuesa**
[1]Department of Aerospace Engineering, University of Michigan
Ann Arbor, MI 48109, USA.
[2]FLOW, Engineering Mechanics, KTH Royal Institute of Technology
114 28 Stockholm, Sweden.
rvinuesa@umich.edu

## ABSTRACT

A physics-guided surrogate-based framework for zero-shot deep reinforcement learning (DRL) control of turbulent boundary layers is presented. Controllers are trained in turbulent channel flows matched to local boundary-layer characteristics and deployed onto the suction side of a NACA4412 wing at $Re_c = 2 \times 10^5$, turbulent conditions, without further training. Thanks to surrogate environments within a multi-agent reinforcement learning framework, the approach dramatically improves data efficiency compared to direct training in spatially developing flows. A hybrid control strategy is adopted, combining DRL in regions of moderate adverse pressure gradient with uniform blowing in the strongest non-equilibrium region. In the wing configuration, DRL further provides approximately 20% reduction in total drag coefficient and up to 60% reduction in skin-friction drag relative to OC performance. The learned policies generate spatially coherent actuation patterns that transfer across configurations and induce broader modifications of near-wall turbulence than OC. These results demonstrate that physics-guided surrogate training, combined with region-specific control strategies, enables robust, scalable, and computationally efficient DRL-based flow control in spatially developing turbulent boundary layers.

## INTRODUCTION

Understanding and controlling turbulent flow over complex geometries is central to many engineering applications, particularly in aerodynamics, where skin-friction drag significantly impacts energy consumption and performance. Comprehensive reviews (Abbas *et al.*, 2017; Fukagata *et al.*, 2024) indicate that turbulence is the primary contributor to skin-friction drag on aircraft wings. However, its control remains challenging due to the high dimensionality, multiscale nature, and chaotic dynamics of turbulence, compounded by geometric complexity and streamwise development.

A canonical closed-loop strategy is opposition control (OC), in which wall-normal velocity is imposed to oppose off-wall velocity fluctuations, thereby attenuating near-wall momentum and reducing skin-friction drag (Choi *et al.*, 1994). A notable feature of OC is its simplicity: using near-optimal parameters (Hammond *et al.*, 1998), it achieves meaningful drag reduction across different flow types and geometries (Wang *et al.*, 2025). This transferability is commonly attributed to similarities in near-wall turbulence statistics across wall-bounded flows (Jiménez *et al.*, 2010). However, OC remains inherently local and reactive, limiting its effectiveness in non-equilibrium boundary layers subjected to strong adverse pressure gradients.

Deep reinforcement learning (DRL) has recently emerged as a promising framework for flow control, enabling the discovery of nonlinear and non-local control strategies. Recent studies demonstrate that DRL-based controllers can surpass OC in canonical configurations, indicating an ability to exploit flow structures beyond the classical near-wall cycle (Guastoni *et al.*, 2023; Suárez *et al.*, 2025). Despite this potential, two major challenges remain: (1) the high computational cost associated with training in high-fidelity turbulent simulations, and (2) limited transferability of learned policies across flows with different characteristics.

Multi-agent reinforcement learning (MARL) alleviates the curse of dimensionality by assigning agents to local subdomains under locally invariant conditions, sharing a common policy (Belus *et al.*, 2019). This approach is particularly effective in canonical flows with homogeneous directions, such as turbulent channel flows (TCFs). In contrast, its efficiency deteriorates in spatially developing boundary layers, where streamwise inhomogeneity reduces the number of available pseudo-environments.

To address these challenges, we propose a physics-guided zero-shot DRL framework for turbulent wing control. The suction-side boundary layer is partitioned into chordwise blocks $\Omega_i$, and for each block a matched TCF surrogate is con-

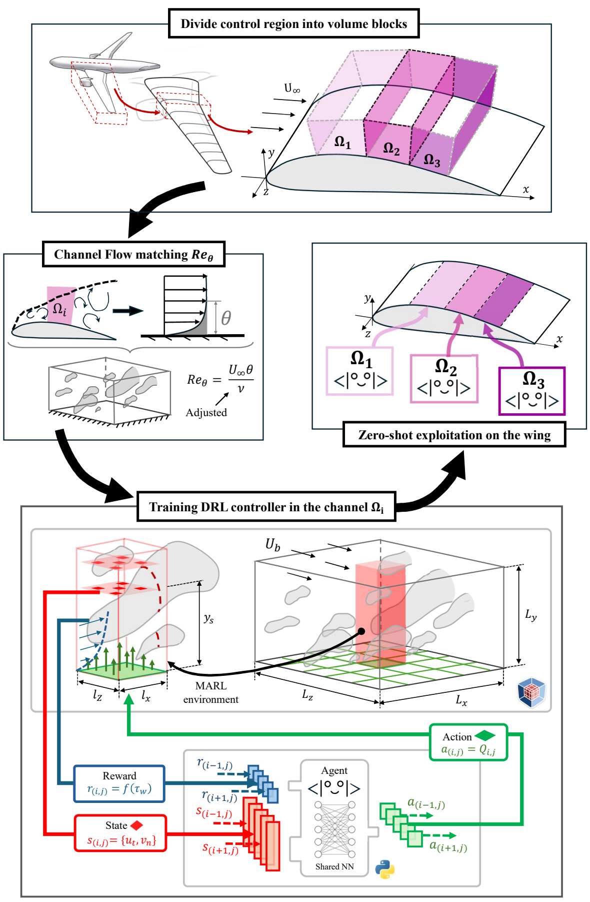


Figure 1. **Physics-guided surrogate-based MARL framework.** The wing boundary layer is partitioned into chordwise blocks $\Omega_i$, each matched to a turbulent channel-flow (TCF) surrogate using $Re_\theta$. Agents share a common policy and are trained in the surrogate environments, enabling efficient exploration. The learned controllers are then deployed on the wing without further training (zero-shot). DRL is applied in $\Omega_1$–$\Omega_3$, while uniform blowing is used in $\Omega_4$.

structed based on local flow characteristics. Controllers are trained exclusively in these surrogate environments and subsequently deployed onto the wing without further training, constituting a zero-shot transfer.

In the present study, we consider the NACA4412 airfoil at $Re_c = 2 \times 10^5$ and $\alpha = 5°$. A hybrid control strategy is adopted, combining DRL-based closed-loop control in regions of mild to moderate adverse pressure gradient with uniform blowing in the most strongly non-equilibrium region. This configuration enables assessing the transferability of learned policies across spatially developing boundary layers while accounting for the varying effectiveness of control mechanisms under increasing adverse pressure gradient.

# METHODOLOGY
## Numerical Setup

The flow is governed by the incompressible Navier–Stokes equations and solved using the spectral-element solver Nek5000 (Fischer *et al.*, 2008), which has been extensively validated for turbulent boundary layers over wing sections (Vinuesa *et al.*, 2018). Large-eddy simulation (LES) is performed for the NACA4412 airfoil at $Re_c = 2 \times 10^5$ and $\alpha = 5°$, with near-wall resolution approaching DNS levels. The computational domain consists of a C-mesh with $L_x = 6c$, $L_y = 4c$, and $L_z = 0.2c$, comprising approximately $2.2 \times 10^8$ grid points. The near-wall resolution satisfies $\Delta x^+ < 18$, $\Delta y^+ < 11$, and $\Delta z^+ < 9$.

Turbulent channel flows (TCFs) are simulated using direct numerical simulation in periodic domains with $L_x = 2\pi h$, $L_y = h$, and $L_z = \pi h$. The grid resolution is matched in viscous units to that of the wing simulations, ensuring consistent representation of near-wall turbulence structures and comparable control resolution across configurations. This consistency is essential for enabling the transferability of the learned control policies.

In particular, matching the near-wall resolution ensures that the effective actuator size and sensing scales remain consistent between surrogate and target environments.

## Surrogate Matching Tree

The suction-side control region ($x_{ss}/c = 0.25$–$0.86$) is partitioned into chordwise blocks $\Omega_i$ based on the Clauser pressure-gradient parameter,

$$\beta = \frac{\delta^*}{\tau_w}\frac{dP}{dx}, \tag{1}$$

where $\delta^*$ is the displacement thickness and $\tau_w$ is the wall-shear stress. The partitioning is guided by the streamwise evolution of $\beta$, aiming to define regions with approximately homogeneous pressure-gradient conditions while retaining the progressive effect of the adverse pressure gradient (APG) along the wing, considered strong for $\beta > 3$. This balance is essential to ensure that each block can be represented by a locally parallel flow while still capturing the relevant non-equilibrium effects.

Each block is associated with a surrogate turbulent channel flow (TCF) matched using the streamwise-averaged momentum-thickness Reynolds number $\langle Re_\theta \rangle_x$, which characterizes the local development state of the boundary layer and governs near-wall turbulence dynamics relevant to skin-friction drag.

The combined use of $\beta$ for spatial partitioning and $Re_\theta$ for surrogate matching provides a physics-guided representation of the wing boundary layer using canonical channel flows. Three surrogate configurations are defined for $\Omega_1$–$\Omega_3$, where DRL-based control is applied. The most downstream region $\Omega_4$, characterized by strong APG, is instead controlled using uniform blowing, reflecting the reduced effectiveness of near-wall sensing-based strategies in highly non-equilibrium conditions.

This separation allows selecting control strategies according to local flow physics, rather than a single approach across the entire boundary layer.

## Opposition Control As Reference

Opposition control (OC) is employed as a baseline closed-loop strategy, originally proposed by (Choi *et al.*, 1994). The method relies on local feedback of the wall-normal velocity, imposing actuation at the wall to oppose off-wall turbulent fluctuations:

$$v(x,0,z,t) = -\alpha\left[v(x,y_s,z,t) - \langle v(x,y_s,z,t)\rangle\right], \tag{2}$$

where $y_s^+ = 15$ denotes the sensing location in viscous units and $\alpha = 1$ is the control gain in the present case.

OC attenuates near-wall coherent structures, such as quasi-streamwise vortices and streaks, by suppressing wall-normal velocity fluctuations. This reduces turbulent momentum transfer towards the wall and leads to a decrease in skin-friction drag. In canonical wall-bounded flows, OC achieves drag reductions of approximately 20–30% depending on the Reynolds number and sensing location (Hammond *et al.*, 1998).

A key advantage of OC is its simplicity, which is commonly attributed to the universality of near-wall turbulence dynamics across different flow configurations (Jiménez *et al.*, 2010). However, OC remains inherently local and reactive, as it relies on instantaneous measurements at a fixed sensing plane. This limits its ability to influence larger-scale turbulent structures and reduces its effectiveness in non-equilibrium boundary layers subjected to strong adverse pressure gradients.

These limitations motivate the exploration of alternative control strategies, such as DRL, as well as hybrid approaches that combine near-wall feedback with physics-informed actuation in regions where local sensing becomes insufficient.

## Reinforcement Learning Framework

Flow control is formulated within a multi-agent reinforcement learning (MARL) framework, in which each wall grid point is treated as an agent operating under locally invariant conditions while sharing a common control policy. This formulation enables a significant reduction in control dimensionality while promoting efficiency by exploring in parallel across a large number of pseudo-environments.

Each agent observes the fluctuating streamwise and wall-normal velocities $(u', v')$ at $y_s^+ = 15$ and outputs a wall-normal actuation velocity bounded by the local friction velocity $u_\tau$. The shared-policy structure allows the controller to learn generalizable strategies based on local flow history and features, rather than memorizing geometry-specific behavior.

A zero-net-mass-flux (ZNMF) constraint is enforced across the control region to ensure physical consistency. The reward is defined as the relative reduction in wall-shear stress,

$$r = 1 - \frac{\tau_w^{\mathrm{ctrl}}}{\tau_w^{\mathrm{ref}}}. \tag{3}$$

Training is performed exclusively in the surrogate TCF environments using the twin delayed DDPG (TD3) algorithm (Fujimoto *et al.*, 2018). The statistical homogeneity of the channel flows maximizes the number of effective pseudo-environments, significantly improving exploration efficiency compared to spatially developing boundary layers. Each configuration is trained for 100 episodes of duration $t^+ = 1{,}500$, after which the learned policies are deployed onto the wing without further training. This is essential for enabling efficient training at high Reynolds numbers, where direct exploration in spatially developing flows would be computationally prohibitive.

This combination of MARL and surrogate environments is key to enabling zero-shot transfer at high Reynolds numbers, where direct on-wing exploration would be computationally prohibitive.

This approach dramatically improves data efficiency thanks to the statistical homogeneity of the surrogate environments. In contrast, direct training on the wing would be significantly more expensive due to the limited number of effective pseudo-environments in spatially developing flows.

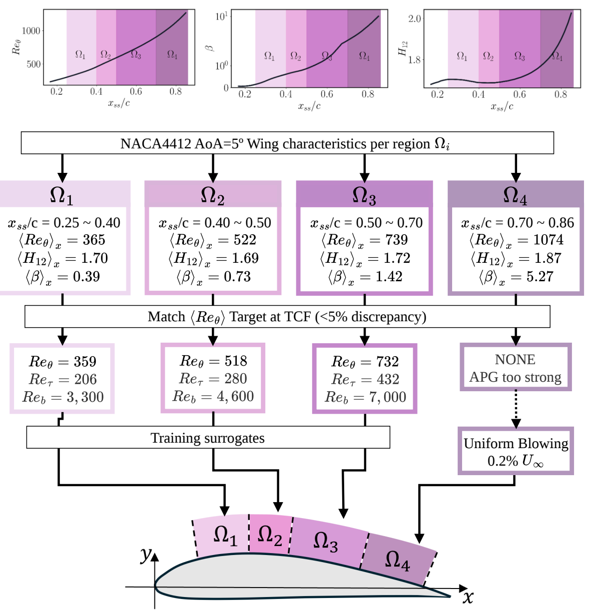


Figure 2. **Physics-guided matching tree of wing regions to surrogate channel flows.** The suction-side control region of the NACA4412 wing is partitioned into chordwise blocks $\Omega_1$–$\Omega_4$ based on the local evolution of $Re_\theta$, $\beta$, and $H_{12}$. For each block, the streamwise-averaged $\langle Re_\theta \rangle_x$ is used to define a matched turbulent channel-flow (TCF) surrogate, yielding three surrogate environments for $\Omega_1$–$\Omega_3$. In the most downstream region $\Omega_4$, no surrogate is defined because the adverse pressure gradient is too strong; this region is instead controlled using uniform blowing at $0.2\% U_\infty$. The diagram summarizes the full workflow from wing characterization to surrogate training and final deployment on the wing.

# RESULTS

## Training In Surrogate Flows

Figure 2 shows the chordwise evolution of the boundary-layer statistics and the resulting partition into control blocks. The close agreement between the wing flow and the matched turbulent channel-flow (TCF) configurations confirms that the surrogate environments accurately reproduce the local development state of the boundary layer within each $\Omega_i$.

Within these surrogate environments, the statistical homogeneity enables clustering more pseudo-environments under the MARL formulation, significantly improving exploration efficiency. As a result, as observed in Figure 3(a), the DRL policies surpass the opposition-control (OC) performance early in training and converge rapidly, indicating that the relevant near-wall dynamics are effectively captured by the TCF surrogates.

Figure 3(b) reports the drag-reduction rate during exploitation. DRL achieves reductions of 34.1%, 30.0%, and 27.0% in $\Omega_1$, $\Omega_2$, and $\Omega_3$, respectively, compared with 25.3%, 22.8%, and 22.1% for OC. These OC levels are consistent with prior studies at comparable Reynolds numbers (Stroh *et al.*, 2015), validating the present implementation. The gradual decrease in performance from $\Omega_1$ to $\Omega_3$ reflects the increas-

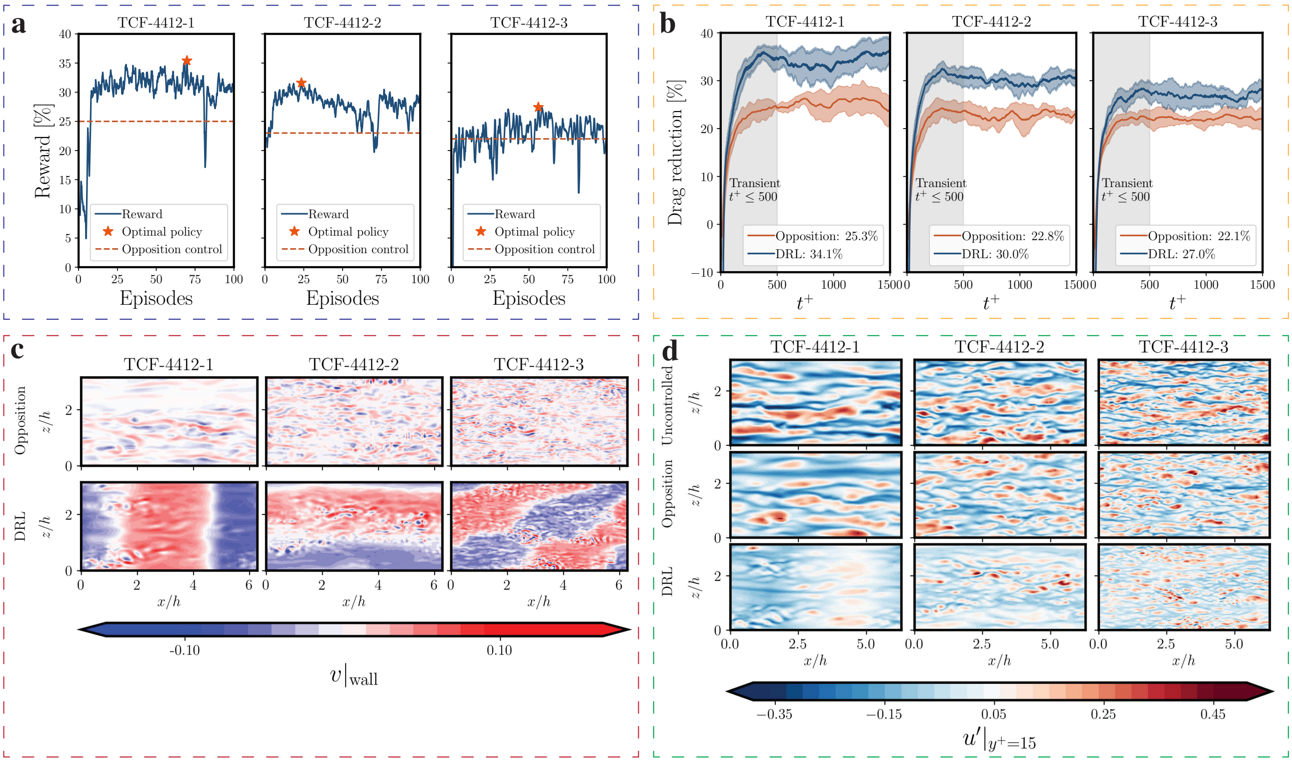


Figure 3. **TCF surrogate results.** (a) Reward evolution during training; stars mark the selected optimal policy, dashed lines indicate the OC reference. (b) Drag-reduction rate during exploitation; shaded band is the standard deviation across six initial conditions; gray region ($t^+ \leq 500$) is the excluded transient. (c) Instantaneous wall actuation of OC and DRL. (d) Instantaneous streamwise velocity fluctuations at $y^+ = 15$.

ing influence of the adverse pressure gradient, which modifies near-wall turbulence structure. Despite this, DRL maintains a consistent advantage over OC across all cases, indicating that the learned policies capture robust control mechanisms rather than flow-specific features. The instantaneous wall actuation (Fig. 3(c)) reveals that DRL develops spatially coherent patterns, in contrast to the localized fluctuation-opposing behavior of OC. These patterns exploit the homogeneous directions of the channel flow, enabling coordinated actuation across agents within the MARL framework.

Consistently, the effect on near-wall turbulence (Fig. 3(d)) shows a strong attenuation of high-speed streaks under DRL, accompanied by a reorganization of their spatial structure. Compared to OC, which primarily suppresses fluctuations locally, DRL induces broader modifications of the flow, suggesting interaction with turbulence beyond the classical near-wall cycle. This behavior suggests a partial disruption of the near-wall regeneration cycle, accompanied by a redistribution of turbulent kinetic energy across scales.

Overall, the surrogate results demonstrate that MARL-trained DRL controllers achieve sustained and enhanced drag reduction in canonical environments, while providing physically consistent actuation patterns that are suitable for zero-shot transfer.

The framework is expected to extend to other aerodynamic configurations, provided that appropriate surrogate matching captures the relevant boundary-layer dynamics.

## Zero-Shot Deployment On The NACA4412

The trained controllers are deployed on the suction side of the NACA4412 wing without further training. A hybrid control configuration is adopted, combining DRL-based control in $\Omega_1$–$\Omega_3$ with uniform blowing in the most downstream region $\Omega_4$. The skin-friction coefficient is defined as $c_f = \tau_w/(\frac{1}{2}\rho U_e^2)$.

Figure 4(a) shows the chordwise distribution of $c_f$ and the corresponding relative drag reduction. The hybrid DRL configuration achieves a mean drag reduction of $\langle R \rangle_{\mathrm{ctrl}} = 31.6\%$, significantly exceeding the 20.3% obtained with opposition control (OC). The improvement is sustained along the chord and becomes more pronounced downstream, where the adverse pressure gradient intensifies and OC performance degrades. This behavior is consistent with the progressive decorrelation between near-wall sensing and wall-shear-stress production under strong adverse pressure gradients.

Integral aerodynamic metrics (Fig. 4(b)) confirm this trend. The DRL-based configuration yields larger reductions in both the skin-friction drag coefficient $C_{d,f}$ and the total drag coefficient $C_d$. Although a slight decrease in lift coefficient is observed, the resulting lift-to-drag ratio remains comparable to that of OC, indicating that the drag-reduction gains are not offset by aerodynamic penalties.

The instantaneous actuation fields (Fig. 4(c)) show that DRL retains the structured patterns learned in the TCF surrogates after deployment, despite deformation due to streamwise development and surface curvature. These patterns remain coherent within each control block, demonstrating that the learned policies encode transferable control mechanisms.

Consistently, the near-wall velocity field (Fig. 4(d)) exhibits a strong attenuation of high-speed streaks and a redistribution of near-wall momentum under DRL. Compared to OC, which primarily suppresses fluctuations locally, DRL induces broader flow modifications that persist along the chord despite increasing adverse pressure gradient. Overall, the results show

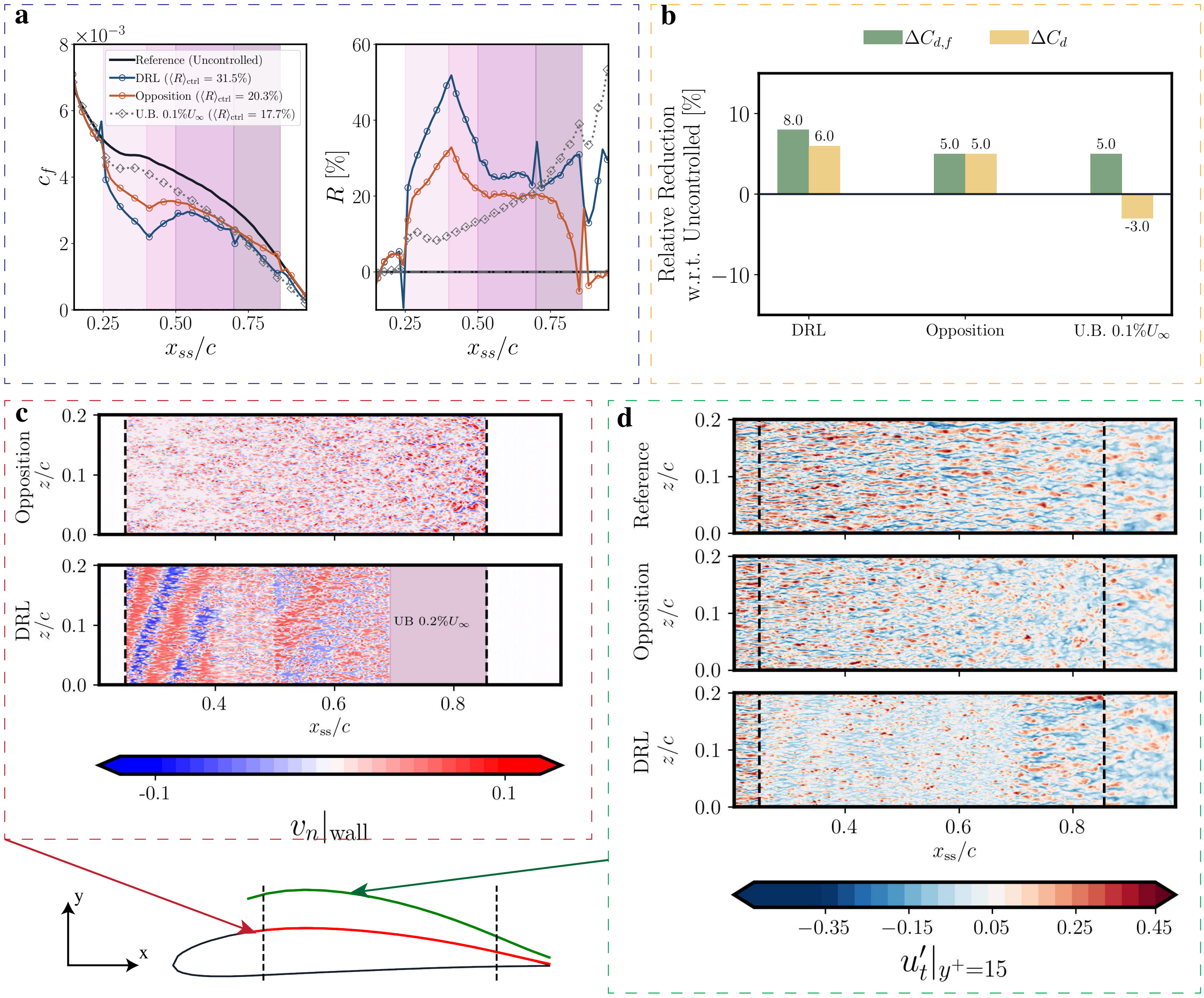


Figure 4. **Zero-shot deployment on NACA4412** ($Re_c = 2\times10^5$, $\alpha = 5°$). (a) Chordwise $c_f$ profiles (left) and relative drag reduction $R$ (right); $\langle R\rangle_{\rm ctrl}$ denotes the area-averaged reduction. (b) Relative changes in lift and drag coefficients and lift-to-drag ratio. (c) Instantaneous wall-normal actuation of OC and DRL. (d) Instantaneous wall-tangential velocity fluctuations at $y_n^+ = 15$.

that the combination of surrogate-trained DRL enables robust drag reduction in spatially developing turbulent boundary layers, maintaining performance across varying pressure-gradient conditions.

## Control Mechanism Interpretation

The improved performance of the DRL-based control can be attributed to its ability to induce non-local and spatially coherent modifications of the near-wall turbulence. While opposition control (OC) operates through local cancellation of wall-normal velocity fluctuations, DRL generates actuation patterns that extend over larger spatial regions and evolve in time.

This difference is evident in the structure of the actuation fields and the resulting flow response. DRL strongly attenuates high-speed streaks while simultaneously modifying their spatial organization, indicating interaction with turbulent structures beyond the classical near-wall cycle. In contrast, OC primarily reduces fluctuations locally, with limited impact on larger-scale motions.

The effectiveness of DRL becomes more pronounced under increasing adverse pressure gradient, where the lift-up mechanism and wall-normal transport are altered. Under these conditions, the correlation between near-wall sensing and wall-shear-stress production decreases, reducing the effectiveness of reactive strategies such as OC. The coherent actuation patterns generated by DRL appear to compensate for this loss of correlation by influencing a broader range of turbulent scales.

In the most downstream region $\Omega_4$, where the boundary layer is strongly non-equilibrium, uniform blowing provides a simple mechanism by directly modifying the mean momentum distribution. This highlights the importance of adapting the control strategy to the local flow physics.

Overall, the results suggest that effective drag reduction in spatially developing boundary layers requires a combination of local and non-local control mechanisms, as achieved by the present hybrid framework.

A physics-guided zero-shot DRL framework has been applied to turbulent boundary-layer control over the suction side of a NACA4412 wing at $Re_c = 2\times10^5$. Controllers are trained only in matched turbulent channel-flow (TCF) surrogates and then deployed directly onto the wing without further training. A hybrid configuration is used, with DRL control in $\Omega_1$–$\Omega_3$ and uniform blowing of $0.2\%U_\infty$ in $\Omega_4$.

The proposed approach outperforms opposition control

(OC) both in the surrogates and on the wing. In zero-shot deployment it reaches $\langle R \rangle_{\mathrm{ctrl}} = 31.6\%$, compared to 20.3% for OC. This corresponds to about 20% more total drag reduction and up to 60% better skin-friction reduction.

The comparison with OC shows clear differences in how control is achieved. OC acts locally by cancelling near-wall velocity fluctuations, mainly targeting the classical near-wall cycle. DRL, instead, produces spatially coherent actuation patterns that modify the flow over larger scales, suggesting interaction with outer-layer structures. These differences become stronger with increasing adverse pressure gradient: OC loses effectiveness, while the hybrid DRL setup keeps a stable level of drag reduction. In addition, DRL can find and use these strategies on its own, without relying on a prescribed model, going beyond what OC can typically achieve.

The fact that similar actuation patterns persist after deployment suggests that the learned policies carry over across different wall-bounded flows. At the same time, the need for uniform blowing in $\Omega_4$ points to the limits of near-wall sensing in strongly non-equilibrium regions.

Overall, the results indicate that controlling spatially developing boundary layers requires adapting the strategy to the local flow conditions.

The surrogate-based framework allows for very efficient training while keeping the relevant near-wall physics, and provides a practical route for scalable DRL-based flow control. Future work should focus on improving performance in strongly non-equilibrium regions, for example by refining the surrogate matching or adding extra flow information from the target case.

## ACKNOWLEDGMENTS

R.V. acknowledges financial support from ERC grant no. '2021-CoG-101043998, DEEPCONTROL'. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (`www.gauss-centre.eu`) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

## REFERENCES

Abbas, A., Bugeda, G., Ferrer, E., Fu, S., Periaux, J., Pons-Prats, J., Valero, E. & Zheng, Y. 2017 Drag reduction via turbulent boundary layer flow control. *Sci. China Technol. Sci.* **60**, 1281–1290.

Belus, V., Rabault, J., Viquerat, J., Che, Z., Hachem, E. & Reglade, U. 2019 Exploiting locality and translational invariance to design effective deep reinforcement learning control of the 1-dimensional unstable falling liquid film. *AIP Adv.* **9** (12), 125141.

Choi, H., Moin, P. & Kim, J. 1994 Active turbulence control for drag reduction in wall-bounded flows. *J. Fluid Mech.* **262**, 75–110.

Fischer, P. F., Lottes, J. W. & Kerkemeier, S. G. 2008 Nek5000 web page. Http://nek5000.mcs.anl.gov.

Fujimoto, S., Hoof, H. & Meger, D. 2018 Addressing function approximation error in actor-critic methods. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 1587–1596. PMLR.

Fukagata, K., Iwamoto, K. & Hasegawa, Y. 2024 Turbulent drag reduction by streamwise traveling waves of wall-normal forcing. *Annu. Rev. Fluid Mech.* **56**, 69–90.

Guastoni, L., Rabault, J., Schlatter, P., Azizpour, H. & Vinuesa, R. 2023 Deep reinforcement learning for turbulent drag reduction in channel flows. *Eur. Phys. J. E* **46** (4), 27.

Hammond, E. P., Bewley, T. R. & Moin, P. 1998 Observed mechanisms for turbulence attenuation and enhancement in opposition-controlled wall-bounded flows. *Phys. Fluids* **10** (9), 2421–2423.

Jiménez, J., Hoyas, S., Simens, M. P. & Mizuno, Y. 2010 Turbulent boundary layers and channels at moderate Reynolds numbers. *J. Fluid Mech.* **657**, 335–360.

Stroh, A., Frohnapfel, B., Schlatter, P. & Hasegawa, Y. 2015 A comparison of opposition control in turbulent boundary layer and turbulent channel flow. *Phys. Fluids* **27** (7), 075101.

Suárez, P., Alcántara-Ávila, F., Miró, A., Rabault, J., Font, B., Lehmkuhl, O. & Vinuesa, R. 2025 Active flow control for drag reduction through multi-agent reinforcement learning on a turbulent cylinder at $Re_D = 3900$. *Flow Turbul. Combust.* **115**, 3–27.

Vinuesa, R., Negi, P. S., Atzori, M., Hanifi, A., Henningson, D. S. & Schlatter, P. 2018 Turbulent boundary layers around wing sections up to $Re_c = 1{,}000{,}000$. *Int. J. Heat Fluid Flow* **72**, 86–99.

Wang, Y., Atzori, M. & Vinuesa, R. 2025 Opposition control applied to turbulent wing sections. *J. Fluid Mech.* **1010**, A29.