Optimization of Tollmien-Schlichting waves control: comparison between a deep reinforcement learning and particle swarm optimization approach

Mohammadikalakoo, B.; Kotsonis, M.; Doan, Nguyen Anh Khoa

DOI
10.2514/6.2024-0930

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the AIAA SCITECH 2024 Forum

AIAA SciTech Forum
8-12 January 2024, Orlando, FL
AIAA SCITECH 2024 Forum

10.2514/6.2024-0930

Optimization of Tollmien-Schlichting waves control: comparison between a deep reinforcement learning and particle swarm optimization approach

B. Mohammadikalakoo*, M. Kotsonis† and N. A. K. Doan‡
Department of Flow Physics and Technology, Section of Aerodynamics, Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, Delft 2629HS, The Netherlands

This work focuses on the suppression of Tollmien-Schlichting (TS) waves in a two-dimensional laminar boundary layer using optimized unsteady suction and blowing jets as an Active Flow Control (AFC) method. The suppression of TS waves via this AFC system is enabled through two artificial intelligence-based optimization methodologies: Single-Step Deep Reinforcement Learning (SDRL) and Particle Swarm Optimization (PSO). The primary aim of this research is to assess the performance of these methods in optimizing the AFC parameters with respect to convergence rate, computational efficiency, and ability to find an optimum control state. The findings demonstrate the success of both methods in finding appropriate control parameters resulting in TS wave attenuation by up to 40 dB in the maximum convective instability amplitude for the linear and nonlinear stages of development. The comparative study in this paper presents the effectiveness of the SDRL algorithm in optimizing the AFC system for TS waves' suppression and demonstrates that it can outperform PSO in terms of convergence rate and computational efficiency alongside a better performance in finding an improved optimum for linear control cases. However, the advantage of the SDRL-based controller over the PSO-based one diminishes in multi-frequency nonlinear control cases where the controller is located downstream and attempting to control highly amplified multi-modal TS waves.

I. Introduction

The control of complex, transitional, and turbulent flows holds a core role in improving the aerodynamic performance of modern transport and energy systems, thus enabling the design of more efficient aircrafts and wind farms. There is a wide range of applications for flow control especially active flow control (AFC) methods encompassing flow manipulation for laminar-to-turbulent transition and drag reduction, wind energy, combustion systems, noise reduction, etc. [1]. Transition is of particular importance as the aerodynamic drag is highly connected with the extent of laminar flow over aerodynamic surfaces. A delay of transition by 50% of the chord length can provide up to 15% reduction of the aerodynamic drag which translates into fuel (and cost) savings for future aircraft in addition to a reduction in greenhouse gas emissions [2].

However, achieving significant transition delay in realistic flight regimes is challenging due to the complex and nonlinear nature of the underlying flow dynamics. Receptivity as the first stage of the transitional process, describes the effect of external disturbances on the creation of the instability waves inside the boundary layer [3]. Various instability types, such as crossflow vortices, Tollmien-Schlichting waves, Kelvin-Helmholtz instabilities, and Rayleigh modes, can arise through this process [3].

Tollmien-Schlichting (TS) waves are primary viscous instabilities, especially dominant in two-dimensional boundary layers leading to laminar-to-turbulent transition in low-speed, low-disturbance flows where pressure gradients are negligible or aligned with the flow direction (unswept geometries) [4]. As such, the suppression of these waves during their early stages can delay the transition, extend laminar flow, and reduce skin friction drag, making them an attractive target for both passive and active flow control methods [5–7]. However, the control and suppression of naturally occurring TS waves in realistic flows is a challenging task due to their multi-frequency characteristics, often leading to a nonlinear behavior.

*PhD candidate, Department of Flow Physics and Technology, Kluyverweg 1, Delft 2629HS, The Netherlands
†Professor, Department of Flow Physics and Technology, Kluyverweg 1, Delft 2629HS, The Netherlands.
‡Assistant Professor, Department of Flow Physics and Technology, Kluyverweg 1, Delft 2629HS, The Netherlands.

To control such instabilities based on the interaction of sensor and actuator, two approaches are utilized to establish a model of the input-output dynamics [4]: (i) model-reduction techniques, which simplify the Navier-Stokes system and extract the input-output dynamics, and (ii) system identification techniques, which directly develop a mathematical model based on measurements, without explicitly accounting for the Navier-Stokes equations. Bagheri et al. [8] have demonstrated that a simplified two-dimensional setup can effectively attenuate disturbances in spatially developing flows. They designed a small optimal feedback controller based on a reduced order model, which successfully reduced the energy of two-dimensional wave packets generated upstream by orders of magnitudes. Bagheri et al. [9] extended this analysis by utilizing blowing/suction actuators instead of body forces. The model reduction approach can also be applied to fully three-dimensional configurations [10, 11]. System identification techniques have been employed in various experimental investigations, such as those conducted by Rathnasingham and Breuer [12]. They used a setup consisting of shear-stress sensors and synthetic jet actuators to successfully explore transition delay in flow control.

Most of these conventional control strategies are effective in attenuating deterministically generated TS waves in an early linear stage of development while failing to deal with the high dimensional and highly nonlinear stages of naturally occurring TS waves. Therefore, modern and responsive flow control techniques able to handle the growing complexity of the fluid dynamics problem are necessary. Towards this direction, Artificial Intelligence (AI) methods including Genetic Programming (GP) and Deep Reinforcement Learning (DRL) algorithms form a promising route for active flow control. For instance, GP-based control of the recirculation area of a backward-facing step was one of the first attempts by Gautier et al. [13] leading to 80% reduction of the recirculation area. Mitigation of separation by reattachment of the turbulent boundary layer to a sharp edge ramp is another application of successful GP-based control [14]. There are also recent attempts of DRL applications for drag reduction on a cylinder in laminar channel flow [15, 16] followed by a similar study in a turbulent regime [17]. These studies demonstrate the success of ML-based AFC as an effective and efficient control strategy [18]. Given the notable advantages of ML-based approaches for handling high nonlinearity and high dimensionality [18], the implementation of ML algorithms in the design of active flow control for TS waves' suppression has the potential to achieve improved performances compared to classical control. This exploration constitutes the main purpose of this work.

Here, the actuator comprises unsteady suction and blowing jets featuring a variable spatial sinusoidal pattern. These jets operate at the same frequency as the TS waves introduced at the inflow. Three parameters of this AFC system, including spatial wavelength (width) (J_L), amplitude (J_A), and phase (J_{ph}), are optimized using both the Single-Step Deep Reinforcement Learning (SDRL) algorithm [19] and the Particle Swarm Optimization (PSO) method [20]. The main objective of this study is to evaluate the effectiveness of these two methods in optimizing the AFC parameters by assessing their performance in terms of convergence rate, computational efficiency, and capability to identify an optimum set of parameters for the AFC.

The paper is organized as follows. Section 2 presents the methodology employed for the study, outlining the solver used for the numerical simulations and two optimization techniques. Building upon the methodology, section 3 delves into the numerical simulation aspects, where there is a description of the computational domain and test cases. Moving forward, section 4 presents the results obtained from numerical simulations, providing a comparative study between the PSO and SDRL methods. Finally, the insights gained from this study are summarized in section 5.

II. Methodology

This section outlines the numerical method and optimization techniques used to simulate the TS instabilities in an incompressible 2D laminar boundary layer and study their interaction with the AFC system. A non-linear Harmonic Navier-Stokes (HNS) solver [21] is implemented for this purpose, motivated by its ability to provide fast feedback for the optimization algorithms and also consider the TS waves in a frequency domain which facilitates the control strategy. A brief description of the HNS solver and a summary of the two methods of optimization (SDRL and PSO) implemented for the AFC are provided in this section.

A. Harmonic Navier-Stokes solver

A common approach for the study of boundary layer instabilities is to consider temporal and spatial harmonic behavior for perturbations based on assumptions of periodicity. This allows opportune simplifications of the Navier-Stokes equations via various assumptions. Employing the HNS solver represents a promising approach for solving Navier-Stokes equations when dealing with the harmonic perturbations defined by the ansatz in equation (1).

$$q'(x, y, z, t) = q(x, y)e^{i(\beta z - \omega t)} + c.c \tag{1}$$

In this method, the perturbation is expanded harmonically in the spanwise z direction and time. $q(x, y)$ is the perturbation shape function in the $x - y$ plane. The $c.c$ refers to the complex conjugate due to the presence of a complex field capturing growth and phase variations of the perturbation in the $x - y$ plane. β as the wavenumber is a non-dimensional complex coefficient related to the growth (decay) and periodicity of the perturbations in z direction and ω is the angular frequency. In this study, β is always real due to the z invariant base flow, and therefore the perturbation is not growing in the spanwise direction.

After inserting equation (1) into the general Navier-Stokes equations, and omitting the exponential terms of the harmonic perturbation, the Harmonic Navier-Stokes equations are derived. These equations possess inherent nonlinearity, and they can be transformed into two distinct forms: the Nonlinear Harmonic Navier-Stokes (NLHNS) equations and the Linear Harmonic Navier-Stokes (LHNS) equations. The details regarding the corresponding LHNS and NLHNS equations are provided in [21] for the interested reader.

The solution of the harmonic Navier-Stokes equations is obtained using an in-house solver developed at the TU Delft Low-Speed Lab [22]. In the discretization process, a uniform grid is employed in the streamwise, x, direction. To approximate derivatives in this direction, a fourth-order central differencing scheme is implemented. In the y direction, the equations are discretized using spectral collocation of Chebyshev polynomial basis functions. This choice of discretization results in a non-uniform grid in the y direction enabling enhanced resolution.

To initiate the flow, the Incompressible Linear Stability Theory (ILST) is utilized for the inflow conditions. However, special attention is given to the outflow conditions to prevent the propagation of unphysical information from the unknown outflow region. In both the Navier-Stokes equations and the perturbation equations systems, backward propagation of information can occur due to the presence of viscous terms and pressure terms [23]. To address the backward propagation problem, a buffering technique proposed by Joslin [24] is applied to the streamwise and spanwise perturbation velocities. This damping mechanism effectively counteracts the amplification of perturbation pressure and viscous terms from the outflow boundary. The buffer region is defined from 90% of the domain length to the end of the domain. At the wall, the no-slip condition is enforced, while the perturbations are forced to decay to zero in the free stream region. To transform the discretized equations, accounting for boundary conditions and derivative calculation procedures, a matrix system is formulated as can be seen in equation (2).

$$[A]q + [B]\frac{\partial q}{\partial x} + [C]\frac{\partial^2 q}{\partial x^2} = r \tag{2}$$

It should be noted that equation (2) includes non-linear modal interactions through the right-hand-side term (r). Solving this system provides the desired perturbation velocity and pressure, enabling further analysis and investigation. The explicit form of the matrices and equations of HNS is available in [21].

The LHNS is used for the first stage of the study to compare the performance of the SDRL and PSO in controlling TS waves during their linear stage of development. Additionally, a nonlinear test case will be carried out with the help of the NLHNS to investigate the effect of nonlinearity on the control performance.

B. Single-Step Deep Reinforcement Learning

SDRL as a policy-based optimization is a degenerative deep reinforcement learning algorithm in which single-step episodes are enough due to the state-independent optimal policy of the neural networks. Usually, this condition is satisfied in optimization cases. The agent in this method uses a policy network to represent the density function for the next evaluations. It employs covariance estimation to guide the process of policy improvement in the correct direction [19].

The agent at the first iteration tries a random policy based on an initial set of parameters (θ_0) and defines the corresponding actions. As can be seen in Figure 1, the agent samples a set of actions (a_t) from the current policy (π_θ) and is motivated to adjust the policy parameters (θ_{t+1}) in a way that the subsequent set of actions generates greater rewards (r_t). A smaller policy network is the main feature of the SDRL compared to other DRL methods due to the fact that the state is a constant single input (s_0), and therefore the state-action relation is not complex anymore.

Based on the schematic of the SDRL action loop, Figure 1, the input of the agent is always constant (s_0) and repeated at each generation while the agent draws actions from a probability distribution function which is a d-dimensional

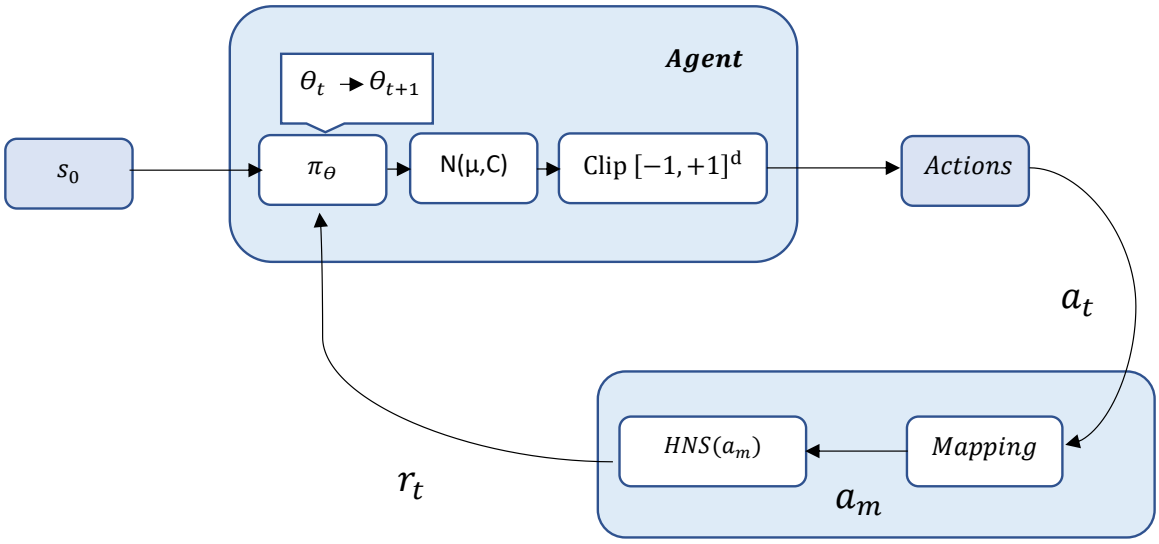

Fig. 1 Schematic of action loop for single-step deep reinforcement learning - policy based optimization [19]

multivariate normal distribution $\mathcal{N}(\mu, C)$. μ is the mean value and C is the full covariance matrix defined in [19]. Three neural networks are implemented to adjust the mean, standard deviation, and also correlation information of that distribution. Then, the actions are mapped (a_m) to the actual range of the AFC parameters and fed into the harmonic Navier-Stokes solver to simulate the TS waves in interaction with the specific AFC introduced by the agent. The reward of each case is calculated based on the maximum amplitude of the TS waves and the energy expenditure of the unsteady suction and blowing jets as will be described in equations 13 and 14. These reward values are fed back to the agent to update the policy with the main aim of maximizing the cost function (reward). This action loop continues until the convergence criterion, described in the results section, is reached.

More details regarding SDRL, its meta-parameters, and its formalism can be found in [19]. The implementation used in the present work is derived from the one available on the GitHub repository [25] related to [19].

C. Particle Swarm Optimization

The other optimization technique used in this work is PSO which is an artificial intelligence technique based on swarm intelligence introduced by Kenndy and Eberhart in 1995 [20]. There is a wide range of applications of this method in various fields of machine learning, adaptive control, etc. [26–28].

The PSO method is a computational method to find an optimal solution to a problem by improving a candidate solution based on its measure of quality (reward or loss function) in an iterative procedure. At the beginning of the algorithm, a population of the particles is initialized randomly in which each of the particles is a candidate solution to the problem. Each of these particles has a specific velocity and position which will be updated in each iterative try based on the local best of the particle itself and the global best of all the particles. So, particles are moving in the direction of the best global position while being affected by their own best local position. Considering a specific number of particles (m) in the search space with d dimensions while the position and velocity of each particle are described in equations (3) and (4) [29]:

$$x_i(t) = [x_{i1}(t), x_{i2}(t), ..., x_{id}(t)]^T \tag{3}$$

$$v_i(t) = [v_{i1}(t), v_{i2}(t), ..., v_{id}(t)]^T \tag{4}$$

the position and velocity of the particle i can be updated based on equations (5)-(6) at iteration $t+1$:

$$v_i(t+1) = \omega v_i(t) + c_1 r_1 (p_{bi}(t) - x_i(t)) + c_2 r_2 (g_b(t) - x_i(t)) \tag{5}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{6}$$

The p_{bi} and g_b are the local best and global best of the population that are defined in equations (7) and (8).

$$p_{bi}(t) = [p_{i1}(t), p_{i2}(t), ..., p_{id}(t)]^T \text{ where } p_{id}(t) = \max [c_{id}(t), c_{id}(t-1), c_{id}(t-2), ..., c_{id}(0)] \tag{7}$$

$$g_b(t) = [g_1(t), g_2(t), ..., g_d(t)]^T \text{ where } g_d(t) = \max [p_{1d}(t), p_{2d}(t), .., p_{id}(t), .., p_{md}(t)] \tag{8}$$

The $c_{id}(t)$ is the cost function in dimension (d) for particle (i) at iteration t. There are some hyperparameters for the PSO algorithm in equation (5) including ω as the inertia weight coefficient, c_1 and c_2 as the personal learning and global learning factors, respectively. $r_1$and r_2 are random variables distributed uniformly between zero and one. For the current work, the set of hyperparameters obtained after a parametric study is set to $(\omega, c_1, c_2) = (0.9, 0.9, 1.2)$.

The PSO algorithm as a metaheuristic method does not guarantee arrival at the global best of the optimization problem, but there is some specific set of hyperparameters that can ensure the convergence of the method and reach one of the best local optimum regions of the exploration area [29]. The dependency of the PSO method on several hyperparameters to achieve reliable and converged results is the main limitation of this method [30]. Due to the nature of the metaheuristic methods, PSO makes few or no assumptions about the optimization problem and does not consider the physics of the problem in the algorithm. Therefore, it is a convenient method of optimization for a wide range of problems, and it can search a large exploration area for optimization as a stochastic method of optimization. Another interesting feature of this method is that PSO does not require a differentiable problem since it does not use the gradient of the problem in the algorithm. This is an advantage with respect to classical optimization techniques such as gradient descent methods.

III. Numerical Simulation

Unsteady suction and blowing jets are chosen as actuation in the active flow control system. This actuation is characterized by the boundary condition applied to the wall with a sinusoidal actuation pattern ($J_A \cdot \sin(2\pi x_s)$) and acts in a wall-normal direction as shown in Figure 2. J_A represents the peak amplitude of actuation set by the controller, while x_s corresponds to the local normalized streamwise coordinate determined by the spatial wavelength of the actuator. The size of the computational domain and a typical configuration of the control components are shown in Figure 2. The computational domain is nondimensionalized with the Blasius length scale at the inflow (δ_0).

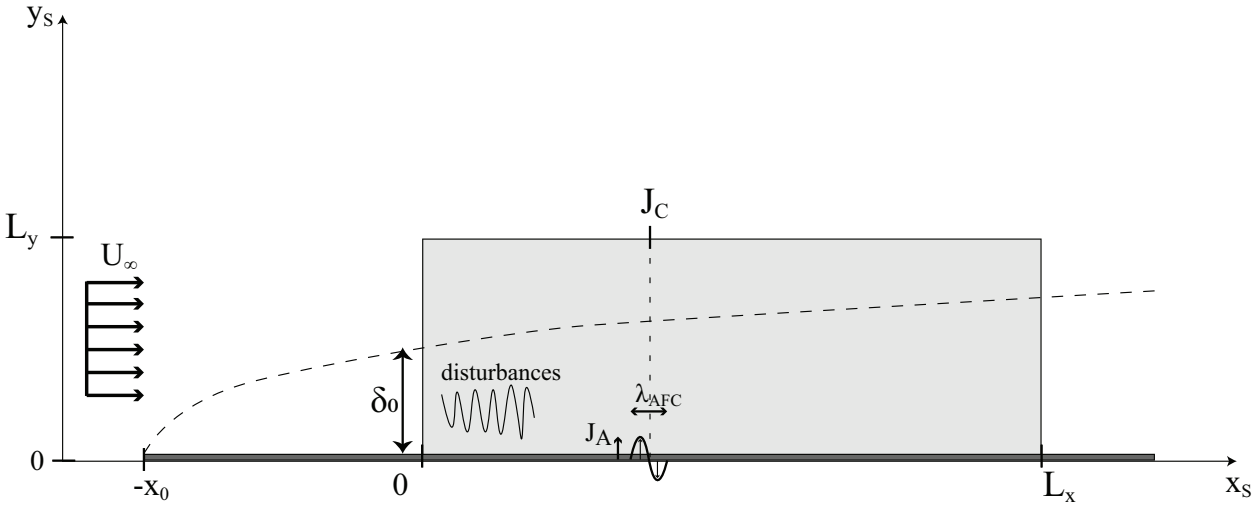

Fig. 2 Schematic view of the computational domain, $\Omega = (0, L_x = 2480) \times (0, L_y = 82)$ represented with gray region. $\lambda_{AFC} = L \times \lambda_{TS}, L \in [0.5, 2]$: Width of AFC while λ_{TS} is the wavelength of the TS wave above AFC. The AFC center is located at $(J_C = 660, 0)$. The inlet streamwise velocity U_∞, has a value of $10\frac{m}{s}$ corresponding to Reynolds number $Re_0 = \frac{U_\infty \delta_0}{\nu_0} = 400$.

As discussed previously, a uniform grid is employed in the streamwise direction and Chebyshev polynomial collocation points for the y direction. Two different resolutions in the streamwise direction are considered for the boundary layer solver (base flow) and the stability solver. The grid for the boundary layer solver consists of 5000 uniform divisions in the x-direction (1200 uniform divisions for the stability solver) and for the y-direction, Chebyshev polynomials with a non-uniform grid of 100 divisions. The positioning of these divisions is such that half of the points are located beneath $y/H = 0.1$, where H represents the domain height. The actual computational domain is the gray box since the boundary layer profile is implemented at the inflow based on Falkner-Skan-Cooke equations initiated at $(x_0 = -400\delta_0)$ to have a developed boundary layer at the inflow of the computational domain.

The numerical simulation of this work includes two different case studies of single-frequency linear and multi-frequency nonlinear control cases. The main focus of the current paper is on the single-frequency linear control case to compare the performance of the two optimization techniques. The study of the multi-frequency nonlinear control case is a preliminary study for the feasibility study of the AI-based active flow control system for suppression of naturally occurring TS waves.

It should be noted that the validation of the HNS has been done for linear and nonlinear TS waves and the results are in good agreement with the reference works [31, 32]. The details of this study are not presented here for the sake of brevity.

A. Test cases

A single-frequency linear control case of TS wave with frequency $\omega_1 = 0.0344$ and an initial root mean square amplitude of 0.25% at $Re_0 = 400$, is considered as the initial case study similar to the conditions described by Herbert [31]. The reference length and velocity for all case studies are the same as described in Figure 2. The test cases considered for the current stage of the work are divided into two categories based on the main optimization target of the two algorithms. The first group comprises cases in which the main target of optimization is the reduction of TS waves' maximum amplitude, see equation (13). The optimization target for the second group is the same as the first one with additional penalization regarding the energy expenditure of the unsteady suction and blowing jets. This energy penalization can be seen in the reward calculation, equation (14). It should be noted that a weight of 10% is assigned to the energy penalization term with the main aim of focusing on the suppression of TS waves as the primary goal.

$$E_{AFC} = \int_0^{L_x} |J_A \cdot \sin(2\pi x)|\, dx \tag{9}$$

$$E_{AFC_{max}} = \int_0^{L_x} |J_{A_{max}} \cdot \sin(2\pi x)|\, dx \tag{10}$$

$$E'(\%) = \left(\frac{E_{\mathrm{AFC}}}{E_{AFC_{max}}} \right) \tag{11}$$

$$\Delta A_{\mathrm{TS}}(\%) = \frac{\max\{A_{\mathrm{TS}}\}_{\mathrm{clean}} - \max\{A_{\mathrm{TS}}\}_{\mathrm{AFC}}}{\max\{A_{\mathrm{TS}}\}_{\mathrm{clean}}} \tag{12}$$

$$\mathrm{R1} = \Delta A_{\mathrm{TS}}(\%) \tag{13}$$

$$\mathrm{R2} = \Delta A_{\mathrm{TS}}(\%) - E'(\%) \times 0.1 \tag{14}$$

Equation (9) represents the mass flow rate of the unsteady suction and blowing jets which indicates the energy expenditure by the AFC. The same strategy has been implemented in equation (10) for the maximum possible mass flow rate which is imposed by the exploration space of the SDRL and PSO. The non-dimensional energy coefficient of E' is defined based on equation (11) as a penalization factor for final reward calculation. The reward function is defined as represented in equation (13) after computing the reduction of the TS maximum amplitude in equation (12). The reduction of TS maximum amplitude is defined as the relative difference of the maximum amplitude of the TS waves (based on maximum streamwise velocity u_{max}) in cases with and without active flow control (the latter being called the clean case).

The optimizer can adjust four key parameters of the AFC: the location (J_C), spatial wavelength (width) (J_L), amplitude (J_A), and phase (J_{ph}) of the unsteady suction and blowing jets. However, for practical purposes in the current study, the number of adjustable parameters has been limited to either three by fixing the location or two by fixing both the location and width of the AFC. All the case studies are presented in Table 1. The code names for these cases follow a specific format: the first two letters denote the optimization method, such as "SD" for SDRL and "PS" for PSO. The initial number in the code signifies the number of parameters under optimization. If the letter "E" is present in the name, it indicates that the reward function is penalized based on the energy expenditure of unsteady suction and blowing jets. Additionally, the number of attempts for each case is denoted by t1, t2, and t3 in the event of three different trials. For instance, SD3Et2 corresponds to the second trial of single-step deep reinforcement learning, focusing on optimizing three parameters while penalizing the reward function with the energy consumption of the actuators.

As can be seen in Table 1, three different trials have been done for each case to have a first assessment of the robustness of the final results to the random initialization of these two stochastic optimization algorithms. It should be noted that the cases with two optimization parameters are not repeated three times due to the lower sensitivity to stochastic initialization.

IV. Results

The test cases introduced in Table 1 are studied to find the optimized AFC system and to compare the performance of single-step DRL with the PSO technique. It is worth recalling that the hyperparameters of the SDRL and PSO are kept fixed for all case studies as explained in the methodology section to have a general baseline for comparison of the optimization techniques. A typical example of the suppression of TS wave with optimized unsteady suction and blowing jet is presented in Figure 3 and then the case studies are considered to have a comparative study.

The suppression of the TS waves can be observed in the reduction of the maximum amplitude of the TS waves in Figure 3 which is the definition of the reward for the optimization algorithms, see equation (13).

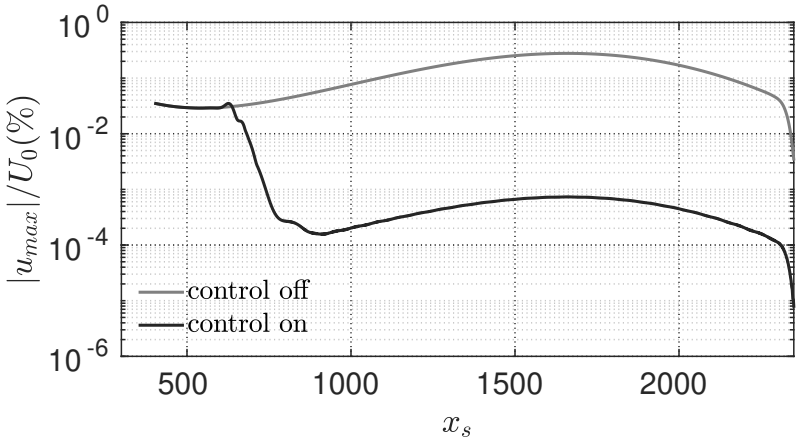

Fig. 3 Normalized amplitude evolution of the TS wave's maximum streamwise velocity (u_{max}) for the clean case (gray), the case with upstream control (SD3t2) (black). x_s is the streamwise coordinate nondimensionalized using the blasius length scale.

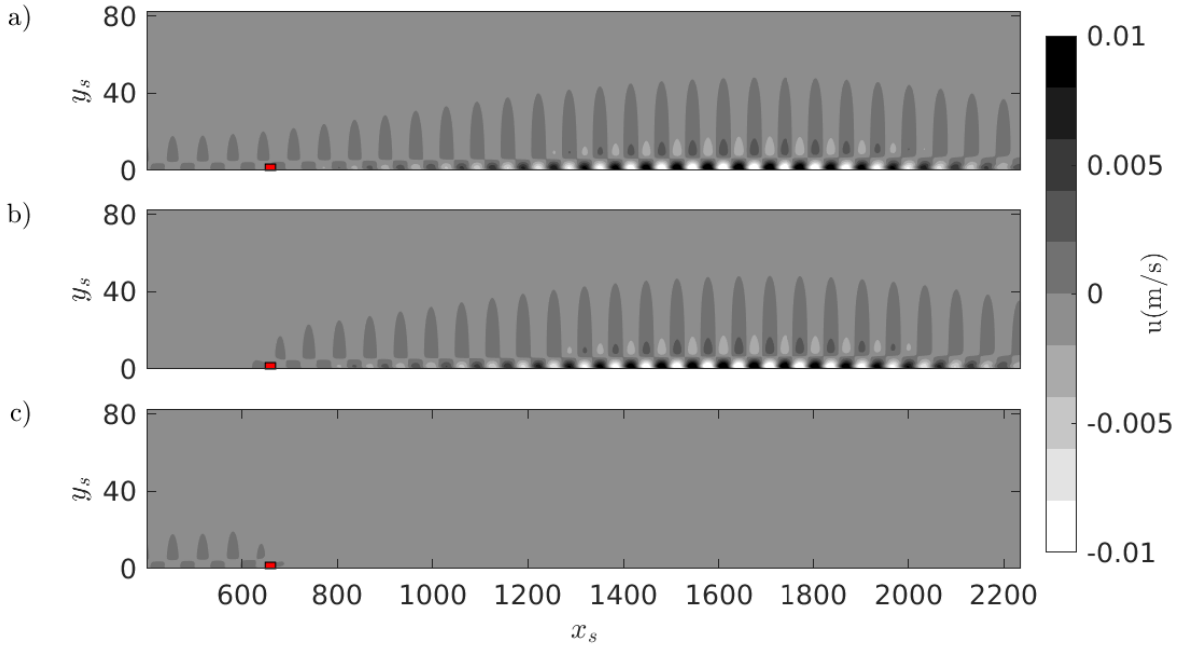

Fig. 4 Contours of streamwise (u) disturbance velocity for the linear single-frequency control case, (a) velocity field in the presence of TS waves triggered at the inflow, (b) velocity field with only control implemented upstream, (c) velocity field of the controlled TS waves. The position of the AFC is illustrated with a red rectangle ($x_s = 660$). Nondimensionalized data using Blasius length scale

Figure 4 shows the perturbation velocity field in the $x-y$ plane for three different situations. The amplitude of the oscillation for the velocity field corresponding to the uncontrolled TS wave reduces after interaction with the AFC wave leading to the suppressed TS wave. The opposition control strategy can be observed by comparing the velocity field of the TS wave in Figure 4 -(a) and the waves generated by the unsteady suction and blowing jets in Figure 4 -(b). The same amplitude and opposite phase of these waves lead to the suppression of the TS wave as illustrated in Figure 4 -(c).

Considering the 16 different case studies defined in section III, it is possible to have a flexible and general enough baseline to compare the performance of a machine learning algorithm with a metaheuristic algorithm in terms of convergence rate, computational efficiency, and ability to find the optimum parameters for the AFC system. The general control law extracted from all the obtained results of the SDRL and PSO shows that these two methods reach opposition control-like law as an effective method of suppressing TS waves similar to the results observed in Figure 4. The optimum parameters for AFC found by the SDRL and PSO algorithms are summarized in Table 1.

Table 1 Active control of TS waves, J is referring to unsteady suction and blowing jets and indices of A, ph, C, and L are referring to amplitude, phase, center's location, and relative width, respectively. $L = \frac{\lambda_{AFC}}{\lambda_{TS}}$, $A_{TS_{max}}$ is the maximum amplitude of controlled TS waves leading to amplitude reduction of $A_{red.}$. R is also the final reward value of the algorithm after reaching convergence at $Episode$.

Dataset	AFC				TS			Convergence
Case	J_A	$J_{ph}(rad)$	$J_C(\delta_0)$	J_L	$A_{TS_{max}}$	$R(\%)$	$A_{red.}(\%)$	Episode
SD3t1	0.000313	2.866719	660	0.618	0.000490	98.23	98.23	348
SD3t2	0.000245	2.747440	660	1.476	0.000010	99.96	99.96	553
SD3t3	0.000303	2.874470	660	0.625	0.000473	98.30	98.29	494
PS3t1	0.000189	1.110100	660	1.252	0.000467	98.32	98.31	824
PS3t2	0.000353	0.980320	660	1.628	0.000835	96.99	96.99	728
PS3t3	0.000430	1.162200	660	0.501	0.000478	98.28	98.28	727
SD3Et1	0.000290	2.711142	660	1.560	0.000011	97.70	97.69	477
SD3Et2	0.000196	2.855238	660	0.930	0.000015	99.04	99.03	456
SD3Et3	0.000183	2.823958	660	1.152	0.000073	98.68	98.68	485
PS3Et1	0.000248	1.0336	660	1.484	0.0000095	98.13	99.96	801
PS3Et2	0.000430	0.90951	660	1.695	0.000015	96.31	99.95	748
PS3Et3	0.000429	0.90955	660	1.695	0.000148	96.20	99.95	748
SD2Et1	0.000188	2.846498	660	1.000	0.000051	98.88	98.87	286
PS2Et1	0.000188	1.1343	660	1.000	0.000050	98.88	99.82	808
SD2t1	0.000186	2.853509	660	1.000	0.000458	98.35	98.35	256
PS2t1	0.000188	1.1343	660	1.000	0.000050	98.88	99.82	748

The overall suppression of TS waves of up to 99.96% (40 dB) based on results of Table 1 show the success of both the PSO and SDRL algorithms in finding a proper control strategy to suppress TS waves. The results reveal some noteworthy trends. Specifically, the selection of SDRL (case: SD3Et2) for having a narrower actuator (relative width: 0.93) with enhanced amplitude, deviates from the principles of opposition control applied to TS waves in some cases.

Comparing the convergence episodes in Table 1, it becomes evident that SDRL consistently outperforms PSO in terms of speed when searching for optimal parameters. A better perception can be achieved by observing the convergence behavior of these two methods in Figure 5. The convergence criterion for these methods is considered by defining ±0.2% bandwidth for the rate of variation of the reward's moving average. The Observing window for the moving average is over 20 episodes. Figure 6 shows this criterion and corresponding results.

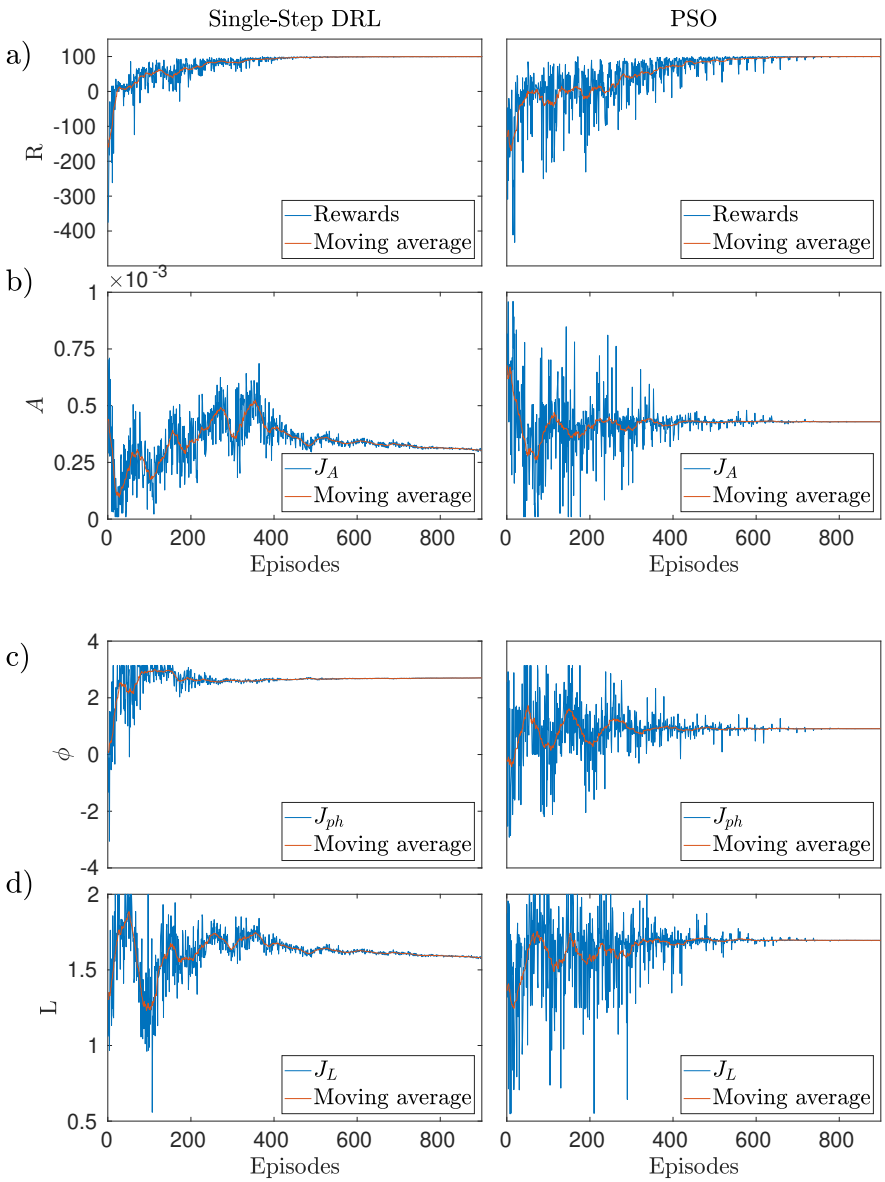

Fig. 5 Convergence of SDRL and PSO algorithm towards optimum control (SD3Et1 vs PS3Et3), a) Reward function, b) Amplitude of unsteady suction and blowing jets (AFC), c) Phase of AFC, d) Relative Width of AFC

By comparing the convergence trends of different optimization parameters in Figure 5 for SDRL and PSO, two different strategies of optimization can be seen which finally leads to a similar control strategy. In the PSO, all the parameters to be optimized are converging at almost the same rate and pass the convergence criterion after a similar number of episodes while in SDRL the rate of convergence is different for the three parameters of the AFC. For instance, in case SD3Et1 shown in Figure 5, the phase is the first parameter that is converged while amplitude and width of AFC are the next parameters with a slower convergence rate. It appears that the SDRL narrows down the exploration space dimension as it approaches convergence. This potential reduction in dimensionality might be the underlying factor behind SDRL's faster overall convergence when compared to the PSO. Nevertheless, more extensive investigations into SDRL's optimization strategy are necessary to gain a comprehensive understanding of the primary mechanism behind its faster convergence.

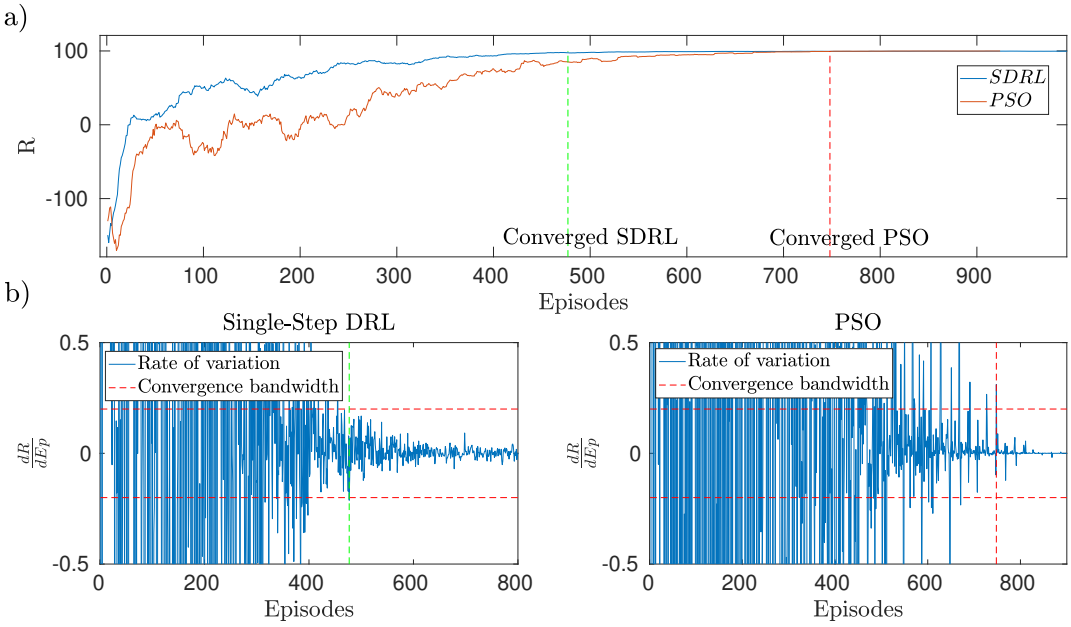

Fig. 6 Evolution of reward function for SDRL and PSO algorithm (SD3Et1 vs PS3Et3) a) Moving average of reward function, b) Rate of variation for reward function based on episodes

The dashed vertical lines indicated in Figure 6 are representing the minimum number of episodes required to reach convergence for both the SDRL and PSO algorithms. The number of episodes for SDRL and PSO is 477, and 748 episodes, respectively which shows a 57% faster performance of SDRL compared to the PSO algorithm. This difference was observed to be even higher for the cases with only two optimization parameters where SDRL was 3 times faster than PSO. The fast convergence of SDRL compared to PSO is a noticeable advantage of this machine-learning algorithm, especially for fluid dynamic problems with high computational cost per episode. For example, in the case of Figure 6, SDRL requires almost 270 fewer episodes than PSO for equal performance. This difference can be translated to the saving of 135 CPU hours for linear control cases which can increase for nonlinear cases assuming that SDRL maintains its performance in a nonlinear context. The test cases are all simulated on a workstation with 16 cores Intel Xeon(R) W-2245 processor and 128 GB of RAM, and half a CPU hour is the cost of running the LHNS solver for each episode on this machine.

The ability of SDRL in finding an improved optimum is also evidence of SDRL outperforming the PSO method: when comparing the maximum suppression of TS waves in the different test cases presented in Table 1, SDRL shows an increased suppression. Although the suppression of the TS waves in all case studies is high, slightly different results can still be seen by comparing the SDRL and PSO algorithms. For instance, the maximum amplitude suppression for the PSO method (case: PS3t1) is 38 dB compared to the 40 dB in the SDRL method (case: SD3t2).

By comparing two different optimization scenarios with and without energy penalization (SD3Et2, and SD3t2) in Table 1, one can say that SDRL tries to find a smaller AFC with smaller amplitude while performing a trade-off between energy expenditure of unsteady jets and TS wave amplitude reduction. Therefore, the suppression of TS wave is slightly

Table 2 Outcome of multi-frequency nonlinear control cases, $NL1$: upstream control, $NL2$: downstream control

Cases #	Modes	Δu_{max} (%)	
		PSO	$SDRL$
$NL1$	M1	99.93	99.22
	M2	98.67	92.12
	M3	95.3	98.97
$NL2$	M1	86.43	82.17
	M2	98.89	77.64
	M3	95.07	97.02

lower when defining the target of optimization with energy penalization.

A. Nonlinearity effects on controller performance

Two additional test cases are considered in this part to study the effect of nonlinear modal interactions on the control performance of SDRL and PSO methods. A nonlinear multi-frequency control case with a combination of the first three harmonics of the TS wave is considered while the frequency of the third harmonic is the same as the previous case studies ($\omega_1 = 0.0344$). The initial amplitude of the first two modes is the root mean square amplitude of 0.5% while the third mode is kept at 0.25% ([31]). The spatial wavelength of the unsteady suction and blowing jets is fixed for both cases ($\lambda_{AFC} = \lambda_{TS}$). The fixed upstream ($C_x = 660\delta_0$) and downstream ($C_x = 1653\delta_0$) locations of the actuator are selected for cases $NL1$ and $NL2$, respectively. This positioning will help to assess the link between the controller performance and the nonlinear evolution of the TS wave.

The outcome of these control cases is provided in Table 2 which indicates similar control performance of SDRL and PSO methods when the actuator is positioned upstream ($NL1$). However, when the actuator is placed downstream ($NL2$), the SDRL-based controller is outperformed by the PSO-based. Specifically, the SDRL-based controller is not able to effectively mitigate the second harmonic of the TS wave when compared to the PSO-based controller.

These findings indicate that the extended exploration space in multi-frequency nonlinear control cases adversely affects the convergence rate of the SDRL-based controller, causing it to converge at a slower pace similar to that of the PSO-based controller. Moreover, the higher complexity of the optimization problem due to the presence of developed nonlinear interactions downstream poses a challenge for the SDRL-based controller resulting in decreased performance.

It should be noted that the results presented in this part are based on a small subset of nonlinear case studies. Drawing a definitive conclusion regarding the controllers' performance in nonlinear control cases requires conducting additional control cases and sensitivity analyses to evaluate how the SDRL-based controller's initial randomization affects its performance.

V. Conclusion

The application of single-step deep reinforcement learning and particle swarm optimization algorithms to find an optimum unsteady suction and blowing jets as active flow control is studied with the main aim of Tollmien-Schlichting waves' suppression. The utilization of SDRL and PSO-based controllers can suppress linearly developed Tollmien-Schlichting waves by up to 99.96% (40 dB). This indicates the effectiveness of these methods in devising suitable control strategies. The strategy identified by the SDRL and PSO involves an opposition control method, which generates an active flow control wave of the same amplitude but in the opposite phase to the TS wave.

The comparative analysis of SDRL and PSO methods has also been done across 16 different test cases. These cases vary in the number of optimization parameters and the definition of the reward function, which either includes or excludes the penalization of performance based on the energy expenditure of the actuator. Faster convergence of SDRL by up to 3 times when compared to the PSO method was observed in single-frequency linear control cases resulting in notable computational power savings. The SDRL-based controller demonstrated superior performance in identifying improved optima compared to PSO. These findings underscore the effectiveness of the SDRL-based

controller in suppressing linearly developed TS waves.

Two additional nonlinear test cases were conducted in a multi-frequency nonlinear control scenario to investigate if the compelling features of SDRL-based controller still exist when dealing with the nonlinearly developed TS waves. The findings reveal that the expanded exploration space in multi-frequency nonlinear control cases adversely affects the convergence rate of the SDRL-based controller, making it converge at a slower pace similar to the convergence rate of the PSO-based controller. The increased complexity resulting from downstream nonlinear interactions poses additional challenges for the SDRL-based controller, resulting in a decline in its performance. Based on the initial investigation into multi-frequency nonlinear control cases, it becomes apparent that the SDRL-based controller no longer demonstrates an advantage over the PSO-based controller in highly nonlinear control scenarios. To fully assess the SDRL-based controller's capabilities in suppressing nonlinearly developed TS waves, more comprehensive studies involving a greater variety of test cases with different levels of nonlinearity are necessary. This will be studied in future works.

References

[1] Pino, F., Schena, L., Rabault, J., and Mendez, M. A., "Comparative analysis of machine learning methods for active flow control," *Journal of Fluid Mechanics*, Vol. 958, 2023. https://doi.org/10.1017/jfm.2023.76.

[2] Niu, X., and Li, J., "Investigation and Design of the Transonic Laminar Flow Characteristics in a Laminar Aircraft," *Applied Sciences*, Vol. 12, No. 22, 2022. https://doi.org/10.3390/app122211820.

[3] Luo, J. S., and Zhou, H., "On the generation of Tollmien-Schlichting waves in the boundary layer of a flat plate by the disturbances in the free stream," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, Vol. 413, 1987, pp. 351–367.

[4] Bagheri, S., and Henningson, D. S., "Transition delay using control theory," *Phil. Trans. R. Soc.*, Vol. 369, 2011, pp. 1365–1381. https://doi.org/10.1098/rsta.2010.0358.

[5] Thomas., A. S., "The Control Of Boundary-Layer Transition Using A Wave-Superposition Principle." *Journal of Fluid Mechanics*, 1983, pp. 233, 250.

[6] Cossu C., L. B., "Stabilization of Tollmien-Schlichting waves by finite amplitude optimal streaks in the Blasius boundary layer." *Phys. Fluids*, Vol. 14, No. L57, 2002.

[7] Kotsonis, M., Giepman, R., Hulshoff, S., and Veldhuis, L., "Numerical study of the control of Tollmien-Schlichting waves using plasma actuators," *AIAA Journal*, Vol. 51, No. 10, 2013, pp. 2353–2364.

[8] Bagheri, S., Brandt, L., and Henningson, D. S., "Input-output analysis, model reduction and control of the flat-plate boundary layer," *Journal of Fluid Mechanics*, Vol. 620, 2009, pp. 263–298. https://doi.org/10.1017/S0022112008004394.

[9] Bagheri, S., Åkervik, E., Brandt, L., and Henningson, D. S., "Matrix-free methods for the stability and control of boundary layers," *AIAA Journal*, Vol. 47, 2009, pp. 1057–1068. https://doi.org/10.2514/1.41365.

[10] Semeraro, O., Bagheri, S., Brandt, L., and Henningson, D. S., "Linear control of 3D disturbances on a flat plate," *7th IUTAM Symposium on Laminar–Turbulent Transition*, Vol. 18, 2009, pp. 373–378.

[11] Semeraro, O., Bagheri, S., Brandt, L., and Henningson, D. S., "Feedback control of three-dimensional optimal disturbances using reduced-order models," *Journal of Fluid Mechanics*, Vol. 677, 2011, pp. 63–102. https://doi.org/10.1017/S0022112011000620, URL https://doi.org/10.1017/S0022112011000620.

[12] Rathnasingham, R., and Breuer, K. S., "Active control of turbulent boundary layers," *Journal of Fluid Mechanics*, Vol. 495, 2003, pp. 209–233. https://doi.org/10.1017/S0022112003006177.

[13] Gautier, N., Aider, J.-L., Duriez, T., et al., "Closed-loop separation control using machine learning," *Journal of Fluid Mechanics*, Vol. 770, 2015, pp. 442–457.

[14] Debien, A., von Krbek, K., Mazellier, N., et al., "Closed-loop separation control over a sharp edge ramp using genetic programming," *Experiments in Fluids*, Vol. 57, 2016, p. 40.

[15] Rabault, J., Kuchta, M., Jensen, A., et al., "Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control," *Journal of Fluid Mechanics*, Vol. 865, 2019, pp. 281–302.

[16] Rabault, J., and Kuhnle, A., "Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach," *Physics of Fluids*, Vol. 31, 2019, p. 094105.

[17] Ren, F., and Tang, H., "Active flow control of flow past a circular cylinder at moderate Reynolds number using deep reinforcement learning," *International Symposium on High-Fidelity Computational Methods and Applications*, Shanghai, China, 2019, p. 141.

[18] Ren, F., Hu, H. b., and Tang, H., "Active flow control using machine learning: A brief review," *Journal of Hydrodynamics*, Vol. 32, No. 2, 2020, pp. 247–253. https://doi.org/10.1007/s42241-020-0026-0.

[19] Viquerat, J., Duvigneau, R., Meliga, P., Kuhnle, A., and Hachem, E., "Policy-based optimization: single-step policy gradient method seen as an evolution strategy," *Neural Computing and Applications*, 2022. https://doi.org/10.1007/s00521-022-07779-0.

[20] Kennedy, J., and Eberhart, R., "Particle swarm optimization," *Proceedings of ICNN'95 - International Conference on Neural Networks*, Vol. 4, 1995, pp. 1942–1948 vol.4. https://doi.org/10.1109/ICNN.1995.488968.

[21] Westerbeek, S., Sumariva, J. A. F., Michelis, T., Hein, S., and Kotsonis, M., "Linear and Nonlinear Stability Analysis of a Three-Dimensional Boundary Layer over a Hump," *AIAA SCITECH Forum*, 2023. https://doi.org/10.2514/6.2023-0678.

[22] Westerbeek, S., Hulshoff, S., Schuttelaars, H., and Kotsonis, M., "DeHNSSo: The Delft Harmonic Navier-Stokes Solver." https://gitlab.tudelft.nl/flow-stability-and-control/DeHNSSo, 2023.

[23] Streett, C., and Macaraeg, M., "Spectral multi-domain for large-scale fluid dynamic simulations," *Applied Numerical Mathematics*, 1989. https://doi.org/10.1016/0168-9274(89)90058.

[24] Joslin, R. D., "Validation of three-dimensional incompressible spatial direct numerical simulation code: A comparison with linear stability and parabolic stability equation theories for boundary-layer transition on a flat plate," *Technical Publication (TP)*, 1992.

[25] Viquerat, J., "Policy-based optimization : single-step policy gradient seen as an evolution strategy," https://github.com/jviquerat/pbo, 2022.

[26] Wenjing, Z., "Parameter identification of lugre friction model in servo system based on improved particle swarm optimization algorithm," *Chinese Control Conference*, 2007.

[27] Bai, Q., "Analysis of particle swarm optimization algorithm," *Computer and information science*, 2010.

[28] Hung C., L. W., "Hybridization of particle swarm optimization with the k-means algorithm for image classification," *IEEE Symposium on Computational Intelligence for Image Processing*, 2009.

[29] Li, Y., and Zhang, Y., "Hyper-parameter estimation method with particle swarm optimization," *arXiv*, 2020. URL http://arxiv.org/abs/2011.11944.

[30] Aote, S. S., Raghuwanshi Principal, M. M., and Latesh Malik HOD, R., "A Brief Review on Particle Swarm Optimization: Limitations & Future Directions," *International Journal of Computer Science Engineering*, 2013.

[31] Herbert., "Special Course on Progress in Transition Modelling." *J. AGARD R793*, Vol. 5, 1994.

[32] Chang, C. L., Malik, M. R., Erlebacher, G., and Hussaini, M. Y., "Compressible stability of growing boundary layers using parabolized stability equations," American Institute of Aeronautics and Astronautics Inc, AIAA, 1991, pp. 1–21. https://doi.org/10.2514/6.1991-1636.