Proceedings of the 1st international Symposium on AI and Fluid Mechanics
Paper No S1 P2

Accelerating Deep Reinforcement Learning-based Active Flow Control on Analogous Geometries through Transfer Learning

Seyed Massoud Mousavi[1]; Mohammad Jebelli[1]; Mehran Masdari[2,*];

[1] *University of Tehran, Tehran, Iran*

[2]*City St George's University of London, London, United Kingdom*

Keywords: Active flow control; Deep reinforcement learning; Transfer learning; bluff body

Abstract: This research explores the application of Deep Reinforcement Learning (DRL) in Active Flow Control (AFC) for flow field around elliptical cylinders of varying aspect ratios (AR). The trained model in the circular cylinder is tested on elliptical cylinders, which are mounted in a channel with two synthetic jets actuation. The DRL model learned from circular geometry may be adequate for elliptical ones by maintaining the original sensor location. However, the research goes on to examine the influence of transfer learning to enhance control performance. Transfer learning can reduce the training time for the agent significantly and provide rewards with less fluctuations, if the sensor placement remains identical. The results highlight the importance of sensor placement for optimal performance in DRL-based AFC. It is also effective in drag reduction, leading to comparable performance with an agent directly trained in the target geometry only if the sensor placement between secondary training and test stage remains identical.

Introduction

AFC appeared as an innovative approach in controlling and optimizing fluid dynamic systems for a great variety of applications. However, this efficient design for AFC is facing some tough challenges; addressing varying flow conditions and geometric configurations. DRL provides a very encouraging framework for AFC, permitting the formation of intelligent control policies without a priori requirement for detailed models of the flow system.

The most well-known work is that of Rabault et al. [1] which reported the first applications of artificial neural networks (ANN) trained with a DRL agent for AFC. In this study, where 2D simulations of a circular cylinder at Re=100 was conducted, the agent effectively learned a control strategy by manipulating the mass flow rates of two jets positioned on either side of the cylinder. Most importantly, studies investigated new variants of different methods as means of increasing the application and overcoming weaknesses; including enhancing/suppressing vortex-induced vibration [2–5], suppressing flow separation on an airfoil[2–5], and optimizing conjugate heat transfer [6,7].

The training of DRL agents is a high-cost computational task that takes a long time to converge, particularly when adapting to new scenarios of flow control. Many studies were conducted by different researchers to mitigate this problem. Rabault et al. obtained speedup by using parallelization[8]. In another study, the authors accelerate the filling of the experience buffer by down-sampling [9]. Inspired by the idea of transfer learning in computer science, some attempts have been made to further reduce the computational cost by reusing or retraining a trained agent for a new task. Wang et al.[10] showed that agents trained at higher Re are still capable of controlling the flow at lower ones. He et al.[11] delved into policy transfer, with their simulations showing that agents successfully trained in 2D environments transferred into 3D environments while maintaining significant drag reduction. Yan et al. [12] focused on transfer learning for a 3D square cylinder. Their results showed a significant improvement with a training cost reduction to half.

The application of transfer learning in DRL-based AFC has been relatively underexplored, especially with the focus on similar geometries. The current study attempts to examine the ability of a trained agent to control the flow in similar geometries and the possible benefits of transfer learning. Among geometries with similar characteristics, cylinders with elliptical cross-sections are taken to be geometries that most closely resemble a circular cylinder, making them good candidates for detailed investigation.

Problem description

In this work, DRL-based AFC is used to control the flow field around elliptical cylinders of various ARs. The study aims to evaluate the capability of a model trained on a circular cylinder to control the flow around elliptical cylinders, reduce training time and computational cost. In DRL-based AFC, the system has two key parts: the environment, which models fluid dynamics and provides state information, and the agent, which learns from observations and rewards to apply control strategies (Figure 1).

In the environment, a cylinder is confined in a channel, and actuation is achieved through two synthetic jets. The jet velocities can be both positive and negative, and the directions are aligned along the cross-flow axis. These synthetic jets having a width of $\omega=\pi/18$ are applied vertically to the cylinder surface at its top and bottom. The flow is considered to be viscous and incompressible, so the governing equations for the momentum and continuity form as equations (1) and (2), respectively.

$$\frac{\partial u}{\partial t} + u.(\nabla u) = -\nabla p + Re^{-1}\Delta u \tag{1}$$

$$\nabla. u = 0 \tag{2}$$

The flow domain is a rectangular channel, with dimensions of 22D in length and 4.1D in width, similar to the benchmark case of Schäfer et al.[15] (Figure 1 – Environment). The inlet is located at the left boundary, which imposes a parabolic profile of flow velocity distribution in the y-direction. The upper and lower boundaries are the channel walls with no-slip condition. The right boundary acts as an outlet with a zero-pressure gradient. The wall of the cylinder, except for the jet hole, is no-slip. The mass flow rates of the two jets are set to zero. The simulations were established on the open-source computer code FEniCS[13] and the grids were generated using Gmsh[14].

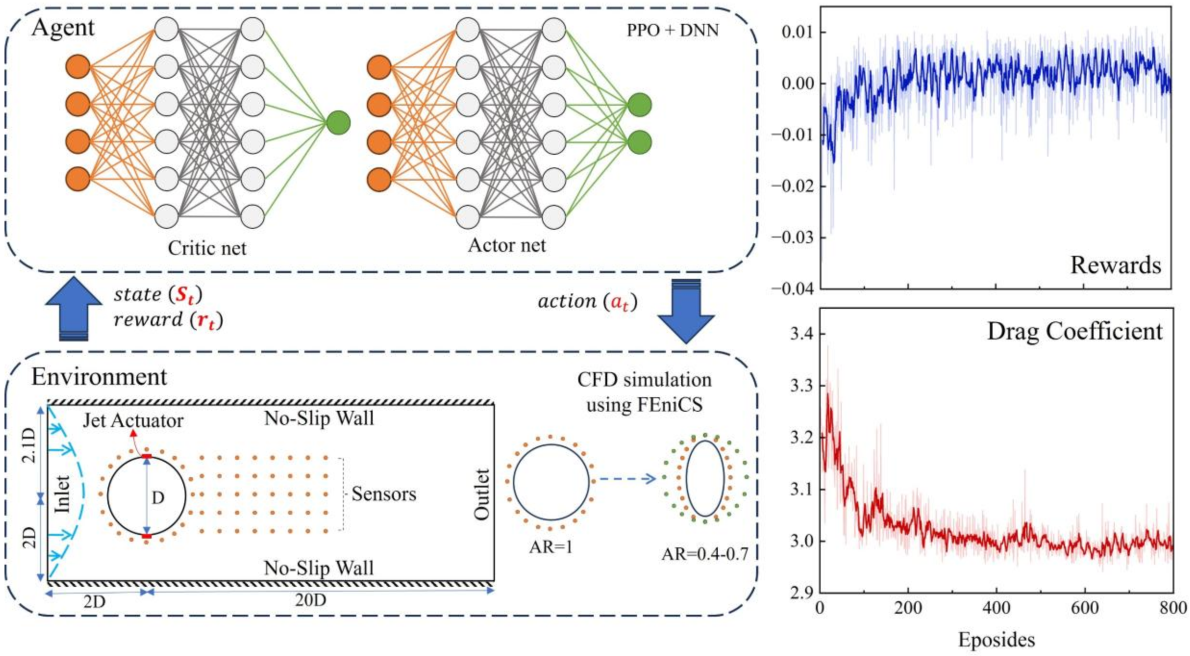

Figure 1) DRL approach and variation of reward function and drag coefficient in training process at AR=1

Based on a temporal dependency study a time step of $dt = 0.002$ is used for numerical simulations. In order to check the validation of numerical approach, the results for the circular cylinder at Re=100 is compared with the data of Schäfer et al. [15] showing a good agreement and indicating the validity of the numerical simulations (Table 1).

Table 1. The validation of numerical method

Re=100	Present study	Schäfer et al.[15]
C_D max	3.22	3.22-3.24
C_L max	1.01	0.99-1.01
St	0.29	0.295-0.305

An agent as the second component of DRL framework (Figure 1 – Agent), receives information (states and rewards) through sensors and actions based on its present state. The agent learns iteratively using data made of states, actions and rewards by interacting with its environment. The agent which learns the control policy using Proximal Policy Optimization (PPO) algorithm which consist of critic and actor networks.

Following the conclusion drawn by Rabault et al.[1] a total of 165 sensors implemented around the cylinder and also the wake flow region can provide sufficient flow information for efficient learning.

The reward function is set in such a way to guide the agent into learning the control policy aimed at drag reduction and avoiding unwanted solutions with large lift fluctuations, defined in equation (3):

$$r = -\langle C_D \rangle_T - 0.2|\langle C_L \rangle_T| \tag{3}$$

In this case, the agent successfully learns an effective control strategy, reducing the cylinder's drag. In the first 300 episodes of training, the agent's behavior is highly stochastic, evaluating various combinations of suction and injection to affect flow control. With each successive training episode, the agent identifies strategies that yield rewards higher than the initial method and keep drag coefficients low (Figure 1 – Rewards and Drag Coefficient).

The effectiveness of DRL in AFC of a circular cylinder has been demonstrated in many studies. In the next section, its capability to control flow in similar geometries, including ellipses with different ARs will be explored from two perspectives. First, the trained agent will be directly tested. Then, the agent will be retrained to determine whether the initial training can shorten the retraining process. In test and transfer learning focused on different geometries, sensor placement has an important implication for agents to sense flow dynamics. This study also looks into sensor placement around elliptical cylinders, especially in when training is conducted under circular geometry with circular sensor placement, and testing or re-training is held in a different geometry.

Results

This study on the trained agent's capability in similar geometries and policy transfer is carried out in two steps. First, try to control flow around elliptical cylinders using a trained agent at circular geometry. This would evaluate how applicable the trained agent is to similar geometries, consequently reducing the need for training on new cases. Secondly, in further training of this agent on elliptical geometry, see whether the aforementioned trained agent accelerates learning and shortens the training time.

In the first step, the trained agent at a circular cylinder is directly applied to control the flow around elliptical cylinders with AR=0.4-0.7, where the AR represents the ratio of the horizontal to vertical diameter. The variation of drag coefficient (Figure 2 – Cylinder Drag Coefficient) point out that the trained agent can successfully reduce drag in elliptical configurations, especially with higher ARs, if the sensor placement remains unchanged with respect to their placement during learning process (circular placement in this case). For AR=0.7, the trained agent significantly reduces the mean drag coefficient and almost eliminates all fluctuations with circular sensor placement. The elliptical placement, adopted to the body's geometry, also leads to mean drag reduction, but is lesser, and the drag fluctuations remain considerable. In AR=0.6, circular placement eliminates drag fluctuations, taking longer time, and the elliptical placement leads to even weaker drag reduction. At AR=0.5, drag reduction is still greater with circular placement; however, in contrast to the last cases, it does not totally eliminate fluctuations of drag. Here, elliptical placement appears to have a more negative effect as it provides a much lesser drop in the average drag. AR=0.4 shows also a similar trend.

The negative effect of elliptical placement for the sensors is clearly visible as the AR reduces (Figure 2 – Cylinder Drag Reduction). In high aspect ratio of AR=0.7, mounting sensors elliptically, leads to a less drag reduction of 8% while Lowering the AR to 0.4 leads to weaker performance falling from 25% in circular placement to only 5% in elliptical one. This trend indicates that as the sensor placement deviates further from that used in the training process, its performance declines more significantly.

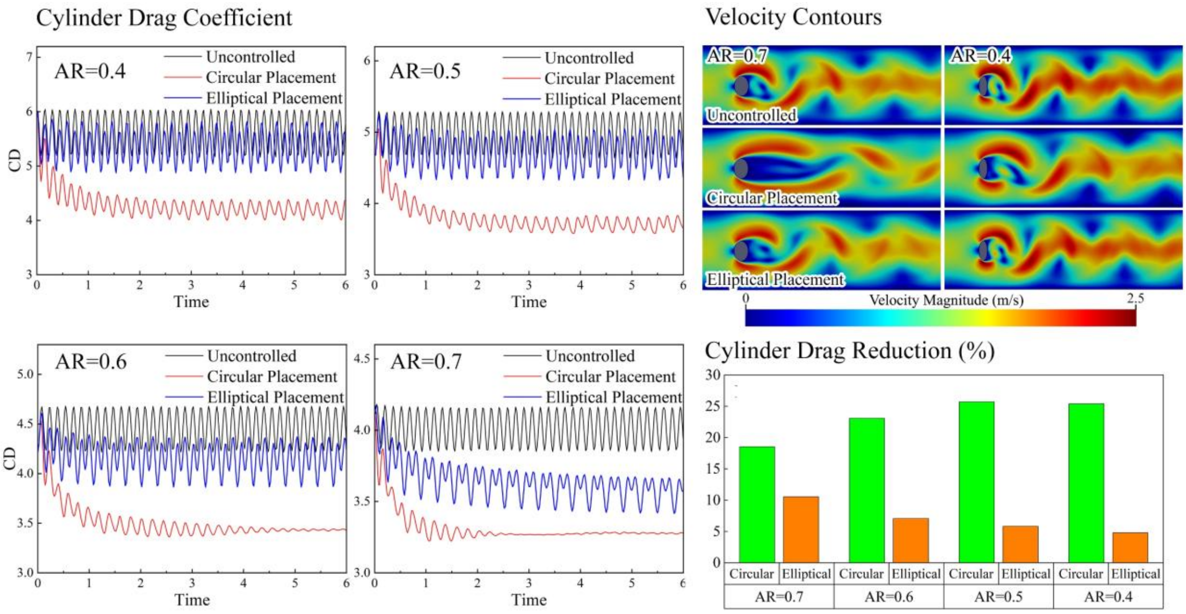

Figure 2) Variation of drag coefficient with velocity contours at AR=04.-0.7 (Agent trained at AR=1)

Comparing the velocity magnitude contours (Figure 2 – Velocity Contours) shows that the agent controls the flow around the elliptical cylinder with AR=0.7 using a circular sensor placement resulting in a longer recirculation zone. An elliptical placement leads to only small changes in flow structure. Contours at AR=0.4 shows only small changes with circular placement. Overall, the results from testing in different ARs suggest that changing the sensor placement for a trained agent weakens the system's control performance, even if the new placement aligns with the geometry of the target body.
To evaluate sensor placement effect, the training and testing process are conducted at AR=0.5 including training with either circular or elliptical sensor placement and testing on the same or opposed configuration. The variation of drag coefficients and velocity contours are presented in Figure 3.

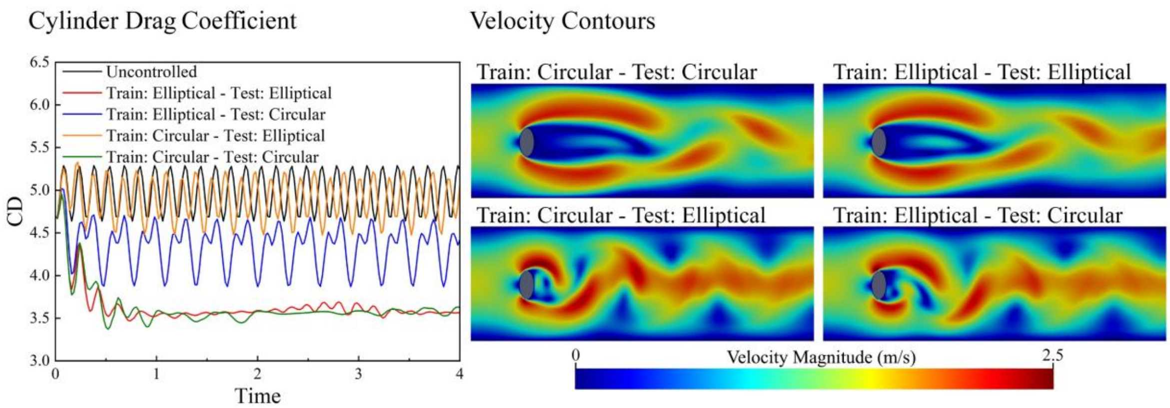

Figure 3) Variation of drag Coefficient and velocity contours at AR=0.5 with different sensor placements

Comparing the drag coefficient shows significant reduction when training and testing placements are defined identical. The lowest mean drag appears in these cases with almost eliminated fluctuations. Training in elliptical and testing with circular placement leads to lower mean drag, while the reverse case results in drag coefficient close to the uncontrolled condition. The instantaneous velocity contours also confirm the higher effectiveness of similar placements with stretched shear layers and postponed vortex shedding (Figure 3 – Velocity Contours). This suggests that the effectiveness of AFC is highly dependent to sensor placement consistency. With the same placements between training and testing, increased stability and drag reduction have been experienced, while the opposite placements lead to less control efficiency.

In the second phase of this study, the training process to control the flow over an elliptical cylinder is studied over two approaches. Direct training and transfer policy. An agent that previously was trained on circular geometry is further trained on an elliptical one with AR=0.5 and then compared with an agent that has been directly trained on this AR. This analysis aims to determine whether policy transfer is beneficial for similar geometries. The sensor placement will also be considered as an effective parameter. The variation of reward functions in different cases are presented in (Figure 4 – Reward Function). Based on the obtained results, two factors have a significant impact: training from scratch and matching the sensor placement between the first training and the transfer policy training. When training starts from scratch on this geometry, the trend in the reward function shows a similar pattern, whether using an elliptical sensor placement or a circular one. In both cases, the maximum reward function value is approached after around 300 episodes. The variations in the reward function applied through policy transfer show a different behavior from two points of view. The first is a shorter convergence time for the reward function, reaching the maximum value in less than 200 episodes as opposed to previous cases. The second notable point concerns the initial reward function values, which are significantly higher in transfer learning. The variations in the reward function applied through policy transfer show a different behavior from two points of view. The first is a shorter convergence time for the reward function, reaching the maximum value in less than 200 episodes as opposed to previous cases. Finally, policy transfer with an unchanged sensor arrangement provides better results. It converges faster than other cases and possesses the least fluctuations in reward function.

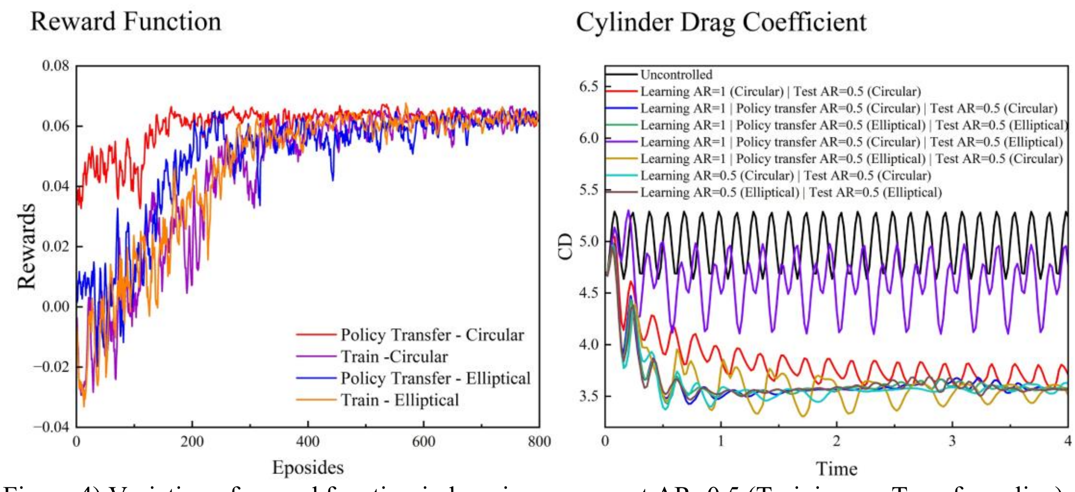

Figure 4) Variation of reward function in learning process at AR=0.5 (Training vs. Transfer policy) and comparing the drag coefficient in different cases.

To analyse the effectiveness of different cases comparisons are conducted on the variation of drag coefficient in testing the trained agent, transfer policy and training from scratch on the target geometry (Figure 4 – Cylinder Drag Coefficient). It can be observed that an agent trained in base geometry (AR=1) is able to control the flow and reduce the drag in target geometry (AR=0.5), by maintaining the sensor placement. However, its performance is lower than the maximum achievable control performance, both in terms of drag reduction and its fluctuations. An alternative is using transfer learning, where all agents are initially trained at AR=1 with circular sensor placement. Of these, the agent is using a transfer policy with circular placement but is tested with an elliptical one seems to perform the worst, given that, during learning, it has never encountered this placement. The reverse case, in which the transfer policy is elliptical and the test is circular allows for a better performance with lower drag approaching the minimum achievable. The negative point of this case would be variations in drag that the agent cannot suppress. The best results are found when the sensor placement during the policy transfer phase is the same as in the tests. This configuration results in a drag coefficient similar to the cases where the agent is directly trained at AR=0.5 with the same sensor placement during testing.

Conclusion

This study explored the use of DRL-based AFC to control flow around elliptical cylinders with different ARs. The primary objective was to evaluate the capability of a DRL agent, trained on a circular cylinder, to control the flow around elliptical cylinders. The results showed that the trained model was able to reduce drag, especially for higher ARs. However, the performance decreased in lower ARs, especially with the changing of sensor placement to follow the new geometry. An alternative solution is using transfer learning where the sensor placement also plays an important role. This approach showed notable performance in the case which has identical sensor placement during the initial learning and transfer learning phases. It significantly reduced training time while maintaining control performance, accelerating convergence compared to direct training. It can be concluded from this analysis that, firstly, the agent trained in a circular geometry can control fluid dynamics in the elliptical geometries, if the sensor placement remains the same. However, for more and optimal performance, rather than direct training in the target elliptical geometry, transfer learning is advised. This process not only enables peak performance with minimum drag but also minimizes the time consumed for the training process.

Acknowledgments

The authors would like to acknowledge the support of the Department of Engineering at City, St George's University of London for funding the conference registration fees. The authors also extend gratitude to colleagues within the department for their encouragement and insightful discussions throughout the development of this work.

References

[1] J. Rabault, M. Kuchta, A. Jensen, U. Réglade, and N. Cerardi, "Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control," J. Fluid Mech. **865**, 281–302 (2019).

[2] Y.-F. Mei, C. Zheng, N. Aubry, M.-G. Li, W.-T. Wu, and X. Liu, "Active control for enhancing vortex induced vibration of a circular cylinder based on deep reinforcement learning," Phys. Fluids **33**(10), (2021).

[3] C. Zheng, T. Ji, F. Xie, X. Zhang, H. Zheng, and Y. Zheng, "From active learning to deep reinforcement learning: Intelligent active flow control in suppressing vortex-induced vibration," Phys. Fluids **33**(6), (2021).

[4] F. Ren, C. Wang, and H. Tang, "Bluff body uses deep-reinforcement-learning trained active flow control to achieve hydrodynamic stealth," Phys. Fluids **33**(9), (2021).

[5] F. Ren, C. Wang, and H. Tang, "Active control of vortex-induced vibration of a circular cylinder using machine learning," Phys. Fluids **31**(9), (2019).

[6] E. Hachem, H. Ghraieb, J. Viquerat, A. Larcher, and P. Meliga, "Deep reinforcement learning for the control of conjugate heat transfer," J. Comput. Phys. **436**, 110317 (2021).

[7] F. Ren, F. Zhang, Y. Zhu, Z. Wang, and F. Zhao, "Enhancing heat transfer from a circular cylinder undergoing vortex induced vibration based on reinforcement learning," Appl. Therm. Eng. **236**, 121919 (2024).

[8] J. Rabault, and A. Kuhnle, "Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach," Phys. Fluids **31**(9), (2019).

[9] S. Shimomura, S. Sekimoto, A. Oyama, K. Fujii, and H. Nishida, "Experimental study on application of distributed deep reinforcement learning to closed-loop flow separation control over an airfoil," in *AIAA Scitech 2020 Forum*, (2020), p. 579.

[10] Y.-Z. Wang, Y. Hua, N. Aubry, Z.-H. Chen, W.-T. Wu, and J. Cui, "Accelerating and improving deep reinforcement learning-based active flow control: Transfer training of policy network," Phys. Fluids **34**(7), (2022).

[11] X.-J. He, Y.-Z. Wang, Y. Hua, Z.-H. Chen, Y.-B. Li, and W.-T. Wu, "Policy transfer of reinforcement learning-based flow control: From two- to three-dimensional environment," Phys. Fluids **35**(5), (2023).

[12] L. Yan, G. Hu, W. Chen, and B.R. Noack, "Deep reinforcement transfer learning for active flow control of a 3D square cylinder under state dimension mismatch," ArXiv Prepr. ArXiv2401.12543, (2024).

[13] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E. Rognes, and G.N. Wells, "The FEniCS project version 1.5," Arch. Numer. Softw. **3**(100), (2015).

[14] C. Geuzaine, and J. Remacle, "Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities," Int. J. Numer. Methods Eng. **79**(11), 1309–1331 (2009).

[15] M. Schäfer, S. Turek, F. Durst, E. Krause, and R. Rannacher, "Benchmark Computations of Laminar Flow Around a Cylinder BT - Flow Simulation with High-Performance Computers II: DFG Priority Research Programme Results 1993–1995," edited by E.H. Hirschel, (Vieweg+Teubner Verlag, Wiesbaden, 1996), pp. 547–566.