arXiv:2305.03647v1 [physics.flu-dyn] 5 May 2023

通过深度强化学习降低钝体的阻力

E. Ballini[#]、A. S. Chiappa[†]、S. Micheletti[#]

2023年5月8日

MOX – Modellistica e Calcolo Scientifico，米兰理工大学数学系，达·芬奇广场32号，I-20133 米兰，意大利
{enrico.ballini,stefano.micheletti}@polimi.it

† 洛桑联邦理工学院，脑与心智研究所，州道，CH-1015，洛桑，瑞士
alberto.chiappa@epfl.ch

摘要

我们提出一种深度强化学习方法，用于解决流体动力学中的经典问题，即降低钝体的阻力。我们将该问题建模为具有连续动作空间的离散时间控制问题：在每个时间步，自主智能体都可以设定位于物体后部的两股流体射流的流量。智能体通过近端策略优化（PPO）进行训练，学会一种有效策略，使射流与尾流中的涡旋相互作用，从而降低阻力。为应对流体动力学仿真的计算复杂度（否则训练过程的成本将高得难以承受），我们在较粗的计算域离散网格上训练智能体。我们提供了数值证据，表明在这一近似环境中训练出的策略迁移到更密的网格后，仍能保持良好性能。仿真结果显示，阻力显著降低，总功率也相应节省。总功率定义为控制系统消耗的功率与阻力功率之和；与不使用任何射流的参考钝体仿真相比，总功率节省达到 40%。最后，我们对神经网络学到的控制策略进行了定性考察。可以看到，该策略通过学习涡旋的形成频率并据此启动射流来实现减阻，从而将涡旋从物体后表面吹离。

1 引言

降低阻力是流体动力学中至关重要的问题，因为这一问题广泛存在于航空、航海和陆路运输应用中。阻力的作用方向与物体运动方向相反，通常是功耗的主要来源。无论是飞机表面等空气动力学物体，还是几何形状紧凑且具有显著回流区域的钝体，阻力问题都已得到深入研究。在这项工作中，我们提出一种利用深度强化学习减阻的方法。我们研究高雷诺数下的二维钝体，并通过非定常计算流体力学（CFD）模拟流场（图 1）。物体的几何形状为矩形，前缘为圆角；此外，后部还增设了两段小曲边，对初始几何形状的影响可以忽略不计（图 2）。钝体的阻力主要由其后方广阔的尾流造成，因为尾流会显著影响物体周围的压力分布。因此，我们的目标是通过与尾流相互作用来降低阻力。具体做法是在后表面设置两个小喷嘴，以期望的流量喷出流体，从而与尾流相互作用。从喷嘴喷出的两股流体射流会吹向位于后部两个角上的小曲边。我们将流体注入速率的控制建模为一个连续控制问题，并通过强化学习求解。自主智能体必须学会一种策略，以最优方式控制两个喷嘴。值得强调的是，该学习过程完全由仿真驱动，无需预先了解系统的控制方程。智能体通过经验学习一个函数，将当前对环境的观测映射为能够最大程度降低阻力的动作。观测由对钝体底面 12 个固定点处压力的实时测量构成。尽管智能体获得的信息有限，训练出的控制策略仍成功降低了钝体的阻力。与不采用主动射流控制的仿真相比，净阻力功率节省达到 40%。

对神经网络学到的控制策略进行分析后，我们发现了射流如何与尾流涡旋相互作用，从而大幅降低阻力。我们发现，智能体会在涡旋生成位置的另一侧点燃射流，将涡旋的形成推向下游。这一机制提高了基底压力，这是减阻的主要原因。

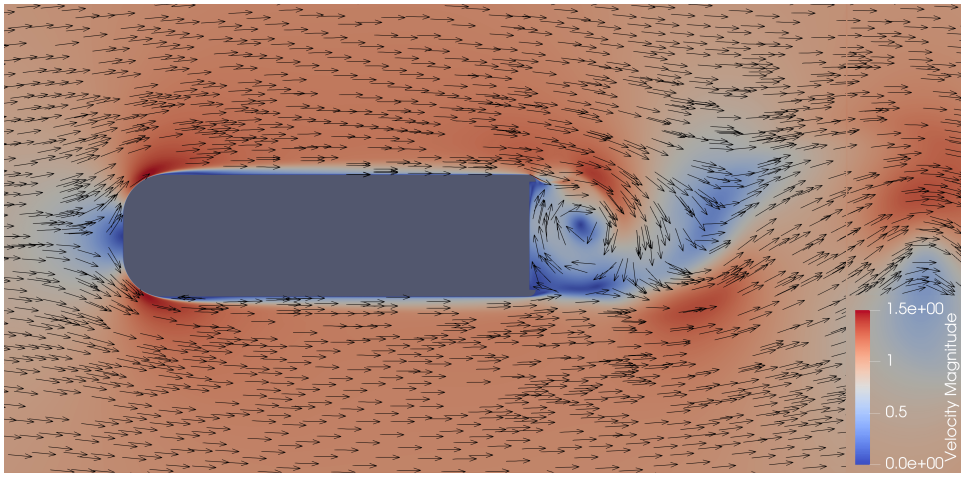

图1：物体周围流动的瞬时快照。箭头表示速度场，颜色图表示相应的幅值。

2 相关工作

人们已经提出了大量降低钝体阻力的方法。我们可以在 [9, 14] 中找到相关综述。一些具体应用的例子包括：通过表面改性控制分离点 [21, 29, 52]，多孔表面 [19, 26]，添加与主体相互作用的其他物体或附属构件 [17, 11, 51, 24, 41, 23, 39]，等离子体执行器 [8, 25]，以及对背部振荡襟翼进行闭环控制 [5]。许多研究关注的是使用位置经过合理设置的流体射流。流体可以以恒定速度喷出 [60], ，也可以以具有非零平均流量的可变速度喷出 [20, 58, 54, 36], ，或以平均流量为零的可变速度喷出 [43], 。关于受控射流的文献则相对较少。一些重要研究涉及通过可变射流控制尾流 [22, 30, 12]。钝体减阻的一项重要实际应用体现在重型车辆领域，因为较低的阻力意味着较低的燃料消耗。事实上，针对这些车辆的前部和后部，均已提出相应的解决方案 [48, 10]。机器学习（ML）也已应用于流体动力学的不同分支。有关综述，请参阅 [47, 6, 27]。特别是，强化学习（RL） [53] 也已应用于流体动力学领域 [18, 45]。一些具体应用包括：偏微分方程控制 [7, 16], 、流动控制 [46, 1, 15], 、对浸没在流体中的物体的运动进行控制，例如类鱼游动体 [35],、微型游动体 [34], 、泽尔梅洛问题 [2], 、无人机和滑翔机 [4, 44], ， 以及形状优化 [57], 等更广泛的应用，还有数值.....

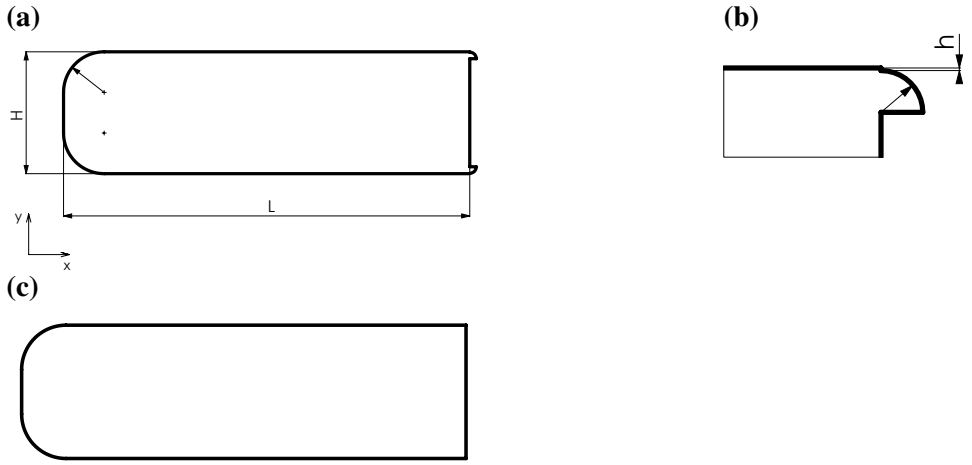

图2：用于流体力学仿真的钝体示意图。**(a)** 带受控射流的钝体的几何形状。**(b)** 曲边和狭缝细节。**(c)** 用作减阻评估参考的洁净钝体的几何形状。

用于近似求解微分方程 [59] 的方法。据作者所知， [61] 中介绍的强化学习应用与我们的方法最为接近。事实上， [61], 的作者通过向强化学习智能体提供流体流动的精确测量，成功减小了浸没在低雷诺数流动中的圆柱体的振荡。

3 流体动力学模型

数值实验对一个高度为$H = 1$、长度为$L = 3.3$的二维钝体周围的流动进行仿真，如图 2 所示。所有量均为无量纲。为避免前缘之后发生大范围分离，钝体前部采用半径为$R = 0.3$ 的圆弧边缘。后部射流由两个高度为$h = 0.03$ 的狭缝喷出。射流作用于两个半径为$r = 0.05$的 90° 圆弧边缘上。以钝体高度为基准的雷诺数为$Re = HU_\infty/\nu = 20000$，其中$U_\infty$为无穷远处速度， ν 为运动黏度。我们将仿真得到的减阻效果与图 2 所示的洁净钝体基准体进行比较。该基准体几乎与所述钝体相同，唯一的区别是其后部没有射流，也没有曲边，因此后部完全平直。流体运动由非定常雷诺平均纳维–斯托克斯方程 [42, 3] 和$k - \omega SST$模型进行封闭 [32, 33]。数值格式配置及常数集

用于 $k-\omega SST$ 的内容见附录A。

3.1 边界条件

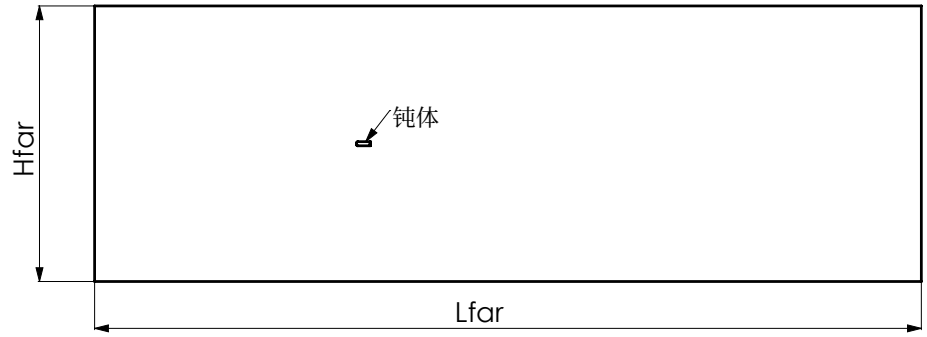

图3：计算域的几何形状。

如图3所示，外部区域设置得足够大，以尽可能减小人工边界的影响。其长度为200个单位，高度为66个单位。钝体位于高度的一半处，以及长度的1/3处。在数值实验中，射流的速度幅值将根据策略网络的输出而变化（见第4节）；但在评估网格与时间收敛性时（见第3.3节），将其固定为一个具有代表性的数值2。在外部边界框上施加远场边界条件。速度U_∞ 设为1，表压P_∞ 设为0。外部边界框上的湍流边界条件由黏性比$r=\frac{\mu_t}{\mu}$ 和湍流强度$I=\sqrt{\frac{3}{2}}k$给定。我们设置$r=10$和$I=0.05$，这两个数值是湍流水平处于平均水平时的典型值。在物体表面施加无滑移边界条件。

3.2 网格

该网格为混合网格：由紧邻物体的结构化区域和其他位置的非结构化区域组成（图 4-5）。结构化区域包含 40 层，增长比为 1.1。这确保能够充分解析边界层以及前部涡黏度的生成。在后部，层数减少至 20 层，因为三角形网格已经足够精细，增加层数并无优势。沿着曲边（图 6），没有设置结构化层，因为网格足够精细，能够捕捉射流与流场之间的相互作用。这意味着通过使用三角形单元，y^+ 可达到足够小的

数值。此外，如果使用矩形单元，穿过单元的高强度对角流将会影响分辨率的收敛速度。

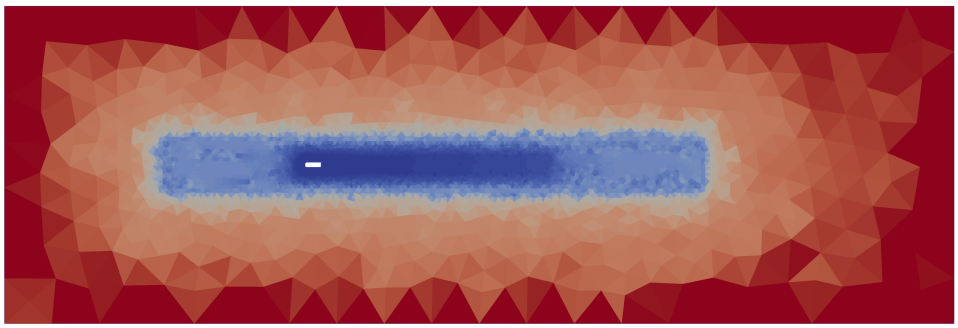

图 4：网格单元尺寸的定性表示。颜色与单元面积成正比。

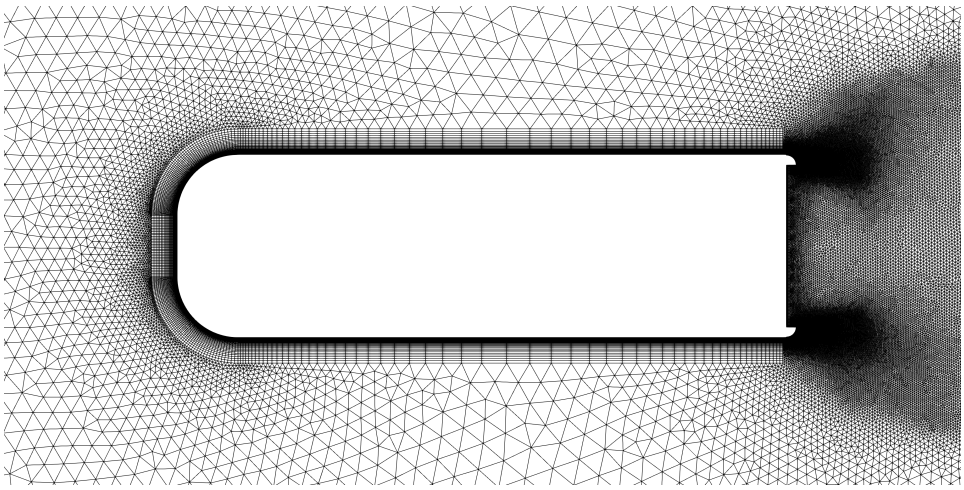

图 5：包含 1.48e5 个单元的网格：钝体细节。

3.3 网格与时间收敛性

网格与时间收敛性（或无关性）分析是评估网格单元的特征尺寸和时间步长大小的一项重要工作。具体过程如下：首先考虑粗网格和较大的时间步长，随后逐步进行细化。该过程持续进行，直至两者均达到收敛。

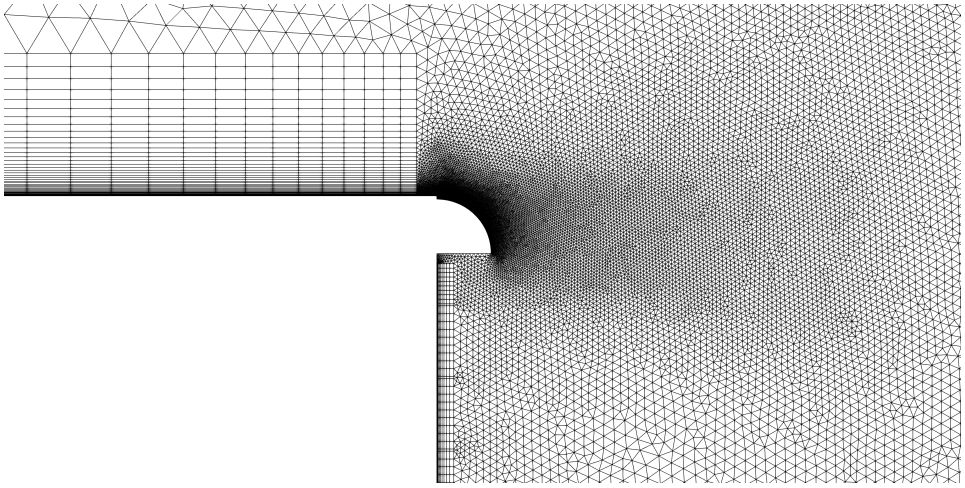

图 6：包含 1.48e5 个单元的网格：曲边细节。

实现网格与时间收敛性。表1展示了采用不同网格和时间步进行仿真所得到的阻力系数。表格应从左上角向右下角阅读，此时网格单元的特征尺寸和时间步逐渐减小，而网格单元总数则相应增加。我们可以观察到，当时间步为 $\Delta t = 2.5\text{e-}4$，且网格包含 $9.44e5$ 个单元时，这两个参数均达到收敛。然而，采用收敛性分析所得参数进行 CFD模拟，其计算复杂度与强化学习训练不相容。因此，我们选择包含 1.48e5 个单元的网格，并采用 4e-3 的时间步，这些设置足以满足训练需求。

表 1：钝体的时间与网格收敛性。

单元数	3.9e4	1.48e5	5.81e5	9.44e5
Δt = 8e-3	0.9145	-	-	-
Δt = 4e-3	0.9219	0.9749	-	-
Δt = 2e-3	0.9242	0.9785	-	-
Δt = 1e-3	-	0.9802	1.035	-
Δt = 5e-4	-	0.9811	1.038	1.045
Δt = 2.5e-4	-	-	-	1.045

4 强化学习

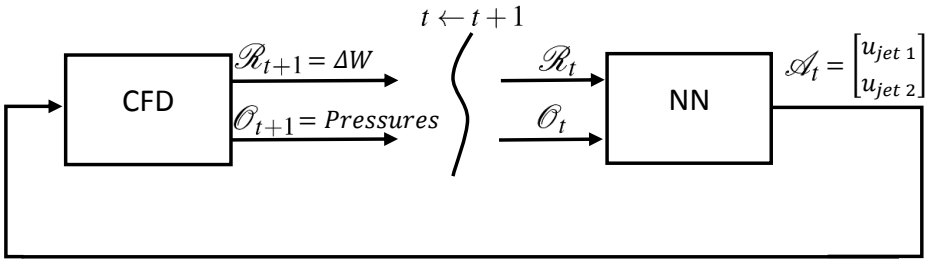

图 7：智能体—环境交互。左侧方框表示环境，右侧方框表示智能体。智能体通过环境的观测值 $\mathscr{O}_t$ 感知环境，这些观测值是在每个时间步对后表面压力的测量。奖励设为功率节省 ΔW。动作 $\mathscr{A}_t$ 是一个向量，其分量为射流的瞬时速度。

我们将喷嘴流量的控制建模为一个部分可观测马尔可夫决策过程 $\mathscr{M} = \langle \mathscr{S}, \mathscr{A}, \mathscr{O}, \mathscr{T}, \mathscr{R}, \gamma \rangle$。图 7 示意了针对我们问题所规定的马尔可夫决策过程的各个要素，现描述如下。在对应于完整流体动力学模拟的完整状态 $\mathscr{S}$ 中，智能体可以访问观测 $\mathscr{O}$，该观测由物体后部的 12 个压力测量值组成（图 8）。每个压力测量值均在两次迭代之间的时间内取平均，以滤除高频波动。压力测量值会在连续 2 次迭代中保留在观测状态中，因此智能体可以访问当前观测和前一次观测（总计 24 个数值）。这使智能体能够获得压力一阶导数的近似值，这是理解涡旋处于形成阶段还是消散阶段所必需的。转移动力学 $\mathscr{T} = P(s_{t+1}|s_t, a_t)$ 由流体动力学模拟经过 $n = 50$ 个模拟时间步后的演化给出。数值 n 对应于大约 $1/5$ 个时间单位。流体动力学现象呈现出振荡行为。因此，我们可以考虑升力系数 CL，并定义其周期，称为C_L-周期。CL-周期约为 4 个时间单位，因此每个C_L-周期大约包含 20 次转移，也就是说每个C_L-周期包含 20 个动作和奖励。数值 n 是在时间分辨率与环境适应时间之间进行折中的结果。更高的数值意味着较弱的时间分辨率，从而导致更粗糙的控制。另一方面，较低的数值也不合适，因为全局现象以及奖励对高频动作相当不敏感，这会使信用分配问题更加困难。动作 $\mathscr{A} = (u^{jet_1}, u^{jet_2}) \in [0,4] \times [0,4]$ 是.....的速度的

喷嘴在下一次状态转移中喷出的流体。参数 γ 是折扣因子。奖励 $\mathscr{R}$ 是净功率节省量，即光滑钝体的阻力功率（见第 3 节）W^{clean} 与带受控射流的钝体的阻力功率 $W^{C_D+\text{射流}}$ 之间的差值：

$$\mathscr{R}_{t+1} = W^{clean} - W_t^{C_D+jets} = W^{clean} - W_t^{C_D} - W_t^{jets}, \tag{1}$$

其中：

$$W_t^{C_D} = drag \cdot U_\infty = \frac{1}{2}\rho U_\infty^2 \left(\frac{1}{\Delta T} \int_t^{t+\Delta T} C_D \right) S U_\infty, \tag{2}$$

$$W^{clean} = drag_{clean} \cdot U_\infty = \frac{1}{2}\rho U_\infty^2 C_D^{clean} S U_\infty, \tag{3}$$

$$W_t^{jets} = \sum_{i=1,2} \frac{1}{\Delta T} \int_t^{t+\Delta T} \int_{slit} \left(\frac{1}{2}\rho |\boldsymbol{u}_{jet_i}|^2 + P_i \right) \boldsymbol{u}_{jet_i} \cdot \hat{n}, \tag{4}$$

其中S 是钝体截面的表面。C_D 由压力和黏性贡献，以及因质量流出物体而产生的小推力组成。为简便起见，计算功率时我们忽略所有可能的效率因子。公式（4）中的压力P_i是相对于无穷远处压力的表压。时间平均值是在一个过渡过程内计算的，因此 ΔT 对应于前文提到的n 个时间步。



图 8：压力测点的位置和编号。

4.1 实验解与算法设置

我们使用近端策略优化（PPO）算法 [50]。主要超参数见附录 C。

演员策略和评论家分别由两个独立的全连接神经网络表示，每个网络包含 2 个隐藏层，每层有 30 个单元。策略网络输出 2 个动作值，分别对应两个射流的流量；评论家网络则输出一个标量，对应于当前状态下的预期累积奖励。网络规模根据 [31, 40] 中的定理选定。参数总数为 1804（演员）+ 1711（评论家），合计 3515 个。激活函数 σ 为修正线性单元（ReLU）。鉴于 CFD 模拟的计算复杂度，以及使用强化学习训练策略所需的大量时间步，对 PPO 超参数和网络架构进行调优超出了本文的研究范围。

为应对 CFD模拟与强化学习相结合所带来的计算复杂度，训练阶段分为两步进行。第一步，在粗网格上运行 CFD模拟。这种近似模拟保留了完整模拟的大部分特性，同时计算成本较低。具体而言，采用包含 2e4 个单元的网格，可将计算时间缩短 92%，而阻力计算误差不到 10%。由于喷嘴的控制策略以随机权重初始化，因此智能体需要与环境进行大量交互，才能找到合适的控制策略。当智能体在这一近似环境中的性能趋于饱和后，第二步便是在采用完整网格的 CFD模拟中对其进行微调。

5 结果

现在我们展示整个模拟过程的主要结果。

5.1 策略学习

图 9 展示了强化学习智能体在第 4.1 节所述两个训练步骤中的学习曲线。可以看到，在第一步训练中，智能体的初始功耗与基线相同，即图 2 所示的光滑钝体；随着经验不断积累，其表现远远超过基线。经过约 170 次训练迭代（相当于 24 小时的 CFD模拟）后，奖励达到平台期。

在训练的第二阶段，采用更细的网格进行训练后，学习曲线图表明，微调对减阻效果的影响相当有限。经过约 60 次训练迭代（相当于约 300 小时的仿真）后，阻力仅进一步降低了 5%，

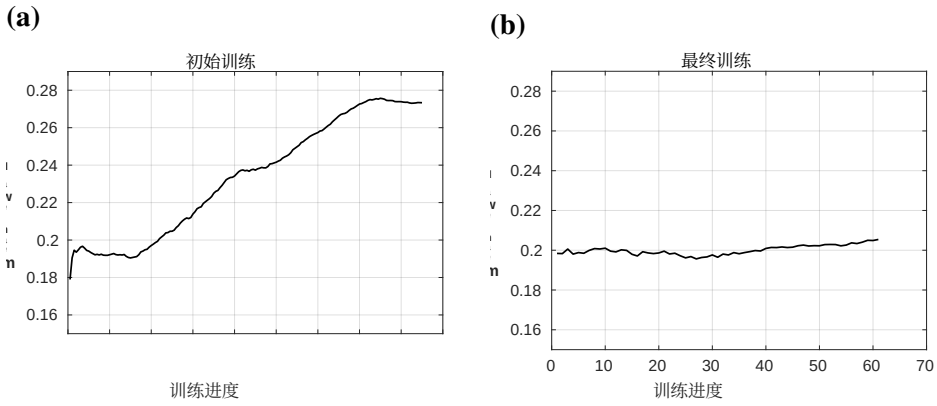

图9：训练进度。**(a)** 使用粗网格进行初始训练，以便快速使神经网络适应环境。图表表示N 个工作进程的平均奖励，并对所有先前回合取平均，随训练进度的变化。**(b)** 使用更细的网格进行最终训练。两个阶段的平均奖励差距源于计算流体力学精度的提升，因此奖励计算也更准确。

因此，大部分功率节省在第一个训练步骤中就已实现。我们强调，两个训练步骤的平均奖励有所不同，这是因为仿真环境不同。事实上，根据我们对网格的收敛性研究，在粗网格中计算出的阻力，以及由此产生的阻力功率，都低于更细的网格中的阻力。我们再次指出，初始训练的唯一目的是使神经网络快速适应环境，从而节省大量计算时间。

最终训练阶段带来的减阻幅度很小，这表明初始训练后得到的神经网络无需任何调整即可成功使用，其性能与经过调整的网络相近。仅经过第一阶段训练后，两个射流便能与涡旋相互作用并实现稳定减阻，尽管其训练所依据的后表面压力是通过粗略模拟得到的，这表明该网络具有良好的迁移学习能力。随后，将训练好的神经网络应用于模拟，以研究针对射流学到的控制策略并评估功率节省。图10报告了瞬时功率W^{C_D}、W_{jet^1}, W_{jet^2}, 和$W^{C_D+}{}_{jets}$。其中，jet_1 和jet_2, 分别表示上方射流和下方射流；而“*jets* ”表示两个射流的总和，如公式（4）所示。表2列出了根据上述模拟计算得出的平均功率，并将其与恒定射流流量和速度下基准控制对应的功率进行了比较。总功率的

带受控射流的钝体的功率比光滑钝体低约 40%，比恒定射流工况低近 20%。

表2：功率节省及其与恒速射流工况的比较。第一行对应于光滑钝体，其中 $u_{jets} = 0$ 。受控射流的功耗与速度等于2的恒速射流的功耗相当，但C_D 的降幅要大得多，从而实现了40%的净功率节省。

u 射流	C_D	W^{C_D}	W 射流	W^{C_D+} 射流	净节省 %
0	1.18	0.589	0	0.589	0
1	1.15	0.577	0.001	0.578	2
2	1.01	0.506	0.020	0.526	11
3	0.83	0.414	0.076	0.490	17
4	0.57	0.283	0.183	0.466	21
5	0.42	0.209	0.364	0.573	3
受控	0.67	0.323	0.028	0.356	40^{+11}_{-12}

根据我们的理解，支配射流与涡旋相互作用的动力学机制如下。涡旋核心处压力较低；在无射流实验中，它们靠近后表面，如图11和图12所示。射流主要有两种作用：首先，它们将涡旋的初生位置推向下游，促使涡旋快速向下游运动，因此使壁面附近的压力升高。事实上，图11(b)显示了涡旋初生位置与壁面之间的距离，以及它们之间的高压间隙。其次，射流会减弱涡旋的强度，这一点可从图11–15中推断得出。

我们来考察图12。这些图展示了一个周期内的若干帧，通过观察涡量可以更好地理解这一现象。在下方涡旋完全形成之前，上方射流启动，从而降低势流并推开正在初生的涡旋。上方射流启动后，上方涡旋随之形成，下方射流也启动。周期接近结束时，沿x 坐标方向有两个未明显分离的涡旋；但片刻之后，正涡移开并消失，负涡则取而代之，重新形成初始状态。此外，这一机制减弱了尾流的垂向振荡（ [12]中也得到了类似结果），使涡旋强度降低且更加不规则。涡脱落的斯特劳哈尔数也受到影响，从0.24降至0.23。

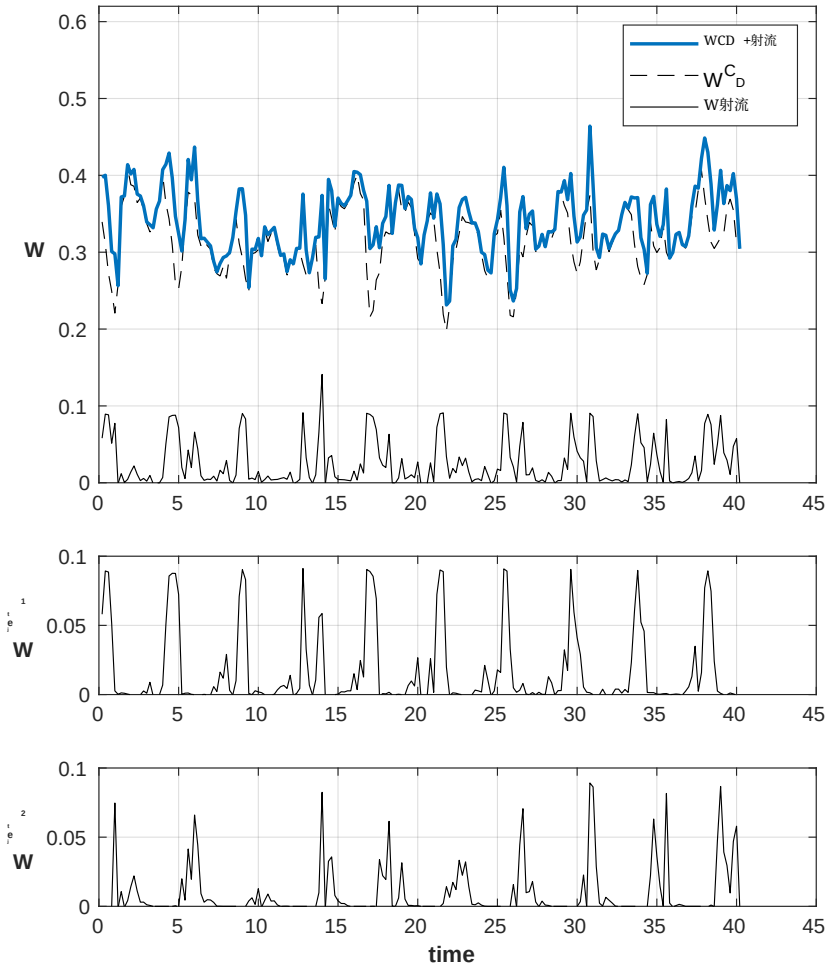

图 10：上图：光滑钝体的功率 W^{C_D+}射流$=W^{C_D}+W$射流、W^{C_D}，以及W射流$=W$射流^1+W射流^2。下图：上方射流和下方射流的功率 W射流^1,W射流^2,。时间跨度约为 $10C_L$ 个周期。

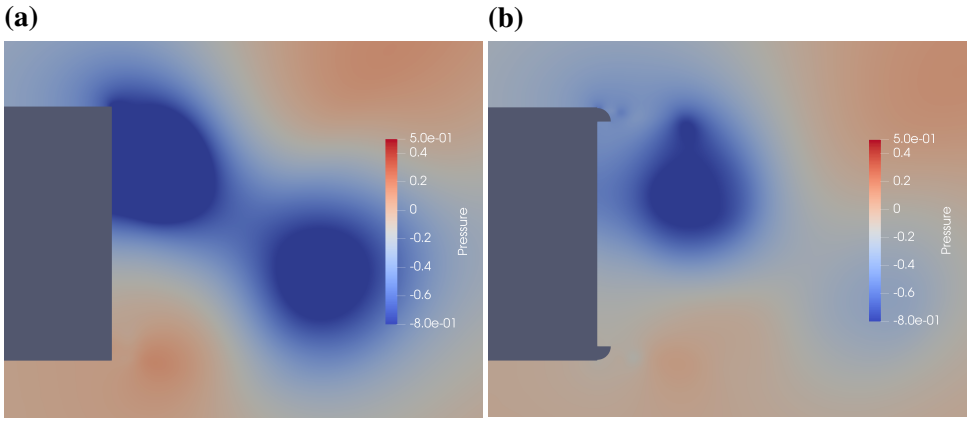

图 11：钝体后方压力分布的快照。（**a**）洁净钝体。涡旋在靠近壁面处形成，产生了显著的低压区域。（**b**）带受控射流的钝体。涡旋在远离壁面处形成。

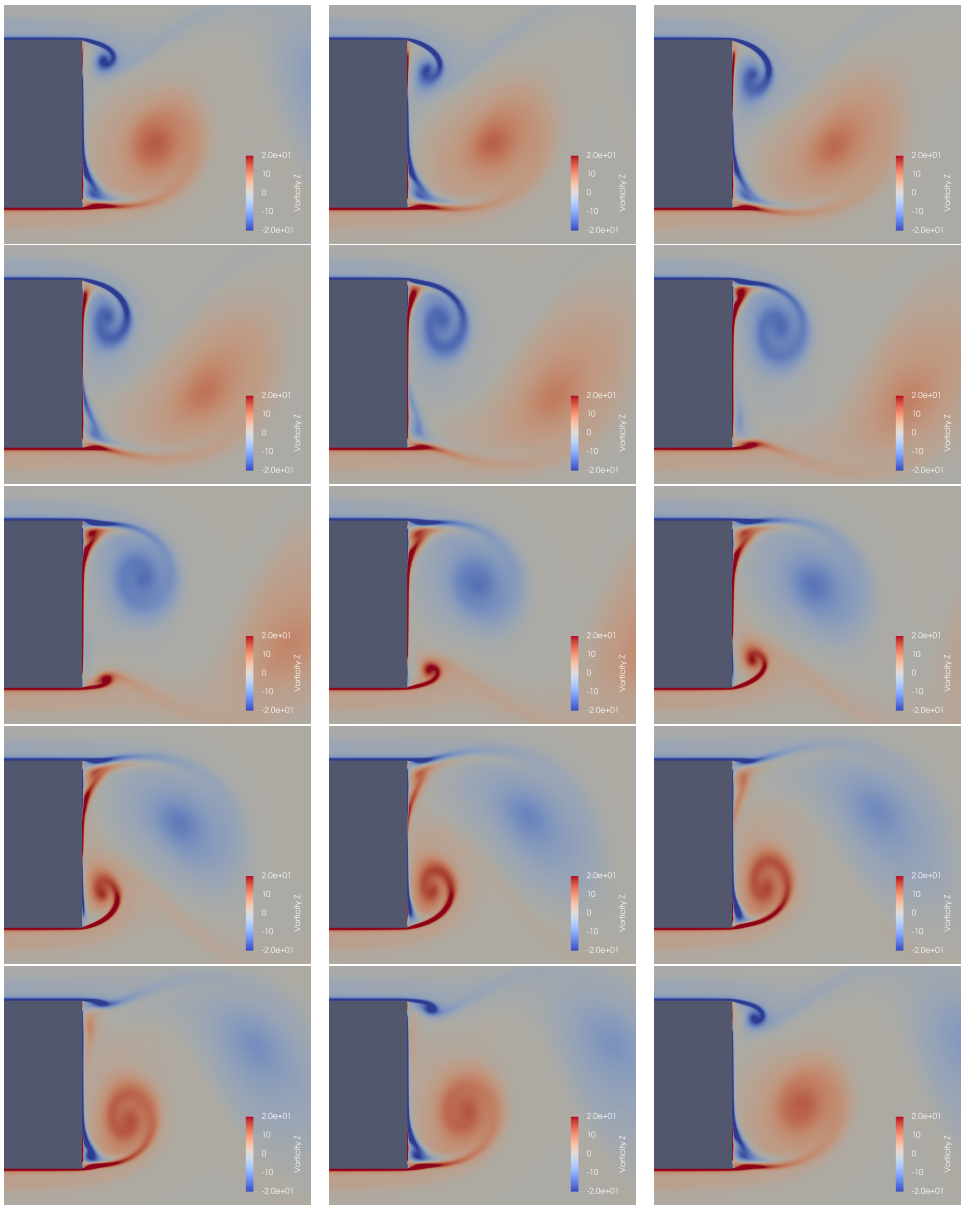

图12：物体后方涡量的时间序列。涡旋形成与脱落规律且呈周期性。

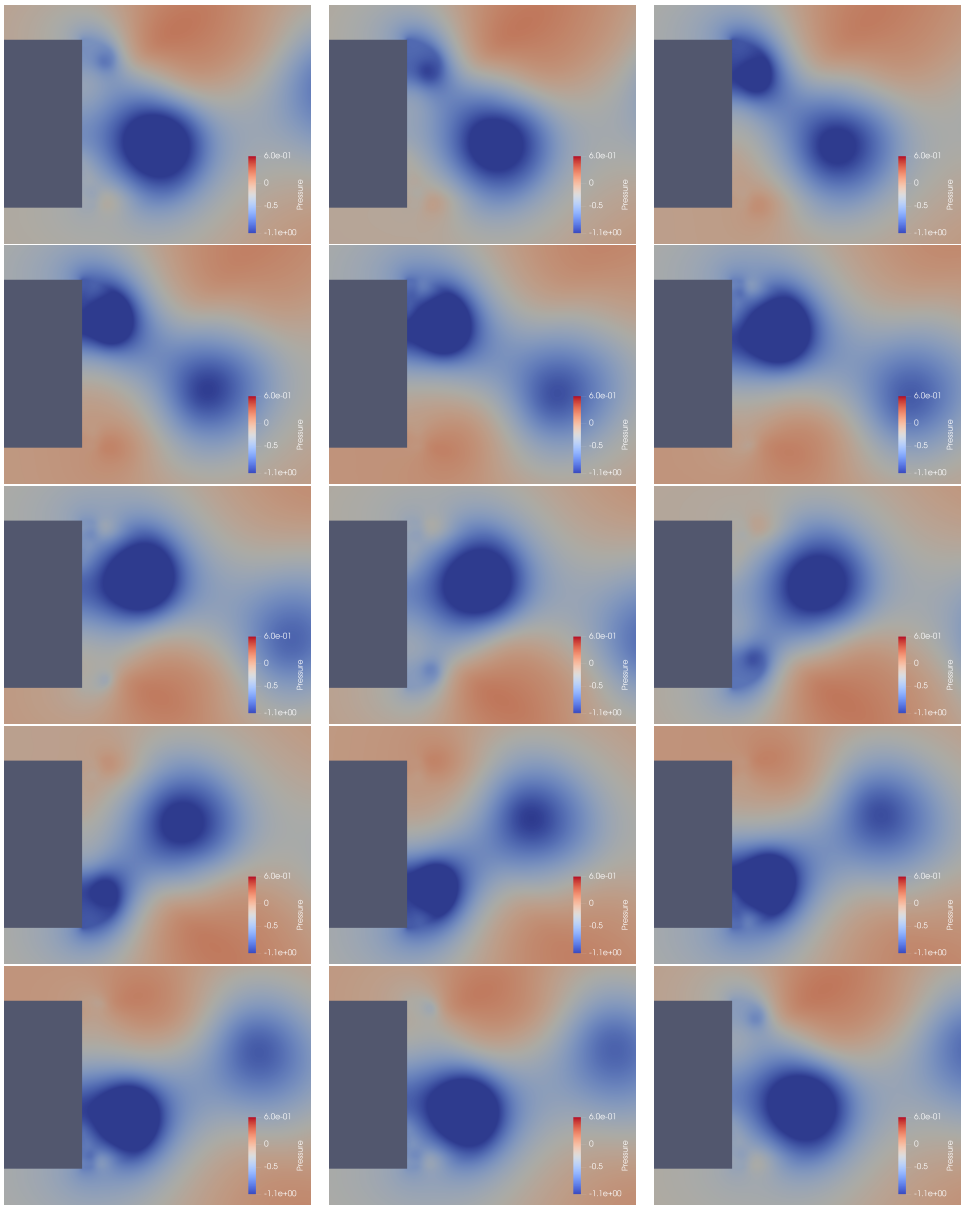

图13：物体后方的压力时间序列。由涡旋核心造成的低压区域接触物体的后表面。

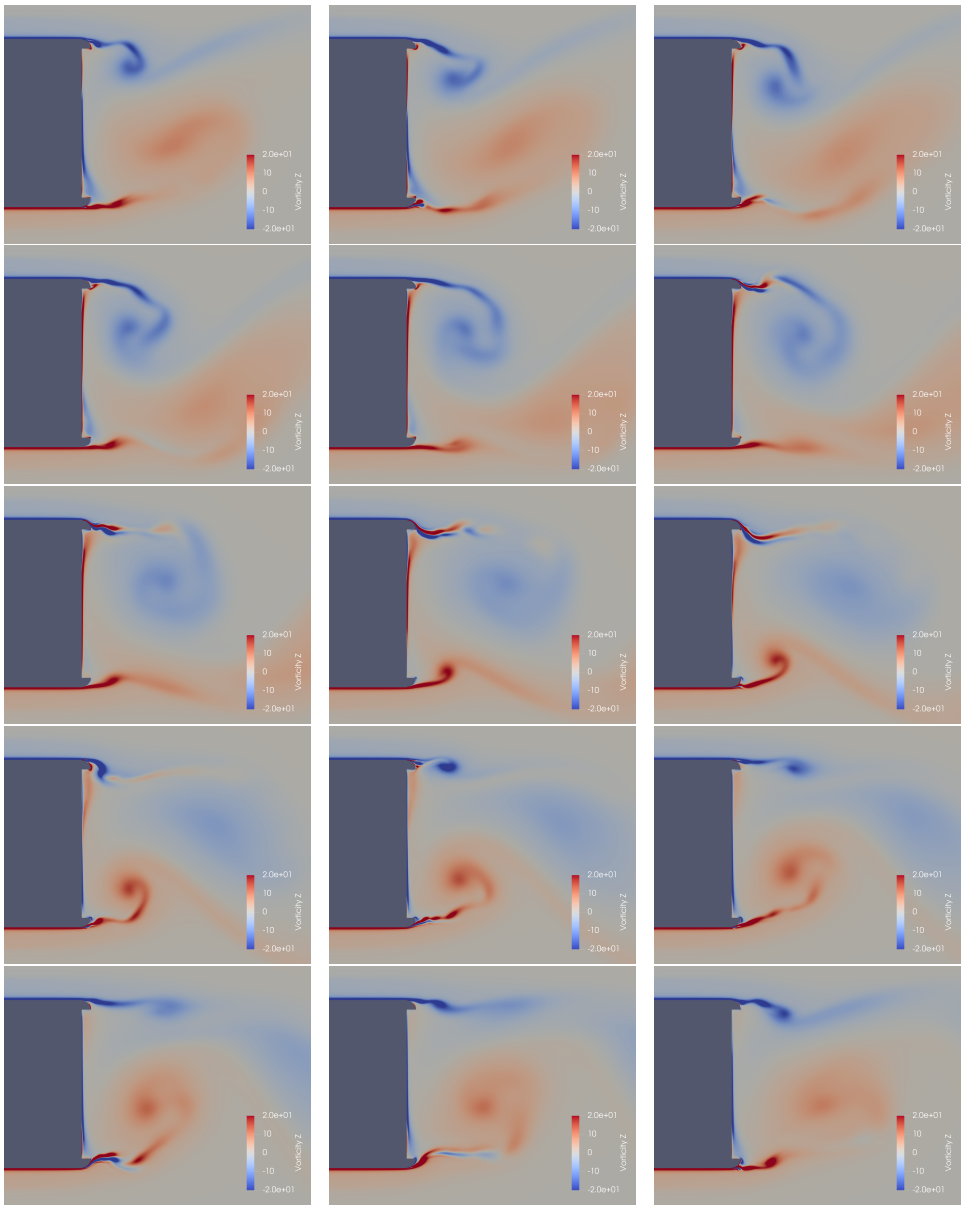

图14：物体后方涡量的时间序列。流体喷射对涡旋有显著影响。

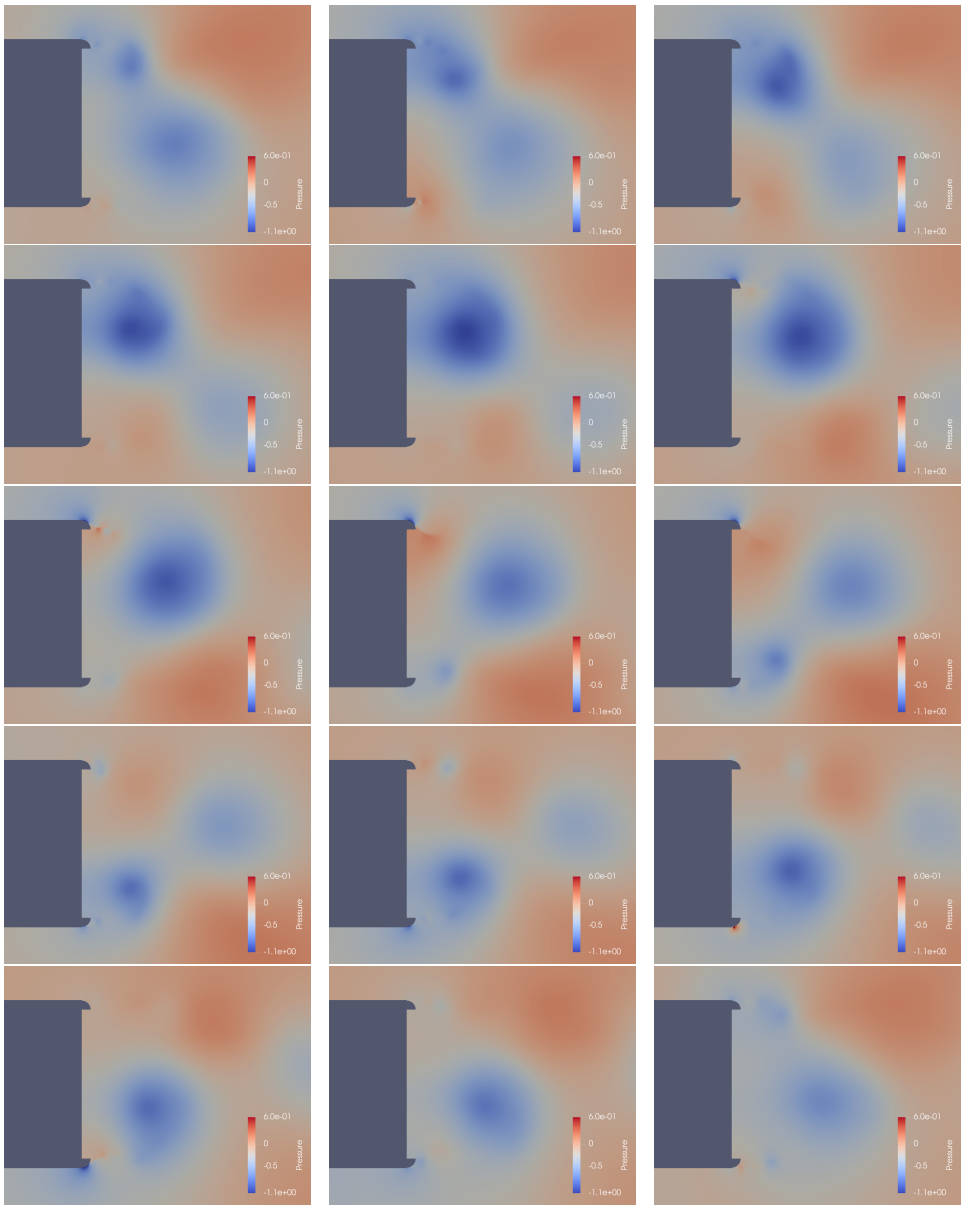

图15：物体后方的压力时间序列。由于流体注入，涡旋核心比光滑钝体的更弱，且距离更远。

6 结论

在本研究中，我们将强化学习应用于钝体减阻这一流体动力学问题。为此，我们专注于控制两个后部射流。

事实证明，强化学习在这一流体动力学背景下同样有效。尽管没有环境模型，且只能有限地观察环境对所采取动作的响应（数据驱动方法），我们仍成功训练出一个神经网络，使其能够与环境，尤其是尾流涡旋，正确交互。这种交互通过将涡旋形成点向下游推移，进而消除由涡旋核心造成的低压区域。事实上，基底压力是减阻的主要影响因素。考虑整个系统，即阻力功率与控制器消耗功率之和，与未受控情况相比，我们实现了可观的节能，达到 40%。

所取得的结果充分证明了继续沿这一方向开展研究的价值。可以探索的一些研究领域包括：
- 进一步提升性能，同时优化几何形状与控制。事实上，在射流与涡旋相互作用的过程中，射流的形状也很重要，而这与曲边的形状有关。优化曲边将提升射流的有效性。这种同步优化应在强化学习的训练阶段进行，且可结合伴随问题所提供的信息来完善这一过程。

- 深入研究获得的最优控制，并将结果与其他控制方法进行比较，以更好地明确强化学习的潜力。
- 研究神经网络的记忆能力，进一步探讨循环神经网络或多级记忆的实用性，即保存当前状态之前的多个历史状态。

- 测试控制的鲁棒性及其对其他边界条件的适应能力。- 提高模拟和/或风洞实验的真实性与准确性，分析更真实的几何形状，并尽可能使其符合实际应用场景。

7 致谢

S. Micheletti 感谢 PRIN 研究资助项目 n. 20204LN5N5，用于多物理场问题的非均质偏微分方程高级多面体离散方法以及 INdAM-GNCS 2022 项目，应用科学的计算约简方法：聚焦复杂系统。

8 利益冲突声明

作者声明不存在任何利益冲突。

A 计算流体动力学设置

SST湍流模型中使用的常数列于表3，可在 [28]中找到。

表3：SST湍流模型所用的常数。

系数	数值
σ_{k1}	0.85
σ_{k2}	1
$\sigma_{\omega 1}$	0.5
$\sigma_{\omega 2}$	0.856
β_1	0.075
β_2	0.0828
γ_1	$\frac{\beta_1}{C_\mu} - \frac{\sigma_{\omega 1}\kappa^2}{\sqrt{C_\mu}}$
γ_2	$\frac{\beta_2}{C_\mu} - \frac{\sigma_{\omega 2}\kappa^2}{\sqrt{C_\mu}}$
C_μ	0.09
κ	0.41
a_1	0.31

流体动力学方程采用有限体积法（FVM）求解，其主要特点如下：

- 对流通量 [13] 采用通量差分分裂法计算，并使用 MUSCL [55] 进行二阶重构。
- 湍流对流通量 [13]不进行二阶重构，以避免强烈的不稳定性。

- 采用 Venkatakrishnan斜率限制器 [56] 。
- 黏性通量 [3, 37, 38, 13]中的速度场导数通过格林-高斯定理求得。

- 时间积分采用一阶隐式欧拉格式。
- 空间离散后产生的线性系统通过采用不完全LU预条件器的双共轭梯度稳定法 [13] issolved。精度设置为使有意义的量（例如阻力系数）的误差在以下范围内准确至其中 the

第四位数字。降低此数值会导致计算时间增加，而结果不会发生明显变化。

B 硬件

以下是所使用硬件的列表：

• Gigat（MOX - 米兰理工大学数学系）：Xeon E5-2640 v4 @2.4GHz，每个节点配备 64GB 内存、20 个核心，最多使用 2 个节点

• Gigatlong（MOX - 米兰理工大学数学系）：Xeon E5-4610 v2 @2.3GHz，每个节点配备 256GB 内存、32 个核心；在本项目期间最多使用 1 个节点。

C 强化学习设置

主要的强化学习设置如下：- rollout 工作器的数量，即并行运行的环境数量，为 4。每个 rollout 工作器选取一个包含 160 个样本的轨迹片段，以填充大小为 $4 \times 160 = 640$ 的训练批次。这些样本并非按顺序排列，而是被联合打乱，以避免局部过拟合。

- 每个训练批次要执行的训练轮数为 30。
- 目标函数的具体表达式可在 [50] 中找到。我们使用广义优势估计器（GAE）[49] 作为优势。
- 学习率为 0.0001，折扣因子 γ 为 0.99。
- KL散度的初始系数为 0。2，而 KL 散度的目标数值为 0。01。

训练数据集的规模和轨迹片段长度是研究对象，因为它们会显著影响训练时间。较小的训练数据集填充所需时间较短，而较大的数据集则需要计算更多样本。从这一角度看，数据集越小越好；但实践证实，由于问题本身的物理特性，较小的数据集可能无法很好地发挥作用：如果环境没有呈现动作的影响，就越来越难以区分好动作和坏动作。这会导致目标函数梯度估计不准确，进而造成权重更新不佳。在计算时间和结果质量之间取得适当平衡的方案，是将 rollout worker 的轨迹片段长度设为 160。轨迹片段的长度约相当于 $8C_L$- 个周期。较短的片段长度，例如 10 个样本（$0.5C_L$- 个周期），

会使训练无法进行；而片段长度设为 80 个样本（$4C_L$- 个周期）则尚可，但无法获得较高奖励。严格来说，我们采用的是回合制设置，每个回合的最大长度为 160 个样本。即使一个回合结束，环境也不会重新初始化，因此新回合会从上一回合最后几个 CFD 时间步开始。回合仅用于将连续轨迹划分成若干段。数值不稳定时则不适用这一规则：一旦发生数值不稳定，CFD模拟就会中断，并从指定的固定初始条件重新启动。这种情况相当少见，因此不会出现对初始条件过拟合的现象。

参考文献

[1] G. BEINTEMA、A. CORBETTA、L. BIFERALE 和 F. TOSCHI，通过强化学习控制瑞利–贝纳德对流，《湍流》, 21 (2020), pp. 585–605.[2] L. BIFERALE、F. BONACCORSO、M. BUZZICOTTI、P. CLARK DI LEONI 和 K. GUSTAVSSON，泽尔梅洛问题：利用强化学习在二维湍流中实现点到点最优导航，Chaos: An Interdisciplinary Journal of Nonlinear Science, 29 (2019), p. 103138.[3] J. BLAZEK，计算流体动力学：原理与应用，Butterworth-Heinemann，牛津，第三版，2015.[4] E. BOHN、E. M. COATES、S. MOE 和 T. A. JOHANSEN，利用近端策略优化对固定翼无人机进行深度强化学习姿态控制，2019年国际无人机系统会议（ICUAS），(2019), pp. 523–533.[5] R. D. BRACKSTON、J. M. GARCÍA DE LA CRUZ、A. WYNN、G. RIGAS 和 J. F. MORRISON，湍流钝体尾流中双稳态的随机建模与反馈控制，《流体力学杂志》，802 (2016年), pp. 726–749.[6] S. L. BRUNTON、B. R. NOACK 和 P. KOUMOUTSAKOS，流体力学中的机器学习，《流体力学年度评论》，52 (2020年), pp. 477–508.[7]M. A. BUCCI、O. SEMERARO、A. ALLAUZEN、G. WISNIEWSKI、L. CORDIER 和 L. MATHELIN，利用深度强化学习控制混沌系统，《英国皇家学会学报A辑：数学、物理与工程科学》，475 (2019), p. 20190351。

[8] Z. CHEN 和 C.-Y. WEN，使用不同介质阻挡放电等离子体激励器控制*D*形钝体流动，《流体与结构期刊》，103 (2021年), p. 103292.[9] H. CHOI、W. P. JEON 和 J. KIM，钝体绕流控制，《流体力学年度评论》，40 (2008), pp. 113–139.[10] H. CHOI、J. LEE 和 H. PARK，重型车辆空气动力学，《流体力学年度评论》，46 (2014), pp. 441–468.[11] M. CICOLIN、O. BUXTON、G. ASSI 和 P. BEARMAN，分离对作用于带控制杆圆柱体上的力的影响，《流体力学杂志》，915 (2021), p. A33.[12] L. DALLA LONGA、A. S. MORGANS 和 J. A. DAHAN，利用线性反馈控制降低*D*形钝体的压差阻力，理论与计算流体动力学，31 (2017年), pp. 567–577.[13] T. D. ECONOMON、F. PALACIOS、S. R. COPELAND、T. W. LUKACZYK 和 J. J. ALONSO，*SU2*：用于多物理场仿真与设计的开源套件，《AIAA期刊》，54 (2016年), pp. 828–846.[14] L. C. EUN、A. S. M. RAFIE、S. WIRIADIDJAJA 和 O. F. MARZUKI，道路车辆钝体被动与主动减阻方法概述，International Journal of Engineering and Technology(UAE), 7 (2018年), pp. 53–56.[15] D. FAN、L. YANG、Z. WANG、M. S. TRIANTAFYLLOU 和 G. E. KARNI-ADAKIS，强化学习在实验和仿真中的钝体主动流动控制，《美国国家科学院院刊》，117 (2020年), pp. 26091–26098.[16] A. M. FARAHMAND、S. NABI 和 D. N. NIKOVSKI，用于偏微分方程控制的深度强化学习，载于2017年美国控制会议（ACC），2017年，pp. 3120–3127.[17]C. GARCÍA-BAENA、J. I. JIMÉNEZ-GONZÁLEZ 和 C. MARTÍNEZ-BAZÁN，通过尾部柔性板重新构形降低钝头体阻力，《流体物理学》，33 (2021年), p. 045102.[18] P. GARNIER、J. VIQUERAT、J. RABAULT、A. LARCHER、A. KUHNLE 和 E. HACHEM，深度强化学习在流体力学中的综述，arXiv:1908.04127, (2021年)。

[19] A. GATTO 和 H. BABINSKY，将被动多孔性作为钝体减阻手段的研究，SAE《国际商用车辆期刊》，11（2018年），第65–73页。[20] Y. HAFFNER、J. BORÉE、A. SPOHN 和 T. CASTELAIN，湍流尾流中的非定常科恩达效应与减阻，《流体力学杂志》，899（2020年），第A36页。[21] M. HASEGAWA 和 H. SAKAUE，用于圆柱体减阻的超细纤维涂层，《流体与结构期刊》，103（2021年），第103287页。[22] L. HENNING 和 R. KING，通过闭环控制减小钝尾缘钝体上分离流的阻力，载于《第44届IEEE决策与控制会议会刊》，2005年，第494–499页。[23] T. HIRST、C. LI、Y. YANG、E. BRANDS 和 G. ZHA，利用喷流船尾被动流动控制实现钝体减阻，SAE《国际商用车辆期刊》，8（2015年），第713–721页。[24]K. KARTHIK、M. VISHNU、S. VENGADESAN 和 S. K. BHATTACHARYYA，利用计算流体力学分析优化钝体，以降低气动阻力和噪声，《风工程与工业空气动力学期刊》，174（2018年），第133–140页。[25] M. KAZEMI、P. GHANOONI、M. MANI 和 M. SAEEDI，利用*SDBD*等离子体致动器实现三维钝体减阻，《机械工程师学会会刊D辑：汽车工程期刊》，235（2021年），第1461–1480页。[26] K. KLAUSMANN 和 B. RUCK，通过在背风侧施加多孔涂层降低圆柱体阻力，《流体力学杂志》，813（2017年），第382–411页。[27] J. N. KUTZ，流体动力学中的深度学习，《流体力学杂志》，814（2017年），第1–4页。[28] 兰利研究中心，湍流建模资源中心。https:// turbmodels.larc.nasa.gov/sst.html。[29] M. F. A. LATIF、M. N. OTHMAN、Q. F. ZAHMANI、N. S. KHASHI’IE、B. E. ZHEN、M. F. ISMAIL 和 A. Y. ISMAIL，通过在钝体顶部添加凹坑网格，优化边界层分离的抑制效果，《流体力学与热科学高级研究期刊》，（2019年），第173–182页。

[30] R. LI、B. R. NOACK、L. CORDIER、J. BORÉE 和 F. HARAMBAT，利用线性遗传编程控制降低汽车模型的阻力，《流体实验》，58（2017年）。[31] Z. LU、H. PU、F. WANG、Z. HU 和 L. WANG，神经网络的表达能力：从宽度的角度看，arXiv:1709.02540，（2017年）。[32] F. R. MENTER，面向工程应用的两方程涡黏性湍流模型，《AIAA期刊》，32（1994年），第1598–1605页。[33] F. R. MENTER 和 C. RUMSEY，跨声速流动两方程湍流模型的评估，1994年。AIAA 1994-2343。流体动力学会议。[34] S. M. NOS LANDIN、K. GHAZI-ZAHEDI 和 F. CICHOS，人工微型游泳器的强化学习，arXiv:1803.06425，（2018年）。[35] G. NOVATI、S. VERMA、D. ALEXEEV、D. ROSSINELLI、W. M. VAN REES 和 P. KOUMOUTSAKOS，通过学习实现两个自主推进游泳器的同步，《生物启发与仿生学》，12（2017年），第036001页。[36] A. R. OXLADE、J. F. MORRISON、A. QUBAIN 和 G. RIGAS，湍流轴对称尾流的高频激励，《流体力学杂志》，770（2015年），第305–318页。[37] F. PALACIOS、J. ALONSO、K. DURAISAMY、M. COLONNO、J. HICKEN、A. ARANAKE、A. CAMPOS、S. COPELAND、T. ECONOMON、A. LONKAR、T. LUKACZYK 和 T. TAYLOR，斯坦福大学非结构化（SU^2）：用于多物理场仿真与设计的开源集成计算环境，2013年。AIAA 2013-287。第51届AIAA航空航天科学会议暨新视野论坛与航空航天博览会。[38]F. PALACIOS、T. D. ECONOMON、A. ARANAKE、S. R. COPELAND、A. K. LONKAR、T. W. LUKACZYK、D. E. MANOSALVAS、K. R. NAIK、S. PADRON、B. TRACEY、A. VARIYAR 和 J. J. ALONSO，斯坦福大学非结构化（*SU2*）：湍流的分析与设计技术，2014年。AIAA 2014-0243。第52届航空航天科学会议。[39] H. PARK、D. LEE、W. P. JEON、S. HAHN、J. KIM、J. KIM、J. CHOI 和 H. CHOI，利用新型被动装置降低钝尾缘二维钝体绕流的阻力，《流体力学杂志》，563（2006年），第389–414页。

[40] S. PARK, C. YUN, J. LEE, AND J. SHIN, 通用近似的最小宽度, arXiv:2006.08859, (2020).[41] A. PINELLI, M. OMIDYEGANEH, C. BRÜCKER, A. REVELL, A. SARKAR, AND E. ALINOVI, *PELSkin* 项目：第四部分——利用毛状细丝控制钝体尾流, Meccanica, 52 (2017), pp. 1503–1514.[42] S. B. POPE, 《湍流》, 剑桥大学出版社, 剑桥, 2000.[43] Z. X. QIAO, G. MINELLI, B. R. NOACK, S. KRAJNOVI`C, AND V. CHERNORAY, 采用遗传算法优化偏航钝体的多频气动控制, 《风工程与工业空气动力学期刊》, 212 (2021), p. 104600.[44] G. REDDY, A. CELANI, T. J. SEJNOWSKI, AND M. VERGASSOLA,学习在湍流环境中翱翔, 《美国国家科学院院刊》, 113 (2016), pp. E4877–E4884.[45] F. REN, H. HU, AND H. TANG, 利用机器学习进行主动流动控制：简要综述, 《水动力学研究与进展》, 32 (2020), pp. 247–253.[46] F. REN, J. RABAULT, AND H. TANG, 将深度强化学习应用于弱湍流条件下的主动流动控制, 《流体物理学》, 33 (2021), p. 037121.[47] I. SADREHAGHIGHI, 用于计算流体动力学的人工智能（*AI*）和深度学习, tech. rep., CFD Open Series, 2021.[48] L. SALATI, 重型卡车空气动力学. http://hdl.handle.net/ 10589/138944, 2018. 博士论文.[49] J. SCHULMAN, P. MORITZ, S. LEVINE, M. JORDAN, AND P. ABBEEL,利用广义优势估计进行高维连续控制, arXiv:1506.02438, (2018).[50]J. SCHULMAN, F. WOLSKI, P. DHARIWAL, A. RADFORD, AND O. KLIMOV, 近端策略优化算法, arXiv:1707.06347, (2017).[51] N. A. SIDDIQUI AND M. A. CHAAB, 一种用于降低*Ahmed*车体阻力的简单被动装置, 《应用流体力学杂志》, 14 (2020), pp. 147–164.

[52] K. SON, J. CHOI, W. P. JEON, AND H. CHOI, 球体表面绊线减阻机制, 《流体力学杂志》, 672 (2011), pp. 411–427.[53] R. SUTTON AND A. BARTO, 强化学习, 麻省理工学院出版社, 伦敦, 第二版, 2018.[54] N. TOUNSI, R. MESTIRI, L. KEIRSBULCK, H. OUALLI, S. HANCHI, AND F. ALOUI, 利用压电致动器对钝体进行流动控制的实验研究, 《应用流体力学杂志》, 9 (2016), pp. 827–838.[55] B. VAN LEER, 迈向终极守恒差分格式：*V. Godunov*方法的二阶后续格式, 《计算物理学杂志》, 32 (1979), pp. 101–136.[56] V. VENKATAKRISHNAN, 限制器的精度与稳态解的收敛性, 1993年. AIAA 1993-880. 第31届航空航天科学会议.[57] J. VIQUERAT, J. RABAULT, A. KUHNLE, H. GHRAIEB, A. LARCHER, AND E. HACHEM, 通过深度强化学习直接进行形状优化, arXiv:1908.09885, (2020).[58] B. WANG, Z. YANG, AND H. ZHU, 利用新型非定常射流对*25° Ahmed*车身进行主动流动控制, 《国际传热与流动杂志》, 79 (2019), p. 108459.[59] Y. WANG, Z. SHEN, Z. LONG, AND B. DONG, 学习离散化：通过深度强化学习求解一维标量守恒律, arXiv:1905.11079, (2020).[60] B. F. ZHANG, K. LIU, Y. ZHOU, S. TO, AND J. Y. TU, 基于稳态吹气对高阻力*Ahmed*车身进行主动减阻, 《流体力学杂志》, 856 (2018), pp. 351–396.[61] C. ZHENG, T. JI, F. XIE, X. ZHANG, H. ZHENG, AND Y. ZHENG,从主动学习到深度强化学习：用于抑制涡激振动的智能主动流动控制, 《流体物理学》, 33 (2021), p. 063607.