arXiv:2602.11378v1 [cs.LG] 11 Feb 2026

迈向自适应非侵入式降阶模型：设计与挑战

Amirpasha Hedayat[1*]，Alberto Padovan[2] 和 Karthik Duraisamy[1]1*密歇根大学航空航天工程系，美国密歇根州安娜堡，MI 48105。[2]新泽西理工学院机械与工业工程系，美国新泽西州纽瓦克，NJ 07102。

*通讯作者。电子邮箱：ahedayat@umich.edu；其他作者：alberto.padovan@njit.edu；kdur@umich.edu；

摘要

基于投影的降阶模型（ROMs）通常作为静态代理模型部署，因此一旦系统偏离训练流形，其实际效用便会受到限制。我们对自适应非侵入式降阶模型进行形式化定义与研究，这类模型能够在线更新潜在子空间和降阶动力学。基于静态非侵入式降阶模型的思路，具体而言，即算子推断（OpInf）和近期提出的基于轨迹优化的非侵入式降阶模型（NiTROM），我们提出了三种方案：自适应 OpInf（序贯基底/算子重拟合）、自适应 NiTROM（对编码器/解码器和多项式动力学进行联合黎曼优化），以及一种以 OpInf 更新进行初始化的混合方案。我们描述了在线数据窗口、自适应窗口和计算预算，并分析了成本缩放。在瞬态扰动的顶盖驱动方腔流中，静态 Galerkin/OpInf/NiTROM 在超出训练范围进行预测时会发生漂移或失稳。相比之下，自适应 OpInf 以适中的成本有效抑制振幅漂移；自适应 NiTROM 在频繁更新时能够实现近乎精确的能量跟踪，但对其初始化和优化深度较为敏感；在状态变化和离线数据极少的情况下，混合方案最为可靠，能够生成物理一致的场并将能量保持在有界范围内。我们认为，对降阶模型的预测能力作出声明时，必须考虑成本并保持透明，明确区分训练/自适应/部署阶段，并明确报告在线预算和全阶模型查询次数。本研究为构建自我校正的非侵入式降阶模型提供了实用范本，使其在动力学演变远超初始流形后仍能保持有效。

关键词： 降阶模型（ROMs）、自适应降阶模型、非侵入式模型降阶、算子推断、流形优化、科学机器学习

1 引言

反复求解大型微分方程组的成本，往往限制了数值模拟在多种工程任务中的适用性，包括优化、不确定性量化、控制和数字孪生。对此，可以用降阶模型（ROM）取代（或补充）昂贵的高保真求解器，从而以原始计算成本的一小部分完成求解 [1–3]。简而言之，模型降阶领域致力于开发能够最大限度提升计算效率、同时尽可能减少准确性损失的模型。尽管降阶模型在研究领域日益受到关注，但其在实际工程应用中的普及仍然有限。

在本研究中，我们探讨基于非侵入式投影的降阶建模问题，以加速物理系统的高成本模拟。本研究有两个目标：首先，阐明当前降阶模型方法在实现预测能力方面面临的挑战；其次，通过自适应非侵入式公式，提出未来的发展方向。具体而言，我们探讨基于最小二乘回归的公式的优势与不足，以及一种新颖范式：联合考量试探子空间、测试子空间与潜空间动力学的动力学感知式自适应。这些公式以算子推断 [4],和近期提出的基于轨迹的非侵入式降阶模型优化方法（NiTROM） [5]为基础。因此，本研究的主要贡献是为现有非侵入式降阶模型开发并系统评估一种实用的在线自适应工作流，同时提出一种融合算子推断与 NiTROM 思路的新模型。本节其余部分将简要回顾现有的静态与自适应、侵入式与非侵入式降阶建模策略。

给定一个状态维数为 $n \gg 1$ 的高保真模型，静态和自适应基于投影的降阶模型通常通过约束状态在维数为 $r \ll n$ 的低维子空间上演化来获得。由于自适应降阶模型利用在线信息，其运行所基于的子空间通常比静态降阶模型的维数更低。该子空间及其上的潜空间坐标分别由所谓的解码器和编码器的选择隐式定义；解码器和编码器分别是从 r 维空间映射到环境 n 维空间，以及反向映射的映射。状态在目标子空间上的演化由一个 r维动力系统支配，其积分成本远低于原始高保真求解器。自适应降阶模型寻求与高保真算子建立接口，以更新潜空间。这些构建模块（即编码器、解码器、潜空间动力学和流形更新）可以使用侵入式或非侵入式算法计算，具体如下文所述。

1.1 侵入式降阶模型

侵入式降阶模型是指：编码器、解码器和潜空间动力学的计算需要直接访问底层高保真算子。在大多数基于投影的降阶模型中，编码器和解码器是定义秩为 r 的线性映射，将系统投影到解码器张成的空间上，而潜空间动力学则通过投影全阶动力学获得。经典示例是通过 正交 投影到与代表性训练数据集 [6–8]相关的本征正交分解（POD）模态 [9, 10]上获得的 Galerkin 模型。这一过程具有“弱”侵入性，因为编码器和解码器的计算完全由数据驱动，而潜空间动力学的计算只需访问高保真动力学。尽管这些模型得到了广泛应用，但众所周知，将系统正交投影到高能量／高方差子空间上，可能会意外截断低能量／低方差特征，而这些特征在动力学上仍然十分重要 [3, 11–13]。这最终可能导致模型的预测准确性较差，尤其是在瞬态轨迹上。

通过适当定义的斜投影构建的 Petrov-Galerkin 模型可以弥补这一不足。采用此类斜投影的代表性降阶模型框架包括以下几种。最小二乘 Petrov–Galerkin（LSPG）方法 [14, 15] 通过以最小二乘意义最小化离散控制方程的残差来构建降阶模型，从而得到稳定的时间离散降阶模型。另一种重要方法是平衡截断（BT） [16, 17]，它通过平衡线性化系统的可控性 Gramian 和可观性 Gramian 来构建降阶模型，从而保留动力学上最重要的模态。此外，平衡 POD（BPOD） [18] 提供了 BT 的一种近似方法，它利用原始模拟和伴随模拟的快照来计算经验 Gramian。

较新的方法包括基于轨迹的斜投影优化（TrOOP） [19]，，该方法将降阶模型构建表述为一个优化问题，旨在寻找最佳斜投影，以最小化训练轨迹上的预测误差；以及利用伴随快照的协方差平衡降阶（CoBRAS） [20, 21]，，该方法平衡由原始快照集合和伴随快照集合导出的协方差算子，从而获得稳健且均衡的降阶模型。这些方法具有很强的侵入性，因为模型的计算与组装需要访问非线性动力学，以及在全阶模型的（可能随时间变化的）解附近对其进行线性化得到的伴随算子。这限制了这些方法在通过黑箱求解器（例如商业及遗留代码）模拟的高维系统中的适用性，因为出于专有原因，源代码访问受限。

1.2 非侵入式降阶模型

非侵入式公式旨在通过开发完全依赖仿真（或实验）数据的算法来解决上述问题，而无需访问高保真求解器的源代码。因此，这些算法广泛适用于使用商业及遗留代码进行仿真的系统。早期示例包括数据驱动的基于模态分解的方法，例如动态模态分解（DMD） [22, 23]，，以及较新的扩展方法，这些方法采用机器学习架构和算子学习 [24–26]。此外，

在非侵入式设定下也可以实现平衡截断 [13, 27]。应用最广泛的非侵入式方法之一是算子推断（OpInf），它将降阶模型辨识重新表述为一个最小二乘回归问题，直接在降阶空间中学习多项式算子 [4, 28]。

OpInf 已应用于复杂非线性动力系统 [29–33] 和参数化问题 [34]，并且还扩展到了二次流形 [35]。然而，该框架寻求位于由高能量/高方差 POD 模态定义的正交投影算子张成空间中的模型，这最终可能导致预测精度较差的模型。近期提出的非侵入式基于轨迹的降阶模型优化（NiTROM）框架 [5]提供了一种非侵入式学习机制，旨在解决这一缺陷。NiTROM 利用现有的高保真轨迹数据，联合优化斜投影算子和降阶多项式动力学。研究表明，该方法能够显著提高模型预测精度，尤其适用于对低能量特征高度敏感且具有大幅瞬态增长的系统（例如对流主导流动和高剪切流）。

1.3 自适应模型降阶

许多研究者在侵入式框架下探索了自适应降阶模型。在这类范式中，自适应通常涉及更新降阶基和超降阶采样位置，而降阶算子则因依赖全阶模型算子而自动更新。Amsallem 和 Farhat 的早期工作 [36–38] 开创了基于插值的降阶模型，通过构建局部降阶基并在矩阵流形上插值算子，实现了跨参数自适应。这些方法针对参数区域的不同区域构造一组降阶子空间，而非依赖单一的全局子空间。此后的方法，例如 Peherstorfer 等人提出的“局部离散经验插值法”（LDEIM） [39], ，将这一思想拓展至超降阶阶段：通过对快照进行聚类并构建多个局部 DEIM空间和采样点，实现了不同动力学状态下高效的非线性模型求值。除此之外，Peherstorfer 及其合作者推动该领域迈向真正的在线自适应框架。在“动态数据驱动降阶模型” [40], 中，他们提出直接利用传感器数据对降阶算子进行低秩更新。随后，自适应DEIM（ADEIM） [41] 及其后续方法 AADEIM [42],进一步发展了这一方向，根据局部残差信息对降阶基和插值点进行实时低秩更新。

其他策略，例如 Singh 等人的前瞻采样技术 [43]，会预测近期的系统状态，以指导数据采集和更新调度。从几何优化的角度来看，Zimmermann 等人 [44] 提出采用“Grassmann 流形秩一更新子空间估计”（GROUSE）方法 [45, 46]，沿 Grassmann 流形上的测地线对降阶基进行连续的秩一更新，从而得到闭式子空间自适应公式。类似地，Huang 和 Duraisamy [47] 提出了一种秩一自适应降阶模型形式，通过增量式修正子空间，确保每个新的系统快照都具有零投影误差。他们的结果表明，即使在混沌反应流系统中，这种高效的一步更新也能保持长期稳定性和物理保真度。最近，

Mohaghegh 和 Huang [48] 开发了特征引导采样方法，该方法引入基于物理的判据，引导在线采样集中于动态活跃区域（在他们的案例中是激波和火焰前锋），从而进一步提升超降阶中的自适应性。除此之外，许多其他研究也探索了自适应降阶模型，包括混合快照与全阶模型/降阶模型耦合方法 [49–51]，用于优化的自适应模型降阶方法 [52]，基底细化 [53, 54]，以及时变基底框架 [55–57]。尽管取得了这些进展，所有这类方法的一个关键局限在于它们都依赖于对全阶模型的侵入式访问。具体而言，它们需要显式计算残差、雅可比矩阵或投影算子，以执行采样和基底更新。在非侵入式场景中，只有状态和输入数据可用，无法获取这些信息，因此任何参数更新都必须完全以数据驱动的方式进行。另一个复杂性来源于：在自适应非侵入式降阶模型中，不仅要更新降阶基，还必须重新识别降阶算子本身；这些算子是直接从数据中学习得到的，而不是通过投影获得的。

近期的一些研究探索了在非侵入式降阶模型中实现有限形式自适应性的方法。例如，Geelen 和 Willcox [58] 提出了一种局部算子推断方法：离线训练多个局部降阶模型，并由在线分类器在预测过程中自适应地选择最合适的局部降阶模型。在该模型中，主要的自适应机制是在预训练的局部降阶模型之间切换，而非在线更新算子。近期，Aretz 和 Willcox [59]提出了嵌套算子推断方法，该方法利用降阶空间中的层次结构，实现高效的离线更新，并对学习得到的算子进行热启动细化。尽管这些工作为构建更灵活的非侵入式降阶模型展现了有前景的方向，但都未能对模型参数进行持续的在线自适应。因此，据我们所知，本文首次提出了一系列构建自适应非侵入式降阶建模通用框架的思路。

我们的框架将在线数据同化和持续学习的原理与降阶建模相结合，使模型能够即时自我改进。基于 OpInf 和 NiTROM 的基础，我们提出了三种具体的自适应公式：（i）自适应 OpInf，对基和算子进行序贯更新；（ii）自适应 NiTROM，执行联合流形优化；（iii）自适应（混合）算子推断–NiTROM，将快速的基于回归的更新与短时的基于优化的细化相结合。需要指出的是，所有提出的模型都假设能够在线定期通过单步全阶模型求解获取高保真状态快照。这一部署假设影响了所提框架的适用范围。通过在一个典型非线性流动问题上开展全面的数值实验，我们评估了这些自适应公式的准确性和实用性。我们的观察表明，当系统保持在训练流形附近时，所有测试过的自适应模型都能防止振幅漂移，并稳定长期预测。此外，当系统发生状态变化或离线训练数据极少时，混合自适应能够实现最稳健的恢复，并取得具有物理意义的结果。这些发现表明，非侵入式在线自适应有望拓展降阶模型的预测范围，标志着迈向

真正 具备预测能力的降阶模型。尽管如此，我们强调，当前框架仍处于早期发展阶段，计算效率问题也尚未解决。作者认为，降阶模型社区（包括我们自己的研究）需要采用更严格且考虑成本的预测声明标准，明确区分训练阶段与部署阶段，并完整报告离线/在线计算预算。在某些情况下，重现已见过的动力学可能有用，但如果没有在给定资源和时间约束下验证流形外性能，就不应将其作为泛化能力的证据。我们希望在这方面提供一些量化分析。

本文其余部分安排如下。第2 节详细介绍问题的表述，以及自适应非侵入式降阶模型框架中的设计选择。第3 节概述用于测试自适应降阶模型的全阶模型、数值设置和评估指标。第4节报告并分析逐步更具挑战性的场景下的结果。最后，第5 节总结主要发现，讨论局限性，并展望未来方向。

2 方法论

考虑一个非线性高维动力系统

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad ; \quad \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)), \tag{1}$$

其状态为$\mathbf{x}(t) \in \mathbb{R}^n$，控制输入为$\mathbf{u}(t) \in \mathbb{R}^m$，测量输出为$\mathbf{y}(t) \in \mathbb{R}^q$。该系统代表全阶模型（FOM），当状态维数 n 非常大时，通常模拟成本过高。降阶模型（ROM）是另一种动力系统，通常由图1所示的流程描述。降阶模型的关键组成部分包括：**(i)** 编码器，它是一个可能为非线性的映射，将 n维空间映射到 r维空间（其中 $r \ll n$）；**(ii)** 解码器，它是一个可能为非线性的映射，将 r维空间映射到 n维空间；以及 **(iii)** 用于演化降阶状态$\mathbf{z}$的潜空间动力学$\mathbf{f}_r$。本文将研究范围限定在线性映射的情形，即编码器和解码器均为线性映射，且编码器是解码器的左逆。于是，编码器可表示为 $\psi = \Psi^\top$，解码器可表示为 $\varphi = \Phi(\Psi^\top\Phi)^{-1}$，其中 $\Phi, \Psi \in \mathbb{R}^{n\times r}$ 分别定义了解码器基（张成试验空间）和编码器基（张成测试空间）。由于编码器是解码器的左逆，编码器/解码器对定义了一个投影算子

$$\mathbb{P} \triangleq \Phi(\Psi^\top\Phi)^{-1}\Psi^\top : \mathbb{R}^n \to \mathbb{R}^n, \quad \mathbb{P}^2 = \mathbb{P}. \tag{2}$$

图1中的降阶模型在 $\mathbb{P}$的值域内近似全阶模型解。当 span(Φ) $\equiv$span(Ψ)时，我们称投影为正交投影；否则称为斜投影。值得注意的是，编码器和解码器也可以定义为非线性函数 [60–63], ，不过要严格确保编码器 ψ 是解码器 φ的左逆，需要谨慎处理。无论如何，这超出了本文的讨论范围。

潜空间动力学$\mathbf{f}_r$ 如图 1所示，可以通过多种方式定义。其中最广为人知的或许是通过（Petrov-）Galerkin 投影来定义：给定一个投影

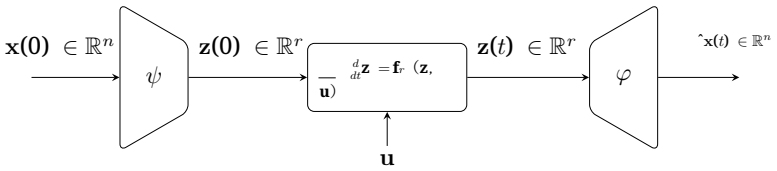


图1：包含编码器 ψ、解码器 φ 和潜空间动力学$\mathbf{f}_r(\mathbf{z},\ \mathbf{u})$。

算子 $\mathbb{P}$ 定义如（2）所示，潜空间动力学通过将全阶动力学 $\mathbf{f}$ 按如下方式投影得到

$$\mathbf{f}_r(\mathbf{z},\mathbf{u}) = \Psi^\top \mathbf{f}(\Phi(\Psi^\top\Phi)^{-1}\mathbf{z},\mathbf{u}). \tag{3}$$

这一过程通常被认为具有（弱）侵入性，因为它需要直接访问全阶动力学 $\mathbf{f}$。另一种方法是使用无需访问 $\mathbf{f}$ 的非侵入式方法，从数据中推断潜空间动力学。例如，OpInf 就是这种情况（将在下一节中详细讨论）：假设动力学具有多项式形式，

$$\mathbf{f}_r(\mathbf{z},\mathbf{u}) = A_r\mathbf{z} + H_r : (\mathbf{z}\mathbf{z}^\top) + B_r\mathbf{u} + \ldots, \tag{4}$$

降阶张量 $A_r \in \mathbb{R}^{r\times r}$、 $H_r \in \mathbb{R}^{r\times r\times r}$、 $B_r \in \mathbb{R}^{r\times m}$ 等通过求解最小二乘优化问题从数据中学习得到。

显然，无论是在静态设置还是自适应设置下，编码器、解码器和潜空间动力学都决定了降阶模型的准确性及其预测系统动力学的能力。本文考虑具有多项式潜空间动力学（如方程（4）所示）和线性编码器/解码器（定义可能为斜交的投影算子，如方程（2）所示）的自适应非侵入式模型。具体而言，我们将考虑三种自适应范式：一种基于 OpInf，一种基于 NiTROM，另一种则结合二者。在介绍自适应方法之前，先回顾它们的静态对应方法很有必要。

2.1 静态降阶模型

2.1.1 算子推断

如前所述，算子推断（OpInf） [4, 28]是一种非侵入式模型降阶框架，通常旨在构建一个在由（正交归一）本征正交分解（POD）模态定义的正交 投影的值域上演化的多项式降阶模型， Φ。按照图 1中的术语，编码器定义为 $\psi = \Phi^\top$，解码器定义为 $\varphi = \Phi$，潜空间动力学$\mathbf{f}_r$ 如公式（4）所示。从数学上说，给定从高保真训练数据集 $\{\mathbf{x}^{(i)}(t_j)\}$ 计算得到的前 r 个正交归一 POD 模态 $\Phi \in \mathbb{R}^{n\times r}$ （其中$\mathbf{x}^{(i)}(t_j) \in \mathbb{R}^n$ 是轨迹 i上时间 t_j的解向量），我们计算潜空间坐标$\mathbf{z} = \Phi^\top\mathbf{x}$ 及其时间导数 $d\mathbf{z}/dt = \Phi^\top d\mathbf{x}/dt$，并求解以下最小二乘优化

问题

$$\min_{\Theta\in\mathcal{M}_{\mathrm{OpInf}}} J=\frac{1}{ML}\sum_{i=0}^{L-1}\sum_{j=0}^{M-1}\left\|\frac{d}{dt}\mathbf{z}^{(i)}(t_j)-\mathbf{f}_r(\mathbf{z}^{(i)}(t_j),\mathbf{u}^{(i)}(t_j))\right\|^2,\quad \Theta=(A_r,H_r,B_r,\ldots), \tag{5}$$

其中 $\mathbf{f}_r$ 的定义同式（4），而 $\mathcal{M}_{\mathrm{OpInf}}=\mathbb{R}^{r\times r}\times\mathbb{R}^{r\times r\times r}\times\mathbb{R}^{r\times m}\times\ldots$。（此处，$L$和 M 分别表示训练轨迹的数量和每条轨迹包含的训练快照数。）尽管 OpInf 是一种优雅的形式，且具有闭式解，但近期研究表明，[5] 它往往难以准确预测经历大幅瞬态的轨迹。这是因为，尽管底层的正交投影在静态重构状态方面是最优的，但将状态投影到 POD 模态上时，可能会无意中截去低能量特征，而这些特征对系统的动力学演化仍然至关重要。

2.1.2 NiTROM

Padovan 等人近期提出的 NiTROM框架 [5] 是一种非侵入式形式，旨在解决通过正交投影到高能量/高方差子空间上所获得的非侵入式模型的不足。具体而言，鉴于已有研究表明，对于表现出大幅瞬态增长的全阶模型，斜投影更适合捕捉其动力学相关机制，NiTROM 同时通过在矩阵流形上求解一个优化问题，寻求最优的编码器 $\psi=\Psi^\top$、解码器$\varphi=\Phi\left(\Psi^\top\Phi\right)^{-1}$以及多项式潜空间动力学（4）。给定训练数据 $\{\mathbf{x}^{(i)}(t_j)\}$，其中，再次地，$\mathbf{x}^{(i)}(t_j)\in\mathbb{R}^n$ 表示轨迹 i上时间 t_j 处的解向量，我们求解以下问题

$$\begin{aligned}
\min_{\Theta\in\mathcal{M}_{\mathrm{NiTROM}}}\quad & J=\frac{1}{ML}\sum_{i=0}^{L-1}\sum_{j=0}^{M-1}\|\mathbf{y}^{(i)}(t_j)-\hat{\mathbf{y}}^{(i)}(t_j)\|^2\\
\text{subject to:}\quad & \frac{d}{dt}\mathbf{z}^{(i)}=\mathbf{f}_r(\mathbf{z}^{(i)},\mathbf{u}^{(i)}),\quad \mathbf{z}^{(i)}(t_0)=\Psi^\intercal\mathbf{x}^{(i)}(t_0)\\
& \hat{\mathbf{y}}^{(i)}=\mathbf{h}\left(\Phi\left(\Psi^\top\Phi\right)^{-1}\mathbf{z}^{(i)}\right)\\
& V=\mathrm{Range}\left(\Phi\right),
\end{aligned} \tag{6}$$

其中 $\mathbf{y}^{(i)}(t_j)=\mathbf{h}\big(\mathbf{x}^{(i)}(t_j)\big)$、参数 $\Theta=(V,\Psi,A_r,H_r,B_r,\ldots)$，以及定义为以下形式的矩阵流形 $\mathcal{M}_{\mathrm{NiTROM}}$a

$$\mathcal{M}_{\mathrm{NiTROM}}=\mathcal{G}_{n,r}\times S_{n,r}\times\mathbb{R}^{r\times r}\times\mathbb{R}^{r\times r\times r}\times\mathbb{R}^{r\times m}\times\ldots. \tag{7}$$

这里，$\mathcal{G}_{n,\ r}$ 是 Grassmann 流形（即 r维子空间组成的空间），而 $\mathbb{R}^n$是 $S_{n,\ r}$ 是 Stiefel 流形（即由 $n\times r$ 正交归一矩阵组成的空间）。这种特定的参数化源于优化问题的性质。具体而言，我们在 Grassmann 流形上对抽象子空间进行优化，因为优化问题（6）是由解码器 V所张成的子空间 $\Phi\left(\Psi^\top\Phi\right)^{-1}$决定的，而不是由矩阵代表元 $\Phi\left(\Psi^\top\Phi\right)^{-1}$ 本身决定的。另一方面，编码器 $\Psi^\top$则在 Stiefel 流形上进行优化，以严格确保必要的

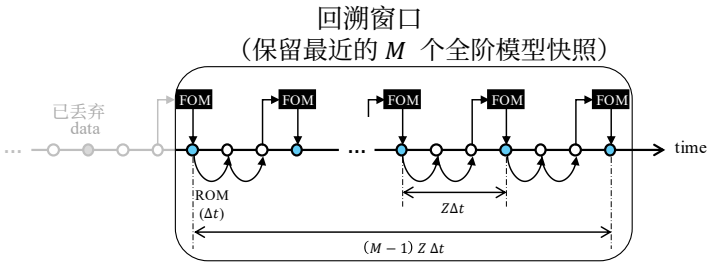


图 2：自适应设置示意图。高亮部分表示移动窗口中保留的全阶模型快照。

条件秩$(\Psi) = r$，从而定义投影 $\mathbb{P}$，如公式（2）所示。定义 $\mathbb{R}^{r\times r}$时所涉及的其他（线性）空间（例如 $\mathcal{M}_{\mathrm{NiTROM}}$ 等）用于参数化降阶张量，以定义公式（4）中的潜空间多项式动力学。

与前文讨论的 OpInf 相比，这种表述有几项优势。首先，允许编码器和解码器跨越不同空间，使模型具有更强的表达能力，因此能够检测对准确预测至关重要的低能量敏感性机制。其次，代价函数衡量的是原始高维空间中的解误差，而不是潜空间向量随时间变化率的偏差，从而有助于提高潜空间动力学的稳定性。第三，所有降阶模型组件（编码器、解码器和动力学）均同时进行优化，从而促进计算动力学感知型投影算子，进而构建能够沿瞬态轨迹预测全阶模型解的模型。然而，这种表述的缺点在于优化问题是非凸的，这意味着存在局部极小值和不唯一的极小化解。此外，与 OpInf 不同，该问题无法得到闭式解。尽管如此，研究表明，可以通过完全非侵入式的伴随方法表述，以闭式形式计算代价函数 J 关于参数 Θ 的梯度 [5] 。这意味着无需依赖可能代价高昂的自动微分，也能高效求解该优化问题。

2.2 自适应降阶模型

自适应模型降阶方法依靠与全阶高保真求解器交互来更新模型参数，并维持所需的准确性水平。本文中采用的自适应设置如图 2所示。这里，Δt 表示全阶模型时间步长，$Z\Delta t$ 表示降阶模型预测窗口的长度。换言之，时间从 $t_j - Z\Delta t$ 推进至 t_j 的过程由降阶模型完成。每经过 $Z\Delta t$个时间单位，降阶模型便与全阶模型交互并收集一个新的高保真数据样本，将其存储在数据矩阵中

$$\mathcal{D}_j = \left\{Q(t_{j-(M-1)Z}),\ \ldots,\ Q(t_{j-Z}),\ Q(t_j)\right\},\quad Q(t_i) = (\mathbf{x}(t_i), \dot{\mathbf{x}}(t_i), \mathbf{u}(t_i))\,. \tag{8}$$

$\mathcal{D}_j$ 是一个移动窗口数据矩阵，因为在任意时刻 t_j，它都包含 M 个高保真样本；一旦有新样本可用，便会丢弃“最旧”的样本。这样，不再能代表系统当前状态的较早样本就会自然淡出，使模型能够侧重于最新且最相关的模式。值得一提的是，每次自适应过程中连续执行多个全阶模型时间步也是可行的，而此处选择单步反映了兼顾成本的设计决策。

每次自适应的目标是将降阶模型参数 Θ_{j-1} 更新为经过优化的参数集 Θ_j，以便更好地表示 $\mathcal{D}_j$中包含的最新观测动力学。这一更新是以下形式的局部学习问题的解：

$$\Theta_j = \arg\min_{\Theta \in \mathcal{M}} J_j(\Theta), \tag{9}$$

其中，J_j 是基于训练数据 $\mathcal{D}_j$构建的某个目标代价函数。例如，当根据基于 OpInf 的方案自适应模型时，代价函数 J_j 由方程（ J 5）中的代价函数给出，其中，且 。类似地，当使用类似 NiTROM 的规则自适应模型时，代价函数由方程（6）给出，同样有 $L = 1$，且 $\mathcal{M} = \mathcal{M}_{\text{NiTROM}}$。

在线积分过程中，会按离散间隔触发自适应。最简单的情况下，每隔 Z 个时间步进行一次更新。这种固定调度方式简化了分析，并能直接控制计算成本。另一种策略是基于残差范数自适应地触发更新，这自然也是未来工作的一个延伸。因此，控制自适应降阶模型整体行为的三个主要超参数为 (M, Z, K)，其中 M 是移动（回溯长度）窗口中保留的高质量快照数量，Z 是连续两次自适应之间的降阶时间步数（自适应窗口），K 是每次自适应执行的优化步数（若采用基于梯度的在线学习，例如 NiTROM 框架）。这些量共同决定了准确性与计算效率之间的权衡。自适应框架的伪代码见算法 1，其中描述了在线预测阶段以及降阶模型与全阶模型交互。

注 *1* 我们强调，对自适应降阶模型预测状态进行解码，得到的状态位于解码器值域内，而该值域通常并不覆盖全阶模型解流形。定期与全阶模型交互的作用，是在累积的模型误差不可逆地增大之前予以纠正，方法是将新观测到的动力学重新注入降阶模型。实践中，这要求在各自适应事件之间，降阶模型预测与真实轨迹保持足够接近；这一条件由自适应窗口控制 Z。第 4节的数值结果表明，只要满足这一条件，从提升状态出发进行一步全阶模型推进，就能得到物理上相容的快照，从而有效实现自适应。

2.2.1 自适应算子推断

自适应 OpInf 将式（9）中的在线更新分解为两个连续阶段：（1）利用移动窗口 $\mathcal{D}_j$ 中包含的新信息更新降阶基；（2）在同一窗口上求解最小二乘问题，重新拟合降阶算子。

算法 1 自适应非侵入式降阶模型框架

要求： 初始训练数据集 $\mathcal{D}_0$，包含在时间 $t \in [t_0$ –采集的 M 个样本 $(M-1)Z\Delta t,\ t_0]$；降阶维数 r；自适应窗口 Z；回溯窗口 M；优化预算 K（如适用）；

$j \leftarrow 0$

$\Theta_j \leftarrow$ 初始化模型$(\mathcal{D}_j,\ r)$

对 $t \in t_0 + \Delta t\{0, 1, 2, \ldots\}$ **执行**

  $\mathbf{z}(t) \leftarrow$ 推进降阶模型一个时间步$(\Theta_j, \mathbf{z}(t-\Delta t))$

  如果 $(t - t_0 + \Delta t) \bmod (Z\Delta t) = 0$ **则**

    解码 $\hat{\mathbf{x}}(t-\Delta t) \leftarrow \Phi_j \left(\Psi_j^\top \Phi_j\right)^{-1} \mathbf{z}(t-\Delta t)$

    $\mathbf{x}(t) \leftarrow$ 推进一个时间步的全阶模型$(\hat{\mathbf{x}}(t-\Delta t))$

    $\mathcal{D}_{j+1} \leftarrow$ 更新数据矩阵$(\mathcal{D}_j, \mathbf{x}(t))$

    $\Theta_{j+1} \leftarrow$ 自适应调整模型$(\Theta_j,\ \mathcal{D}_{j+1},\ Z,\ M,\ K)$

    $j \leftarrow j + 1$

  结束 if

结束 for

阶段 ***1***：基更新。

在第 j次自适应事件中，最近的 M 个高质量快照构成数据窗口 $\mathcal{D}_j$，如式（8）所定义。随着 $\mathcal{D}_j$ 随时间更新，其张成空间也会发生变化，因此用于 OpInf 的 POD 模态 Φ 也应随之更新（详见第 2.1.1 节）。有多种策略可用于更新基。

- **窗口化奇异值分解**：在窗口内计算奇异值分解（SVD），并选取 r 最前面的左奇异向量。这种方法忽略之前的基，每次自适应时都从头重建基。这样可以为当前窗口提供最优秩为r 的子空间（按 ℓ_2 意义而言），但每次自适应都需要进行完整奇异值分解。尽管存在更高效的技术（见下文），我们在整个实验过程中都采用此方法来建立可靠的基线。

- **增量奇异值分解（*iSVD*）**：使用递归低秩更新，利用新的快照$\mathbf{x}(t_j)$更新基。这提供了一种高效的增量技术，无需重新计算奇异值分解（例如，参见 Brand的iSVD算法 [64]）。

- **秩*1*更新**：Huang & Duraisamy（2023）的方法 [47] 对旧基进行高效的秩一修正，从而有效消除新快照的投影误差。这种方法相当激进，因为旧快照不会直接参与构建新基。当每个时间步都对基进行空间局部更新时，该方法效果显著；而这通常需要在区域内选定的位置侵入式访问全阶模型算子。在此处考虑的非侵入式设定中，这种操作并不可行，因此本文不探讨秩一更新。

阶段 *2*：算子重新拟合。
基 Φ_j 更新后，通过求解第 $\Theta_j = (A_{rj}, H_{rj}, B_{rj})$ 节概述的优化问题，重新计算降阶算子。2.1.1 其中 $L = 1$。自适应 OpInf 的顺序两阶段更新提供了一种轻量且有效的机制，用于在线学习模型参数。

备注 *2* 在 OpInf 设置中，最小二乘解的唯一性要求回归数据矩阵列满秩，这通常意味着快照数量相对于算子系数数量必须达到某个下限。在当前的自适应设置中，移动窗口中保留的快照数量可能无法满足列满秩条件。为改善这类情况下的条件性，我们采用正则化最小二乘，从而得到适用于短时域预测的适定解。

2.2.2 自适应 NiTROM
自适应 NiTROM 将离线 NiTROM 优化框架扩展到在线流式场景。与将基和降阶算子视为彼此独立实体的自适应 OpInf 不同，自适应 NiTROM 会执行联合优化，如第 2.1.2 节所述。这样便能在单个优化步骤中同时修正降阶子空间和潜在动力学，确保基和算子与不断演变的数据保持相互一致。形式上，在每个自适应事件中，我们针对 2.1.2 节中的训练数据 $\mathcal{D}_j$ 求解 NiTROM 问题。优化以先前参数 Θ_{j-1} 作为热启动，并在乘积流形 $\mathcal{M}_{\text{NiTROM}}$ 上通过黎曼梯度步骤进行自适应更新。

2.2.3 自适应（混合）OpInf–NiTROM
虽然自适应 NiTROM 为轨迹级自适应提供了最有原则的框架，但其收敛性和成本在很大程度上取决于初始化点 Θ_{j-1}。如果自上次自适应以来系统动力学发生了显著变化，直接从 Θ_{j-1} 开始可能会陷入较差的局部极小值，或导致收敛缓慢。另一方面，自适应 OpInf 可以利用新数据窗口上的最小二乘法快速计算近似更新，但缺少基与算子之间的几何感知耦合。我们称之为自适应 OpInf–NiTROM 的混合方法，其核心思想是将快速的 OpInf 更新用作后续流形优化的起跳点，从而提供高质量的初始化，使优化算法接近代价函数景观中的有利区域。

2.3 计算方面的考量
自适应降阶模型的计算成本是评估其实用性的关键因素。模型降阶旨在加快预测，但自适应性会在在线阶段带来额外开销。下面我们概述所提出的自适应框架中各主要组成部分的计算复杂度标度。

在两次连续自适应之间，降阶模型完全在维数为 r 的降阶空间中进行积分。对于具有 α 次多项式非线性的降阶模型，α 个变量中的不同 r 次单项式数量为 $s = \binom{r+\alpha-1}{\alpha} = \mathcal{O}(r^\alpha)$，因此，式（4）右端项在每个时间步的计算量随以下规模增长

$\mathcal{O}(r^{\alpha+1})$，其中最高次项占主导。由于 $r \ll n$，与高保真求解器和后续自适应步骤相比，这一计算开销可以忽略不计。

在每次自适应事件中，都必须获取一个新的高质量快照。在数字孪生场景中，这对应于传感器数据，因此不会产生计算成本。然而，在基于仿真的环境中，新快照需要通过将降阶模型预测提升至全空间，并执行一步全阶模型推进来获取。提升成本按 $\mathcal{O}(nr)$ 进行缩放，而全阶模型推进通常按 $\mathcal{O}(n^{\beta})$ 进行缩放，其中，对于隐式格式，$\beta > 1$，具体取决于求解器类型和系统矩阵的稀疏结构。在实际的计算流体力学求解器中，这一步全阶模型查询通常占据一个自适应周期的总成本的主要部分。

获取新快照后，使用前文讨论的一种自适应公式更新降阶模型参数。对于自适应 OpInf，每次更新的总成本包括基更新、全空间与降阶空间之间的投影，以及降阶算子的最小二乘重新拟合。基更新可以采用多种策略：窗口化奇异值分解的成本为 $\mathcal{O}(nM^2)$，增量奇异值分解（iSVD）的成本为 $\mathcal{O}(nr)$，或者采用与增量奇异值分解（iSVD）具有相近缩放关系的秩一更新。将数据投影至降阶空间以及从降阶空间投影回来的成本为 $\mathcal{O}(Mnr)$（每个窗口）。对于一个次数为 α 的多项式模型，在 M 个快照上组装算子推断数据矩阵需要 $\mathcal{O}(Mr^{\alpha})$的计算量。随后，最小二乘重新拟合的近似上界为 $\mathcal{O}(Mr^{3\alpha+1})$，这反映了 α 中的指数增长（参见 [4]的第 3.5 节）。综合这些部分，单次自适应 OpInf 更新（采用窗口化奇异值分解）的计算量按以下方式缩放：

$$\mathcal{O}(nM^2 + Mnr + Mr^{3\alpha+1}).$$

相比之下，自适应 NiTROM 模型涉及联合流形优化，每次自适应需要进行 K 次黎曼迭代。每次迭代都需要在包含 M 个时间步的时间窗口内，对降阶模型进行一次正向积分和一次伴随积分。相应的成本按 $\mathcal{O}(KMr^{\alpha+1})$ 变化，其中 α 是降阶模型的多项式次数。此外，还需要在 Stiefel 流形和 Grassmann 流形上计算涉及编码器基和解码器基的矩阵-向量乘积(Φ, Ψ)，每次更新大约增加 $\mathcal{O}(Knr^2)$ 的成本（见 [5] 第 2.4 节）。因此，自适应 NiTROM 每次自适应的总复杂度为

$$\mathcal{O}(Knr^2 + KMr^{\alpha+1}).$$

自适应 OpInf–NiTROM 方法将这两种机制结合起来：先快速执行 OpInf 更新，再进行少量 NiTROM 精化步骤。其总成本可估算为

$$\mathcal{O}(nM^2 + Mnr + Mr^{3\alpha+1}) + \mathcal{O}(Knr^2 + KMr^{\alpha+1}).$$

在实践中，成本主要取决于 NiTROM 迭代次数，而该次数通常较少（$K \sim 5$–10），因为 OpInf 阶段提供了近乎最优的初始化。

表 1 总结了本文所考虑的所有自适应模型的渐近单次自适应成本。请注意，仅保留随完整维度n 缩放的项，因为其他所有项都可忽略不计。第 4 节给出了定量比较（见表 2），其中列出了每个步骤的实际运行时间。总体而言，

表 1：每次自适应的渐近计算成本。其中，n 为完整维度，r 为降阶维数，α 为降阶算子的多项式次数，M 为移动窗口中的快照数量，K 为优化迭代次数。

模型	操作	Cost
自适应 OpInf	基 + 算子更新	$\mathcal{O}(Mnr)$
自适应 NiTROM	联合流形更新	$\mathcal{O}(Knr^2)$
自适应 OpInf–NiTROM	混合	$\mathcal{O}(n(M^2 + Mr + Kr^2))$
全阶模型单步推进	高保真求解	$\mathcal{O}(n^\beta)$ $\beta > 1$

相较于高保真模拟，自适应带来的计算开销仍然较小；但对于频繁更新或需要执行多个优化步数的情况，这一开销不可忽略，因此构成了实时部署的一项实际限制。

3 实验设置

3.1 全阶模型（FOM）

我们的测试问题是经典的雷诺数为 $\mathrm{Re} = 8300$ 的顶盖驱动方腔流，与 [5] 第 5 节中的问题完全相同。该系统由无量纲形式的二维不可压缩 Navier–Stokes方程支配

$$\begin{aligned} \frac{\partial \mathbf{V}}{\partial t} + \mathbf{V} \cdot \nabla \mathbf{V} &= -\nabla P + \mathrm{Re}^{-1} \nabla^2 \mathbf{V}, \\ \nabla \cdot \mathbf{V} &= 0, \end{aligned} \tag{10}$$

其中 $\mathbf{V} = (U, V)$ 是二维速度向量，P是压力，求解域为单位正方形 $D = [0,1] \times [0,1]$，除 $U = 1$ 所在的顶盖外，所有壁面上的速度均为零。空间离散采用 100×100 全交错均匀网格上的二阶有限体积格式，通过投影步骤自动施加无散约束，并从状态变量中消除压力。由此得到的半离散系统包含$n = 2 \times 10^4$ 个未知量。在我们的实验中，观测算子 $\mathbf{h}$（见式（1））选为恒等算子，因此输出为完整速度场 $\mathbf{v}$。

系统由一个空间局部化输入驱动，该输入作用于 x-动量方程，

$$B(x,y)\, w(t) = \exp\!\big\{ -5000\big((x - x_c)^2 + (y - y_c)^2\big)\big\}\, w(t), \quad x_c = y_c = 0.95, \tag{11}$$

其中 B（x，y）定义了输入映射，而 w（t）是一个标量强迫函数。时间积分采用两步亚当斯–巴什福思格式，步长为 $\Delta t = 0.00\underline{2}5$，以保持稳定性。该流动在此雷诺数下存在一个稳定的稳态$\mathbf{V}$，但受到扰动时会表现出显著的瞬态增长。我们从稳态初始化模拟，并施加正弦强迫来激励系统，

$$w(t) = 0.1 \, \sin(4t), \tag{12}$$

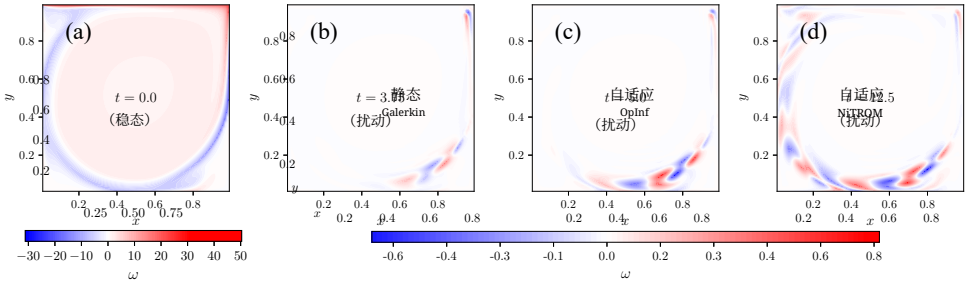


图 3：瞬态二维顶盖驱动方腔流。最左侧的场图展示了系统在 $Re = 8300$ 时的稳态解所对应的涡量场，其余图则展示涡量波动，体现了动力学随时间的演变。

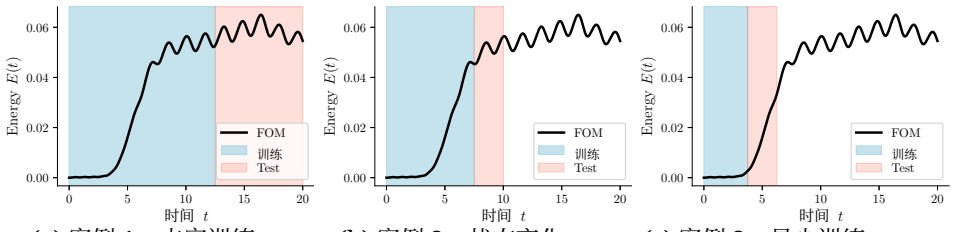
该强迫在腔体右上角附近注入动量，位置为 $x_c = y_c = 0.95$，如式（11）所述。这种连续强迫会产生丰富的瞬态和振荡行为（图 3），为测试自适应降阶模型提供了合适的基准。

备注 *3* 选择这个问题，是因为它展现了一系列特征（例如瞬态增长和类行波结构），这些特征会出现在较不典型的流动构型中，并给开发精确的降阶模型带来挑战。针对不同训练/测试机制开展的消融研究和大量数值实验，揭示了所提出的非侵入式自适应降阶模型方法的优势与不足，并阐明了在这些降阶模型能够成功应用于更大规模问题之前必须解决的研究问题。

3.2 训练和测试窗口

在本文余下部分，我们将在三种离线训练窗口逐渐缩短的不同工况下测试自适应框架。当我们使用方程（11）和（12）中的强迫信号使流动从静止状态开始时，流动会经历剧烈的瞬态能量增长（这是由动力学的非正规性所致），随后进入一个（受迫的）时间周期极限环。这一行为在图 4 中清晰可见，其中能量 $E(t)$ 将在后文定义。以下将讨论这三种训练/测试机制，其中训练时间区间为 $[t_0, t_1]$，测试区间为 $(t_1, t_2]$。

1. **案例 1（丰富训练）：** 离线时间的最终时刻 t_1 位于振荡状态的中段，提供了丰富的数据集，其中已包含后瞬态动力学的大部分信息（见图4a）。2. **案例 2（状态变化）：** 离线时间的最终时刻 t_1 恰好处于后瞬态阶段的起始处，因此降阶模型需要在在线预测期间适应新的振荡模式（见图4b）。此处，“状态变化”指的是从瞬态动力学向持续振荡动力学的转变，而离线训练窗口中并未包含后者。



(a) 案例 1：丰富训练　(b) 案例 2：状态变化　(c) 案例 3：最少训练

图 4：三种训练/测试窗口配置的示意图。每个面板展示系统能量的时间演化，蓝色阴影区域表示离线（训练）窗口 $[t_0, t_1]$ ，红色阴影部分表示在线（测试）窗口$(t_1, t_2]$。训练区间从案例 1到案例 3逐渐缩短。

3. **案例 3（最少训练）：** t_1 位于瞬态起始之前，对应于极其有限且能量很低的训练数据；此时，自适应必须从头恢复未见动力学（见图4c）。

评注 *4* 最终测试时间 t_2 针对每种案例分别选取，以保持在线预测阶段与可用离线训练数据之间的比例相当。在案例2和案例3中，离线训练窗口结束得明显更早，若将 t_2 固定为案例 1中所用的值，将导致外推时域不成比例地延长。此外，由于在线自适应依赖于由升维ROM预测初始化的FOM查询，预测误差在较长时域上不可避免地累积，在具有挑战性的工况中尤为如此。因此，我们在这些案例中报告较短但仍极具挑战性的预测区间上的结果。

质疑这些测试对静态降阶模型是否公平是合理的，因为静态降阶模型顾名思义依赖于更大的训练数据集。然而，我们在更复杂问题上使用更广泛训练数据的实验（例如 [65, 66]）表明，这些测试案例代表了静态降阶模型在强瞬态、输运主导物理情形下所面临的挑战，尤其是在每个快照都独一无二且处于流形外的混沌问题中。

3.3 离线/静态降阶模型训练

静态模型（Galerkin、OpInf 和 NiTROM）使用来自$[t_0, t_1]$的快照进行训练。静态 Galerkin ROM 使用与静态 OpInf 模型相同的 POD基构建。在所有案例中，我们都设计了一个如式（4）所示的二次ROM，因为纯线性形式被发现过于受限。事实上，它无法再现训练期间观察到的周期性能量行为，而二次模型则能准确捕捉这些分布内振荡（图 5）。确实，降阶动力学中存在二次项在物理上是合理的，因为式（10）所示的控制方程 Navier–Stokes方程包含二次对流项。降阶维数 r 选为 $r = 10$，以捕捉能量主导模态，同时避免过高的计算成本。在经典静态降阶模型中， r 通常根据累积POD能量含量来选择；而在自适应

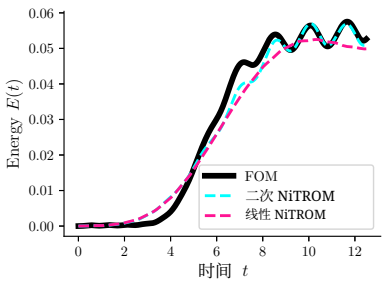


图 5：静态 NiTROM 的10维线性模型与二次模型相对于训练轨迹的性能。我们注意到，线性模型无法跟踪能量的振荡运动，而二次ROM能够捕捉真实轨迹。

情形下，这种限制不那么重要，因为模型参数会随时间演变，并能逐渐吸收新信息。对于静态 NiTROM，我们在流形上执行固定数量的离线优化步数，以确保准确地拟合训练轨迹。图 6 展示了本研究所考虑的三个案例中，离线NiTROM优化的收敛行为。案例 1 大约经过 60 次迭代后收敛，案例 2 需要 50 次迭代，而案例 3 仅经过 10 步便收敛。随后，这三个静态 NiTROM模型分别用作各案例中对应自适应降阶模型的初始条件。

3.4 在线/自适应降阶模型设置

在测试窗口$(t_1,\ t_2]$的在线阶段，我们启用第2.2 节和算法1中描述的自适应机制。每两个自适应事件之间，降阶模型向前推进 Z 个时间步。随后触发自适应，模型与全阶模型交互以获取新的快照，然后使用长度为 M 的数据窗口 $\mathcal{D}_j$ 更新其参数。我们通过扫描主要超参数$(M,\ Z,\ K)$开展消融研究，以评估自适应频率、窗口长度和优化预算对稳定性和预测准确性的影响。

3.5 评估指标

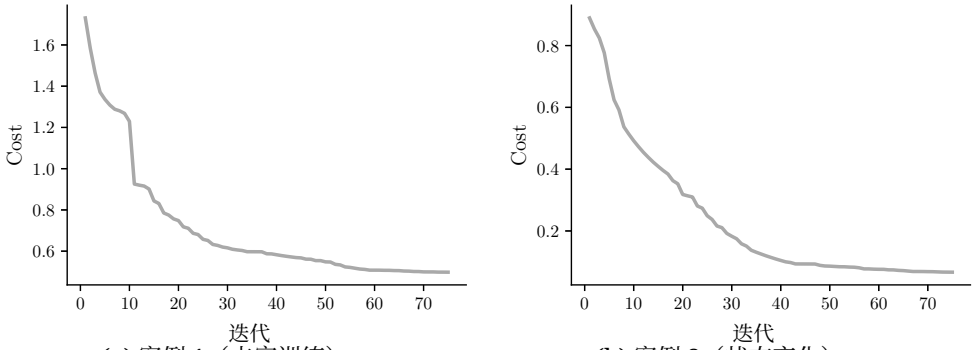
我们采用定量和定性指标评估每个降阶模型的预测质量。主要诊断指标是扰动能量，

$$E(t) = \|\mathbf{v}\|_2^2,$$

它直接衡量速度波动的振幅 $\mathbf{v}$ 相对于图中所示的底层稳态速度分布 $\mathbf{V}$ 3 a。除能量外，我们还计算绝对场误差

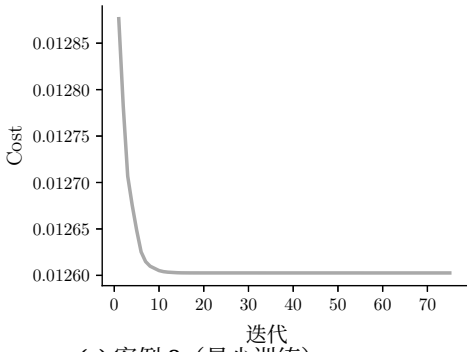
$$e(t) = \|\mathbf{v}_{\mathrm{FOM}}(t) - \mathbf{v}_{\mathrm{ROM}}(t)\|_2$$

它提供了空间域内准确性的全局度量。为补充这些全局指标，我们考察位于底壁上方 $y = 0.05\,L_y$ （此处为 $L_y = 1$）处的水平切片上的 u-速度剖面，该处的大部分主导



(a) 案例 1（丰富训练）

(b) 案例 2（状态变化）



(c) 案例 3（最少训练）

图 6：用于训练每个算例的 staticROM 的离线NiTROM优化收敛历史。

流动活动都发生于此。这为动力学的局部振幅和相位提供了一种简洁的定量表征，与二维场图中观察到的情况高度吻合。最后，我们通过目视检查全局涡量和速度场，判断预测的流动结构在相位和振幅上是否与全阶模型一致。

4 结果与讨论

4.1 算例 1：丰富训练

在此算例中，模型在 $[0, 12.5]$ 上进行训练，然后用于预测后续的$(12。5，\ 20]$ 区间（图4a）。训练窗口包含流动的整个瞬态过程和若干振荡周期，这意味着模型已经观测到大多数相关的时空结构。因此，在线阶段相对简单，只需进行小幅振幅与相位修正。

4.1.1 静态降阶模型

图 7 展示了静态 Galerkin、OpInf 和 NiTROM 模型的能量演变及相应的场误差。三个模型都再现了全阶模型的一般振荡模式，但很快便出现偏离，表明振幅

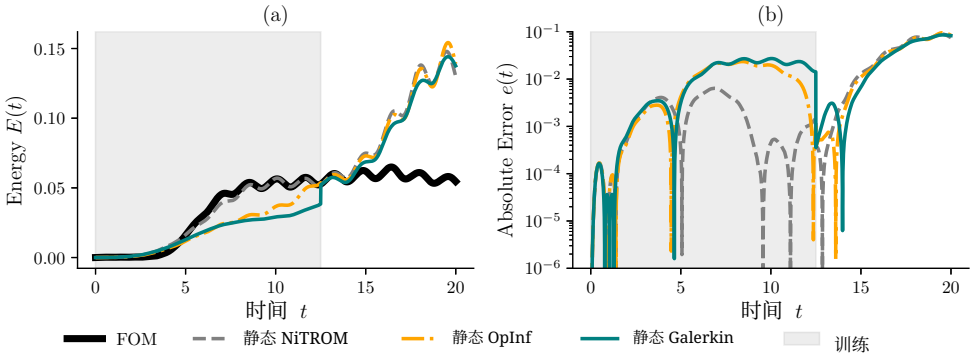


图 7：案例 1（丰富训练）中静态降阶模型的能量和场误差演化。所有静态模型（Galerkin、OpInf、NiTROM）都能匹配总体振荡模式，但会逐渐高估能量，导致振幅不断增大，并在超出训练范围后失去稳定性。请注意，在训练窗口内，NiTROM 的表现远优于其他静态模型。

过度增长；若继续积分，这很可能会导致不稳定。请注意，Galerkin 投影通过将全阶控制方程投影到该基上，以侵入式方式推导降阶算子；而 OpInf 则通过回归从数据中识别降阶动力学，因此，尽管二者使用相同的基，得到的降阶算子仍不相同，通常也会表现出不同的预测行为。在训练窗口内，静态 NiTROM 对数据的拟合效果最佳，体现了离线流形优化和斜投影所带来的优势。然而，在预测窗口内，所有静态模型的表现都大致相当。这表明，通过离线优化获得的外推准确性有限，因此需要在线自适应机制来维持准确性。对预测时域末端涡量场的检查也证实了这一趋势，见图 10 。尽管主要流动结构和相位模式均得到了正确捕捉，但预测场的振幅明显大于全阶模型的场，这与观测到的能量过冲一致。附录 A 包含本文所有测试方法在所有案例中的（扰动）速度分量场 u 和 v。

4.1.2 频繁更新下的自适应降阶模型

接下来，我们考察自适应窗口为 $Z = 10$ 、回溯窗口包含 $M = 100$ 个快照的自适应模型。对于自适应 OpInf，基使用窗口化奇异值分解进行更新；对于自适应 NiTROM，每次更新执行 $K = 10$ 次优化迭代。观察图 8中的能量和误差图可见，两种方法都成功抑制了静态模型中出现的非物理能量增长。自适应 OpInf 的能量略有衰减，表明存在轻微的过度阻尼，在极长时间范围内可能导致振幅逐渐减弱。相比之下，自适应 NiTROM 几乎完美地匹配了全阶模型的能量曲线。速度切片（图 9）和涡量场（图 10）的对比表明，两种自适应模型都能准确再现动力学并保持振幅控制，而自适应 NiTROM 则引入了

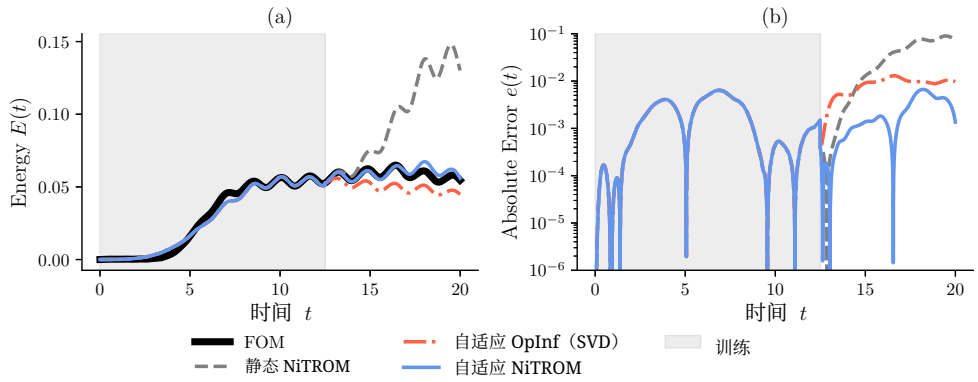


图 8：案例 1（丰富训练）中自适应降阶模型（$Z = 10$、 $M = 100$）的能量与场误差演化。两种自适应模型都抑制了静态降阶模型的非物理能量增长。自适应 OpInf 表现出轻微阻尼，而自适应 NiTROM（使用 $K = 10$）的能量几乎与全阶模型一致。

相对于全阶模型的轻微相位滞后。原则上，可以通过在 NiTROM 的损失函数中加入相位对齐项来减轻这种滞后。

4.1.3 更新频率较低时的自适应降阶模型

为考察降低自适应频率的影响，我们将自适应窗口增大至 $Z = 50$，并将回溯长度减小至 $M = 20$。能量和误差图如图11所示。在此配置下，自适应 OpInf 仍能控制能量增长，但与前一种情况一样，整体振幅会随时间减小。自适应 NiTROM 现在每次更新执行 $K = 20$ 步优化，却难以恢复正确的能量轨迹。这是由其优化特性以及基于先前学习到的参数进行热启动初始化所致。当系统在两次自适应事件之间发生显著变化时，新的最优参数在优化景观中可能与先前参数相距甚远，导致算法收敛缓慢或陷入局部极小值。这一现象凸显了纯粹基于优化的自适应方法的一项关键局限。为解决这一问题，我们测试了自适应 OpInf–NiTROM 方法，该方法先快速进行 OpInf 重新拟合，再执行 $K = 10$步 NiTROM 优化。正如预期，这种配置的表现优于单独使用任一种自适应方法。OpInf 阶段根据当前窗口快速重新学习算子，从而有效提供良好的初始化；NiTROM 精化则沿流形微调参数。图11 中的能量轨迹与全阶模型高度吻合，证实这种混合自适应能够有效应对系统在更新之间发生显著变化的情况。不过，这种配置尚未充分展现混合方法的潜力。从图12 中的速度切片和图13中的定性场对比可以看出，自适应 OpInf 已能准确再现流动结构，因此尽管能量曲线略有改善，额外的流形优化步骤仍没有太大的提升空间。接下来的案例（2 和 3）将引入

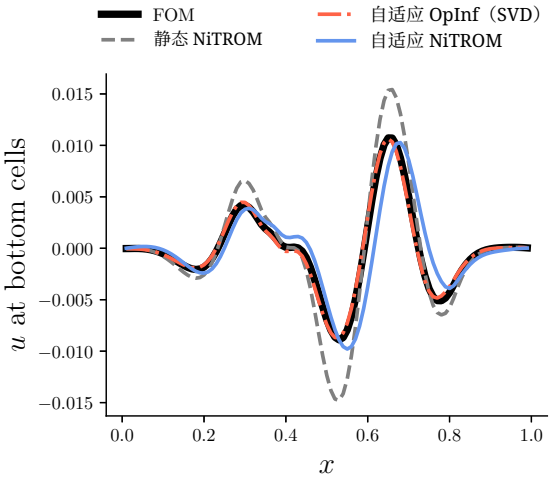


图 9：案例 1（丰富训练）中，位于底壁上方$y = 0.05\,L_y$ 处的水平位置$y = 0.05\,L_y$ 的自适应降阶模型（$Z = 10$, $M = 100$）的速度切片，在 $t = 20$时的结果。自适应 OpInf 和自适应 NiTROM 均能准确再现近壁速度的振幅，证实它们能够在保持物理一致性的同时控制能量增长。自适应 OpInf 还能正确捕捉相位，而自适应 NiTROM 则会造成轻微的相位偏移。

更具挑战性的条件，在这些条件下，混合方法相较于单独使用各方法的优势将清晰显现。

4.1.4 消融研究

我们进行了详细的消融研究，以评估自适应模型对其超参数和设计选择的敏感性。完整的实验结果见图14–16。以下结论概述了主要发现：

- **自适应窗口的影响（Z）：** 自适应 OpInf 在很宽的 Z 取值范围内仍保持稳健，即使更新间隔较大也能维持稳定的预测（见图 14a）。自适应 NiTROM 在较小的 Z 时表现最佳，此时优化器的初始位置接近当前最小值；随着 Z增大，其准确性会下降（见图15a）。请注意，在图14a和15a中，针对每个 M 值选择窗口大小 Z，以确保所有模型覆盖相同的时间区间。这样可以确保观察到的差异不会受到回溯范围不匹配的影响；更多细节将在下文讨论。
- **回溯窗口的影响（M）：** 回溯窗口的选择应与 Z保持一致。一旦 Z 被选定，量 $(M-1)\,Z\Delta t$ 便有效决定了每次自适应中所考虑的系统过去的历史的时间跨度（图2）。 M 的适当选择因此取决于具体问题，并应依据底层的特征物理时间尺度

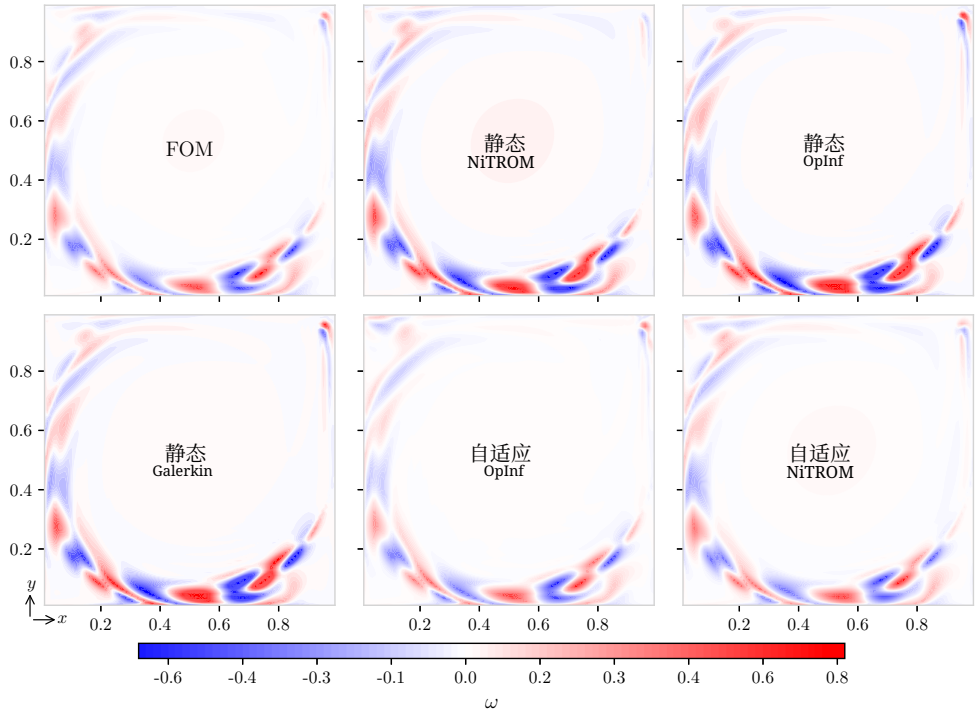


图10：案例 1 中在线窗口（$t = 20$）结束时的真实值（全阶模型）和预测值（降阶模型）涡量场。所有模型均为二次模型，维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 100$，每次自适应优化预算为 $K = 10$。

动力学。例如，我们在图14b 和15b 中看到，较小的 M 无法捕捉到足够的时间变化，因而无法实现有效学习。

- **基更新策略：**在本研究中，我们采用窗口化奇异值分解作为默认的基更新策略，因为它提供了清晰且一致的基线，不依赖额外的算法选择。然而，图 14c中的消融结果表明，增量奇异值分解可能是值得未来探索的方向：它在案例 1 中的性能与窗口化奇异值分解相近，同时计算成本显著更低。如前所述，在这种非侵入式设定下，秩一基更新效果不佳，因为无法对基进行局部侵入式修改。

- **NiTROM 中的优化步数（K）：**增加 K 通常会提高适应精度（如图 16所示），但过大的值会带来不必要的计算负担。在动力学快速演变的工况下，K 的值过大可能会因过分强调局部窗口中、已无法充分代表当前动力学的数据而降低响应能力；而在缓慢变化的工况下，则不会观察到这种影响。

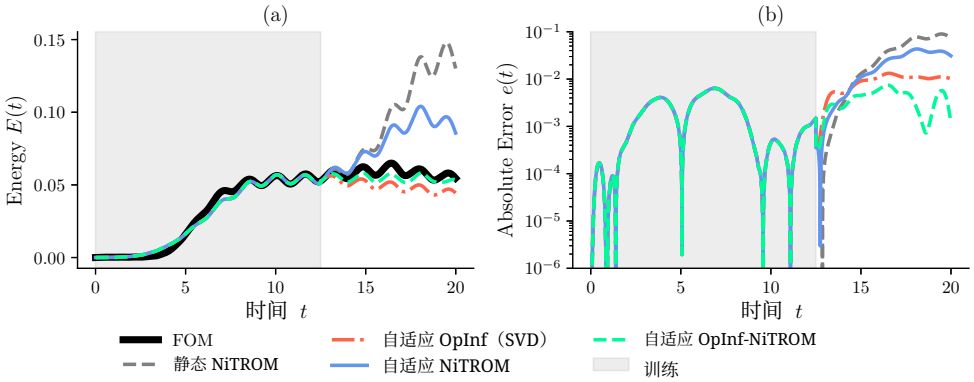


图11：案例 1（丰富训练）中自适应降阶模型（$Z = 50$， $M = 20$）的能量与场误差演化。自适应 OpInf 继续稳定该解，但对能量略有过度阻尼。自适应 NiTROM（含 $K = 20$）因其热启动敏感性而表现不佳，而自适应 OpInf–NiTROM（含 $K = 10$）则恢复了近乎完美的能量跟踪。

4.2 案例 2：状态变化

我们在另一种设置下测试了该模型：离线窗口覆盖 $t = (0,7.5)$，在线预测覆盖随后的 $t = (7.5,10)$，对应 1000个外推步数（图4b）。因此，这些模型仅在低能量阶段和瞬态增长阶段接受训练，必须外推至随后的振荡状态。对于自适应运行，我们将自适应窗口固定为 $Z = 10$，回溯窗口固定为 $M = 10$，并将基于 NiTROM 的更新的优化预算设为每次自适应 $K = 10$次迭代。自适应 OpInf 使用窗口化奇异值分解进行基更新。

4.2.1 静态降阶模型

图17表明，所有静态降阶模型均无法复现新的振荡状态。静态 Galerkin 与静态 OpInf 在能量上迅速发散。静态 NiTROM 能够跟踪最初的一次振荡，但随后能量不受控制地增长，解偏离物理吸引子，与其他静态模型类似。观察自适应 NiTROM 在在线阶段结束时的预测场（见图20），可见静态模型无法忠实地表示真实动力学，从而证实一旦系统离开训练分布，静态预测便具有非物理性质。

4.2.2 自适应降阶模型

采用 $Z = 10$ 和 $M = 10$后，图 18 中的自适应 OpInf 曲线显示，在线窗口内能量明显下降，表明动力学逐渐消退。虽然这使解趋于稳定，但它低估了真实振荡的振幅，因此无法定量捕捉状态变化。自适应 NiTROM 使用$K = 10$ 每次更新的迭代次数，与静态 NiTROM 相比几乎没有改进，因为第一个周期后能量同样增长。这与观察到的行为一致

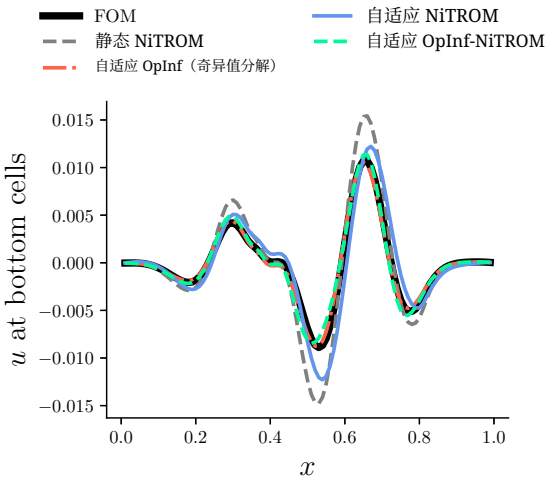


图12：在案例 1（丰富训练）的 $t = 20$时刻，位于底壁上方$y = 0.05\,L_y$ 水平位置处的自适应降阶模型（$Z = 50$、 $M = 20$） u速度切片。与自适应 OpInf 类似，混合模型保留了振荡的相位和振幅。自适应 NiTROM 会导致轻微的相位偏移和振幅增长。

在采用稀疏更新的案例 1 中，这也符合 NiTROM 热启动流形优化的机制。

相比之下，自适应 OpInf–NiTROM 配置在整个预测窗口内保持了总体能量水平。尽管混合曲线未能完美匹配细致的振荡行为，但总能量得以维持，而非像自适应 OpInf 那样衰减，或像自适应 NiTROM 那样增强。

速度切片（图 19）和场可视化（图 20）清晰地展示了这些差异。自适应 OpInf 虽然抑制了能量，却在振荡运动中引入了明显的相位误差，并且在区域的部分位置出现了轻微的数值振荡。显著的相位偏移在 u 切片中也很明显。我们还发现，与静态版本相比，单独使用自适应 NiTROM 并未带来明显改善，这与我们从能量图中观察到的情况一致。

混合方法得到的场最为令人信服。我们观察到位置和强度都准确的清晰涡结构，且 u 切片相较于全阶模型展现出良好的振幅控制和相位对齐。值得注意的是，经过简短的 NiTROM 精化后，自适应 OpInf 预测中出现的轻微振荡大幅减少，不过相同区域仍残留着微弱的振荡。这是一个典型案例，说明单纯依赖数值指标可能会产生误导。尽管混合模型的能量曲线未能完美追踪振荡运动，但定性的场评估表明，混合模型对状态变化的重构最为忠实。

4.3 案例 3：最小训练

在这一最具挑战性的配置中，模型使用以下区间的数据进行训练： $t = (0,\ 3.75)$ 随后使用以下区间的数据进行测试： $t = (3.75,\ 6.25)$ (图 4c)。训练数据对应于初始

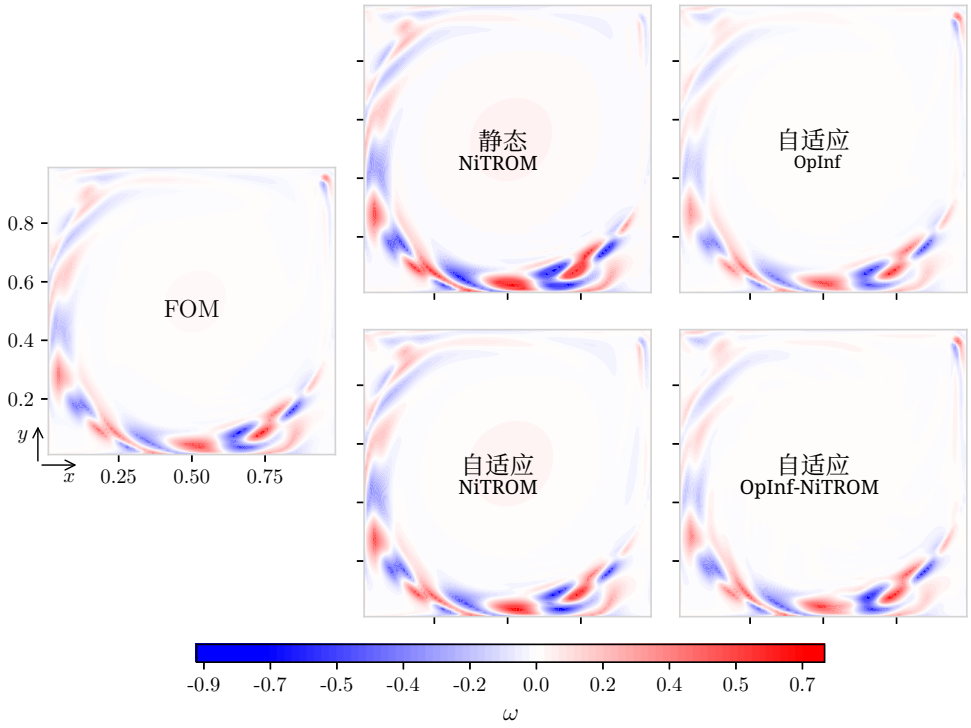
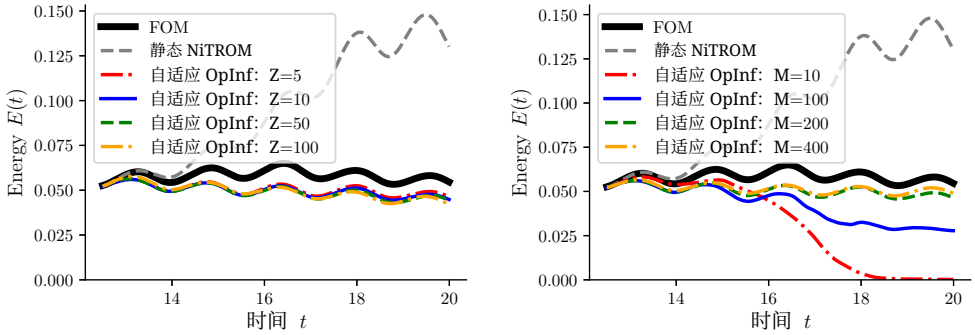


图13：案例 1 中在线窗口结束时（$t = 20$）真实值（全阶模型）与预测（降阶模型）的涡量场。所有模型均为二次模型，维数为 $r = 10$。对于自适应模型，我们设置自适应窗口为 $Z = 50$，回溯窗口为 $M = 20$。对于自适应 NiTROM 和自适应 OpInf–NiTROM，每次自适应优化预算分别设为 $K = 20$ 和 $K = 10$。

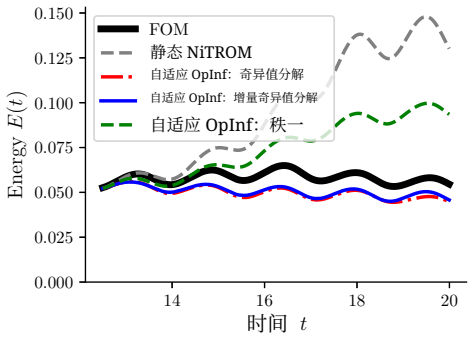
流动的低能量状态，其中几乎没有动力学演化。训练窗口结束后，全阶模型的能量迅速增长，因为流动转入瞬态阶段。因此，这种设置对任何降阶建模方法都构成了极限考验：其基和算子几乎未经历任何变化，而后续动力学在振幅和空间结构上都迅速演变。自适应参数固定为 $Z = 10$、$M = 10$ 和 $K = 10$。较短的回溯窗口确保模型仅依赖最新信息，而非过时且缺乏代表性的状态。

4.3.1 静态降阶模型

与训练数据不同，静态降阶模型在预测期间并未保持低能量状态。图 21显示，在线阶段的能量几乎立即开始增长，并在预测窗口结束时偏离全阶模型。检查在线窗口末尾静态 NiTROM 的流场（见图 24）可知，预测的动力学仍局限于右壁附近，因为基是根据有限的低能量训练数据构建的。由于该基从未



(a) 改变 Z 的影响（ M 可变，针对 (b) 改变 M 的影响（ $Z = 10$ 固定）的每个 Z，以确保窗口跨越 1,000 个时间步）



(c) 基适应方法的影响（使用冻结的 $Z = 10$ 和 $M = 100$）

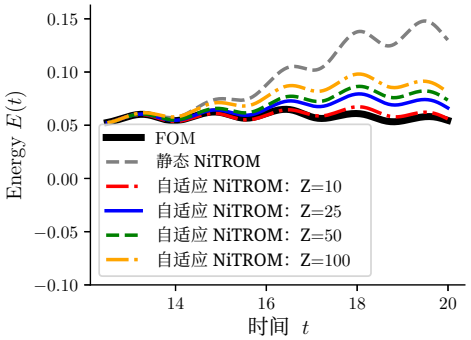
图 14：将自适应 OpInf 应用于案例 1 的消融研究。此处， Z 和 M 分别表示自适应窗口和回溯窗口。所示为仅在在线（预测）窗口内的性能，因为所有模型的训练过程都相同。

观测到更广泛的瞬态演化，预测的涡旋未能传播至域内部。

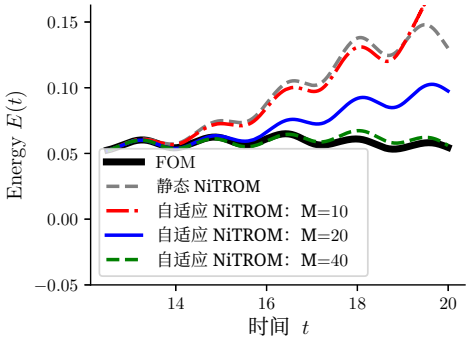
4.3.2 自适应降阶模型

快速变化也对该案例中的自适应模型构成了挑战。如图所示22，自适应 OpInf 的能量出现了明显衰减，与其在案例 2 中的表现相似。模型保持稳定，但低估了能量振幅，因为窗口内有限的历史数据限制了算子推断所用的时间上下文。尽管自适应 NiTROM 采用了表达能力更强的优化方法，但与静态 NiTROM 相比仍未显示出显著改善。其能量曲线呈现出相似的增长模式，未能再现瞬态动力学。

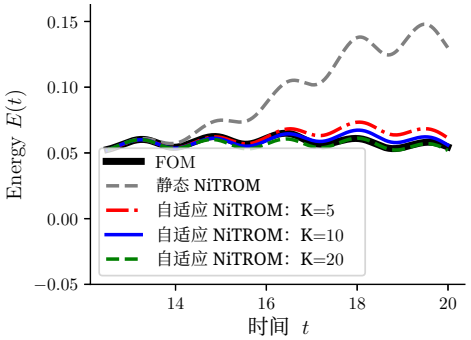
相比之下，自适应 OpInf–NiTROM 模型再次取得了最稳定且具有物理意义的结果。能量水平保持有界，并与全阶模型一致，既未衰减至零，也未发散。



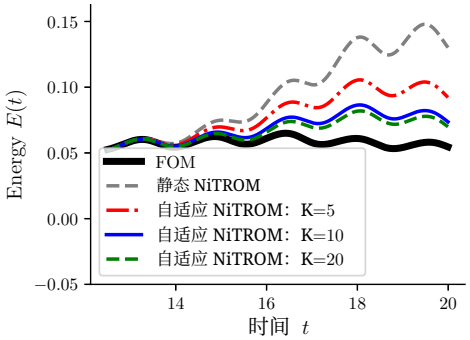
（a）改变 Z 的影响（固定 $K = 10$，并针对每个 Z 使用可变的 M，以确保窗口涵盖 1,000 个时间步）



（b）改变 M 的影响（固定 $Z = 10$ 和 $K = 10$）。请注意，NiTROM 会收敛与真实值相比，
窗口大小显著小于 OpInf。



（c）改变 K 的影响（冻结 $Z = 10$ 和 $M = 40$）



（d）改变 K 的影响（冻结 $Z = 50$ 和 $M = 10$）

图15：应用于案例 1 的自适应 NiTROM 消融研究。其中，Z、M 和 K 分别表示自适应窗口、回溯窗口和每次自适应的流形优化步数。此处仅展示在线（预测）窗口内的性能，因为所有模型的训练均相同。

接下来，我们考察图 23中的速度切片对比和图 24 中的场可视化，以确认这些定量趋势。仔细观察自适应 OpInf 的预测场可以发现，它仅捕捉到了右壁附近的局部增强，未能再现内部涡旋传播。另一方面，自适应 NiTROM 相较于其静态版本没有明显改善，这与能量图中呈现的趋势一致。然而，自适应 OpInf–NiTROM 方法展现出清晰的涡旋形成过程，其空间模式也与全阶模型高度吻合。短暂的 NiTROM 精化再次“清理”了 OpInf 阶段产生的小幅振荡伪影，使场更加平滑且符合物理规律。 u 切片同时捕捉到了振幅增长以及大致的相位

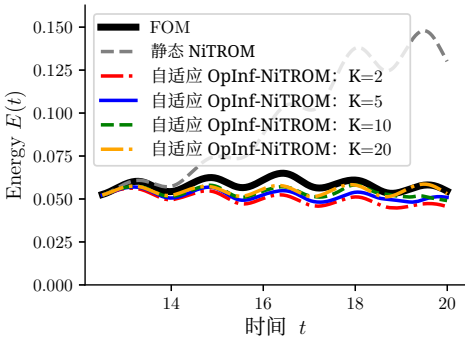


图 16：冻结 $Z = 50$ 和 $M = 20$ 时，自适应（混合）算子推断–NiTROM 应用于案例 1 的每次自适应流形优化步数 K 的影响。图中仅展示在线（预测）窗口内的性能，因为所有模型的训练过程都相同。

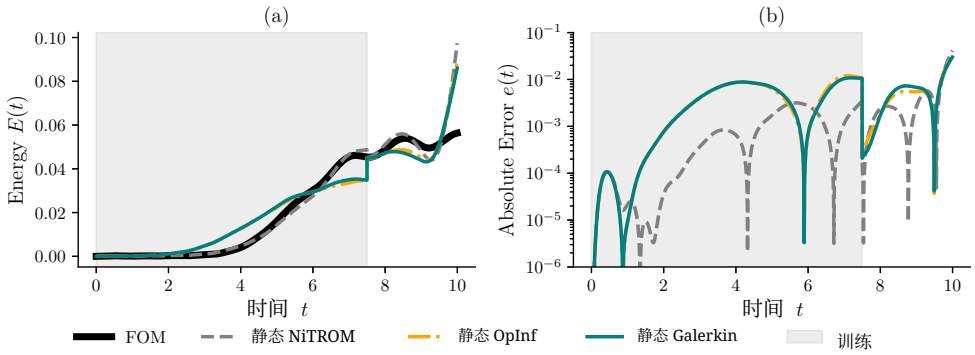


图 17：案例 2（状态变化）中静态降阶模型的能量和场误差演化。所有静态模型均无法再现状态变化。静态 NiTROM 虽然捕捉到了第一次振荡，但很快便发生偏离，这凸显了静态降阶模型无法泛化到未见过的动力学状态。

瞬态振荡，表明即使离线模型几乎没有见过相关数据，混合方法仍能跟踪动力学。

4.4 主要观察与要点

这三个案例共同全面展现了不同自适应非侵入式形式在难度逐步增加时的表现：从训练良好的模型中进行小幅修正，到基于几乎不含特征的数据进行完全外推。这些实验揭示了若干一致的趋势和见解。

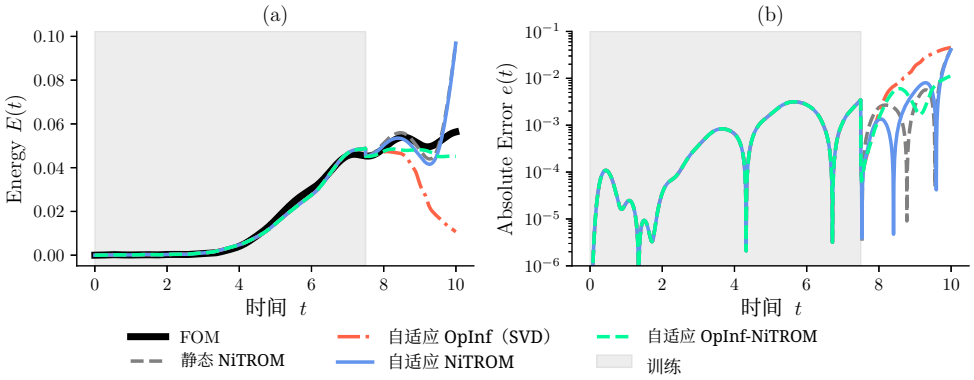


图 18：案例 2（状态变化）中自适应降阶模型（$Z = 10$、 $M = 10$）的能量与场误差演化。自适应 OpInf 稳定了轨迹，但低估了振荡幅度。自适应 NiTROM（采用 $K = 10$）相对于静态 NiTROM 并无改进，而自适应 OpInf–NiTROM（采用 $K = 10$）方法则保持了有界且物理一致的能量演化。

i. 静态降阶模型受限于训练覆盖范围。
在所有案例中，静态 Galerkin、OpInf 和 NiTROM 模型只有在训练窗口已包含主导流动动力学时才能表现良好。一旦超出这一工况，它们便会迅速失去物理一致性。这一局限凸显了在线校正机制的必要性，尤其是在系统演化超出离线数据所代表的分布之后。

ii. 自适应 ***OpInf*** 稳健且高效。
在具有丰富训练数据的情形中（案例 1），自适应 OpInf 能够稳定预测并防止长期能量漂移。在更具挑战性的设置中（案例 2 和 3），只要仔细调节其超参数，它便能够跟踪定性的状态变化。其主要优势在于简单、计算成本低，以及对较大自适应间隔 Z具有显著的容许度。然而，其性能对窗口大小 M 的选择高度敏感，并且偶尔会在场中引入局部数值伪影。

iii. 自适应 ***NiTROM*** 受限于优化敏感性与成本。
自适应 NiTROM 中的联合流形优化能够在更新频繁且优化充分收敛（ Z较小、 K充足）时实现高度准确的修正。在这种情况下，能量曲线几乎可以与全阶参考解完全吻合。然而，随着自适应间隔增大，该方法可能无法消除累积误差。在分布外案例（案例2和案例3）中，优化景观会变得过于复杂，与静态模型相比，该方法几乎没有优势。这种敏感性凸显了在快速变化的动力系统中进行在线轨迹优化的难度。此外，与自适应 OpInf 相比，流形优化会带来显著的计算开销，这

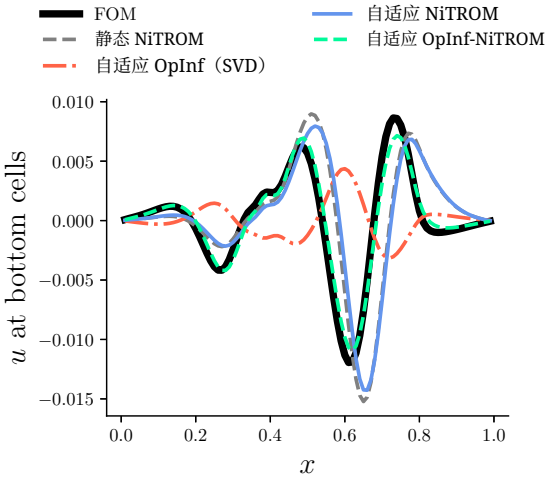


图1： u：案例 2（状态变化）下， $t = 10$处底壁上方水平位置$y = 0.05\,L_y$ 的自适应降阶模型速度切片（$Z = 10$， $M = 10$）。自适应 OpInf 未能捕捉真实剖面；与静态对应模型相比，自适应 NiTROM 的修正幅度很小，而混合方法的振幅和相位均与全阶模型一致。

限制了该方法用于在线自适应的适用性。表2 报告了每种操作单次执行的平均实际运行时间。

iv. 自适应 ***OpInf–NiTROM*** 兼具两者的优势。
自适应（混合）OpInf–NiTROM 方法在鲁棒性与准确性之间实现了实用的平衡。初始的 OpInf 更新会根据最新数据重新学习算子，从而有效地为新工况重新初始化模型。随后进行的 NiTROM 精化则强制保持几何一致性，并消除仅使用 OpInf 的预测中出现的细微数值伪影。在案例2和案例3中，混合配置始终生成物理意义最明确、视觉上最连贯的场。这一观察表明，将快速的基于回归的更新与简短的流形精化相结合，是在具有挑战性的环境中保持准确性的有效策略，不过仍需进一步研究，以提高流形优化的效率。

v. 在线 ***NiTROM*** 更新的收敛行为。
对于基于 NiTROM 的自适应降阶模型，其收敛分析与离线设置存在根本差异，因为在在线预测期间，每次发生自适应事件时，优化都会重新初始化并反复求解。为了评估在线优化的行为，我们考察了所有自适应事件中混合 OpInf–NiTROM 更新的收敛历史。图 25 展示了案例 1–3 中各自适应事件对应的归一化目标值随优化迭代次数的变化。尽管各事件

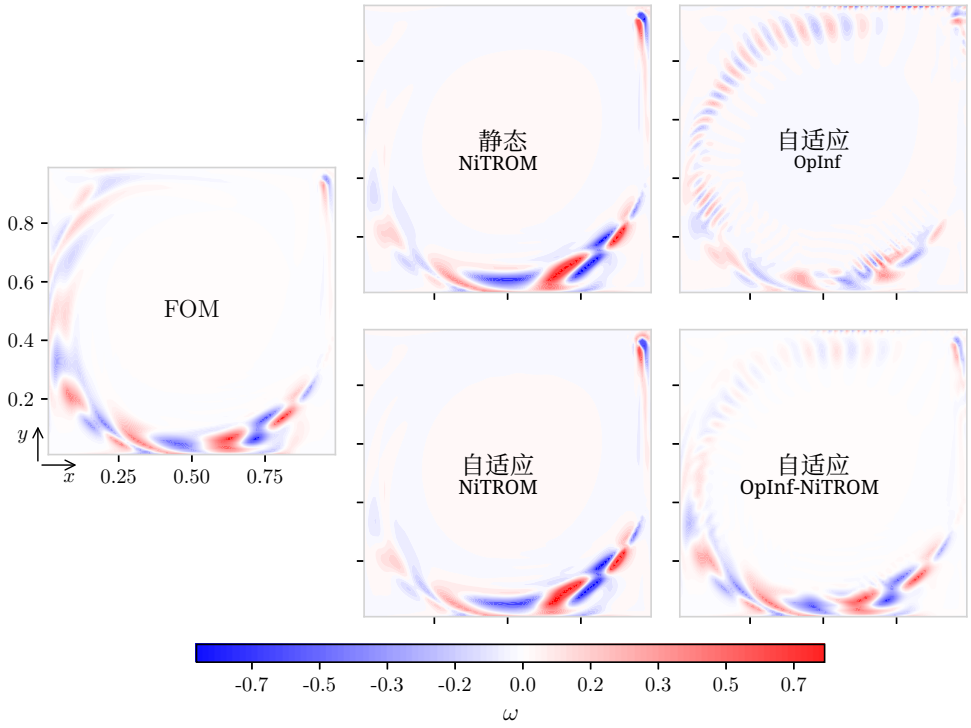


图 20：案例 2 中在线窗口结束时（$t = 10$）的真实值（全阶模型）与预测值（降阶模型）涡量场。所有模型均为二次模型，且维数为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 10$，每次自适应优化预算为 $K = 10$。

呈现出不同的收敛速率，但在所有案例中，目标值都能在少量迭代内持续下降。这表明，只需固定且适中的优化步数即可获得有效的在线更新，无需针对不同事件专门调整收敛参数。

vi. 更广泛的影响。
总体而言，结果表明，在线自适应可显著扩展非侵入式降阶模型的预测范围。当系统处于或接近原始子空间内时（案例 1），自适应更新能够完全校正振幅漂移。当动力学演变到超出该子空间时（案例2和案例3），自适应仍能实现部分恢复，尤其是在采用混合更新时。表 3 总结了在全部三个案例中观察到的定性性能趋势。

vii. 表 *2* 中关于实际运行时间的说明。
目前，由于我们使用的是未经修改的离线版本的奇异值分解、OpInf 和 NiTROM 构造方法，这些方法的开销较大。我们可以通过多种方式

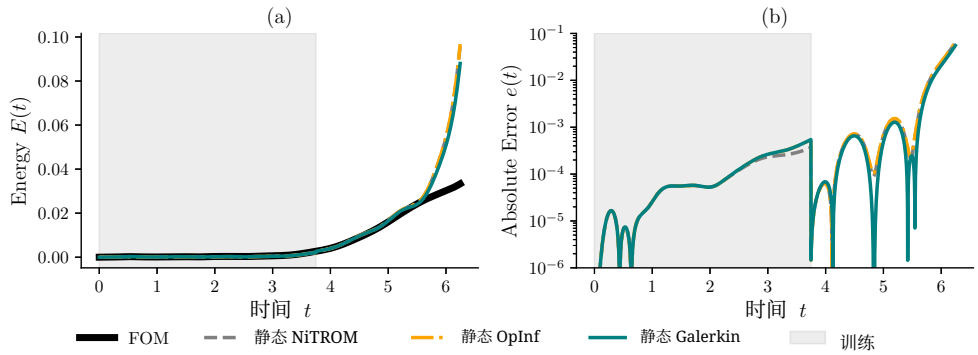


图 21：案例 3（最少训练）中静态降阶模型的能量和场误差演化。尽管训练数据的能量较低，静态降阶模型仍会出现人为的能量增长。

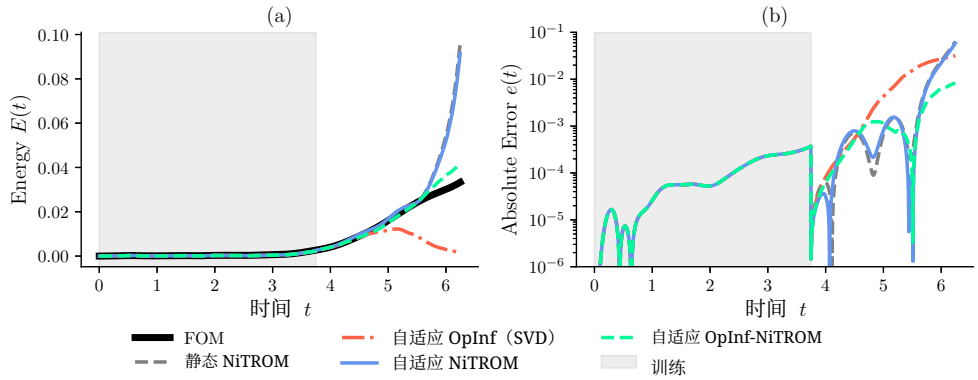


图 22：案例 3（最少训练）中自适应降阶模型（$Z = 10,\ \ M = 10$）的能量和场误差演化。自适应 OpInf 能够稳定系统，但会低估能量；自适应 NiTROM（使用 $K = 10$）的表现与其静态版本相似。自适应 OpInf–NiTROM（使用 $K = 10$）能够维持有界能量，并正确跟踪瞬态增长趋势。

加速降阶模型：a) 将算法改进为适用于在线增量场景；b) 利用降阶模型能够采用比全阶模型更大时间步的事实。我们还要指出，本文研究的方腔流并非特别适合凸显计算节省。事实上，空间离散的模板宽度较小、采用显式时间推进、 100×100 网格相对较小，而且使用了 Python 向量化，这些因素使全阶模型求解器的速度极快（如表中所示）。对于需要隐式时间推进、单个时间步计算成本更高（例如，采用激波捕捉逻辑的可压缩流求解器）以及网格精细得多的复杂问题，更适合用来展示成本降低。不过，正如前文所述，本研究的一个主要目标是量化实现预测性模型降阶所面临的挑战。这

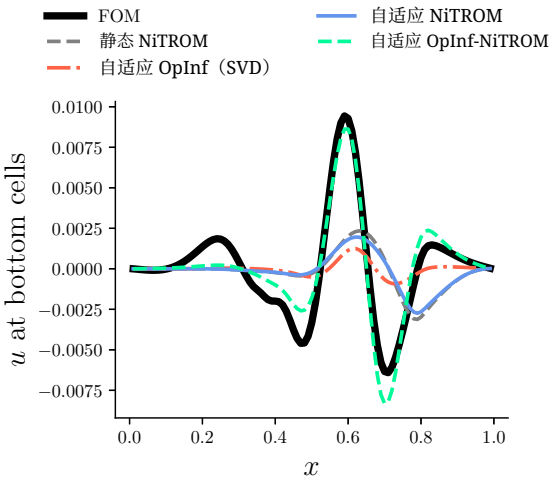


图 23：u-速度切片，展示案例 3（最少训练）中水平位置$y = 0.05\,L_y$（底壁上方）处的自适应降阶模型（$Z = 10, M = 10$），时间为 $t = 6.25$。混合方法能够捕捉振幅和相位的变化，体现了自适应 OpInf–NiTROM 的外推能力。

选择这个特定问题，是因为全阶模型求解器效率很高（允许我们运行数百次），而且甚至 这种流动构型也呈现出降阶模型面临的复杂动力学特征，并使我们能够确定有前景的研究方向，最终推动自适应降阶模型应用于瞬态及平流主导流动问题。

尽管如此，该表仍有助于识别趋势。特别是，显而易见，基于 NiTROM 的自适应比 OpInf 的计算成本高得多，因为优化需要通过环境空间矩阵-向量及矩阵-矩阵运算，利用现有数据对测试子空间与试探子空间进行优化。鉴于 NiTROM 精化在 OpInf 自适应模型中展现出的优势，未来工作将着重解决这一问题，并降低矩阵流形优化的相关成本。

5 结论与展望

本研究提出了一个通用框架和一套设计规则，用于构建自适应、非侵入式降阶模型（ROM），最终目标是将传统数据驱动降阶模型从静态设置拓展至能够持续学习的流式代理模型。我们的动机是设计出不仅能够压缩和近似复杂动力系统，而且还能随系统演化的降阶模型。为此，我们提出了自适应非侵入式降阶模型的完整表述，并给出三种具体实现：自适应算子推断（OpInf）、自适应非侵入式基于轨迹的降阶模型优化（NiTROM），以及融合两者各自优势的自适应 OpInf–NiTROM 混合模型。这些

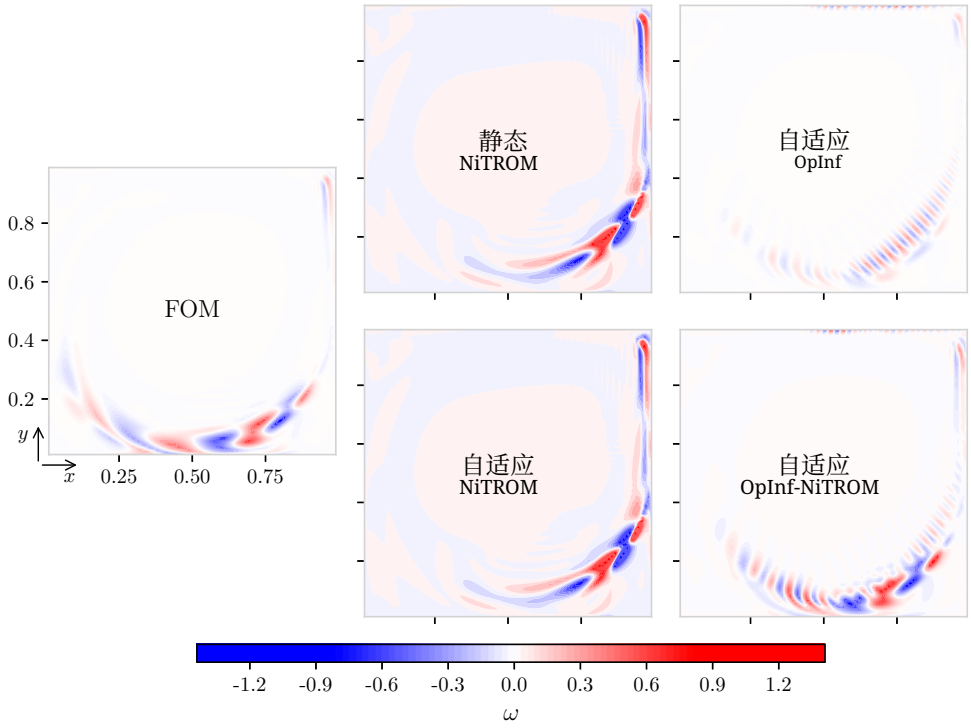


图 24：案例 3 中在线窗口 ($t = 6.25$) 结束时的真实值（全阶模型）和预测值（降阶模型）涡量场。所有模型均为二次模型，维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 10$，每次自适应优化预算为 $K = 10$。

方法仅利用高质量快照即可实现在线参数校正，无需侵入式访问完整阶模型的算子。

针对二维顶盖驱动方腔流开展的系统实验表明，与传统静态降阶模型相比，所提出的框架能够显著提升长期稳定性和准确性。当模型在接近其训练范围的工况下运行时，自适应更新可有效消除能量漂移，并在较长时间跨度内保持振幅与相位对齐。在更具挑战性的外推情形下，系统会转入此前未见的工况，此时混合模型始终能给出最符合物理规律且最稳定的结果。总之，这些发现证明了真正能够自我修正的非侵入式降阶模型具有可行性，使其能够实时适应不断演变的动力学。

尽管取得了这些令人鼓舞的成果，但仍存在若干局限。虽然基于 NiTROM 的模型在中等程度的自适应场景中表现出较高的准确性，但其流形优化步骤计算开销较大，并且对自适应窗口大小、回溯窗口大小以及在线优化步数等超参数较为敏感。尽管这些方法的成本低于全阶模拟，

表 2：自适应非侵入式降阶模型框架中各项操作单次执行的平均实际运行时间（毫秒）。所有实际运行时间均在配备 Apple M3 Pro 芯片的本地计算机上测量，并取 10 次运行的平均值，以减小测量噪声。所报告的数值对应于回溯窗口为 $M = 10$ 个快照的情况。请注意，各方法均从同一离线阶段初始化，因此离线训练成本在各方法间相同，故此处未予列出。

操作	描述	实际运行时间 [毫秒]
降阶模型时间步	一个降阶模型时间步	0.322
全阶模型时间步	一个全阶模型时间步	2.33
SVD*	对整个窗口进行一次完整的奇异值分解	2.85
OpInf*	对降阶算子进行一次最小二乘重新拟合	7.49
NiTROM*	一次流形优化迭代	31.2

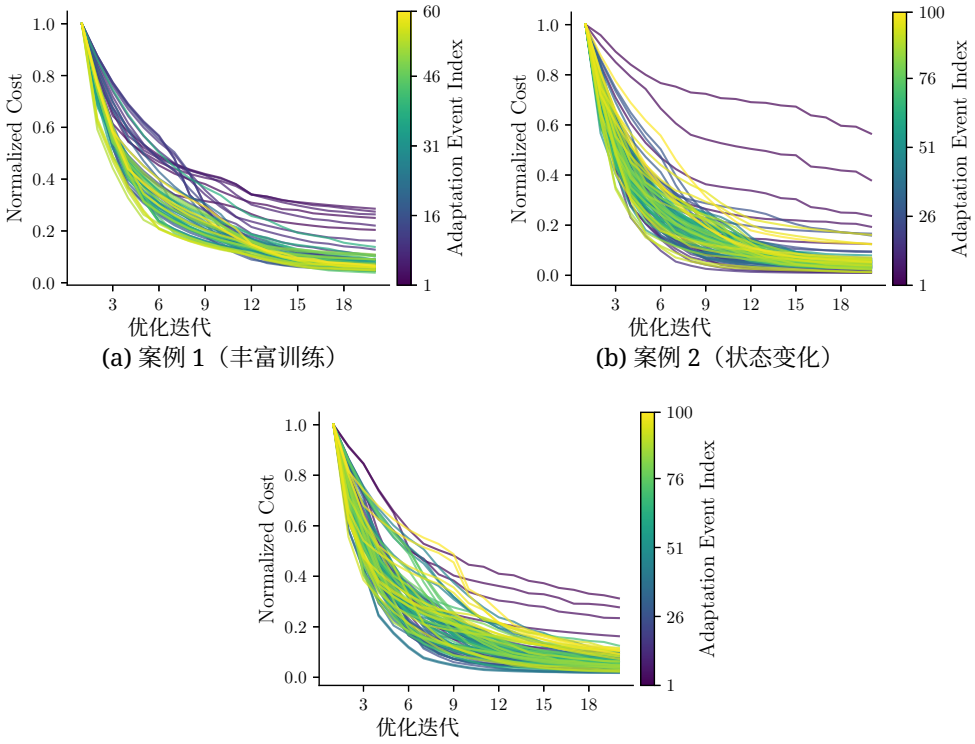
* 表示这些算法未经过修改，仍为其离线版本。

表3：各案例中降阶模型性能的定性总结。

模型	案例 1	案例 2	案例 3
静态 Galerkin/OpInf/NiTROM	部分成功	失败	失败
自适应 OpInf	鲁棒性极佳	部分失效	失效
自适应 NiTROM	对较小的 Z准确	失效	失败
自适应 OpInf–NiTROM	**总体最佳**	**总体最佳**	**总体最佳**

但目前的成本和敏感性使其难以部署于数字孪生类场景。我们也谨慎指出，仍需在更广泛的测试问题和数据场景中进行严格测试与评估。目前，该框架仍属于概念验证或求解器加速器。要解决这些局限，仍需持续开展算法开发，以提升鲁棒性和效率。**展望：** 未来研究自然会朝几个方向展开。首先，如文中所述，所提算法的最大加速比严格受自适应频率的限制 Z，因为每隔 Z 个时间步就要更新一次完整的高保真算子。可以像经典侵入式降阶模型那样，对高保真算子进行采样并插值，从而加快这一过程。其次，开发自适应的、基于误差的触发机制，可以用由残差指标驱动的智能决策取代固定的自适应窗口。第三，通过增量式或低秩算子更新加速自适应 OpInf 形式，有望使该方法更接近实时应用，尤其是在高频自适应的情况下。最后，在真正的数字孪生环境中测试该框架，让自适应直接由传感器数据而非全阶反馈驱动，是迈向实际应用的重要一步。

总体而言，本研究首次对自适应非侵入式降阶建模进行了详细探索，并证明持续的数据驱动自适应是



(a) 案例 1（丰富训练）

(b) 案例 2（状态变化）

(c) 案例 3（最少训练）

图 25：混合 OpInf–NiTROM 模型自适应更新期间在线 NiTROM 优化的收敛历史。每条曲线对应一个自适应事件，显示归一化目标值随优化迭代次数的变化。

训练真正具有预测能力的模型所必需的。这些结果为新一代自我校正、低维代理模型奠定了基础，使其能够在复杂的非平稳系统中进行长期且符合物理规律的预测。我们预计，此类自适应降阶模型将在未来面向高维动力系统的数字孪生、基于模型的控制和科学机器学习发展中发挥核心作用。

A 附加流场可视化

为补充第4节中的涡量可视化，本附录提供了所有案例对应的 u-和 v-速度场（图 A1–A8）。这些图展示了第4节讨论的相同定性趋势：所有自适应模型都能有效抑制案例 1 中非物理的能量增长，而在更具挑战性的情形（案例2和案例3）中，混合模型是最有效的方法。

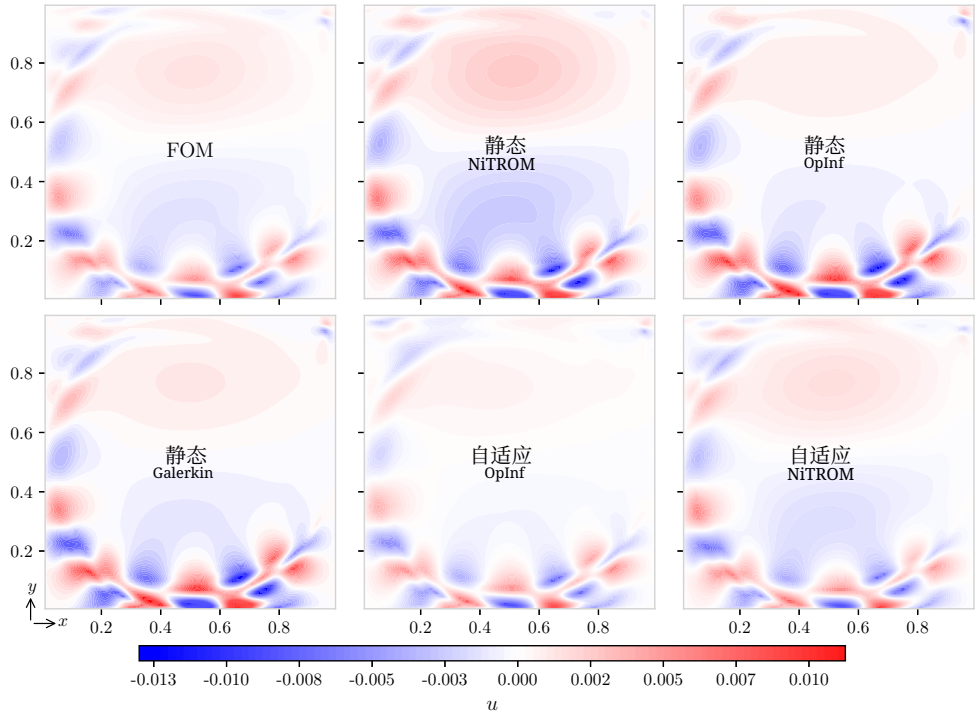


图 A1：案例 1 中在线窗口 ($t = 20$) 结束时的真实值（全阶模型）和预测（降阶模型）u-速度场。所有模型均为二次模型，且维数为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 100$，每次自适应优化预算为 $K = 10$。

声明

资金

A.H. 和 K.D. 获得了美国空军科学研究办公室资助（编号 #FA9550-17-10195）以及负责研究与工程的国防部副部长办公室资助（编号 #N00014-21-1-295）。A.P. 感谢新泽西理工学院纽瓦克工程学院提供的启动经费支持。

作者贡献

A.H. 参与了方法论、软件开发、数值实验以及撰写——初稿和审阅与编辑。**A.P.** 参与了方法论、软件开发、监督以及撰写——审阅与编辑。**K.D.** 参与了概念化、资金筹集、资源、监督以及撰写——审阅与编辑。

结果复现（数据可用性）

复现结果所需的源代码和数据将在向通讯作者提出请求后提供。

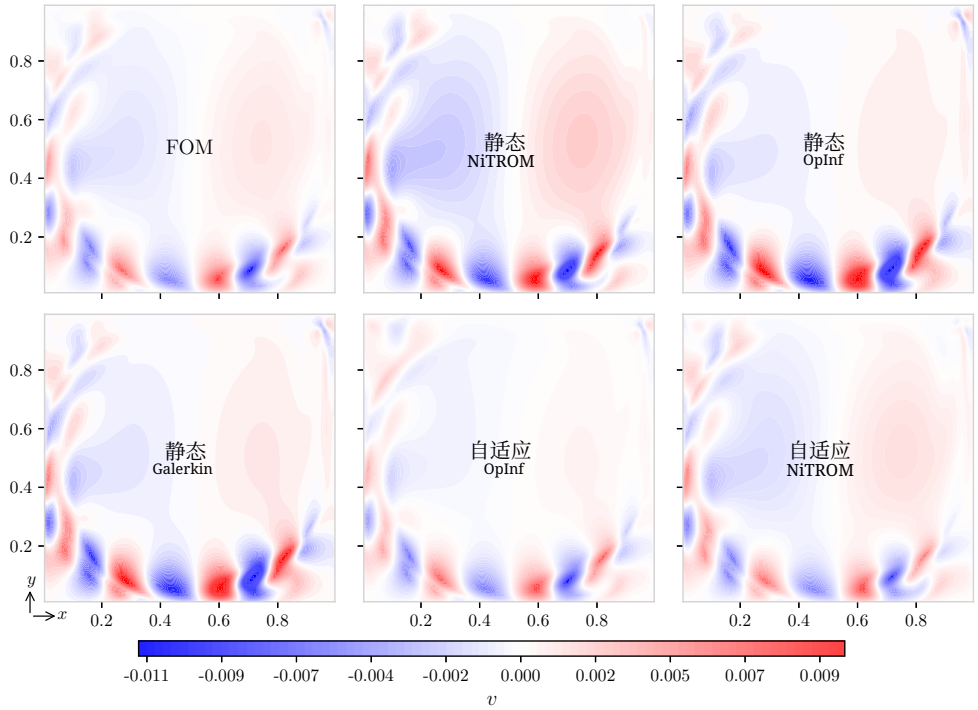


图 A2：案例 1 中在线窗口（$t = 20$）结束时真实值（全阶模型）和预测（降阶模型）的 v-速度场。所有模型均为二次且维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 100$，每次自适应优化预算为 $K = 10$。

利益冲突

通讯作者代表所有作者声明，与本研究相关的利益冲突不存在。

伦理审批和参与知情同意

不适用。

致谢

作者感谢密歇根大学的 Laura Balzano 教授提供的真知灼见。

参考文献

[1] Benner, P., Gugercin, S., Willcox, K.：参数化动力系统的基于投影的模型降阶方法综述。《SIAM Review》**57**(4), 483–531 (2015) https://doi.org/10.1137/130932715

[2] Antoulas, A.C., Beattie, C.A., Güğercin, S.：《模型降阶的插值方法》。工业与应用数学学会，费城，宾夕法尼亚州

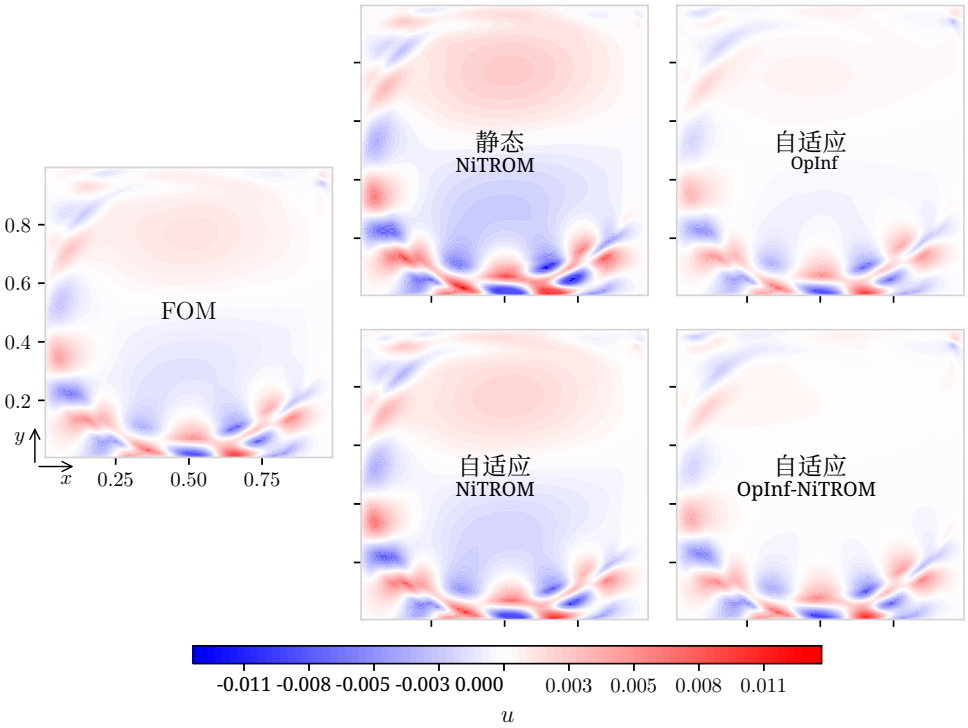


图 A3：案例 1 中在线窗口结束时的真实值（全阶模型）和预测（降阶模型） u-速度场（$t = 20$）。所有模型均为二次模型，且维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 50$，回溯窗口为 $M = 20$，每次自适应优化预算为 $K = 10$。

（2020）。https://doi.org/10.1137/1.9781611976083[3] Rowley, C.W., Dawson, S.T.M.：用于流动分析与控制的模型降阶。《流体力学年度评论》**49**（第49卷，2017年），387–417（2017）https://doi.org/10.1146/annurev-fluid-010816-060042[4] Peherstorfer, B., Willcox, K.：用于非侵入式基于投影的模型降阶的数据驱动算子推断。《计算方法与应用力学和工程》**306**，196–215（2016）https://doi.org/10.1016/j.cma.2016.03.025[5]Padovan, A., Vollmer, B., Bodony, D.J.：通过对投影算子和降阶动力学进行非侵入式优化实现的数据驱动模型降阶。《SIAM 应用动力系统期刊》**23(4)**，3052–3076（2024）https://doi.org/10.1137/24M1628414

[6] Lumley, J.L.：非均匀湍流的结构，166–178（1967）

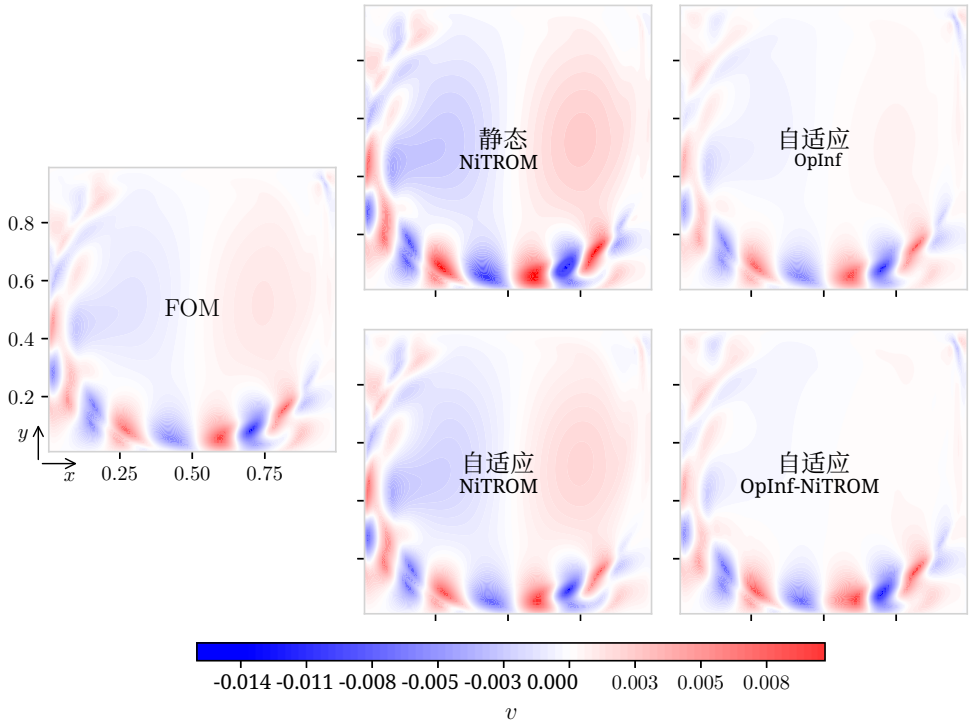


图 A4：案例 1 中在线窗口末端（$t = 20$）的真实值（全阶模型）和预测（降阶模型）v-速度场。所有模型均为二次且维度为 $r = 10$。对于自适应模型，我们设置了自适应窗口 $Z = 50$、回溯窗口 $M = 20$ 和每次自适应优化预算 $K = 10$。

[7]Sirovich, L.：湍流与相干结构的动力学。第一部分：相干结构。Quarterly of Applied Mathematics **45**(3)，561–571 (1987) https://doi.org/10.1090/qam/910462[8] Holmes, P., Lumley, J.L., Berkooz, G.：《湍流、相干结构、动力系统与对称性》。剑桥大学出版社，英国剑桥 (1996)。https://doi.org/10.1017/CBO9780511622700[9] Rowley, C.W., Colonius, T., Murray, R.M.：使用 POD 和 Galerkin 投影对可压缩流进行模型降阶。Physica D：非线性现象 **189**(1-2)，115–129 (2004) https://doi.org/10.1016/j.physd.2003.03.001[10] Barone, M.F., Kalashnikova, I.,Segalman, D.J., Thornquist, H.K.：用于线性化可压缩流的稳定 Galerkin 降阶模型。计算物理学杂志 **228**(6)，1932–1946 (2009) https://doi.org/10.1016/j.jcp.2008.11.015[11] Chomaz, J.-M.：空间发展流中的全局不稳定性：非正规性

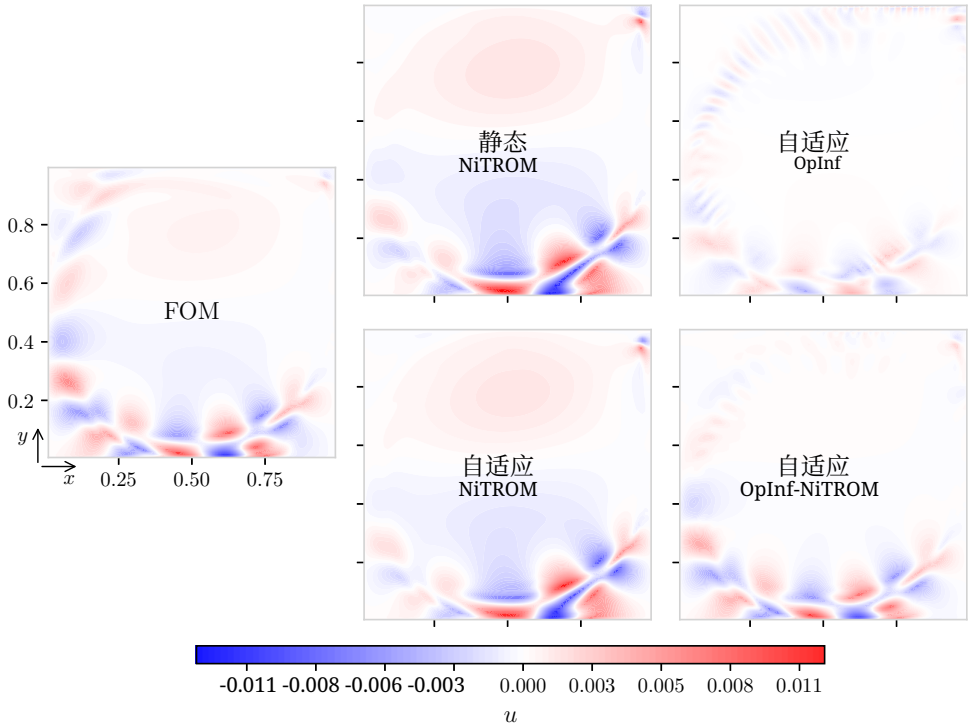


图 A5：案例 2 中在线窗口末尾（$t = 10$）的真实值（全阶模型）和预测（降阶模型）u-速度场。所有模型均为二次模型，维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 10$，每次自适应优化预算为 $K = 10$。

与非线性。《流体力学年度评论》**37**，357–392 (2005) https://doi.org/10.1146/annurev.fluid.37.061903.175810[12] Ahmed, S.E., Pawar, S., San, O., Rasheed, A., Iliescu, T., Noack, B.R.：降阶模型的闭合模型——从第一性原理到机器学习方法的谱系。《流体物理学》**33**(9) (2021) https://doi.org/10.1063/5.0061577[13] Rezaian, E., Huang, C., Duraisamy, K.：对具有轻阻尼脉冲响应的高度刚性系统进行非侵入式平衡变换。Philosophical Transactions of the Royal Society A**380**(2229)，20210202 (2022) https://doi.org/10.1098/rsta.2021.0202[14] Carlberg, K., Bou-Mosleh, C., Farhat, C.：通过最小二乘 Petrov–Galerkin 投影和压缩张量近似实现高效的非线性模型降阶。《国际工程数值方法杂志》**86**(2)，155–181 (2011) https://doi.org/10.1002/nme.3050

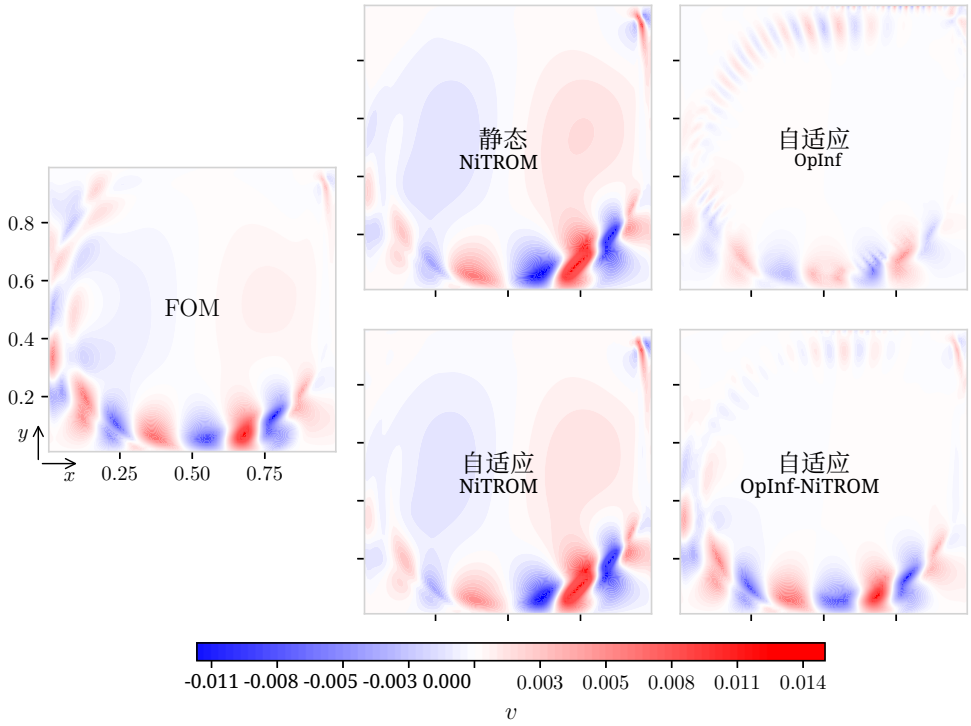


图 A6：案例 2 中在线窗口（$t = 10$）结束时的真实值（全阶模型）和预测值（降阶模型） v-速度场。所有模型均为二次模型，维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为 $M = 10$，每次自适应优化预算为 $K = 10$。

[15] Carlberg, K., Barone, M., Antil：非线性模型降阶中的 Galerkin 投影与最小二乘 Petrov–Galerkin 投影。计算物理学杂志**330**，693–734（2017）https://doi.org/10.1016/j.jcp.2016.10.033[16] Moore, B.：线性系统中的主成分分析：可控性、可观测性与模型降阶。IEEE 自动控制汇刊**26(1)**，17–32（1981）https://doi.org/10.1109/TAC.1981.1102568[17]Willcox, K., Peraire, J.：通过本征正交分解进行平衡模型降阶。AIAA 期刊**40(11)**，2323–2330（2002）https://doi.org/10.2514/2.1570[18] Rowley, C.W.：利用平衡本征正交分解对流体进行模型降阶。国际分岔与混沌杂志**15(03)**，997–1013（2005）https://doi.org/10.1142/S0218127405012429

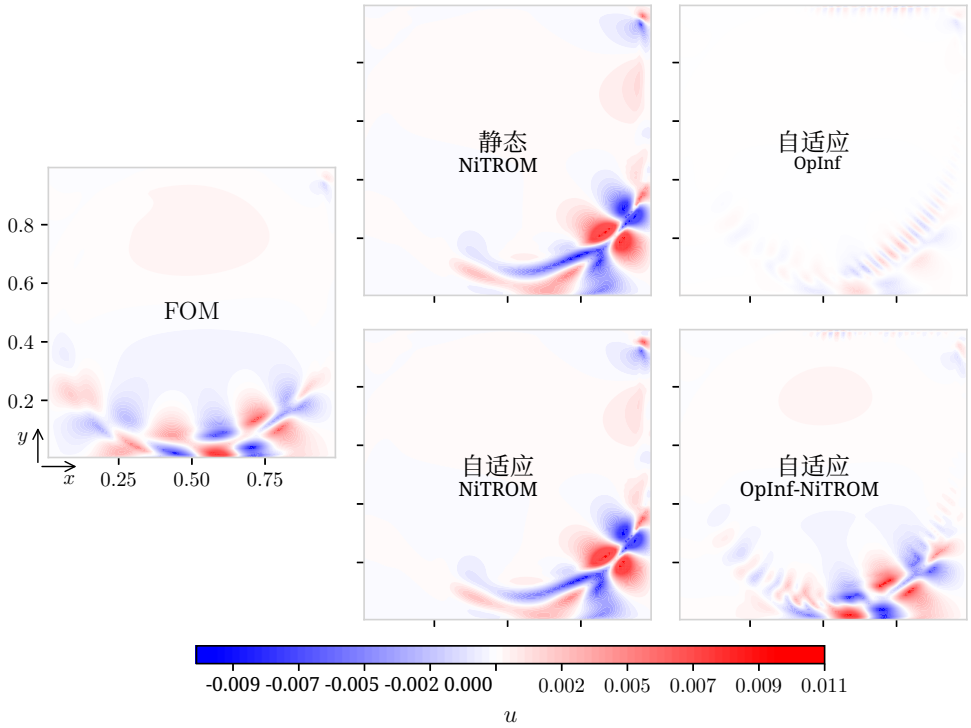


图 A7：案例 3 中在线窗口 ($t = 6.25$) 结束时的真实值（全阶模型）和预测（降阶模型） u-速度场。所有模型均为二次模型，且维数为 $r = 10$。对于自适应模型，我们设置自适应窗口 $Z = 10$、回溯窗口 $M = 10$ 和每次自适应优化预算 $K = 10$。

[19] Otto, S.E., Padovan, A., Rowley, C.W.：利用轨迹优化非线性系统的斜投影。SIAM Journal on Scientific Computing**44**(3)，1681–1702（2022）https://doi.org/10.1137/21M1425815[20] Otto, S.E., Padovan, A., Rowley, C.W.：通过对状态和梯度协方差进行平衡截断，实现非线性系统的模型降阶。SIAM Journal on Scientific Computing**45**(5)，2325–2355（2023）https://doi.org/10.1137/22M1513228[21] Zanardi, I., Padovan, A., Bodony, D.J., Panesi, M.：热化学非平衡气体混合物的 Petrov-Galerkin 模型降阶。计算物理学杂志**533**，113999（2025）https://doi.org/10.1016/j.jcp.2025.113999[22]Schmid, P.J.：数值数据和实验数据的动态模态分解。流体力学杂志**656**，5–28（2010）https://doi.org/10.1017/S0022112010001217[23] Rowley, C.W., Mezic, I., Bagheri, S., Schlatter, P., Henningson, D.S.：谱

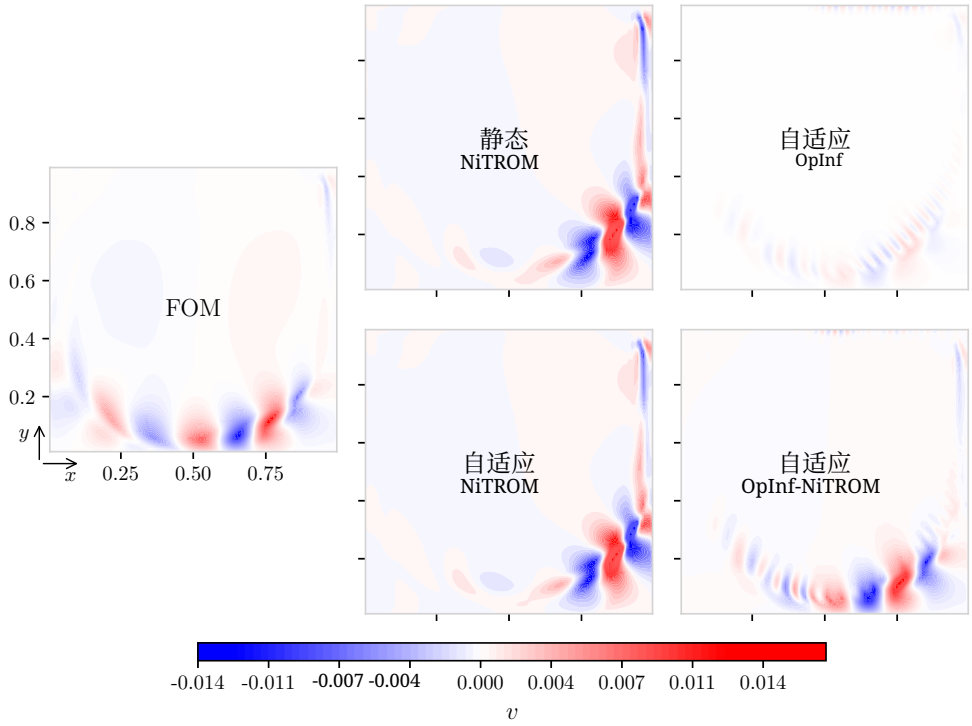


图 A8：案例 3 中在线窗口结束时的真实值（全阶模型）和预测（降阶模型）v速度场（$t = 6.25$）。所有模型均为二次模型，且维度为 $r = 10$。对于自适应模型，自适应窗口为 $Z = 10$，回溯窗口为$M = 10$，每次自适应优化预算为 $K = 10$。

非线性流动分析。《流体力学杂志》**641**, 115–127 (2009) https://doi.org/10.1017/S0022112009992059[24] Hesthaven, J.S., Ubbiali, S.：利用神经网络对非线性问题进行非侵入式降阶建模。《计算物理学杂志》**363**, 55–78 (2018) https://doi.org/10.1016/j.jcp.2018.02.037[25] Bhattacharya, K., Hosseini, B., Kovachki, N.B., Stuart, A.M.：参数化偏微分方程的模型降阶与神经网络。SMAI计算数学期刊**7**, 121–157 (2021) https://doi.org/10.5802/smai-jcm.74[26]Xu, J., Duraisamy, K.：用于时空动力学参数化预测的多层卷积自编码器网络。《计算机方法在应用力学与工程中的应用》**372**, 113379 (2020) https://doi.org/10.1016/j.cma.2020.113379[27] Rezaian, E., Duraisamy, K.：用于预测模型的 data-driven 平衡截断

气动声学响应的降阶。《AIAA 期刊》**61**(10), 4524–4545 (2023) https://doi.org/10.2514/1.J063116

[28] Kramer, B.、Peherstorfer, B.、Willcox, K.E.：利用算子推断从数据中学习非线性降阶模型。《流体力学年度评论》**56**（第 56 卷，2024 年），521–548（2024）https://doi.org/10.1146/annurev-fluid-121021-025220[29] Qian, E.、Kramer, B.、Peherstorfer, B.、Willcox, K.：Lift & learn：面向大规模非线性动力系统的物理信息机器学习。Physica D：非线性现象**406**，132401（2020）https://doi.org/10.1016/j.physd.2020.132401[30] Qian, E.、Kramer, B.、Marques, A.N.、Willcox, K.E.：Transform & Learn：一种数据驱动的非线性模型降阶方法。https://doi.org/10.2514/6.2019- 3707

[31] McQuarrie, S.A.,Huang, C., Willcox, K.E.: 基于正则化算子推断的单喷油器燃烧过程数据驱动降阶模型。《新西兰皇家学会学报》**51**(2), 194–211 (2021) https://doi.org/10.1080/03036758.2020.1863237

[32] Swischuk, R., Kramer, B., Huang, C.,Willcox,K.: 单喷油器燃烧过程的基于物理的降阶模型学习。AIAA 期刊**58**(6), 2658–2672 (2020) https://doi.org/10.2514/1.J058943

[33] Benner, P., Goyal,P., Kramer, B., Peherstorfer, B., Willcox, K.: 非多项式非线性项系统的非侵入式模型降阶算子推断。《计算机方法在应用力学与工程中的应用》**372**, 113433 (2020) https://doi.org/10.1016/j.cma.2020.113433

[34] McQuarrie, S.A., Khodabakhshi,P., Willcox,K.: 基于数据驱动算子推断的参数偏微分方程非侵入式降阶模型。SIAM Journal on Scientific Computing **45**(4), 1917–1946 (2023) https://doi.org/10.1137/21M1452810

[35] Geelen, R., Wright, S., Willcox, K.: 基于二次流形的非侵入式模型降阶算子推断。《计算机方法在应用力学与工程中的应用》**403**,115717 (2023) https://doi.org/10.1016/j.cma.2022.115717[36] Amsallem, D., Farhat, C.: 用于调整降阶模型并应用于气动弹性的插值方法。AIAA 期刊**46**(7), 1803–1813 (2008) https://doi.org/10.2514/1.35374[37] Amsallem, D., Farhat, C.: 一种在线插值线性参数降阶模型的方法。SIAM Journal on Scientific Computing **33**(5), 2169–2198 (2011) https://doi.org/10.1137/100813051

[38] Amsallem, D., Zahr, M.J., Farhat, C.：基于局部降阶基的非线性模型降阶

。《国际工程数值方法杂志》**92**(10), 891–916 (2012) https://doi.org/10.1002/nme.4371[39] Peherstorfer, B., Butnaru, D., Willcox, K., Bungartz, H.-J.：局部离散经验插值法。SIAM Journal on Scientific Computing **36**(1), 168–192 (2014) https://doi.org/10.1137/130924408[40] Peherstorfer, B., Willcox, K.：动态数据驱动降阶模型。Computer Methods in Applied Mechanics and Engineering **291**, 21–41 (2015) https://doi.org/10.1016/j.cma.2015.03.018[41] Peherstorfer, B., Willcox, K.：通过低秩更新实现非线性系统的在线自适应模型降阶。SIAM Journal on Scientific Computing **37**(4), 2123–2150 (2015) https://doi.org/10.1137/140989169[42] Peherstorfer, B.：通过在线自适应基和自适应采样实现输运主导问题的模型降阶。SIAM Journal on Scientific Computing**42**(5), 2803–2836 (2020) https://doi.org/10.1137/19M1257275[43]Singh, R., Uy, W.I.T., Peherstorfer, B.：用于输运主导问题在线自适应模型降阶的前瞻性数据采集策略。Chaos: An Interdisciplinary Journal of Nonlinear Science **33**(11), 113112 (2023) https://doi.org/10.1063/5.0169392[44] Zimmermann, R., Peherstorfer, B., Willcox：几何子空间更新及其在在线自适应非线性模型降阶中的应用。SIAM Journal on Matrix Analysis and Applications**39**(1), 234–261 (2018) https://doi.org/10.1137/17M1123286[45] Balzano, L., Nowak, R., Recht, B.：基于高度不完整信息的子空间在线识别与跟踪。载于：2010年第48届Allerton Conference on Communication, Control, and Computing (Allerton)，第704–711页 (2010)。https://doi.org/10.1109/ALLERTON.2010.5706976[46] Balzano,L., Wright,S.J.：关于GROUSE和增量奇异值分解。载于：2013年第5届IEEE多传感器自适应处理计算进展国际研讨会 (CAMSAP)，第1–4页 (2013)。https://doi.org/10.1109/CAMSAP.2013.6713992[47] Huang, C., Duraisamy, K.：利用自适应采样投影对混沌多尺度问题进行预测性降阶建模。计算物理学杂志**491**, 112356 (2023) https://doi.org/10.1016/j.jcp.2023.112356[48] Mohaghegh, A., Huang, C.：用于对流主导问题自适应模型阶数约化的特征引导采样策略。计算物理学杂志**545**, 114468 (2026) https://doi.org/10.1016/j.jcp.2025.114468[49]Zucatti, V., Zahr, M.J.：一种自适应免训练降阶模型

基于混合快照的对流主导问题。《流体数值方法国际期刊》**96**(2),189–208 (2024) https://doi.org/10.1002/fld.5240

[50] Bai, F., Wang, Y.：一种基于嵌入 GNAT 的混合快照模拟的降阶建模方法。《数学与计算机模拟》**199**, 100–132 (2022)https://doi.org/10.1016/j.matcom.2022.03.006

[51]Feng, L., Fu, G., Wang, Z.：一种用于加速数值模拟的全阶模型/降阶模型混合方法。《科学计算杂志》**89**(3), 61 (2021) https://doi.org/10.1007/s10915-021-01668-9

[52]Yano, M., Huang, T., Zahr, M.J.：一种利用即时模型降阶加速拓扑优化的全局收敛方法。《计算机方法：应用力学与工程》**375**,113635(2021) https://doi.org/10.1016/j.cma.2020.113635

[53] Carlberg, K.：降阶模型的自适应 h-细化。《国际工程数值方法杂志》**102**(5), 1192–1210（2015）https://doi.org/10.1002/nme.4800[54]Etter, P.A.、Carlberg, K.T.：通过向量空间筛分实现降阶模型的在线自适应基底细化与压缩。《计算机方法在应用力学与工程中的应用》**364**，112931（2020）https://doi.org/10.1016/j.cma.2020.112931[55] Patil, P.、Babaee, H.：采用时变基底对具有随机边界条件的偏微分方程进行降阶建模。《SIAM/ASA 不确定性量化杂志》**11**(3)，727–756（2023）https://doi.org/10.1137/21M1468097[56]Jung, K.S.、Lacey, C.E.、Babaee, H.、Chen, J.H.：利用具有时变基底的降阶建模加速化学反应流的高保真模拟。《计算机方法在应用力学与工程中的应用》**437**，117758（2025）https://doi.org/10.1016/j.cma.2025.117758[57]Ramezanian, D.、Nouri, A.G.、Babaee, H.：通过时变流形对惰性与反应性物种进行在线降阶建模。《计算机方法在应用力学与工程中的应用》**382**，113882（2021）https://doi.org/10.1016/j.cma.2021.113882[58] Geelen, R.、Willcox, K.：算子推断框架中的局部非侵入式降阶建模。《皇家学会哲学汇刊 A 辑：数学、物理与工程科学》**380**(2229)，20210206（2022）https://doi.org/10.1098/rsta.2021.0206[59] Aretz, N.、Willcox, K.：用于自适应数据驱动学习降阶模型的嵌套算子推断（2025）。https://doi.org/10.48550/arXiv.2508.11542

[60] Lee, K.、Carlberg, K.T.：使用深度卷积自编码器在非线性流形上对动力系统进行模型降阶。计算物理学杂志**404**，108973（2020）https://doi.org/10.1016/j.jcp.2019.108973[61] Fresca, S.、Dede’，L.、Manzoni, A.：针对非线性时变参数化偏微分方程的全面深度学习降阶建模方法。《科学计算杂志》**87**(2)，61（2021）https://doi.org/10.1007/s10915-021-01462-7[62] Conti, P.、Gobat, G.、Fresca, S.、Manzoni, A.、Frangi, A.：通过自编码器和稀疏非线性动力学识别方法对参数化系统进行降阶建模：周期解的延拓。《计算机方法在应用力学与工程中的应用》**411**，116072（2023）https://doi.org/10.1016/j.cma.2023.116072[63] Otto, S.E.、Macchio, G.R.、Rowley, C.W.：使用约束自编码器学习动力系统的非线性投影以进行降阶建模。Chaos: An Interdisciplinary Journal of Nonlinear Science **33**(11)，113130（2023）https://doi.org/10.1063/5.0169688[64] Brand, M.：含缺失值不确定数据的增量奇异值分解。载于：Heyden, A.、Sparr, G.、Nielsen, M.、Johansen, P.（编）：计算机视觉——ECCV 2002，第707–720页。Springer，柏林、海德堡（2002）[65] Huang, C.、Wentland, C.R.、Duraisamy, K.、Merkle, C.：通过变量变换和保持模型形式的最小二乘投影，对多尺度输运问题进行模型降阶。计算物理学杂志**448**，110742（2022）https://doi.org/10.1016/j.jcp.2021.110742[66] Arnold-Medabalimi, N.、Huang, C.、Duraisamy, K.：燃气轮机模型燃烧室的大涡模拟及基于投影的降阶建模所面临的挑战。《国际喷雾与燃烧动力学杂志》**14**(1-2)，153–175（2022）https://doi.org/10.1177/175682772211006