arXiv:2601.04482v1 [math.NA] 8 Jan 2026

分数阶 Burgers 方程非线性参数化求解器

HAOJUN QIN*、ZHIWEI GAO*、JINYE SHEN*、和 GEORGE E. KARNIADAKIS†

摘要。 由于非局部性与激波形成非线性动力学的综合影响，分数阶 Burgers 方程给经典数值方法带来了重大挑战。特别是，谱方法、有限差分方法或间断 Galerkin 方法等线性近似框架往往会出现 Gibbs 型振荡，或需要精心调节稳定化机制，而这些机制在输运占主导和长时间积分区域中的效果会减弱。本文提出一种用于求解分数阶 Burgers 方程的时间序列非线性参数化（STNP），涵盖具有分数阶拉普拉斯算子或非局部非线性通量的模型。解由非线性参数化试探函数表示，并在每个时间步通过正则化最小二乘形式，将控制动力学投影到参数流形的切空间上，从而获得参数演化。这一方法产生适定且稳定的时间推进格式，既保持因果性，又避免全时间域优化。我们对所得投影动力学进行理论分析，包括稳定性估计，以及一个 $aposteriori\ error$界限，该界限将总误差明确分解为初始条件拟合、投影残差和分数阶算子离散所造成的误差。我们的分析阐明了正则化的稳定作用，并量化其与非局部离散误差之间的相互影响。针对两类分数阶 Burgers 模型的数值实验表明，STNP 能够无振荡地分辨激波，并准确捕捉长时间动力学。该方法的表现始终优于结合谱消失黏性的高阶谱格式，同时所需自由度显著更少，也无需临时稳定化。

1. 引言。 分数阶偏微分方程（FPDEs）在刻画各类系统的非局部行为方面展现出巨大潜力，例如多孔介质中的反常扩散 [1] 以及黏弹性材料中的反常扩散 [11]。尤其是，作为整数阶 Burgers 方程的推广，空间分数阶 Burgers 方程因其在气体过驱爆轰、非线性马尔可夫过程的混沌传播等领域的应用而备受关注[30, 23]。

具体而言，经典 Burgers 方程主要有两种推广方式。首先，采用分数阶拉普拉斯算子，即 $(-\Delta)^{\frac{\alpha}{2}}$，替代标准扩散项，从而得到

$$u_t + \frac{1}{2}(u^2)_x = -\varepsilon(-\Delta)^{\frac{\alpha}{2}}u, \quad 1 < \alpha < 2, \tag{1.1}$$

这称为含分数阶拉普拉斯算子的分数阶 Burgers 方程（FBEFL）。在此，非局部算子与线性项相关，因此充当超扩散项，使解得到平滑。对于 $\alpha > 1$，文献 [2, 15] 已证明正则解的存在性与唯一性。另一种推广称为具有非局部非线性的分数阶 Burgers 方程（FBENN），它通过用两个左侧 Caputo 导数的复合算子替代非线性项而得到，因此是经典 Burgers 方程的一种非局部推广，即

$$u_t + \frac{1}{2}{}_a^C D^{\beta}({}_a^C D^{1-\beta}u)^2 = \varepsilon u_{xx}, \tag{1.2}$$

*Haojun Qin、Zhiwei Gao 和 Jinye Shen 贡献相同，均为共同第一作者。

†布朗大学应用数学系。电子邮箱：haojun qin@brown.edu（H. Qin）、zhiwei gao@brown.edu（Z. Gao）、jinye shen@brown.edu（J. Shen）、george_karniadakis@brown.edu（G. E. Karniadakis）。通讯作者：G. E. Karniadakis。

本手稿仅供审阅。

其中 $0 \le \beta \le 1$。通过对 β 作不同选择，式（1.2）可以还原为经典 Burgers 方程或非线性扩散方程。更重要的是，采用 Hopf–Cole 变换的分数阶推广后，可以显式地得到解[24], 即

$$u(x,t) = -2\varepsilon\, {}_{a}^{C}D^{\beta}\left[\log\left(\frac{1}{\sqrt{4\pi\varepsilon t}}\int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4\varepsilon t}} e^{-\frac{1}{2\varepsilon}\, {}_{a}I^{\beta}u_0(y)}\,\mathrm{d}y\right)\right]. \tag{1.3}$$

相关记号及其他一些性质将在第 2 节中介绍。

对于经典 Burgers方程，即使初始数据光滑，也可能形成激波 [18] ，进而在标准的线性参数化方案中引发吉布斯现象 [17] 和伪振荡，例如谱展开、有限差分（FD）以及间断Galerkin（DG）离散。在整数阶情形下，通常采用局部激波捕捉技术来缓解这些问题，例如迎风通量 [8] 和 ENO/WENO重构 [31], ，通过添加局部耗散或进行非线性重构，有效稳定底层线性表示。

对于空间分数阶 Burgers 方程，由于分数阶算子固有的非局部性，线性参数化面临的困难更加突出。分数阶拉普拉斯算子和 Caputo型导数会引入全局耦合，使得每个线性自由度的演化都受到整个定义域上解的影响。因此，线性离散化往往会导致稠密且全局耦合的算子。此外，局部稳定化机制无法直接迁移过来，因为激波附近产生的数值扰动可能会非局部传播，而为抑制振荡添加耗散又很容易降低精度。

若干方法可视为用于分数阶模型的稳定化线性参数化。Droniou [9]在严格的收敛框架下分析了一大类 FD 格式。Xu和Hesthaven [36] 将LDG方法扩展至分数阶对流–扩散方程，在引入分数阶算子的同时保留了 DG 型线性表示。最近，Mao和Karniadakis [23]将谱消失黏性（SVV）引入谱和谱元LDG近似，以抑制 Gibbs 型振荡，同时保持高阶精度 [22]。尽管取得了成功，这些稳定化线性参数化在输运主导区域仍可能计算开销较大。要解析尖锐的层结构，通常需要大量自由度；而且，为了在控制振荡与避免激波涂抹和精度损失之间取得平衡，往往需要仔细调节稳定化参数，尤其是在长时间积分中。

受线性参数化这些局限性的启发，我们提出一种按时间顺序推进的非线性参数化（STNP），作为求解两个分数阶 Burgers 方程 FBEFL 和 FBENN 的高效求解器。STNP 通过参数化试探形式 $\hat{u}(x,\ t)\ =\ \Phi(x;q(t))$ 表示解，并通过对偏微分方程动力学进行切空间投影来演化参数。在每个时间步，$\dot{q}(t)$ 均通过一个正则化最小二乘问题求得，随后进行标准时间积分。由此得到适定的更新，并提供一种稳定机制，以控制离散非局部项所导致的误差增长。我们进一步证明了稳定性结果，并推导出相应投影动力学的后验误差估计。数值实验表明，STNP 能够准确且高效地求解 FBEFL 和 FBENN，同时解析激波且不产生伪振荡。

从数值分析的角度来看，STNP 可解释为一种非线性 Galerkin 型方法，其中解在低维非线性流形上演化，并通过切空间投影施加控制动力学。

本手稿仅供审阅。

所提出框架的一个关键特点是引入了正则化切空间投影，它具有双重作用：既能稳定投影动力学，又能支持严谨的 后验误差分析。总之，本文的新贡献如下：

- 介绍 STNP 作为一种非线性求解器（而非学习框架）。
- 正则化切空间投影的稳定性界。
- 后验误差分解（初始拟合 / 投影 / 离散化）。
- 关于具有激波的分数阶 Burgers 方程的数值证据。

本文其余部分安排如下。第 2 节回顾分数阶导数的基本定义以及本文采用的数值离散。第 3 节给出 FBEFL 和 FBENN 的正则性结果。第 4 节介绍 STNP 算法及其实现细节。第 5 节论证正则化的必要性，并推导稳定性和 后验误差估计。第 6 节给出数值实验，以验证本文的分析并展示所提方法的性能。

2. 背景。 本节介绍分数阶微积分的相关记号，包括黎曼–刘维尔导数、卡普托导数和分数阶拉普拉斯算子。随后，我们介绍相应的数值离散。

2.1. 分数阶微积分。首先，阶数为 β 的左黎曼–刘维尔积分定义为

$$({}_aI_x^{\beta}u)(x) = \int_a^x \omega_\beta(x-s)u(s)\,\mathrm{d}s, \tag{2.1}$$

其中，对于 $\beta > 0$，$\omega_\beta(x) = \frac{x^{\beta-1}}{\Gamma(\beta)}$，而 $\Gamma(x)$ 表示标准伽马函数。阶数为 β 的右黎曼–刘维尔积分定义为

$$({}_xI_a^{\beta}u)(x) = \int_x^a \omega_\beta(s-x)u(s)\,\mathrm{d}s. \tag{2.2}$$

阶数为 $\alpha \in (0, 1)$ 的左卡普托导数定义为：

$$({}_a^C D_x^{\alpha}u)(x) = ({}_aI_x^{1-\alpha}u')(x) = \frac{1}{\Gamma(1-\alpha)}\int_a^x \frac{u'(s)}{(x-s)^{\alpha}}\,\mathrm{d}s. \tag{2.3}$$

类似地，阶数为 $\alpha \in (0, 1)$ 的左黎曼–刘维尔导数定义为：

$$({}_a^{RL}D_x^{\alpha}u)(x) = ({}_aI_x^{1-\alpha}u)'(x) = \frac{1}{\Gamma(1-\alpha)}\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x \frac{u(s)}{(x-s)^{\alpha}}\,\mathrm{d}s. \tag{2.4}$$

阶数为 $\alpha \in (0, 1)$ 的右黎曼–刘维尔导数为

$$({}_x^{RL}D_a^{\alpha}u)(x) = ({}_xI_a^{1-\alpha}u)'(x) = \frac{1}{\Gamma(1-\alpha)}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\right)\int_x^a \frac{u(s)}{(s-x)^{\alpha}}\,\mathrm{d}s. \tag{2.5}$$

基于这些定义，我们可以将分数阶拉普拉斯算子定义为两个黎曼–刘维尔导数之和，即：

$$(-\Delta)^{\frac{\alpha}{2}}u(x) = \frac{1}{2\cos\left(\frac{\alpha\pi}{2}\right)}\left({}_{-\infty}^{RL}D_x^{\alpha}u(x) + {}_x^{RL}D_{+\infty}^{\alpha}u(x)\right). \tag{2.6}$$

本文稿仅供审阅。

值得一提的是，分数阶拉普拉斯算子有几种等价的定义方式 [19]。本文仅采用基于黎曼–刘维尔导数的定义，因为这种定义更便于构造有限差分格式。当 $\alpha = 2$ 时，分数阶拉普拉斯算子与经典拉普拉斯算子一致。要对分数阶拉普拉斯算子进行离散化，只需对黎曼–刘维尔导数进行离散化。本文采用移位的 Gr¨unwald–Letnikov 格式对黎曼–刘维尔导数进行离散化，该格式将在下一小节中介绍。

2.2. 分数阶导数的离散化。由于分数阶导数具有非局部性，并且依赖于整个定义域上的取值，因此无法使用标准的局部有限差分格式进行离散化。对于定义在 $[a, b]$ 上的卡普托导数 ${}_a^C D^\alpha$ with$\alpha \in (0,1)$，最常用的方法之一是基于分段线性插值 [33] 的 L1 格式。对于固定的 N，令 $h = \frac{b-a}{N}$。网格点为 $x_k = a + kh$，其中 $k = 0, 1, 2, \ldots, N$。在 x_n 处的卡普托导数近似为

(2.7)
$$ {}_a^C D^\alpha u(x)\big|_{x=x_n} \approx \frac{h^{-\alpha}}{\Gamma(2-\alpha)} \left[a_0^{(\alpha)} u(x_n) - \sum_{k=1}^{n-1} \big(a_{n-k-1}^{(\alpha)} - a_{n-k}^{(\alpha)}\big) u(x_k) - a_{n-1}^{(\alpha)} u(x_0) \right], $$

with

$$ a_l^{(\alpha)} = (l+1)^{1-\alpha} - l^{1-\alpha}. \tag{2.8} $$

L1 格式的截断误差为 $\mathcal{O}(h^{2-\alpha})$ [33]。

对于分数阶拉普拉斯算子，由于它可以表示为两个黎曼–刘维尔导数之和，我们采用移位的 Gr¨unwald–Letnikov 公式对其进行近似。首先，定义

$$ A_{h,p}^\alpha u(x) = h^{-\alpha} \sum_{k=0}^{\infty} g_k^{(\alpha)}\, u\big(x - (k-p)h\big), \tag{2.9} $$

其中，p 表示平移，且

$$ g_k^{(\alpha)} = (-1)^k \binom{\alpha}{k}, \quad \binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}. \tag{2.10} $$

则一阶近似 [34] 为

$$ {}_{-\infty}^{RL} D_x^\alpha u(x) = A_{h,p}^\alpha u(x) + \mathcal{O}(h). \tag{2.11} $$

类似地，对于 $p \neq q$，二阶近似为 [34]

$$ {}_{-\infty}^{RL} D_x^\alpha u(x) = \lambda_1 A_{h,p}^\alpha u(x) + \lambda_2 A_{h,q}^\alpha u(x) + \mathcal{O}(h^2), \tag{2.12} $$

其中

$$ \lambda_1 = \frac{\alpha - 2q}{2(p-q)}, \qquad \lambda_2 = \frac{2p - \alpha}{2(p-q)}. \tag{2.13} $$

关于其他高阶格式，请读者参阅 [13, 39]。我们强调，与局部有限差分格式不同，上述分数阶离散化会产生卷积型算子。实际应用中，$A_{h,\ p}^\alpha$ 中的无穷级数会在

本文仅供审阅。

有限网格上截断，由此得到的离散算子可视为作用于 u 网格值的稠密矩阵。权重 $\{g_k^{(\alpha)}\}$ 按代数规律衰减，并可通过递推关系 $g_0^{(\alpha)} = 1$ 和 $g_k^{(\alpha)} = \left(1 - \frac{\alpha+1}{k}\right) g_{k-1}^{(\alpha)}$ 高效生成，因此无需显式构造稠密矩阵即可实现离散化。这凸显了分数阶导数的关键特性——非局部性，并促使我们在数值格式中采用移位的 Gr¨unwald–Letnikov 公式。

3. 分数阶 Burgers 方程。本节回顾 FBEFL 和 FBENN 的定义，并讨论它们的理论性质，包括正则性、渐近行为以及与其他方程的联系。

3.1. FBEFL。带分数阶拉普拉斯算子的 Burgers 方程定义如下：

$$\begin{cases} u_t + \frac{1}{2}(u^2)_x = -\varepsilon(-\Delta)^{\frac{\alpha}{2}} u & \text{in } \mathbb{R} \times \mathbb{R}^+, \\ u(x,0) = u_0(x) & \text{in } \mathbb{R}, \end{cases} \tag{3.1}$$

其中 $1 < \alpha < 2$ 和 $\varepsilon > 0$ 是扩散率。已证明，对于周期初值，$1 < \alpha \leq 2$ 的情形是次临界的，此时实解析解整体存在。其证明与经典情形 $\alpha = 2$类似。这里我们引用[15]中的结果：

THEOREM 3.1. *[15]* 假设 $1 < \alpha \leq 2$，且初值 $u_0(x) \in H^s(\mathbb{S}^1)$，其中 $s > \max(\frac{3}{2} - 2\alpha, 0)$。那么，方程 *(3.1)* 存在唯一整体解 $u(x,t)$，满足 u 属于 $C([0,\infty)\ H^s(\mathbb{S}^1))$。此外，$u(x,t)$ 在 $\mathbb{R}$ 中对 $t > 0$是实解析的。对于更一般的初值 $u_0 \in H^1(\mathbb{R})$，*Biler* [4] 已证明，当 $\frac{3}{2} < \alpha \leq 2$时，解仍属于 $H^1(\mathbb{R})$。[4] 中的结果如下：

THEOREM 3.2. *[4]* 假设 $\frac{3}{2} < \alpha \leq 2$、$T > 0$ 以及初始数据 $u_0 \in H^1(\mathbb{R})$，方程 *(3.1)* 的解是唯一的，即

$$u \in L^\infty((0,T); H^1(\mathbb{R})) \cap L^2((0,T); H^{1+\frac{\alpha}{2}}(\mathbb{R})), \tag{3.2}$$

对每个 $T > 0$ 均成立。并且解满足以下衰减性质：

$$\lim_{t\to\infty} \|u(\cdot,t)\|_{L^\infty} = 0. \tag{3.3}$$

证明采用抛物型方程中能量估计的标准步骤，并结合 Sobolev 嵌入定理。

3.2. FBENN. 具有非局部非线性项的分数阶 Burgers 方程定义如下：

$$\begin{cases} u_t + \frac{1}{2}\,{}_a^C D^\beta ({}_a^C D^{1-\beta} u)^2 = \varepsilon u_{xx} & \text{in } \mathbb{R} \times \mathbb{R}^+, \\ u(x,0) = u_0(x) & \text{in } \mathbb{R}, \end{cases} \tag{3.4}$$

其中 $0 \leq \beta \leq 1$。这里，${}_a^C D^\alpha$ 表示关于空间变量 x的 α 阶 Caputo 分数阶导数，参数 $\beta \in [0,1]$ 则在局部与非局部对流机制之间进行插值。具体而言，非线性通量通过分数阶积分—导数对 ${}_a^C D^\beta({}_a^C D^{1-\beta} u)^2$,构造，其中 ${}_a^C D^{1-\beta}$ 在平方运算前对 u 进行非局部平均，而 ${}_a^C D^\beta$ 则进一步对所得通量进行非局部微分。其极限情形可还原为

本文仅供审稿使用。

熟知的模型：当 $\beta = 1$时，对流项化为经典的 Burgers 非线性 $\frac{1}{2}(u^2)_x$；而当 $\beta = 0$ 时，它则变为$\frac{1}{2}(u_x)^2$,，对应于 Hamilton–Jacobi 型非线性。全文中，$\varepsilon > 0$ 表示黏性系数，而扩散项 εu_{xx} 则提供正则化，并确保初值问题适定。

对于上文讨论的两个 Burgers 型方程，直接采用基于线性函数参数化的经典求解器（例如谱方法或间断 Galerkin（DG）方法）并非易事。从解的正则性来看，Burgers 动力学可能会形成陡峭梯度和激波状过渡层；在这种情况下，若不采用谨慎的稳定化和限制过程，高阶多项式逼近容易产生 Gibbs 型振荡。此外，分数阶项引入的非局部性进一步加大了离散化的复杂性和计算成本。鉴于这些挑战，我们采用非线性参数化——具体而言，使用神经网络——来逼近解，相关内容将在下一节介绍。尽管该框架适用于一般的非线性参数化，包括张量分解 [26]和径向基展开 [35], ，但我们着重采用神经网络，因为它们具有强大的表达能力，且易于求导。

4. 方法论。 本节首先介绍非线性参数化逼近的基础知识。随后，我们采用时间序列技术推导参数的时间演化，最后说明如何将所得框架应用于上文介绍的两个 Burgers 型方程。

4.1. 非线性参数化逼近。本部分考虑依赖于参数的非线性参数化。令 q 表示参数，$\Phi(\cdot; q)$ 表示参数化逼近。令 $\Phi'(\cdot; q)$ 表示 Φ 对 q 的导数；Φ 为非线性等价于 $\Phi''(\cdot; q) \neq 0$。非线性逼近的例子包括深度神经网络 [12],基于张量的分解[25, 26] 以及多高斯逼近 [29]。

现在，我们介绍如何将非线性参数化用作时变偏微分方程的数值求解器。核心思想是用参数化试探解 $\hat{u}(x, t) = \Phi(x; q(t))$ 表示解，并随时间演化参数 $q(t)$，使试探解近似满足控制方程。我们不在整个时空域上最小化一个全局时间目标函数，而是在局部时间上施加偏微分方程约束：在每个时间层，选择对 $q(t)$ 的更新，使其在最小二乘意义下尽可能匹配偏微分方程右端项。由此得到一个逐步执行的过程：反复针对 $\dot{q}(t)$ 求解一个小型回归问题，然后使用标准时间积分器推进 q。

4.2. 时间序列训练。与物理信息神经网络（PINNs）及其变体中采用的全局时间训练相比， [27, 28], 本节将讨论时间序列训练方法。该方法将时间变量分开处理，以保持因果性并减少参数数量[16, 5]。

考虑以下时变偏微分方程，

$$\frac{\partial}{\partial t}u(x,t) = f(t,x,u,Du), \quad \text{with } u(x,0) = u_0(x), \tag{4.1}$$

其中 $x \in \Omega \subset \mathbb{R}$, $u : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ 是关于时间和空间的标量函数，而 f是依赖于 t, x, u 的微分算子，更一般地，它可以是任意类型的

本稿仅供审阅。

u 的分数阶导数。接下来，我们寻求一种非线性参数化逼近

$$u(x,t) \approx \hat{u}(x,t) = \Phi(x; q(t)), \tag{4.2}$$

其中 $q(t) \in \mathcal{Q} \subseteq \mathbb{R}^{N_P}$ 表示时变参数。

现在，我们开始介绍时间序列训练的公式。根据链式法则，数值解的时间演化定义为

$$\frac{\partial \hat{u}}{\partial t} = \frac{\partial \hat{u}}{\partial q}\dot{q}. \tag{4.3}$$

在每个时间步，通过求解以下最小二乘问题，得到参数的时间导数，

$$\frac{\partial q}{\partial t} = \operatorname{argmin}_{\gamma \in \mathcal{Q}} \mathcal{J}(\gamma), \tag{4.4}$$

其中代价函数为

$$\mathcal{J}(\gamma) = \left\| \frac{\partial \hat{u}}{\partial q}\gamma - f(t, \cdot, \hat{u}, D\hat{u}) \right\|_{L^2}^2. \tag{4.5}$$

一阶最优性条件为

$$\nabla_\gamma \mathcal{J}(\gamma^*) = \left(\int_\Omega \frac{\partial \hat{u}}{\partial q} \otimes \frac{\partial \hat{u}}{\partial q} \, \mathrm{d}x \right) \gamma^* - \left(\int_\Omega \frac{\partial \hat{u}}{\partial q} \otimes f(t, x, \hat{u}, D\hat{u}) \, \mathrm{d}x \right) = 0, \tag{4.6}$$

其中 $\otimes$ 表示关于参数的梯度的外积。实际应用中，我们选取一些配点 $\{x_1, x_2, ..., x_{N_C}\} \subseteq \Omega$。这些配点将连续最小二乘问题 (4.5) 转化为离散最小二乘近似。值得注意的是，使用足够多的配点，可以使离散近似达到任意精度。第 6 节所示的数值实验将进一步验证这一观察结果。现在我们回到离散最小二乘系统的构造。在选定的配点上，通过求解以下问题来近似最优解 γ^*

$$J^T J \hat{\gamma}^* = J^T \vec{f}, \tag{4.7}$$

其中， J 是梯度矩阵， f 是在这些点处求得的向量场，即 $\vec{}$ ，

$$(J)_{ij} = \frac{\partial \hat{u}^i}{\partial q_j}, \quad (\vec{f})_i = f(t, x_i, \hat{u}^i, D\hat{u}^i), \tag{4.8}$$

并且， $i \in [N_C]$ 是配点的索引，而 $j \in [N_P]$ 是参数的索引。

一般来说，梯度矩阵 J 可能具有任意小的奇异值，这会导致数值稳定性问题，甚至使最小二乘问题 (4.7) 存在多个解。因此，有必要引入额外的正则化项来构造一个适定问题，即

$$\mathcal{J}(\gamma) = \left\| \frac{\partial \hat{u}}{\partial q}\gamma - f(t, \cdot, \hat{u}, D\hat{u}) \right\|_{L^2}^2 + \lambda^2 \|\gamma\|_{\mathcal{Q}}^2, \tag{4.9}$$

本文仅供审阅。

其中， λ 是正则化参数， $\|\cdot\|_{\mathcal{Q}}$ 表示 $\mathcal{Q}$上的范数，该范数选为 $\mathbb{R}^{N_P}$上的标准 ℓ^2 内积。相应的一阶最优性条件的离散形式为

$$(J^T J + \lambda^2 I)\gamma^* = J^T \vec{f}, \tag{4.10}$$

其中 I 表示单位矩阵。求解 (4.10) 可采用直接求逆方法或迭代方法。从几何角度看，这一过程可解释为：相对于由 L^2 范数和正则化诱导的混合内积，将向量场投影到参数流形的切空间上。除了数值稳定化作用外，正则化在理论分析中也发挥着根本作用。这一点将在第 5 节中详细讨论。在概述了按时间顺序训练框架的公式和解释后，我们将通过说明初始条件、边界条件和时间积分策略的处理方式，完成对该算法的描述。

初始条件通过求解优化问题或直接插值来满足。本文采用深度神经网络作为非线性参数化。初始参数 q_0 通过优化以下损失函数获得，即

$$q_0 = \arg\min_{q\in\mathcal{Q}} \sum_{i=1}^{N_u} |\hat{u}(x_i) - u_0(x_i)|^2, \tag{4.11}$$

使用 Adam [14] 和 LBFGS [20]等优化器。

边界条件的处理遵循基于神经网络的方法中的标准做法，具体实现细节请参阅相关文献。对于周期边界条件，将输入坐标映射为三角特征（如 sin(x) 和 cos(x)），从而通过构造保证周期性；例如，参见 [37]。对于狄利克雷边界条件，我们采用常见方法，将网络输出乘以到边界的距离函数，或引入适当的插值多项式，以精确满足给定的边界值，参见 [3, 21]。随后，可采用不同的时间积分格式进行时间推进。本文主要采用显式龙格–库塔方法，例如 RK3 和 RK45 格式。

对于本文研究的两个分数阶 Burgers 方程，通过指定算子 f 并相应地组装式 (4.8) 中的向量 f，即可应用上述过程。具体而言，在每个龙格–库塔阶段和每个配点 x_i处，我们通过结合以下两部分来计算 $f(t,\ x_i,\ \hat{u}^i,\ D\hat{u}^i)$：(i) 通过自动微分从参数代理模型 $\hat{u} = \Phi(\cdot; q)$ 得到的整数阶导数（例如 u_x， u_{xx}）；以及 (ii) 使用本文前文介绍的有限差分格式进行数值计算的分数阶／非局部算子。

具体而言，对于分数阶 Burgers 方程中出现的分数阶拉普拉斯算子项，我们采用第 2.2 节所述的移位 Grünwald–Letnikov 近似（参见（2.11）–（2.12））。对于（3.4）中的非局部非线性项，其中涉及复合 ${}_a^C D^{\beta}\big({}_a^C D^{1-\beta} u\big)^2$，我们先计算内层 Caputo 算子 ${}_a^C D^{1-\beta} u$，然后对平方结果应用外层 Caputo 算子 ${}_a^C D^{\beta}$，所采用的仍是前文介绍的相同 Caputo 离散格式（参见（2.7））。通过这种方式，非局部贡献被纳入 f，而整体的按时间顺序最小二乘框架（4.10）–（4.11）保持不变。

本手稿仅供审阅。

4.3. 算法。本节基于我们介绍的非线性参数化公式和不同的离散化方案，详细介绍分数阶 Burgers 方程算法的步骤；见 4.1。

算法 4.1： 分数阶 Burgers 方程非线性参数求解器

输入： 初始条件 $u_0(x)$、配点 $\{x_i\}_{i=1}^{N_u}$、初始时间步长 Δt，终止时间 T、正则化参数 λ。

输出： 通过参数集 $q(t)$得到的近似解 $\hat{u}(x,t)$。

通过插值或最小化 (4.11) 来初始化参数化 $q(0) \in \mathcal{Q}$。；

当 $t < T$ **时循环**

  根据 (2.7)、(2.11)、(2.12) 计算空间分数阶导数，以及通过自动微分计算整数阶导数。；

  根据 (4.8) 构造梯度矩阵 J 和向量场 $\vec{f}$。；

  求解正则化最小二乘问题 (4.10)。；

  采用龙格-库塔方法进行时间推进，并得到新的时间步 Δt.；

  $t \leftarrow t + \Delta t$.；

结束

返回 $\hat{u}(x,\ t)$ 表示为 $\Phi\big(x;\quad q(t)\big)$.；

5. 理论分析。本节将分析非线性参数求解器的稳定性，并通过跟踪每个最小二乘子问题的缺陷，推导后验误差估计。最后，我们将考察分数阶导数有限差分近似的阶数如何影响结果。首先，我们将介绍分析中所用的一些基本记号。

由于向量场 f 包含分数阶导数，我们用有限差分格式对其进行近似，因此引入离散向量场 f^h，其中 h 表示网格大小。相应的截断误差为

$$R^h(t,x) = f(t,x,\hat{u},D\hat{u}) - f^h(t,x,\hat{u},D\hat{u}). \tag{5.1}$$

我们的目标是量化这种离散化误差以及将动力学限制在参数流形上所产生的近似误差的累积影响。为此，令

$$\mathcal{M} = \{\Phi(\cdot;q) \mid q \in \mathcal{Q}\}$$

表示由参数化诱导的流形，并令 $T_q\mathcal{M}$ 表示其在 q处的切空间。在每个时刻 t，我们通过正则化的 L^2 意义下的 $T_{q(t)}\mathcal{M}$ 中的一个元素来近似向量场 $f^h(t,\ \cdot,\ \hat{u},\ D\hat{u})$。具体而言，我们为有序对 $(u,\ q) \in L^2 \times \mathcal{Q}$ 赋予混合内积

$$((u_1,q_1),(u_2,q_2)) := (u_1,u_2)_{L^2} + \lambda^2 (q_1,q_2)_{\mathcal{Q}}, \tag{5.2}$$

其中 $\lambda > 0$ 是一个正则化参数。诱导投影 P_q^λ到 $T_q\mathcal{M}$ 上通过以下 Tikhonov 正则化最小二乘问题实现：

$$\dot{q}(t) = \arg\min_{\gamma \in \mathcal{Q}} \left\| \frac{\partial \hat{u}}{\partial q}(t)\,\gamma - f^h(t,\cdot,\hat{u},D\hat{u}) \right\|_{L^2}^2 + \lambda^2 \|\gamma\|_{\mathcal{Q}}^2. \tag{5.3}$$

采用此处的 $\dot{q}(t)$ 选择，投影向量场为

$$P_{q(t)}^\lambda f^h(t,x,\hat{u},D\hat{u}) := \frac{\partial \hat{u}}{\partial q}(t,x)\,\dot{q}(t) \in T_{q(t)}\mathcal{M}. \tag{5.4}$$

本稿仅供评审使用。

离散向量场与其切空间近似之间的不匹配由缺陷函数衡量

$$d^{\lambda,h}(t,x) := f^h(t,x,\hat{u},D\hat{u}) - P^{\lambda}_{q(t)} f^h(t,x,\hat{u},D\hat{u}), \tag{5.5}$$

并将相应的最小二乘残差定义为

$$\delta^{\lambda,h}(t) := \left(\left\| \frac{\partial \hat{u}}{\partial q}(t)\,\dot{q}(t) - f^h(t,\cdot,\hat{u},D\hat{u}) \right\|^2_{L^2} + \lambda^2 \|\dot{q}(t)\|^2_{\mathcal{Q}} \right)^{\frac{1}{2}}. \tag{5.6}$$

由于 $\delta^{\lambda,h}(t)$包含缺陷的平方 L^2-范数，我们立即得到

$$\|d^{\lambda,h}(t,\cdot)\|_{L^2} \le \delta^{\lambda,h}(t). \tag{5.7}$$

因此，跟踪 $\delta^{\lambda,h}(t)$可提供一个可计算的事后指标，用以衡量参数动力学对离散化向量场 f^h 的跟随程度。

最后，根据构造，数值解沿着流形按照投影动力学演化

$$\hat{u}_t(t,x) = P^{\lambda}_{q(t)} f^h(t,x,\hat{u},D\hat{u}) = f^h(t,x,\hat{u},D\hat{u}) - d^{\lambda,h}(t,x). \tag{5.8}$$

一旦指定投影算子 P^{λ}_q，动力学 (5.8) 就完全确定。然而在实践中，雅可比矩阵 $\partial \hat{u}/\partial q$ 可能条件不佳，使得未正则化的 L^2 投影容易受到离散化和建模误差的影响。为解决这一问题，我们引入正则化投影 P^{λ}_q，并分析其稳定作用。下一小节将通过比较 L^2 投影与正则化 L^2 投影，对这一点进行形式化说明。

5.1. 正则化。本节计算正则化投影的算子范数，并阐明正则化在所提出的算法稳定性中的作用。

在每个时间步，我们求解最小二乘问题 (5.3)，其解可写为

$$\dot{q} = \left(\left(\frac{\partial \hat{u}}{\partial q} \right)^* \left(\frac{\partial \hat{u}}{\partial q} \right) + \lambda^2 I \right)^{-1} \left(\frac{\partial \hat{u}}{\partial q} \right)^* f^h, \tag{5.9}$$

其中 $\frac{\partial \hat{u}}{\partial q}\colon \mathcal{Q} \to L^2$表示参数化的雅可比矩阵，而 $\left(\frac{\partial \hat{u}}{\partial q}\right)^*$是其$L^2$-伴随。相应的投影向量场为

$$P^{\lambda}_q f^h = \left(\frac{\partial \hat{u}}{\partial q} \right) \left(\left(\frac{\partial \hat{u}}{\partial q} \right)^* \left(\frac{\partial \hat{u}}{\partial q} \right) + \lambda^2 I \right)^{-1} \left(\frac{\partial \hat{u}}{\partial q} \right)^* f^h. \tag{5.10}$$

令 $\{\sigma_i\}_{i=1}^{N_P}$ 表示 $\frac{\partial \hat{u}}{\partial q}$ 的奇异值，等价地说，它们是 $\left(\frac{\partial \hat{u}}{\partial q}\right)^* \left(\frac{\partial \hat{u}}{\partial q}\right)$ 的非零特征值的平方根。进一步地，分别用 $\sigma_{\min}$和 $\sigma_{\max}$表示最小和最大的奇异值。那么，正则化投影的谱满足

$$\operatorname{spec}(P^{\lambda}_q) = \left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}, \quad i = 1, \dots, N_P \right\}, \tag{5.11}$$

本稿件仅供审阅。

其在 L^2 上诱导的算子范数满足如下界

$$\|P_q^\lambda\| \le \frac{\sigma_{\max}^2}{\sigma_{\max}^2+\lambda^2} < 1. \tag{5.12}$$

因此，正则化投影算子 P_q^λ在 L^2上是压缩映射，这一性质将在稳定性估计中用到，即定理 5.1。

尽管正则化投影 P_q^λ在 L^2上具有良好的压缩性质，但这并不意味着参数演化不会表现出强烈的敏感性，下面我们将对此进行量化。如第 4.9 节所述，最小二乘子问题可能严重病态，对应于 $\sigma_{\min}\approx 0$。式 (5.9) 中从L^2 到参数空间 $\mathcal{Q}$ 的诱导算子范数满足

$$\left\| \left(\left(\frac{\partial \hat{u}}{\partial q}\right)^* \left(\frac{\partial \hat{u}}{\partial q}\right) + \lambda^2 I \right)^{-1} \left(\frac{\partial \hat{u}}{\partial q}\right)^* \right\| = \max_{1\le i\le N_P} \frac{\sigma_i}{\sigma_i^2+\lambda^2} \le \frac{1}{2\lambda}. \tag{5.13}$$

相比之下，在未正则化情形 $\lambda = 0$下，相应的算子满足

$$\left\| \left(\left(\frac{\partial \hat{u}}{\partial q}\right)^* \left(\frac{\partial \hat{u}}{\partial q}\right) \right)^{-1} \left(\frac{\partial \hat{u}}{\partial q}\right)^* \right\| = \frac{1}{\sigma_{\min}}, \tag{5.14}$$

当 $\sigma_{\min}\to 0$时，它将变得无界。

因此，向量场中的任何扰动——尤其是由分数阶导数的有限差分近似所产生的离散化误差——都会导致

$$\|\delta \dot{q}\| \lesssim \frac{1}{\lambda}\,\|\delta f^h\|, \tag{5.15}$$

而在没有正则化的情况下，放大因子则按 $1/\sigma_{\min}$的量级增长。因此，病态性不会影响正则化投影P_q^λ的 L^2-算子范数，而是通过映射 $f^h \mapsto \dot{q}$产生影响，使参数演化对向量场中的微小扰动高度敏感。正则化通过一致地控制这种敏感性并防止离散化误差在参数动力学中被大幅放大，从而稳定了算法。

5.2. 稳定性。 由于 $(-\Delta)^{\frac{\alpha}{2}}$ 的耗散效应，FBEFL 解的 L^2 能量在齐次狄利克雷边界条件下将单调衰减。我们的数值解将近似遵循耗散律。然而，由于投影是基于正则化的 L^2 投影，而不是 L^2 投影，这将引入一些取决于 h 和 λ的能量漂移效应。

THEOREM 5.1（L_2 稳定性估计）。 设 $\hat{u}(x,t) = \Phi(x;\ q(t))$ 为 *FBEFL* (3.1) 的参数化数值解。如果 $\hat{u}(x,t) \in H_0^{\alpha/2}$，则

$$\|\hat{u}(\cdot,t)\|_{L^2} \le e^{-\lambda_\alpha^* \varepsilon t}\|\hat{u}(\cdot,0)\|_{L^2} + \int_0^t e^{-\lambda_\alpha^*\varepsilon(t-s)} \left(\delta^{\lambda,h}(s) + \|R^h(s,\cdot)\|_{L^2}\right) \mathrm{d}s,$$

其中，$\lambda_\alpha^* > 0$ 是具有齐次狄利克雷边界条件的 $(-\Delta)^{\frac{\alpha}{2}}$ 的最小特征值。

证明。 在 FBEFL 中，设

$$f(\hat{u}, D\hat{u}) = -\varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - \hat{u}\hat{u}_x \tag{5.16}$$

本文仅供审阅使用。

为相应的向量场。根据演化方程 (5.8)，我们有

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|\hat{u}(\cdot,t)\|_{L^2}^2 &= \langle \hat{u}(\cdot,t), \partial_t \hat{u}(\cdot,t)\rangle \\ &= \langle \hat{u}, P_{q(t)}^{\lambda} f^h\rangle \\ &= \langle \hat{u}, P_{q(t)}^{\lambda}(f - R^h)\rangle \\ &= \left\langle \hat{u}, -\varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - \hat{u}\hat{u}_x - d^{\lambda,h}\right\rangle - \langle \hat{u}, P_{q(t)}^{\lambda} R^h\rangle. \end{aligned} \tag{5.17}$$

利用（5.12），并注意到在齐次狄利克雷边界条件下，$\langle \hat{u},\ \hat{u}\hat{u}_x\rangle = 0$，可得

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|\hat{u}(\cdot,t)\|_{L^2}^2 \le -\lambda_\alpha^* \varepsilon \|\hat{u}(\cdot,t)\|_{L^2}^2 + \|\hat{u}(\cdot,t)\|_{L^2}\|d^{\lambda,h}(t,\cdot)\|_{L^2} \\ + \|\hat{u}(\cdot,t)\|_{L^2}\|R^h(t,\cdot)\|_{L^2},\end{aligned} \tag{5.18}$$

其中我们还利用了（$-\Delta$）$^{\frac{\alpha}{2}}$ [6]的谱性质，即 $-$

$$\langle \hat{u}, (-\Delta)^{\frac{\alpha}{2}}\hat{u}\rangle \ge \lambda_\alpha^* \|\hat{u}\|_{L^2}^2. \tag{5.19}$$

最后，由（5.7）可得

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|\hat{u}(\cdot,t)\|_{L^2}^2 \le -\lambda_\alpha^* \varepsilon \|\hat{u}(\cdot,t)\|_{L^2}^2 + \|\hat{u}(\cdot,t)\|_{L^2}\,\delta^{\lambda,h}(t) \\ + \|\hat{u}(\cdot,t)\|_{L^2}\|R^h(t,\cdot)\|_{L^2},\end{aligned} \tag{5.20}$$

利用格朗沃尔不等式，证明完成，

$$\|\hat{u}(\cdot,t)\|_{L^2} \le e^{-\lambda_\alpha^*\varepsilon t}\|\hat{u}(\cdot,0)\|_{L^2} + \int_0^t e^{-\lambda_\alpha^*\varepsilon(t-s)}\left(\delta^{\lambda,h}(s) + \|R^h(s,\cdot)\|_{L^2}\right)\mathrm{d}s. \quad \square \tag{5.21}$$

右端的第一项直接源于（$-\Delta$）$^{\alpha/2}$的耗散效应，而第二项则计及能量漂移，该漂移由以下因素驱动

- 可计算投影缺陷 $\delta^{\lambda,\ h}$，
- 空间有限差分截断误差 $\|R^h\|_{L^2}$。

因此，考虑极限情形 $\lambda \to 0$ 和 $h \to 0$颇有启发意义。在这种情况下，空间扰动项消失，正则化投影退化为精确L^2 投影。因此，可以得到严格能量耗散律，这为我们的格式提供了基线一致性检验，并表明当 λ 足够小时且网格足够精细时，$\lambda > 0$ 所观察到的漂移应当很小。然而，要恢复精确的能量耗散恒等式，仅依赖投影的正交性是不够的。还必须确保数值解本身属于投影算子的值域。这需要对参数化 Φ作出结构性假设，我们将在下文给出。

Assumption 5.1（参数线性结构）。存在一个参数子集 $a \in \mathbb{R}^m$（例如最后一层权重），满足 $m < N_p$，以及其余参数 θ，使得

$$\Phi(x; a, \theta) = \sum_{j=1}^{m} a_j \psi_j(x;\theta) \tag{5.22}$$

本文仅供审稿使用。

对于每个 θ，都有 $\psi(;\theta) \in L^2$。那么，对于 $q = (a,\theta)$ 和所有 t，数值解$\hat{}u(\cdot,t) = \Phi(\cdot;q(t))$ 属于切空间，即，

$$\hat{u}(,t) \in T_q\mathcal{M}. \tag{5.23}$$

因此，对于正交（未正则化）投影 $P_q := P_q^{\lambda=0}$，有 $P_{q(t)}\hat{}u(t) = \hat{}u(t)$。

REMARK 5.1. 事实上，只要最后一层是线性的，假设 *5.1* 就适用于广泛的参数化方式*[38],*，包括前馈深度神经网络。因此，在我们的定理中作出这一假设是合理的。

这一假设确保，在未正则化的连续极限下，即 $\lambda = 0$和 $h \to 0$，投影算子对数值解的作用是平凡的。因此，当用 $\hat{}u$ 测试时，投影动力学与完整向量场一致。这一性质对于恢复精确的能量耗散恒等式至关重要，正如下一个定理所示。

THEOREM 5.2（精确耗散 L^2 在未正则化极限下）。假设参数化方式 $\Phi(\cdot;q)$ 满足假设 *5.1*。设 $\hat{}u(x,t) = \Phi(x;\, q(t))$ 为 *FBEFL (3.1)* 的参数化数值解。考虑极限情形 $\lambda = 0$ 和 $h \to 0$，以及投影动力学

$$\hat{u}_t = P_{q(t)} f(t,x,\hat{u},D\hat{u}), \quad with\, f(t,x,\hat{u},D\hat{u}) = -\varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - \hat{u}\hat{u}_x, \tag{5.24}$$

其中 P_q 是到 $T_q\mathcal{M}$的 L^2 投影。若 $\hat{}u(x,t) \in H_0^{\alpha/2}$，则

$$\|\hat{u}(\cdot,t)\|_{L^2} \le e^{-\lambda_\alpha^* \varepsilon t}\|\hat{u}(\cdot,0)\|_{L^2},$$

证明。证明采用标准的能量论证；例如参见 [38]。由于 P_q 是 L^2 投影，因此它是自伴的，并满足

$$\langle P_q u, v\rangle_{L^2} = \langle u, P_q v\rangle_{L^2}. \tag{5.25}$$

由投影动力学 (5.24)，可得

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|\hat{u}(\cdot,t)\|_{L^2}^2 &= \langle \hat{u}(\cdot,t), \partial_t\hat{u}(\cdot,t)\rangle \\ &= \langle \hat{u}, P_{q(t)}[-\varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - \hat{u}\hat{u}_x]\rangle \end{aligned} \tag{5.26}$$

由 (5.23) 和自伴性 (5.25)，可得

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|\hat{u}(\cdot,t)\|_{L^2}^2 &= \langle P_{q(t)}\hat{u}, -\varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - \hat{u}\hat{u}_x\rangle \\ &= \langle \hat{u}, -\varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - \hat{u}\hat{u}_x\rangle, \end{aligned} \tag{5.27}$$

利用分部积分和 (5.19)，可得所需结果。 ∎

在进行误差分析之前，我们先总结框架中两个关键量的作用：最小二乘残差 $\delta^{\lambda,h}$ 和有限差分截断误差 R^h。

在上述稳定性估计中，这些量作为扰动源，用于量化数值流与精确耗散动力学之间的偏差。重要的是，$\delta^{\lambda,h}$ 可在每一步根据迭代值完全计算得出，而随着 h 逐步细化，R^h 则具有已知的收敛行为。因此，它们为诊断能量漂移和潜在的稳定性丧失提供了实用的后验指标。

本手稿仅供审阅。

在下一小节中，我们将证明，同一对量（$\delta^{\lambda,h}$, R^h）也会出现在数值误差的演化中，并由此得到一个后验 L^2-误差界。这带来一个统一的视角：这些量不仅能解释稳定性分析中的能量漂移，也能控制数值误差的增长。

5.3. 误差估计。在我们的算法中，向量场分两步近似。首先，我们使用有限差分格式近似分数阶导数，记为 f^h，它近似原始的 f。然后，我们将 f^h 投影到切空间 $T_q\mathcal{M}$ 上。在误差估计中，我们考虑这两个步骤引入的误差。基于这一方法，下面的定理给出了数值解的误差估计。

设 FBEFL (3.1) 的精确解满足

$$u_t = f(t,x,u,Du), \tag{5.28}$$

且相应的数值解满足 $\hat{u}_t = P^{\lambda}_{q(t)} f_h$，其中 $f_h = f + R^h$ 和 $\|d^{\lambda,h}(t)\|_{L^2} \le \delta^{\lambda,h}(t)$ 如 (5.7) 和 (5.8) 所示。

Theorem 5.3（误差估计）。假设 $\exists M>0$，使得 $\|\hat{u}(\cdot,t)\|_{W^{1,\infty}} \le M$ 且 $\|u(\cdot,t)\|_{W^{1,\infty}} \le M$。若 $\hat{u}(x,t) \in H_0^{\alpha/2}$，则有

$$\|u(\cdot,t)-\hat{u}(\cdot,t)\|_{L^2} \le e^{(\frac{3}{2}M-\lambda_\alpha^*\varepsilon)t}\|u(\cdot,0)-\hat{u}(\cdot,0)\|_{L^2} + \int_0^t e^{(\frac{3}{2}M-\lambda_\alpha^*\varepsilon)(t-s)}\delta^{\lambda,h}(s)\,\mathrm{d}s + \int_0^t e^{(\frac{3}{2}M-\lambda_\alpha^*\varepsilon)(t-s)}\|R^h(s,\cdot)\|_{L^2}\,\mathrm{d}s.$$

证明。由 (5.8) 和 (5.28)，可得

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u-\hat{u}\|_{L^2}^2 &= \langle u-\hat{u}, \partial_t u - \partial_t \hat{u}\rangle \\
&= \langle u-\hat{u}, f(t,\cdot,u,Du) - P^{\lambda}_{q(t)} f^h(t,\cdot,\hat{u},D\hat{u}^h)\rangle \\
&= \langle u-\hat{u}, f(t,\cdot,u,Du) - f(t,\cdot,\hat{u},D\hat{u})\rangle \\
&\quad + \langle u-\hat{u}, f(t,\cdot,\hat{u},D\hat{u}) - f^h(t,\cdot,\hat{u},D\hat{u})\rangle \\
&\quad + \langle u-\hat{u}, f^h(t,\cdot,\hat{u},D\hat{u}) - P^{\lambda}_{q(t)} f^h(t,\cdot,\hat{u},D\hat{u}^h)\rangle,
\end{aligned} \tag{5.29}$$

其中，在最后一个等式中我们使用了加减技巧。我们将分别估计这三项。第一项通过标准的分部积分进行估计，

$$\begin{aligned}
\langle u-\hat{u}, f(t,\cdot,u,Du) - f(t,\cdot,\hat{u},D\hat{u})\rangle &= \langle \hat{u}-u, \hat{u}\hat{u}_x + \varepsilon(-\Delta)^{\frac{\alpha}{2}}\hat{u} - uu_x - \varepsilon(-\Delta)^{\frac{\alpha}{2}}u\rangle \\
&\le -\lambda_\alpha^*\varepsilon\|u-\hat{u}\|_{L^2}^2 + |\langle \hat{u}-u,\ \hat{u}\hat{u}_x - uu_x\rangle|.
\end{aligned} \tag{5.30}$$

根据假设，$\|\hat{u}(\cdot,t)\|_{W^{1,\infty}}$ 和 $\|u(\cdot,t)\|_{W^{1,\infty}}$ 均由 M 一致有界，

$$\begin{aligned}
|\langle \hat{u}-u, \hat{u}\hat{u}_x - uu_x\rangle| &\le |\langle \hat{u}-u, \hat{u}\hat{u}_x - \hat{u}u_x\rangle| + |\langle \hat{u}-u, \hat{u}u_x - uu_x\rangle| \\
&= |\langle \frac{1}{2}(\hat{u}-u)^2, \hat{u}_x\rangle| + |\langle (\hat{u}-u)^2, u_x\rangle| \\
&\le (\frac{1}{2}\|\hat{u}_x\|_{L^\infty} + \|u_x\|_{L^\infty})\|\hat{u}-u\|_{L^2}^2 \\
&\le \frac{3}{2}M\|\hat{u}-u\|_{L^2}^2.
\end{aligned} \tag{5.31}$$

本稿件仅供审阅使用。

根据定义 (5.5)、(5.1)、不等式 (5.7) 和柯西-施瓦茨不等式，第二项和第三项可估计为

$$|\langle u-\hat{u}, f(t,\cdot,\hat{u},D\hat{u})-f^h(t,\cdot,\hat{u},D\hat{u})\rangle| \le \|u-\hat{u}\|_{L^2}\|R^h\|_{L^2}, \tag{5.32}$$

and

$$|\langle u-\hat{u}, f^h(t,\cdot,\hat{u},D\hat{u})-P^{\lambda}_{q(t)}f^h(t,\cdot,\hat{u},D\hat{u}^h)\rangle| \le \|u-\hat{u}\|_{L^2}\delta^{\lambda,h}(t). \tag{5.33}$$

结合 (5.30)、(5.31)、(5.32)、(5.33)，并应用 Gr¨onwall 不等式，可得最终的误差估计，

$$\begin{aligned}\|u(\cdot,t)-\hat{u}(\cdot,t)\|_{L^2} \le e^{(\frac{3}{2}M-\lambda^*_\alpha\varepsilon)t}\|u(\cdot,0)-\hat{u}(\cdot,0)\|_{L^2} + \int_0^t e^{(\frac{3}{2}M-\lambda^*_\alpha\varepsilon)(t-s)}\delta^{\lambda,h}(s)\,\mathrm{d}s \\ + \int_0^t e^{(\frac{3}{2}M-\lambda^*_\alpha\varepsilon)(t-s)}\|R^h(s,\cdot)\|_{L^2}\,\mathrm{d}s. \end{aligned} \tag{5.34}$$

□

上述定理表明，误差可分解为三个部分：

- 初始条件拟合误差 $\|u(\cdot,0)-\hat{u}(\cdot,0)\|_{L^2}$，可通过增加参数数量或优化迭代次数来减小。
- 最小二乘残差 $\delta^{\lambda,h}(s)$，取决于切空间$T_q\mathcal{M}$对右端项的逼近程度。如果不必担心稳定性，减小正则化项也有助于降低这一误差。
- 有限差分近似的截断误差 $R^h(t,x)$，随 $h\to 0$ 呈现出明显的收敛特性，可通过采用更高阶格式或使用更精细的网格来减小。

下一节中的实验表明，一旦初始条件拟合足够准确，且有限差分误差占主导地位，提高有限差分格式的阶数和细化网格都有助于降低总体误差。然而，一旦最小二乘残差占主导地位，误差便会趋于平稳，模型的拟合能力也会达到饱和。

6. 数值实验。本节中，我们使用所提出的方法求解两类分数阶 Burgers 方程，以展示该方法的能力和准确性。为凸显该方法相较于传统有限差分法和谱方法的优势，我们还对比了采用不同方法所得的结果。此外，为验证前述理论分析，我们进一步考察数值解误差与有限差分近似截断误差之间的关系，以确认理论结果的正确性。

在以下所有测试中，我们采用多层感知机进行非线性逼近，并使用双曲正切函数作为激活函数。初始参数采用 Xavier 初始化方案进行初始化。通过使用 L-BFGS 优化器最小化 (4.11)，使其满足初始条件。具体的边界条件通过网络结构 [10] 的设计加以施加。我们将在具体示例中说明该方法的实现。所有计算均在一台独立的 NVIDIA GPU RTX 6000 上完成。

为了比较不同方法的性能，我们将使用定义如下的相对L^2 误差

$$e=\sqrt{\frac{\sum_{i=1}^{N_t}|u(t_i,x_i)-\hat{u}(t_i,x_i)|^2}{\sum_{i=1}^{N_t}|u(t_i,x_i)|^2}}, \tag{6.1}$$

其中，N_t 是空间和时间定义域内测试点的数量，$(t_i,\ x_i)$表示第 i个测试点。

本文仅供审阅。

6.1. FBEFL。在本部分中，我们使用 STNP 求解 (3.1)。我们进行了两次数值实验。在第一个实验中，我们将 STNP 与高分辨率激波捕捉参考求解器（WENO–Roe）以及基准中心差分离散方法进行比较，以验证 STNP 的正确性。在第二个实验中，我们考察 STNP 的收敛行为与分数阶拉普拉斯算子有限差分近似阶数之间的关系。

我们首先考虑 $[-1,1]$ 上的 (3.1)，其初始条件和边界条件为

$$u_0(x) = -\sin(\pi x), \qquad u(\pm 1, t) = 0. \tag{6.2}$$

为突出 STNP 抑制伪振荡的能力，我们选取较小的黏性系数 $\varepsilon = 0.01$，并将分数阶阶数固定为 $\alpha = 1.6$。对于 STNP，我们采用架构为 $[1, 10, 10, 1]$, 的全连接神经网络，即包含两个宽度为 10 的隐藏层，共有 141 个可训练参数，并在空间中使用 300 个均匀分布的配点。作为参考求解器，我们按照 [7] 实现 WENO–Roe格式，使用 201 个自由度。两种方法均采用三阶强稳定保持龙格–库塔积分器（SSP-RK3） [32]，时间步长为 $\Delta t = 10^{-3}$。我们绘制 $t = 0.5$ 和 $t = 1.0$ 时刻的预测解，并在图 1 中将 STNP 与中心差分格式和 WENO–Roe格式进行比较。

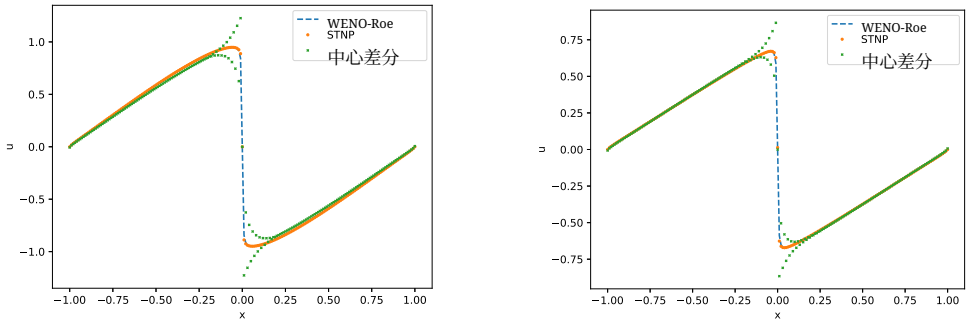


FIG. 1。 *Comparison of numerical solutions computed by STNP and central-difference schemes. Both STNP and WENO–Roe suppress spurious oscillations whereas the central-difference scheme exhibits pronounced oscillations near the shock.*

WENO–Roe,,

在这些设置下，中心差分基准方法由于数值耗散不足，在激波附近产生了明显的数值振荡。相比之下，STNP 生成了无振荡解，其结果与 WENO–Roe 高度吻合，同时避免了 WENO类方法所需的非线性重构过程。这一比较证实了 STNP 的准确性及其在低黏性区间中的稳健性。

接下来，我们评估 STNP 在网格加密以及分数阶拉普拉斯算子的有限差分近似阶数提高时的收敛特性。为将截断误差与其他误差来源分离开来，我们采用一个构造的光滑解，

$$u(x,t) = e^{-t}x^3(1-x)^3, \qquad x \in [0,1],$$

并设置 $\varepsilon = 1.0$，使解保持足够的正则性。相应的受迫问题形式如下

$$u_t + \frac{1}{2}(u^2)_x = -\varepsilon(-\Delta)^{\frac{\alpha}{2}}u + f(x,t), \tag{6.3}$$

本文稿仅供审阅使用。

其中，强迫项 $f(x,t)$ 的选取使得精确解由上面的表达式给出：

$$
\begin{aligned}
f(x,t) = -e^{-t}(x^3(1-x)^3) + e^{-2t}(3-6x)x^5(1-x)^5 \\
+ \frac{\varepsilon}{2\cos(\frac{\alpha\pi}{2})}\Big(\frac{\Gamma(4)}{\Gamma(4-\alpha)}\left(x^{3-\alpha}+(1-x)^{3-\alpha}\right) \\
- \frac{3\Gamma(5)}{\Gamma(5-\alpha)}\left(x^{4-\alpha}+(1-x)^{4-\alpha}\right) \\
+ \frac{3\Gamma(6)}{\Gamma(6-\alpha)}\left(x^{5-\alpha}+(1-x)^{5-\alpha}\right) \\
- \frac{\Gamma(7)}{\Gamma(7-\alpha)}\left(x^{6-\alpha}+(1-x)^{6-\alpha}\right)\Big).
\end{aligned}
$$

我们使用 RK3 和 $\Delta t = 10^{-3}$。我们采用不同阶数的有限差分公式近似分数阶拉普拉斯算子，并通过增加配点的数量来加密网格

$$N \in \{5, 10, 20, 40, 50, 100, 200, 400, 800, 1000\}.$$

为避免网络近似误差主导结果，我们将网络宽度固定为 10，并将深度增加到 3、5 和 7。这一选择降低了初始条件失配和残差误差，使得观测到的收敛性主要由有限差分近似的截断误差驱动。相对 L^2 误差见图 2。

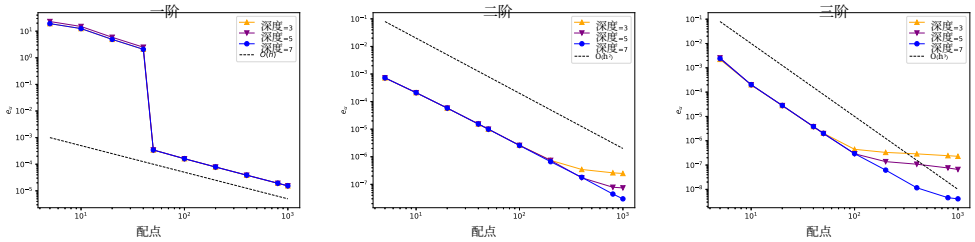


FIG. 2。 *Relative L^2 error versus the number of collocation points N using first-order (left), second-order (middle) and third-order (right) finite-difference approximations of the fractional , , 、、 Laplacian for networks of fixed width* 10 *and depths* 3 5 *and* 7.

正如预期的那样，当 N 较小时，总体误差由截断误差主导，不同网络深度对应的曲线几乎重合。随着 N 增加，截断误差减小，更丰富的网络参数化所带来的优势开始显现：更深的网络降低了初始条件失配和残差误差，从而减小了最终误差，并展现出每种有限差分阶数所预期的收敛趋势。

6.2. FBENN。 本节中，我们使用 STNP 求解（3.4）。初始条件为

$$u_0(x) = -\sin(\pi x), \tag{6.4}$$

边界条件由根据 (1.3) 得到的精确解给定。扩散系数固定为 $\varepsilon = \frac{1}{150\pi}$。我们进行了三组测试。我们

本文仅供审阅。

首先，将 STNP 与采用谱消失黏性（SVV） [23] 的谱方法进行基准比较，以验证其精度和激波捕捉性能。随后，通过比较固定步长格式（RK3）与自适应格式（RK45），研究时间积分的影响。最后，考察 STNP 的长期行为及其对随机初始化的敏感性。

我们首先使用 STNP 和谱-SVV方法求解 (3.4)。对于 STNP，我们采用架构为 $[1, 10, 10, 1]$ 的全连接网络。对于谱方法，我们使用 128 阶雅可比多项式。两种求解器均采用三阶 Runge– Kutta 方法（RK3），时间步为 $\Delta t = 10^{-3}$。图 3 比较了 $\beta = 0.6$ 和 $\beta = 0.8$分别对应的 $t = 0.5$ 和 $t = 1.0$ 处的数值解与精确解。谱方法需要相对较大的 SVV 来抑制吉布斯现象，但这不可避免地会模糊间断，并降低激波附近的精度。相比之下，STNP 能更清晰地解析激波位置，并更贴近精确解，从而显著提高精度。

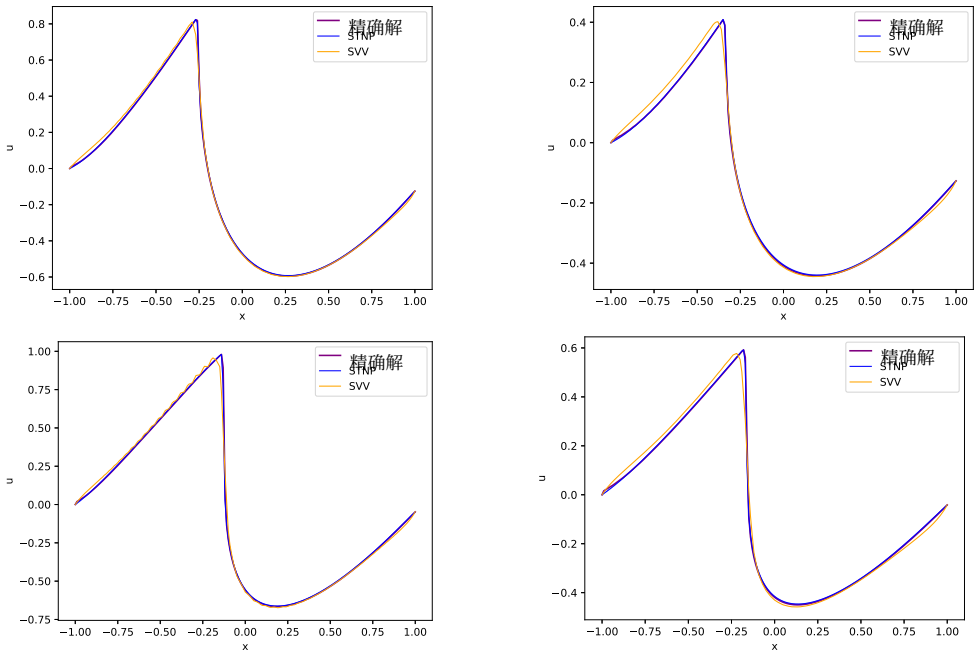


FIG. 3 *Comparison of spectral-SVV STNP and exact solutions. The first column shows* $t = 0.5$*, the second column shows* $t = 1.0$*; the top row corresponds to* $\beta = 0.6$ *and the bottom row to* $\beta = 0.8$*. Both methods suppress spurious oscillations but STNP achieves higher accuracy near the shock.*

接下来，我们评估时间积分对 STNP 动力学的影响。我们考虑 $\beta = 0.8$，并将 RK3 与自适应龙格–库塔方法（RK45）进行比较。该方法根据嵌入式局部截断误差估计调整步长。如图4中间面板所示，RK45 的相对 L^2误差略小于固定步长的 RK3。更重要的是，RK45 能提供更鲁棒的轨迹，因为当解变化迅速时（本例中发生在 $t \approx 0.3$ 附近），它会自动缩短步长。图4右侧面板展示了每个积分步骤所达到的累计时间，并说明了这种自适应步长缩减。值得注意的是，RK45 平均使用约 451 个时间步，而 RK3 始终使用 1000 个，这表明自适应时间步进能够提升我们框架的鲁棒性和

本手稿仅供审阅。

效率。

最后，我们研究 STNP 的长期性能。我们将系统积分至 $t = 5$，并使用十个相互独立的随机初始化重复实验。所得的 1σ 置信带很窄，表明对初始化的敏感性可以忽略不计。图4左图展示了 STNP 和谱黏性值方法的相对L^2 误差随时间的演变，表明在整个时间范围内，STNP 始终保持更高的精度。

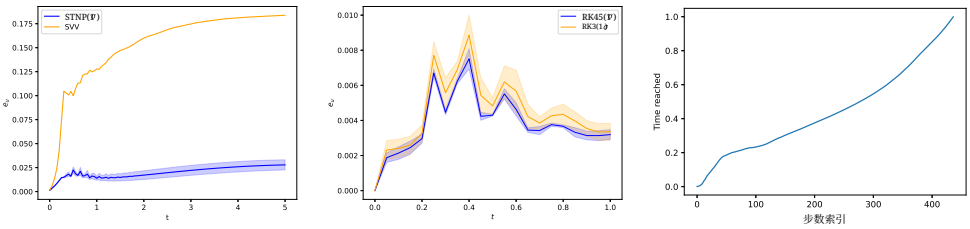


FIG. 4 *Left: temporal evolution of the relative L^2 error for STNP and spectral-SVV; STNP . , remains more accurate over the entire time horizon. Middle: relative L^2 error of RK3 versus RK45 showing improved robustness with adaptive time stepping. Right: cumulative time reached at each integration step; RK45 automatically reduces the step size around $t \approx 0.3$.*

总之，STNP 始终能够生成准确且无伪振荡的解，性能优于谱黏性值方法基准，尤其是在激波附近。此外，它在长时间积分下仍保持稳定，并且对随机初始化几乎不敏感，表明其在实际应用中具有很强的鲁棒性。

7. 结论。本文针对分数阶 Burgers 方程提出了一种时间序贯非线性参数化方法（STNP），并研究了其稳定性和精度。该方法将偏微分方程的动力学投影到神经流形的切空间上，并在每个时间步通过求解正则化最小二乘问题来计算参数演化。这种设计能够确保更新适定，并准确追踪解的动力学。我们针对由此产生的投影动力学，建立了 FBEFL 和 FBENN 的稳定性与收敛性估计，数值结果与理论相符。具体而言，STNP 能够捕捉激波而不产生伪振荡；在参数数量相当的情况下，其性能与 WENO–Roe 和谱黏性值方法基准相当或更优；对于光滑构造解，其相对误差达到 10^{-8}，同时呈现出预测的收敛率。此外，自适应 RK45 时间积分能够通过在参数流快速变化时自动减小步长，提高方法的鲁棒性。

该框架的若干方面对未来研究至关重要。对于无黏双曲守恒律，自适应配点以及在激波附近进行残差驱动的加密，对于准确解析参数动力学似乎至关重要。此外，将守恒约束或熵约束直接纳入切空间投影仍是一个有待探索的方向。我们预计，这些进展将进一步确立非线性参数化作为处理非局部、输运主导偏微分方程的实用工具。

致谢。本研究得到 MURI（多学科大学研究计划）资助（FA9550-20-1- 0358），并获美国能源部高级科学计算研究计划通过 SEA-CROGS 项目提供支持（DE-SC0023191）。此外，支持 MURI 研究及后续工作的神经偏微分方程与神经算子 GPU 集群也提供了额外资助，资助编号为 #FA9550-23-1-0671。

本稿件仅供审阅。

参考文献

[1] A. Albinali, *Analytical solution for anomalous diffusion in fractured nanoporous reservoirs*, 科罗拉多矿业学院, 2016 年。[2] N. Alibaud, J. Droniou, and J. Vovelle, *Occurrence and non-appearance of shocks in fractal burgers equations* 4 2007 479 499 ,《双曲微分方程杂志》, ( ), 第 – 页。[3] J. Berg and K. Nyström, *A unified deep artificial neural network approach to partial differential equations in complex geometries* 317 2018 28 41 ,《神经计算》, ( ), 第 – 页。[4] P. Biler, T. Funaki, and W. A. Woyczynski, *Fractal burgers equations*,《微分方程杂志》, 148 (1998), 第 9–46 页。[5] J. Bruna, B. Peherstorfer, and E. Vanden-Eijnden, *Neural galerkin schemes with activelearning for high-dimensional evolution equations*,《计算物理学杂志》, 496 (2024), 第 112588 页。[6] Z.-Q. Chen and R. Song, *Two-sided eigenvalue estimates for subordinate processes in domains*,《泛函分析杂志》, 226 (2005), 第 90–113 页。[7] B. Cockburn, C.-W. Shu, C. Johnson, E. Tadmor, and C.-W. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws* 1998 , 施普林格, 。

[8] R. Courant, E. Isaacson, and M. Rees, *On the solution of nonlinear hyperbolic differential equations by finite differences*,《纯粹与应用数学通讯》, 5 (1952), 第 243–255 页。[9] J. Droniou, *A numerical method for fractal conservation laws*,《计算数学》, 79 (2010), 第 95–124 页。[10] Y. Du and T. A. Zaki, *Evolutional deep neural network*,《物理评论E》, 104 (2021), 第 045303 页。[11] M. Fukunaga and N. Shimizu, *Fractional derivative constitutive models for finite deformation of viscoelastic materials*,《计算与非线性动力学杂志》, 10 (2015), 第 061002 页。[12] I. Goodfellow, *Deep learning*, 2016 年。[13] Z.-p. Hao, Z.-z. Sun, and W.-r. Cao, *A fourth-order approximation of fractional derivatives with its applications* 281 2015 787 805 ,《计算物理学杂志》, ( ), 第 – 页。[14] D. P. Kingma, *Adam: A method for stochastic optimization*, arXiv 预印本 arXiv:1412.6980, (2014)。[15] A. Kiselev, F. Nazarov, and R. Shterenberg, *Blow up and regularity for fractal burgers equation* 0804 3549, 2008 , arXiv 预印本 arXiv: 。( )。[16] O. Koch and C. Lubich, *Dynamical low-rank approximation*,《SIAM 矩阵分析与应用杂志》, 29 (2007), 第 434–454 页。[17] P. D. Lax, *Gibbs phenomena*,《科学计算杂志》, 28 (2006), 第 445–449 页。[18] R. J. LeVeque and R. J. Leveque, *Numerical methods for conservation laws*, 第 132 卷, 施普林格, 1992 年。[19] A. Lischke, G. Pang M. Gulian F. Song C. Glusa X. Zheng Z. Mao W. Cai M. M. , , , , , , , , , , Meerschaert M. Ainsworth et al., *What is the fractional laplacian? a comparative review with new results* 404 2020 109009 ,《计算物理学杂志》, ( ), 第 页。[20] D. C. Liu and J. Nocedal, *On the limited memory bfgs method for large scale optimization*,《数学规划》, 45 (1989), 第 503–528 页。[21] T. Luo and H. Yang, *Two-layer neural networks for partial differential equations: Optimization and generalization theory* 25, 2024, , 载于《数值分析手册》, 第 卷, 爱思唯尔, 第 515–554 页。[22] Y. Maday and E. Tadmor, *Analysis of the spectral vanishing viscosity method for periodic conservation laws* 26 1989 854 870 ,《SIAM 数值分析杂志》, ( ), 第 – 页。[23] Z. Mao and G. E. Karniadakis, *Fractional burgers equation with nonlinear non-locality: Spectral vanishing viscosity and local discontinuous galerkin methods*,《计算物理学杂志》, 336 (2017), 第 143–163 页。[24] P. Miškinis et al., *A generalization of the hopf-cole transformation*, SIGMA. 对称性、可积性与几何: 方法与应用 9 (2013), 第 016 页。[25] R. Orús, *A practical introduction to tensor networks: Matrix product states and projected entangled pair states* 349 2014 117 158 ,《物理学年鉴》, ( ), 第 – 页。[26] I. V. Oseledets, *Tensor-train decomposition*,《SIAM 科学计算杂志》, 33 (2011), 第 2295–2317 页。[27] G. Pang, L. Lu, and G. E. Karniadakis, *fpinns: Fractional physics-informed neural networks*,《SIAM 科学计算杂志》, 41 (2019), 第 A2603–A2626 页。[28] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Physics-informed neural networks: A deep*

本文仅供审阅。

learning framework for solving forward and inverse problems involving nonlinear partial differential equations 378 2019 686 707 ,《计算物理学杂志》，（），第 – 页。[29] G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, *Quantum dynamics simulations using gaussian wavepackets: the vmcg method*,《国际物理化学评论》，34（2015），第 269–308 页。[30] J.-y. Shen, J. Ren, and S. Chen, *A second-order energy stable and nonuniform time-stepping scheme for time fractional burgers' equationimage 1*,《计算机与应用数学》，123（2022），第 227–240 页。[31] C.-W. Shu, *High order eno and weno schemes for computational fluid dynamics*,载于《计算物理中的高阶方法》，施普林格，1999 年，第 439–582 页。[32] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock-capturing schemes* 77 1988 439 471 ,《计算物理学杂志》，（），第 – 页。[33] Z.-z. Sun and X. Wu, *A fully discrete difference scheme for a diffusion-wave system*,《应用数值数学》，56（2006），第 193–209 页。[34] W. Tian, H. Zhou, and W. Deng, *A class of second order difference approximations for solving space fractional diffusion equations* 84 2015 1703 1727 ,《计算数学》，（），第 – 页。[35] H. Wendland, *Scattered data approximation*,第 17 卷，剑桥大学出版社，2004 年。[36] Q. Xu and J. S. Hesthaven, *Discontinuous galerkin method for fractional convection-diffusion equations* 52 2014 405 423 ,《SIAM 数值分析杂志》，（），第 – 页。[37] A. Yazdani, L. Lu, M. Raissi, and G. E. Karniadakis, *Systems biology informed deep learning for inferring parameters and hidden dynamics*,《公共科学图书馆计算生物学》，16（2020），第 e1007575 页。[38] H. Zhang, Y. Chen, E. Vanden-Eijnden, and B. Peherstorfer, *Sequential-in-time training of nonlinear parametrizations for solving time-dependent partial differential equations*,arXiv 预印本 arXiv:2404.01145，（2024）。[39] H. Zhou, W. Tian, and W. Deng, *Quasi-compact finite difference schemes for space fractional diffusion equations* 56 2013 45 66 ,《科学计算杂志》，（），第 – 页。

本手稿仅供审阅。