arXiv:2601.01462v1 [math.NA] 4 Jan 2026

有界区域内分数阶扩散方程物理信息神经网络的收敛分析

ELIE ABDO*、LIHUI CHAI†、RUIMENG HU‡、和XU YANG§

摘要。 我们建立了有界区域上随时间变化的分数阶扩散方程的物理信息神经网络（PINNs）收敛性。分数阶拉普拉斯算子的存在引入了非局部行为和正则性约束，而标准的神经网络近似无法自然地满足相关的谱边界条件。为应对这一挑战，我们提出了一种谱定义的光滑化策略，在确保边界相容性的同时保留非局部算子的结构。这使得我们能够在索伯列夫空间中推导出严格的能量估计。我们的结果依赖于偏微分方程理论中的分析工具，凸显了物理信息神经网络近似与非局部方程经典能量估计之间的相容性。我们证明了物理信息神经网络近似在任意时空索伯列夫范数 H^k（其中 $k \in \mathbb{N}$）下的收敛性。分析表明，光滑化残差对于为非局部偏微分方程的基于神经网络的求解器提供理论保证发挥着重要作用。

1. 引言。 分数阶偏微分方程（偏微分方程）已成为有效模型，可用于描述各种物理与工程系统中的反常输运、长程相互作用和记忆效应。其应用包括地球物理学中的地下流 [6], 图像处理中的边缘检测 [14], 湍流扩散 [22], 以及数理金融中的期权定价 [10]。在这些模型中，分数阶扩散方程以分数阶拉普拉斯算子取代经典拉普拉斯算子，能够捕捉由重尾跳跃分布或列维飞行支配的输运过程的非局部特性´。与标准扩散相比，这种非局部性会引发定性上截然不同的动力学，并在分析和计算方面带来重大挑战。

在有界区域中使用分数阶拉普拉斯算子，会带来分析和计算上的复杂性。其中最主要的是算子的非局部特性以及边界条件的恰当处理。与局部算子不同，分数阶扩散方程的解会受到整个定义域上函数行为的影响，而施加狄利克雷条件则需要采用非标准形式。在有界区域中，分数阶拉普拉斯算子存在多种解释，包括谱定义、受限定义和区域定义，每种定义在分析和数值计算方面都有不同的影响 [8, 12]。本文重点研究狄利克雷分数阶拉普拉斯算子的谱定义，该定义自然源于经典拉普拉斯算子的特征结构，并广泛应用于物理模型。关于狄利克雷分数阶拉普拉斯算子及其在有界区域非局部主动标量方程中的作用的最新研究，请参见 [4]。

物理信息神经网络（PINNs）被提出作为一种灵活的无网格方法，通过结合神经网络的表达能力与基于物理的损失函数来求解偏微分方程 [3, 11, 13, 17–20, 23, 25, 27]。尽管物理信息神经网络在多种情形下展现出良好的性能 [9, 24, 26], , 但严格的理论保证仍然有限，尤其是在非局部和分数阶问题中。近期

*贝鲁特美国大学数学系，贝鲁特，1107 2020，黎巴嫩。（ea94@aub.edu.lb）。

†中山大学数学学院，广州，510275，中国。（chailihui@mail.sysu.edu.cn）。

‡加利福尼亚大学圣巴巴拉分校数学系、统计与应用概率系，美国加利福尼亚州圣巴巴拉，邮编 93106-3080。（rhu@ucsb.edu）。§加利福尼亚大学圣巴巴拉分校数学系，美国加利福尼亚州圣巴巴拉，邮编 93106-3080。（xuyang@math.ucsb.edu）。

的研究 [3, 5, 16, 21] 在强光滑性假设和理想化优化设置下，分析了局部偏微分方程的物理信息神经网络近似误差。然而，关于底层算子为非局部算子且定义在具有非平凡边界条件的有界区域上时物理信息神经网络收敛性的研究相对较少。具体而言，谱狄利克雷分数阶拉普拉斯算子带来了两项主要困难。首先，标准网络架构无法自然满足谱边界条件。其次，该非局部算子具有全局作用，这使经典能量方法的应用更为复杂。类似问题也出现在周期性情形中，在网络架构中纳入边界条件能够显著提升稳定性和准确性 [15]。

近期的一项工作 [2], 为物理信息神经网络近似动理学方程（具体而言，是平衡态附近的玻尔兹曼方程）建立了收敛理论。虽然该研究涉及速度空间中的非局部性，但本文关注的是有界区域中的空间非局部性，这带来了本质不同的分析挑战。我们的目标是将严格的收敛分析拓展到有界区域中的分数阶扩散算子；由于非局部相互作用和不规则边界，标准数值方法在这类问题中往往会变得更加复杂。

为应对这些挑战，我们提出一种基于谱定义的光滑化策略，通过修改残差损失，使其与狄利克雷边界条件相容。该方法使我们能够在索伯列夫空间中进行严格的收敛分析。具体而言，我们为物理信息神经网络近似在任意时空索伯列夫范数 H^k （其中 $k \in \mathbb{N}$）下建立了误差界。我们的结果有助于奠定物理信息神经网络的数学基础，并为其应用于有界区域上的非局部扩散方程提供理论依据。尽管我们的工作受到非局部偏微分方程数值方法的启发，本文所发展的主要技术立足于偏微分方程分析和泛函估计工具，而非离散化或算法方案。因此，我们的贡献主要属于分析性质。

我们首先回顾有界区域上谱狄利克雷分数阶拉普拉斯算子的定义和性质。

1.1. 狄利克雷分数阶拉普拉斯算子。 设 Ω 为$\mathbb{R}^d$中的一个有界光滑定义域。我们将具有齐次狄利克雷边界条件的拉普拉斯算子记为 Δ_D 。注意， $-\Delta_D$ 定义在 $D(-\Delta_D) = H^2(\Omega) \cap H_0^1(\Omega)$ 上，且在 $L^2(\Omega)$ 中为正自伴算子。因此，存在一个由 $L^2(\Omega)$的特征函数 $\{w_j\}_{j=1}^{\infty} \subset H_0^1(\Omega)$ 构成的正交归一基，满足 $-\Delta_D$

$$-\Delta_D w_j = \lambda_j w_j, \tag{1.1}$$

其中，特征值 λ_j 满足 $0 < \lambda_1 \le \ldots \le \lambda_j \le \ldots \to \infty$。对于 $s \in \mathbb{R}$，我们将阶数为 s的分数阶拉普拉斯算子记为 Λ_D^s，定义如下：

$$\Lambda_D^s h = \sum_{j=1}^{\infty} \lambda_j^{\frac{s}{2}} (h, w_j)_{L^2} w_j, \tag{1.2}$$

其定义域为

$$\mathcal{D}(\Lambda_D^s) = \left\{ h : \|\Lambda_D^s h\|_{L^2}^2 := \sum_{j=1}^{\infty} \lambda_j^s (h, w_j)_{L^2}^2 < \infty \right\}. \tag{1.3}$$

特别地，当 $s > 0$ 时，将空间 $D(\Lambda_D^{-s})$ 理解为 $D(\Lambda_D^s)$ 的对偶空间。显然，只要 $s_1 \le s_2$，就有 $D(\Lambda_D^{s_2}) \subset D(\Lambda_D^{s_1})$。对于 $s \in [0, 1]$，我们将

定义域 $\mathsf{D}(\Lambda_D^s)$ 与通常的索博列夫空间对应如下:

(1.4)

$$\mathcal{D}(\Lambda_D^s) = \begin{cases} H^s(\Omega), & \text{if } s \in [0, \frac{1}{2}), \\ H_{00}^{\frac{1}{2}}(\Omega) = \left\{ h \in H_0^{\frac{1}{2}}(\Omega) : h/\sqrt{d(x)} \in L^2(\Omega) \right\}, & \text{if } s = \frac{1}{2}, \\ H_0^s(\Omega), & \text{if } s \in (\frac{1}{2}, 1], \end{cases}$$

其中， $H_0^s(\Omega)$ 是 $H^s(\Omega)$ 中边界迹为零的元素构成的希尔伯特子空间，而 $d(x)$ 是到边界的距离函数。

我们回顾如下恒等式

$$\lambda^{\frac{s}{2}} = c_s \int_0^\infty t^{-1-\frac{s}{2}} (1 - e^{-t\lambda}) dt, \tag{1.5}$$

该恒等式对 $s \in (0,2)$ 成立，其中 c_s 定义为

$$1 = c_s \int_0^\infty t^{-1-\frac{s}{2}} (1 - e^{-t}) dt. \tag{1.6}$$

利用后者，我们得到积分表示

$$(\Lambda_D^s f)(x) = c_s \int_0^\infty [f(x) - e^{t\Delta_D} f(x)] t^{-1-\frac{s}{2}} dt, \tag{1.7}$$

对于 $f \in \mathsf{D}(\Lambda_D^s)$ 和 $s \in (0,2)$。这里，热算子 $e^{t\Delta_D}$ 定义为

$$(e^{t\Delta_D} f)(x) = \int_\Omega H_D(x, y, t) f(y) dy, \tag{1.8}$$

其核 $H_D(x, y, t)$ 定义为

$$H_D(x, y, t) = \sum_{j=1}^\infty e^{-t\lambda_j} w_j(x) w_j(y). \tag{1.9}$$

对于 $\epsilon \in (0,\ 1)$，令 J_ϵ 为基于热半群 $e^{t\Delta_D}$ 定义的谱正则化算子

$$J_\epsilon \theta(x) = \frac{-1}{\ln \epsilon} \int_\epsilon^{\frac{1}{\epsilon}} \frac{e^{t\Delta_D} \theta(x)}{t} dt. \tag{1.10}$$

这一族正则化算子不仅能平滑 L^2 函数，还能确保其所有高阶拉普拉斯算子在 Ω 的边界上均为零。也就是说，给定函数$\theta \in L^2(\Omega)$，有 $J_\epsilon \theta \in \mathcal{D}(\Lambda_D^k)$ 对任意 $k \in \mathbb{N}$ 成立（见命题 2.1）。

1.2. 残差与误差。令 $\alpha \in [0, 2]$。我们关注对流扩散方程解的近似

$$\partial_t \psi + u \cdot \nabla \psi + \Lambda_D^\alpha \psi = f, \tag{1.11}$$

满足边界条件

$$\psi|_{\partial\Omega} = 0, \tag{1.12}$$

和初始数据

$$\psi(x, 0) = \psi_0(x), \tag{1.13}$$

由物理信息神经网络给出。这里 $u := u(x,t)$ 是一个给定的光滑无散向量场，满足

$$u \cdot n|_{\partial\Omega} = 0, \tag{1.14}$$

其中， n 是指向 $\partial\Omega$的外向单位法向量， $f := f(x,\ t)$ 是给定的光滑函数，满足 $(-\Delta)^k f|_{\partial\Omega} = 0$ 对任意 $k \in \mathbb{N}$

神经网络不一定在 Ω 的边界上消失，因此当 $\alpha \geq 1/2$ 时，它们不一定属于 $\mathrm{D}(\Lambda_D^\alpha)$。为克服这一挑战，我们使用磨光子 J_ϵ 来定义以下偏微分方程残差

$$\mathcal{R}_i[\theta](x,t) = \partial_t J_\epsilon \psi_\theta + u \cdot \nabla J_\epsilon \psi_\theta + \Lambda_D^\alpha J_\epsilon \psi_\theta - f, \tag{1.15}$$

初始残差

$$\mathcal{R}_t[k;\theta] = \sum_{i=1}^{k} [(\partial_t^{(i)} \psi_\theta) - (\partial_t^{(i)} \psi)](x,0), \tag{1.16}$$

以及边界残差

$$\mathcal{R}_b[\theta](x,t) = \psi_\theta|_{\partial\Omega}. \tag{1.17}$$

这些残差是良定义的，因为 $J_\epsilon \psi_\theta \in \mathcal{D}(\Lambda_D^k)$ 对所有 $k \in \mathbb{N}$均成立。我们指出，初始残差定义中的时间导数项 $(\partial_t^{(i)}\psi)(x,\ 0)$ 仅取决于给定初始数据的空间导数 ψ_0 ，因为 ψ 是偏微分方程（1.11）的精确光滑解。对于正则性指标 $\ell,\ k \in \mathbb{N}$，我们定义总误差 $\mathcal{E}[\ell, k;\ \theta]$ 为

$$\mathcal{E}[\ell,k;\theta]^2 = \int_0^T \sum_{i=0}^{k} \|\Lambda_D^\ell \partial_t^{(i)} (\psi - J_\epsilon \psi_\theta)\|_{L^2}^2 dt. \tag{1.18}$$

我们注意到，$\mathrm{E}[\ell, k;\theta]$ 是定义良好的，因为对于任意 $j \in \mathbb{N}$，解 ψ 都属于 $\mathrm{D}(\Lambda_D^j)$（见附录A），且对于任意 $j \in \mathbb{N}$， $J_\epsilon \psi_\theta \in \mathrm{D}(\Lambda_D^j)$。此外，$\mathrm{E}[k,k;\ \theta]$ 支配 $\psi - J_\epsilon \psi_\theta$的索博列夫 $H^k([0,T]\times\Omega)$范数（见命题 4.1），因此，$\mathrm{E}[\ell, k;\theta]$ 衡量了真实解与神经网络在所有时空索伯列夫空间中的距离。

对于时间正则性指标 $k \in \mathbb{N}$ 和空间正则性指标 $\ell \in \mathbb{N}$，我们将泛化误差 $\mathrm{E}_G[\ell, k;\theta]$ 定义为

$$\mathcal{E}_G[\ell,k;\theta]^2 = \mathcal{E}_G^i[\ell,k;\theta]^2 + \mathcal{E}_G^t[\ell,k;\theta]^2 + \mathcal{E}_G^b[\theta]^2, \tag{1.19}$$

其中

$$\mathcal{E}_G^i[\ell,k;\theta]^2 = \begin{cases} \int_0^T \sum_{j=0}^{k} \|(-\Delta)^{\frac{\ell}{2}} \partial_t^{(j)} \mathcal{R}_i[\theta]\|_{L^2}^2 dt, & \text{if } \ell \text{ is even} \\ \int_0^T \sum_{j=0}^{k} \|\nabla(-\Delta)^{\frac{\ell-1}{2}} \partial_t^{(j)} \mathcal{R}_i[\theta]\|_{L^2}^2 dt, & \text{if } \ell \text{ is odd} \end{cases}, \tag{1.20}$$

$$\mathcal{E}_G^t[\ell,k;\theta]^2 = \begin{cases} \|(-\Delta)^{\frac{\ell}{2}} \mathcal{R}_t[k;\theta]\|_{L^2}^2, & \text{if } \ell \text{ is even} \\ \|\nabla(-\Delta)^{\frac{\ell-1}{2}} \mathcal{R}_t[k;\theta]\|_{L^2}^2, & \text{if } \ell \text{ is odd} \end{cases}, \tag{1.21}$$

and

$$\mathcal{E}_G^b[\theta]^2 = \int_0^T \int_{\partial\Omega} \mathcal{R}_b[\theta]^2 d\sigma(x). \tag{1.22}$$

我们注意到，由于残差是光滑函数，且所涉及的算子都是局部的，因此这些误差也是定义良好的。

1.3. 论文组织结构。 本文其余部分的组织结构如下。在第 2节中，我们介绍光滑化策略，并建立谱定义正则化算子的基本解析性质；这些性质对于确保物理信息神经网络残差满足边界相容性起着关键作用。第 3节专门讨论泛化误差分析，并在目标解满足适当光滑性假设的条件下，推导物理信息神经网络残差的上界。在第 4节中，我们结合逼近分析与泛化误差分析，得到物理信息神经网络解在任意时空索 贝列夫范数 H^k （其中 $k \in \mathbb{N}$）下的总误差估计。第 5 节给出总结，并讨论未来工作的方向。最后，在附录 A中建立有界光滑区域上分数阶反应扩散方程唯一光滑解的存在性。

2. 正则化算子的性质。本节将研究磨光算子族 J_ϵ 的关键性质。这些算子在后续章节所建立的物理信息神经网络收敛分析中发挥着核心作用。

算子 J_ϵ 在 $\mathrm{D}(\Lambda_D^s)$ 的范数下是一致有界的，并且能够将 L^2 中的任意函数光滑化：

PROPOSITION 2.1. 设 s 为实数， $\epsilon \in$ （0，1） 为一个较小的正数。则存在一个仅取决于 s 的正数 C ，使得

$$\|\Lambda_D^s J_\epsilon \theta\|_{L^2} \le C \|\Lambda_D^s \theta\|_{L^2}, \tag{2.1}$$

只要 $\theta \in \mathcal{D}$（Λ_D^s）。此外，对于任意实数 $s \ge 0$ ，都有

$$\|\Lambda_D^s J_\epsilon \theta\|_{L^2} \le C \epsilon^{-\frac{s}{2}} \|\theta\|_{L^2}, \tag{2.2}$$

对于 $\theta \in L^2$（Ω）。

命题 2.1 的证明与 [1] 中的引理1的证明基本相同，因此为简洁起见，此处略去。

接下来，我们陈述分数阶拉普拉斯算子 Λ_D^s与正则化算子 J_ϵ之间的一项关键交换性质，它在我们的分析中起着核心作用。这一结果确保了平滑化过程与问题的非局部结构相容。具体而言，算子 Λ_D^s 和 J_ϵ 在 $\mathrm{D}(\Lambda_D^s)$ 上可交换：

PROPOSITION 2.2. 设 $\epsilon \in (0,1)$， $s > 0$，且 $f \in \mathcal{D}(\Lambda_D^s)$。则

$$\Lambda_D^s J_\epsilon f(x) = J_\epsilon \Lambda_D^s f(x),$$

对几乎处处的 $x \in \Omega$成立。

详细证明请参见 [4] 中的命题2.2。

我们还注意到，拉普拉斯算子的整数次幂与算子 J_ϵ 可交换，且无需任何边界假设。特别是，对于 H^{2k} 函数，该交换性质仍然成立，其中 $k \in \mathbb{N}$，而这些函数不一定满足齐次边界条件。下面我们陈述并证明这一结果。

PROPOSITION 2.3. 设 $\epsilon \in (0,1)$。设 $f \in H^2(\Omega)$。则

$$\Delta J_\epsilon f(x) = J_\epsilon \Delta f(x), \tag{2.3}$$

对于几乎处处的 $x \in \Omega$。因此，如果 $k \in \mathbb{N}$ 且 $f \in H^{2k}(\Omega)$，那么

$$(-\Delta)^k J_\epsilon f(x) = J_\epsilon (-\Delta)^k f(x), \tag{2.4}$$

对于几乎处处的 $x \in \Omega$。

证明。交换子 $T_\epsilon = J_\epsilon\Delta - \Delta J_\epsilon$ 在 $C_0^\infty(\Omega)$ 上消失，因为 J_ϵ 和 $-\Delta = \Lambda_D^2$ 在 $\mathrm{D}(\Lambda_D^2)$ 上可交换（见命题2.2）。鉴于 $C_0^\infty(\Omega)$ 在 $L^2(\Omega)$ 中稠密， T_ϵ 可唯一延拓为 $L^2(\Omega)$ 上的有界线性算子 $\tilde{T}_\epsilon$ ，使得对于任意 $h \in L^2(\Omega)$，$\tilde{T}_\epsilon h = 0$ 。然而，算子 T_ϵ 在 $H^2(\Omega)$ 上定义良好。此外，有

$$\|T_\epsilon f\|_{L^2} \le C_\epsilon \|f\|_{L^2}. \tag{2.5}$$

为了证明后者，一方面，我们有

$$\|\Delta J_\epsilon f\|_{L^2} \le C_\epsilon \|f\|_{L^2}, \tag{2.6}$$

由于（2.1）。另一方面， $J_\epsilon\Delta f$ 具有如下特征函数展开，

$$J_\epsilon \Delta f = \sum_{j=1}^{\infty} \left(-\frac{1}{\ln \epsilon} \int_\epsilon^{1/\epsilon} \frac{e^{-t\lambda_j}}{t} dt \right) (\Delta f, \omega_j)_{L^2} \omega_j, \tag{2.7}$$

因此，其 L^2 范数满足

$$\|J_\epsilon \Delta f\|_{L^2}^2 = \sum_{j=1}^{\infty} \left(-\frac{1}{\ln \epsilon} \int_\epsilon^{1/\epsilon} \frac{e^{-t\lambda_j}}{t} dt \right)^2 (\Delta f, \omega_j)_{L^2}^2. \tag{2.8}$$

由于 $t \ge \epsilon$，有 $e^{-t\lambda_j} \le e^{-\epsilon\lambda_j}$，因此，

$$\|J_\epsilon \Delta f\|_{L^2}^2 \le \sum_{j=1}^{\infty} \left(-\frac{1}{\ln \epsilon} \int_\epsilon^{1/\epsilon} \frac{1}{t} dt \right)^2 e^{-\epsilon\lambda_j} (\Delta f, \omega_j)_{L^2}^2 = \sum_{j=1}^{\infty} 4 e^{-\epsilon\lambda_j} (\Delta f, \omega_j)_{L^2}^2. \tag{2.9}$$

鉴于算子 $(-\Delta_D)^{-1}\Delta$ 从 $L^2(\Omega)$ 到 $L^2(\Omega)$ 的有界性，我们有 e

$$\begin{aligned} |(\Delta f, \omega_j)_{L^2}| &= |(-\Delta_D)(-\Delta_D)^{-1}\Delta f, \omega_j)_{L^2} = |(-\Delta_D)^{-1}\Delta f, -\Delta_D \omega_j)_{L^2}| \\ &\le \|(-\Delta_D)^{-1}\Delta f\|_{L^2} \|\Delta \omega_j\|_{L^2} \le C\lambda_j \|f\|_{L^2}, \end{aligned} \tag{2.10}$$

分部积分后，并利用 ω_j 和 $(-\Delta_D)^{-1}\Delta f$ 在Ω的边界上消失这一事实。因此，

$$\|J_\epsilon \Delta f\|_{L^2}^2 \le C\|f\|_{L^2}^2 \sum_{j=1}^{\infty} \lambda_j^2 e^{-\epsilon\lambda_j} \le C_\epsilon \|f\|_{L^2}^2. \tag{2.11}$$

将 (2.6) 和 (2.11) 合并，得到 (2.5)。这表明 T_ϵ 是从 $(H^2(\Omega)\, \|\cdot\|_{L^2})$ 到 $(L^2(\Omega)\, \|\cdot\|_{L^2})$ 的有界线性算子。根据延拓的唯一性，我们推得

$$\tilde{T}_\epsilon = J_\epsilon \Delta - \Delta J_\epsilon, \tag{2.12}$$

在 $(H^2(\Omega)\|\cdot\|_{L^2})$ 上。但 $\tilde{T}_\epsilon$ 在 $L^2(\Omega)$ 上为零，因此

$$J_\epsilon \Delta - \Delta J_\epsilon = 0, \tag{2.13}$$

在 $H^2(\Omega)$ 上。 □

我们为正则化算子 J_ϵ 在 $\mathrm{D}(\Lambda_D^s)$ 的范数下的收敛推导出了新的定量界。

PROPOSITION 2.4. 设 $\epsilon \in (0,1)$ 和 $s \geq 0$ 为实数。设 $\psi \in \mathcal{D}(\Lambda_D^{2s+d})$。存在一个取正值的实函数 $\kappa(\epsilon)$，使得 $\kappa(\epsilon) \to 0$ 当 $\epsilon \to 0$ 时，以及

$$\|\Lambda_D^s(J_\epsilon f - f)\|_{L^2} \leq \kappa(\epsilon)\|f\|_{L^2}^{\frac{1}{2}}\|\Lambda_D^{2s+d} f\|_{L^2}^{\frac{1}{2}}. \tag{2.14}$$

证明。 $J_\epsilon f - f$ 的特征函数展开为

$$J_\epsilon f - f = \sum_{j=1}^{\infty}(J_\epsilon f - f, \omega_j)_{L^2}\omega_j, \tag{2.15}$$

因此,

$$\|\Lambda_D^s(J_\epsilon f - f)\|_{L^2}^2 = \sum_{j=1}^{\infty}\lambda_j^s(J_\epsilon f - f, \omega_j)_{L^2}^2. \tag{2.16}$$

利用 J_ϵ 的积分表示公式以及恒等式

$$\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{1}{2t}dt = -1, \tag{2.17}$$

可得

$$\begin{aligned}(J_\epsilon f - f, \omega_j)_{L^2} &= \left[-\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{e^{-t\lambda_j}}{t}dt - 1\right](f,\omega_j)_{L^2} \\ &= \left[\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{1-2e^{-t\lambda_j}}{2t}dt\right](f,\omega_j)_{L^2},\end{aligned} \tag{2.18}$$

由此得到

$$\|\Lambda_D^s(J_\epsilon f - f)\|_{L^2}^2 = \sum_{j=1}^{\infty}\left[\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{1-2e^{-t\lambda_j}}{2t}dt\right]^2\lambda_j^s(f,\omega_j)_{L^2}^2. \tag{2.19}$$

应用柯西-施瓦茨不等式可得

$$\|\Lambda_D^s(J_\epsilon f - f)\|_{L^2}^2 \leq \left(\sum_{j=1}^{\infty}\frac{1}{\lambda_j^d}\left[\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{1-2e^{-t\lambda_j}}{2t}dt\right]^4\right)^{\frac{1}{2}}\left(\sum_{j=1}^{\infty}\lambda_j^{2s+d}(f,\omega_j)_{L^2}^4\right)^{\frac{1}{2}}. \tag{2.20}$$

令

$$\kappa(\epsilon) := \left(\sum_{j=1}^{\infty}\frac{1}{\lambda_j^d}\left[\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{1-2e^{-t\lambda_j}}{2t}dt\right]^4\right)^{\frac{1}{4}}. \tag{2.21}$$

由于

$$\left|\frac{1}{\ln\epsilon}\int_\epsilon^{1/\epsilon}\frac{1-2e^{-t\lambda_j}}{2t}dt\right| \leq \frac{1}{|\ln\epsilon|}\int_\epsilon^{1/\epsilon}\frac{1}{2t}dt \leq 1, \tag{2.22}$$

并且对于任意 $j \in \mathbb{N}$, $\lambda_j \geq cj^{\frac{2}{d}}$ 可得

$$\kappa(\epsilon)^4 \leq \sum_{j=1}^{\infty}\frac{1}{\lambda_j^d} \leq C\sum_{j=1}^{\infty}\frac{1}{j^2} < \infty, \tag{2.23}$$

因此 $\kappa(\epsilon)$ 有良好定义。此外，由于

$$\lim_{\epsilon\to 0}\frac{1}{\ln\epsilon}\int_{\epsilon}^{1/\epsilon}\frac{1-2e^{-t\lambda_j}}{2t}dt=0, \tag{2.24}$$

由勒贝格控制收敛定理可知，当 $\epsilon\to 0$ 时， $\kappa(\epsilon)\to 0$。最后，由于对任意 $j\in\mathbb{N}$, $\|\omega_j\|_{L^2}=1$，我们有

$$\begin{aligned}\sum_{j=1}^{\infty}\lambda_j^{2s+d}(f,\omega_j)_{L^2}^4 &\le \sum_{j=1}^{\infty}\lambda_j^{2s+d}(f,\omega_j)_{L^2}^2\|f\|_{L^2}^2\|\omega_j\|_{L^2}^2\\ &\le \|f\|_{L^2}^2\sum_{j=1}^{\infty}\lambda_j^{2s+d}(f,\omega_j)_{L^2}^2\le\|f\|_{L^2}^2\|\Lambda_D^{2s+d}f\|_{L^2}^2.\end{aligned} \tag{2.25}$$

因此，我们得到所需的定量界

$$\|\Lambda_D^s(J_\epsilon f-f)\|_{L^2}^2\le\kappa(\epsilon)^2\|f\|_{L^2}\|\Lambda_D^{2s+d}f\|_{L^2}. \tag{2.26}$$

□

3. 泛化误差估计。本节中，我们针对在有界区域上逼近分数阶扩散方程解的物理信息神经网络，建立其泛化误差的界。我们的主要结果表明，在适当的正则性假设下，可以构造泛化误差任意小的神经网络。分析中的一个关键要素是深度神经网络在索博列夫范数下逼近光滑函数的能力。以下命题改编自经典的通用逼近定理，给出了神经网络在 $W^{k,\infty}$ 范数下逼近光滑目标函数的定量速率。

PROPOSITION 3.1. 设 $\Omega\subset\mathbb{R}^d$ 是一个光滑有界区域。设 $T>0$ 任意且 $k\in\mathbb{N}$。设 $f\in C^{k+1}([0,T]\times\bar{\Omega})$。那么，存在一个由复杂度为 N 的深度神经网络表示的函数 ϕ_N，使得

$$\|f-\phi_N\|_{W^{k,\infty}([0,T]\times\bar{\Omega})}\le\frac{1}{N^{\frac{1}{d+1}}}\|f\|_{W^{k+1,\infty}([0,T]\times\bar{\Omega})}. \tag{3.1}$$

证明。由于 $[0,T]\times\bar{\Omega}$ 是 $\mathbb{R}^{d+1}$ 的紧子集，且 $f\in C^k([0,T]\times\bar{\Omega})$，我们可以应用 [7] 中的定理 2.1，推得存在一个深度神经网络 ϕ_N，其复杂度为 N，使得

$$\|D^\alpha f-D^\alpha\phi_N\|_{L^\infty([0,T]\times\bar{\Omega})}\le\frac{C}{N^{\frac{k-|\alpha|}{d+1}}}\sup_{|(x,t)-(y,t)|\le N^{\frac{-1}{d+1}}}|D^\beta f(x,t)-D^\beta f(y,t)|, \tag{3.2}$$

对于任意 $|\alpha|\le k$ 和 $|\beta|=k$。由于 $N\ge 1$ 和 $k\ge|\alpha|$，可得

$$\|D^\alpha f-D^\alpha\phi_N\|_{L^\infty([0,T]\times\bar{\Omega})}\le C\sup_{|(x,t)-(y,t)|\le N^{\frac{-1}{d+1}}}|D^\beta f(x,t)-D^\beta f(y,t)|, \tag{3.3}$$

对于任意 $|\alpha|\le k$ 和 $|\beta|=k$。因此，

$$\begin{aligned}\|D^\alpha f-D^\alpha\phi_N\|_{L^\infty([0,T]\times\bar{\Omega})}&\le C\sup_{|(x,t)-(y,t)|\le N^{\frac{-1}{d+1}}}\left[\frac{|D^\beta f(x,t)-D^\beta f(y,t)|}{|(x,t)-(y,t)|}\cdot|(x,t)-(y,t)|\right]\\ &\le\frac{1}{N^{\frac{1}{d+1}}}\|f\|_{W^{k+1,\infty}([0,T]\times\bar{\Omega})},\end{aligned} \tag{3.4}$$

对于任意 $|\alpha|\le k$，由此得到所需的估计（3.1）。 □

THEOREM 3.1. 令 $T>0$ 为任意正时间。令 $\psi_0 \in \bigcap_{m\geq 0} \mathcal{D}(\Lambda_D^m)$ 和 ψ 为 (1.11) 在齐次狄利克雷边界条件下对应的光滑解。令 $\delta>0$。则存在一个函数 ψ，它由一个复杂度为 N 的深度神经网络表示，使得 $\mathcal{E}_G \leq \delta$。

证明。令 ℓ 和 k 为非负整数。根据命题 3.1，存在一个神经网络 ψ 使得

$$\|\psi-\widehat{\psi}\|_{W^{\max\{2+\ell,k+1\},\infty}([0,T]\times\overline{\Omega})} \leq CN^{-\frac{1}{d+1}}, \tag{3.5}$$

其中 C 是一个依赖于初始数据的常数。选择索博列夫空间 $W^{\max\{2+\ell,k+1\},\infty}([0,T]\times\Omega)$ 是受下述估计的启发。

步骤 1. 对 E_G^i. 的估计我们首先将偏微分方程残差改写如下，

$$\begin{aligned}\mathcal{R}_i &= \partial_t(J_\epsilon\widehat{\psi}-\psi)+u\cdot\nabla(J_\epsilon\widehat{\psi}-\psi)+\Lambda_D^\alpha(J_\epsilon\widehat{\psi}-\psi)\\ &= \partial_t J_\epsilon(\widehat{\psi}-\psi)+\partial_t(J_\epsilon\psi-\psi)+u\cdot\nabla J_\epsilon(\widehat{\psi}-\psi)+u\cdot\nabla(J_\epsilon\psi-\psi)\\ &\quad+\Lambda_D^\alpha J_\epsilon(\widehat{\psi}-\psi)+\Lambda_D^\alpha(J_\epsilon\psi-\psi).\end{aligned} \tag{3.6}$$

固定一个整数 j，表示时间导数的阶数，使得 $0\leq j\leq k$。若 ℓ 为偶数，则对 R_i 应用 $(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}$，并分别估计每一项。若 ℓ 为奇数，则改为对 R_i 应用 $\nabla(-\Delta)^{\frac{\ell-1}{2}}\partial_t^{(j)}$。

假设 ℓ 为偶数。我们首先估计时间导数项。事实上，根据命题 2.3，J_ϵ 和 $(-\Delta)^{\frac{\ell}{2}}$ 可交换，因此我们有

$$\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}\partial_t J_\epsilon(\widehat{\psi}-\psi)\|_{L^2}=\|J_\epsilon(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j+1)}(\widehat{\psi}-\psi)\|_{L^2}, \tag{3.7}$$

其可由以下各项界定

$$\begin{aligned}\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}\partial_t J_\epsilon(\widehat{\psi}-\psi)\|_{L^2} &\leq C\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j+1)}(\widehat{\psi}-\psi)\|_{L^2}\leq C\|\widehat{\psi}-\psi\|_{H^\beta([0,T]\times\overline{\Omega})}\\ &\leq C\|\widehat{\psi}-\psi\|_{W^{\beta,\infty}([0,T]\times\overline{\Omega})}\leq CN^{-\frac{1}{d+1}},\end{aligned} \tag{3.8}$$

由于关于 ϵ 的一致有界性 J_ϵ 在 L^2 中成立，且根据估计式（3.5），这里 β =最大值 $\{\ell,\ k+1\}$。鉴于命题2.4中得到的定量收敛界，可得

$$\begin{aligned}&\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}\partial_t(J_\epsilon\psi-\psi)\|_{L^2}^2=\|(-\Delta)^{\frac{\ell}{2}}(J_\epsilon\partial_t^{(j+1)}\psi-\partial_t^{(j+1)}\psi)\|_{L^2}^2\\ &\leq\kappa(\epsilon)^2\|\partial_t^{(j+1)}\psi\|_{L^2}\|\Lambda_D^{2\ell+d}\partial_t^{(j+1)}\psi\|_{L^2}\leq C\kappa(\epsilon)^2.\end{aligned} \tag{3.9}$$

现在我们估计涉及狄利克雷拉普拉斯算子分数次幂的扩散项。由于 Λ_D^α 与 $\partial_t^{(j)}$ 可交换，且 J_ϵ 与 $\partial_t^{(j)}$ 可交换，因此有

$$\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}\Lambda_D^\alpha J_\epsilon(\widehat{\psi}-\psi)\|_{L^2}=\|\Lambda_D^{\ell+\alpha}J_\epsilon(\partial_t^{(j)}\widehat{\psi}-\partial_t^{(j)}\psi)\|_{L^2}. \tag{3.10}$$

令 γ 为大于或等于 $\frac{\ell+\alpha}{2}$ 的最小整数。由于 $\mathcal{D}(\Lambda_D^{2\gamma})$ 在 $\mathcal{D}(\Lambda_D^{\ell+\alpha})$中的连续嵌入，$J_\epsilon$ 与拉普拉斯算子的整数次幂可交换，且 ϵ 一致有界，J_ϵ 在 L^2中成立，因此恒等式（3.10）给出

$$\begin{aligned}&\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}\Lambda_D^\alpha J_\epsilon(\widehat{\psi}-\psi)\|_{L^2}\leq C\|(-\Delta)^\gamma J_\epsilon(\partial_t^{(j)}\widehat{\psi}-\partial_t^{(j)}\psi)\|_{L^2}\\ &=\|J_\epsilon(-\Delta)^\gamma(\partial_t^{(j)}\widehat{\psi}-\partial_t^{(j)}\psi)\|_{L^2}\leq C\|(-\Delta)^\gamma\partial_t^{(j)}(\widehat{\psi}-\psi)\|_{L^2}\\ &\leq C\|\widehat{\psi}-\psi\|_{H^{\max\{\ell+2;k\}}([0,T]\times\overline{\Omega})}\leq\|\widehat{\psi}-\psi\|_{W^{\max\{\ell+2;k\},\infty}([0,T]\times\overline{\Omega})}\leq CN^{-\frac{1}{d+1}}.\end{aligned} \tag{3.11}$$

需要指出的是，当 J_ϵ 和 $\Lambda_D^{\ell+\alpha}$不为整数时，二者不可交换，因为$\partial_t^{(i)}\widehat{\psi}$ 及其高阶拉普拉斯算子在 Ω 的边界上不一定为零；但是 J_ϵ 和 $(-\Delta)^\gamma$ 可交换，如命题2.3所示，因为 γ 是整数。对于不依赖神经网络的扩散项，我们有

$$\begin{aligned}&\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}\Lambda_D^{\alpha}(J_\epsilon\psi-\psi)\|_{L^2}^2=\|\Lambda_D^{\ell+\alpha}(J_\epsilon\partial_t^{(j)}\psi-\partial_t^{(j)}\psi)\|_{L^2}^2\\&\le\kappa(\epsilon)^2\|\partial_t^{(j)}\psi\|_{L^2}\|\Lambda_D^{2\ell+2\alpha+d}\partial_t^{(j)}\psi\|_{L^2}\le C\kappa(\epsilon)^2,\end{aligned}\tag{3.12}$$

由于命题 2.4 中研究的收敛性质。最后，我们处理涉及 u 的平流项。根据莱布尼茨定理，有

$$\begin{aligned}&\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}(u\cdot\nabla J_\epsilon(\widehat{\psi}-\psi))\|_{L^2}\\&\quad=\sum_{v=0}^{j}\binom{j}{v}\|(-\Delta)^{\frac{\ell}{2}}\left(\partial_t^{(v)}u\cdot\nabla\partial_t^{(j-v)}J_\epsilon(\widehat{\psi}-\psi)\right)\|_{L^2}\\&\quad\le C\sum_{v=0}^{j}\|\partial_t^{(v)}u\|_{W^{\ell,\infty}(\Omega)}\|\partial_t^{(j-v)}J_\epsilon(\widehat{\psi}-\psi)\|_{H^{1+\ell}(\Omega)}.\end{aligned}\tag{3.13}$$

由于 $\mathrm{D}(\Lambda_D^{2+\ell})$ 连续嵌入 $H^{1+\ell}(\Omega)$ 中， J_ϵ 与 $(-\Delta)^{1+\frac{\ell}{2}}$ 可交换（因为 $1+\frac{\ell}{2}$ 是整数），且 J_ϵ 在 L^2中有界，我们估计

$$\begin{aligned}&\|\partial_t^{(j-v)}J_\epsilon(\widehat{\psi}-\psi)\|_{H^{1+\ell}(\Omega)}\le C\|(-\Delta)^{1+\frac{\ell}{2}}J_\epsilon\partial_t^{(j-v)}(\widehat{\psi}-\psi))\|_{L^2(\Omega)}\\&\le C\|(-\Delta)^{1+\frac{\ell}{2}}\partial_t^{(j-v)}(\widehat{\psi}-\psi)\|_{L^2}^2\le C\|\widehat{\psi}-\psi\|_{W^{\max\{2+\ell,j\},\infty}([0,T]\times\overline{\Omega})}\le CN^{-\frac{1}{d+1}}.\end{aligned}\tag{3.14}$$

将后者代回（3.13）中，我们得到

$$\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}(u\cdot\nabla J_\epsilon(\widehat{\psi}-\psi))\|_{L^2}\le C\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}N^{-\frac{1}{d+1}}.\tag{3.15}$$

至于 u 中的最后一项，我们再次应用莱布尼茨法则，并利用收敛性质（2.4）可得

$$\begin{aligned}&\|(-\Delta)^{\frac{\ell}{2}}\partial_t^{(j)}(u\cdot\nabla(J_\epsilon\psi-\psi))\|_{L^2}^2\\&\le C\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}^2\sum_{v=0}^{j}\kappa(\epsilon)^2\|\partial_t^{(j-v)}\psi\|_{L^2}\|\Lambda_D^{2\ell+2+d}\partial_t^{(j-v)}\psi\|_{L^2}\\&\le C\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}^2\kappa(\epsilon)^2.\end{aligned}\tag{3.16}$$

这表明

$$\mathcal{E}_G^i\le C\left(1+\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}^2\right)(N^{-\frac{2}{d+1}}+\kappa(\epsilon)^2),\tag{3.17}$$

当 ℓ 为偶数时。

假设 ℓ 为奇数。鉴于 $\mathrm{D}(\Lambda_D)$ 到 $H^1(\Omega)$ 的连续嵌入以及 $\mathrm{D}(\Lambda_D^2)$ 到 $\mathrm{D}(\Lambda_D)$ 的连续嵌入， J_ϵ 与 $(-\Delta)^{\frac{\ell+1}{2}}$ 可交换（因为 $\frac{\ell+1}{2}$ 是整数），且 J_ϵ 在 L^2中有界，我们可以将时间导数项估计如下，

$$\begin{aligned}&\|\nabla(-\Delta)^{\frac{\ell-1}{2}}\partial_t^{(j)}\partial_t J_\epsilon(\widehat{\psi}-\psi)\|_{L^2}\le C\|\Lambda_D\Lambda_D^{\ell-1}\partial_t^{(j+1)}J_\epsilon(\widehat{\psi}-\psi)\|_{L^2}\\&\le C\|\Lambda_D^2\Lambda_D^{\ell-1}\partial_t^{(j+1)}J_\epsilon(\widehat{\psi}-\psi)\|_{L^2}\le C\|(-\Delta)^{\frac{\ell+1}{2}}\partial_t^{(j+1)}(\widehat{\psi}-\psi)\|_{L^2}\\&\le\|\widehat{\psi}-\psi\|_{W^{\max\{\ell+1,k+1\},\infty}([0,T]\times\Omega)}\le CN^{\frac{1}{d+1}}.\end{aligned}\tag{3.18}$$

如偶数情形所示，有

$$\|\nabla(-\Delta)^{\frac{\ell-1}{2}}\partial_t^{(j)}\partial_t(J_\epsilon\psi-\psi)\|_{L^2}^2 \le C\|\Lambda_D^\ell\partial_t^{(j+1)}(J_\epsilon\psi-\psi)\|_{L^2}^2 \le C\kappa(\epsilon)^2. \tag{3.19}$$

至于其余各项，我们采用相同的思路，将 $\mathrm{D}(\Lambda_D)$ 嵌入 $H^1(\Omega)$，并逐字沿用偶数情形中的计算过程。为避免重复，我们省略这些细节。因此，我们得到

$$\mathcal{E}_G^i \le C\left(1+\|u\|^2_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}\right)(N^{-\frac{2}{d+1}}+\kappa(\epsilon)^2), \tag{3.20}$$

当 ℓ 为奇数时。

步骤 2. 对 E_G^t. 的估计我们希望在 k 为偶数时，对 $\mathrm{E}_G^t[\ell,k;\ \theta]$ 进行良好控制。k 为奇数的情形类似，此处略去。当 k 为偶数时，我们有 (3.21)

$$\begin{aligned}\mathcal{E}_G^t &\le C\sum_{j=1}^k\|(-\Delta)^{\frac{\ell}{2}}(\partial_t^{(j)}(\widehat{\psi}-\psi))(x,0)\|^2_{L^2(\Omega)} \le C\sum_{j=1}^k\|\partial_t^{(j)}(\widehat{\psi}-\psi)(x,0)\|^2_{H^\ell(\Omega)}\\ &\le C\sum_{j=1}^k\|\partial_t^{(j)}(\widehat{\psi}-\psi)\|^2_{H^{\ell+1}([0,T]\times\overline{\Omega})} \le C\|\widehat{\psi}-\psi\|^2_{W^{\max\{\ell+1,k\},\infty}([0,T]\times\overline{\Omega})} \le CN^{-\frac{2}{d+1}},\end{aligned}$$

根据迹定理。

步骤 3. 对 E_G^b. 的估计利用迹定理以及解 ψ 在 Ω 上为零这一事实，我们有

$$\mathcal{E}_G^b \le \|\widehat{\psi}-\psi\|^2_{L^2([0,T]\times\partial\Omega)} \le C\|\widehat{\psi}-\psi\|^2_{H^1([0,T]\times\overline{\Omega})} \le C\|\widehat{\psi}-\psi\|^2_{W^{1,\infty}([0,T]\times\overline{\Omega})} \le CN^{-\frac{2}{d+1}}. \tag{3.22}$$

步骤 4. 结论。 综合以上所有估计，我们可推得

$$\mathcal{E}_G \le C\left(1+\|u\|^2_{W^{\max\{\ell,k,1\},\infty}([0,T]\times\Omega)}\right)(N^{-\frac{2}{d+1}}+\kappa(\epsilon)^2). \tag{3.23}$$

最后，取 N 充分大且 ϵ 充分小，可推得 $\mathrm{E}_G \le \delta$. □

4. 总误差估计。 本节将推导分数阶扩散方程的真实解与其神经网络近似之间总误差的界。总误差同时考虑了泛化误差，以及前文引入的正则化策略的一致性。本分析的一个关键步骤是控制平滑化神经网络输出在索博列夫范数下的近似误差。以下命题证明， H^k 范数下的误差可以用第 2 节定义的修正能量泛函来界定。这为我们的主要结果奠定了基础：总误差可以由第 3 节推导的泛化误差来控制，从而将训练表现与物理信息神经网络近似的整体精度联系起来。

PROPOSITION 4.1. 设 $\epsilon \in (0,1)$。设 ψ 是方程 (1.11) 的解，且 ψ_θ 是对 ψ的神经网络近似。则有

$$\|\psi - J_\epsilon\psi_\theta\|^2_{H^k([0,T]\times\Omega)} \le C\mathcal{E}[k,k;\theta]^2, \tag{4.1}$$

对于某个通用常数 $C>0$。

证明。 鉴于 $\mathrm{D}(\Lambda_D^j)$ 连续嵌入到 $H^j(\Omega)$ 中，且 $\partial_t^{(j)}(\psi - J_\epsilon\psi_\theta)$ 在 $\partial\Omega$ 上为零，我们有

$$\begin{aligned}\|\psi - J_\epsilon\psi_\theta\|^2_{H^k([0,T]\times\Omega)} &= \sum_{i=0}^{k}\sum_{j=0}^{k}\|D^j\partial_t^{(i)}(\psi - J_\epsilon\psi_\theta)\|^2_{L^2([0,T]\times\Omega)}\\ &\le C\sum_{i=0}^{k}\sum_{j=0}^{k}\|\Lambda_D^j\partial_t^{(i)}(\psi - J_\epsilon\psi_\theta)\|^2_{L^2([0,T]\times\Omega)}\\ &\le C\sum_{i=0}^{k}\|\Lambda_D^k\partial_t^{(i)}(\psi - J_\epsilon\psi_\theta)\|^2_{L^2([0,T]\times\Omega)}\\ &\le C\mathcal{E}[k,k;\theta]^2.\end{aligned} \tag{4.2}$$

□

Theorem 4.1. 令 $T>0$ 为任意正时间。令 $\psi_0 \in \bigcap_{m\ge 0}\mathcal{D}(\Lambda_D^m)$ 和 ψ 为满足齐次狄利克雷边界条件的方程 (1.11) 的相应光滑解。令 $\widehat{\psi}$ 为逼近 ψ 的神经网络。令 ℓ 和 k 为非负整数。则有

$$\mathcal{E}[\ell,k;\theta]^2 \le C[u]\mathcal{E}_G[\ell+k,k;\theta]^2, \tag{4.3}$$

其中常数 $C[u]$ 取决于 T、ℓ、k 以及 $W^{\max\{\ell,k\},\infty}$ 范数的 u。

证明。 我们对 k 进行归纳证明。对于基步（$k=0$），总误差为

$$\mathcal{E}[\ell,0;\theta] = \int_0^T \|\Lambda_D^\ell(\psi - J_\epsilon\widehat{\psi})\|^2_{L^2}dt. \tag{4.4}$$

从 ψ 所满足的偏微分方程中减去残差 $\mathrm{R}_i := \mathrm{R}_i[\theta]$，得到

$$\partial_t(\psi - J_\epsilon\widehat{\psi}) + u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}) + \Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi}) = -\mathcal{R}_i. \tag{4.5}$$

需要指出的是，$\mathrm{R}_i \notin \mathrm{D}(\Lambda_D^\ell)$，因为它在 $\partial\Omega$ 上不一定为零。因此，我们不能对 $\psi - J_\epsilon\widehat{\psi}$ 所满足的方程（4.5）应用 Λ_D^ℓ。解决方法是利用以下事实：$\Lambda_D^\ell = (-\Delta_D)^{\frac{\ell}{2}}$，当 ℓ 为偶数时这是一个局部算子；还要利用以下事实：当 ℓ 为奇数时，$\Lambda_D^\ell = \Lambda_D\Lambda_D^{\ell-1}$ 相当于 $\overline{L}^2$ 中的 $\nabla(-\Delta_D)^{\frac{\ell-1}{2}}$。由于 ∇ 和 Δ 是局部算子，只要 R_i 足够光滑，我们就可以将它们应用于 R_i。我们用这一技巧处理以下两种情况：

情况 1。 ℓ **为偶数**。我们将 $(-\Delta)^{\frac{\ell}{2}}$ 应用于（4.5），得到

$$\partial_t(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) + (-\Delta)^{\frac{\ell}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi}) + (-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi})) = -(-\Delta)^{\frac{\ell}{2}}\mathcal{R}_i. \tag{4.6}$$

将后者乘以 $(-\Delta)^{\frac{\ell}{2}}$，并在 Ω 上积分。由于 $\psi - J_\epsilon\widehat{\psi} \in \mathrm{D}(\Lambda_D^\ell)$ 和 $\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi}) \in \mathrm{D}(\Lambda_D^\ell)$，因此有

$$(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) = (-\Delta_D)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) = \Lambda_D^\ell(\psi - J_\epsilon\widehat{\psi}), \tag{4.7}$$

and

$$(-\Delta)^{\frac{\ell}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi}) = (-\Delta_D)^{\frac{\ell}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi}) = \Lambda_D^\ell\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi}), \tag{4.8}$$

yielding
(4.9)
$$\frac{1}{2}\frac{d}{dt}\|\Lambda_D^{\ell}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 + \|\Lambda_D^{\ell+\frac{\alpha}{2}}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2$$
$$= -\int_\Omega (-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}))(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) - \int_\Omega (-\Delta)^{\frac{\ell}{2}}\mathcal{R}_i(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}).$$

记交换子为

$$[(-\Delta)^{\frac{\ell}{2}}, u\cdot\nabla](\psi - J_\epsilon\widehat{\psi}) := (-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi})) - u\cdot\nabla(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}), \tag{4.10}$$

并利用消去律

$$\int_\Omega u\cdot\nabla(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi})\cdot(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) = 0 \tag{4.11}$$

该消去律源于 u所满足的无散条件，因此我们可以将速度项写成如下形式，

$$\begin{aligned}\mathcal{U} &:= -\int_\Omega (-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}))(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi})\\ &= -\int_\Omega [(-\Delta)^{\frac{\ell}{2}}, u\cdot\nabla](\psi - J_\epsilon\widehat{\psi})(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}).\end{aligned} \tag{4.12}$$

利用交换子估计

$$\|[(-\Delta)^{\frac{\ell}{2}}, u\cdot\nabla](\psi - J_\epsilon\widehat{\psi})\|_{L^2} \le C\|u\|_{W^{\ell,\infty}}\|\psi - J_\epsilon\widehat{\psi}\|_{H^\ell}, \tag{4.13}$$

以及椭圆正则性估计

$$\|\psi - J_\epsilon\widehat{\psi}\|_{H^\ell} \le C\|\Lambda_D^{\ell}(\psi - J_\epsilon\widehat{\psi})\|_{L^2} \tag{4.14}$$

该估计源于 $\mathsf{D}(\Lambda_D^{\ell})$ 连续嵌入 H^ℓ，我们可推得

$$\mathcal{U} \le C\|u\|_{W^{\ell,\infty}}\|\Lambda_D^{\ell}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2. \tag{4.15}$$

因此，我们得到微分不等式

$$\begin{aligned}&\frac{d}{dt}\|\Lambda_D^{\ell}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 + \|\Lambda_D^{\ell+\frac{\alpha}{2}}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2\\ &\quad\le C\|(-\Delta)^{\frac{\ell}{2}}\mathcal{R}_i\|_{L^2}^2 + C(1 + \|u\|_{W^{\ell,\infty}})\|\Lambda_D^{\ell}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2.\end{aligned} \tag{4.16}$$

根据Gronwall不等式，我们推得

$$\mathcal{E}[\ell, 0; \theta]^2 \le C_T[u]\mathcal{E}_G[\ell, 0; \theta]^2, \tag{4.17}$$

其中， $C_T[u]$ 是一个仅取决于 T和 u的 $W^{\ell,\infty}$范数的常数。

情形 2。 ℓ 是奇数。应用 $\nabla(-\Delta)^{\ell}$ 至（4.5），我们得到

$$\begin{aligned}&\partial_t\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi}) + \nabla(-\Delta)^{\frac{\ell-1}{2}}\Lambda_D^{\alpha}(\psi - J_\epsilon\widehat{\psi})\\ &\quad + \nabla(-\Delta)^{\frac{\ell-1}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi})) = -\nabla(-\Delta)^{\frac{\ell-1}{2}}\mathcal{R}_i.\end{aligned} \tag{4.18}$$

将后者与 $\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})$ 取 L^2 内积，得到（4.19—
）

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 + \int_\Omega \nabla(-\Delta)^{\frac{\ell-1}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi})\cdot\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\\
&\quad = -\int_\Omega \nabla(-\Delta)^{\frac{\ell-1}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}))\cdot\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\\
&\qquad - \int_\Omega \nabla(-\Delta)^{\frac{\ell-1}{2}}\mathcal{R}_i\cdot\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi}).
\end{aligned}$$

由于 $\psi - J_\epsilon\widehat{\psi}\in \mathrm{D}(\Lambda_D^\ell)$，我们有

$$\frac{d}{dt}\|\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 = \frac{d}{dt}\|\Lambda_D\Lambda_D^{\ell-1}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 = \frac{d}{dt}\|\Lambda_D^\ell(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2. \tag{4.20}$$

由于对于任意 $k\in\mathbb{N}$都有 $\psi - J_\epsilon\widehat{\psi}\in \mathrm{D}(\Lambda_D^k)$，我们有 $(-\Delta)^{\frac{\ell-1}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi})|_{\partial\Omega} = 0$，因此可以通过分部积分得到恒等式

$$\begin{aligned}
&\int_\Omega \nabla(-\Delta)^{\frac{\ell-1}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi})\cdot\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\\
&\quad = -\int_\Omega (-\Delta)^{\frac{\ell-1}{2}}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi})\Delta(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\\
&\quad = \int_\Omega \Lambda_D^{\ell-1}\Lambda_D^\alpha(\psi - J_\epsilon\widehat{\psi})\Lambda_D^2\Lambda_D^{\ell-1}(\psi - J_\epsilon\widehat{\psi})\\
&\quad = \|\Lambda_D^{\ell+\frac{\alpha}{2}}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2.
\end{aligned} \tag{4.21}$$

对于 u项，我们采用与 ℓ 为偶数时类似的技巧，并利用交换子推导出以下估计，

$$\begin{aligned}
&-\int_\Omega \nabla(-\Delta)^{\frac{\ell-1}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}))\cdot\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\\
&\quad = -\int_\Omega [\nabla(-\Delta)^{\frac{\ell-1}{2}}, u\cdot\nabla](\psi - J_\epsilon\widehat{\psi})\cdot\nabla(-\Delta)^{\frac{\ell-1}{2}}(\psi - J_\epsilon\widehat{\psi})\\
&\quad \le C\|u\|_{W^{\ell,\infty}}\|\psi - J_\epsilon\widehat{\psi}\|_{H^\ell}^2 \le C\|u\|_{H^\ell}\|\Lambda_D^\ell(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2.
\end{aligned} \tag{4.22}$$

结合这些估计并应用 Gronwall不等式，我们推得

$$\mathcal{E}[\ell,0;\theta]^2 \le C_T[u]\mathcal{E}_G[\ell,0;\theta]^2, \tag{4.23}$$

当 ℓ为奇数时。这里， $C_T[u]$是一个仅依赖于 T和 $W^{\ell,\infty}$范数的常数。现在我们假设存在一个仅依赖于 T、 k 和$\|u\|^2_{W^{\max\{\ell,k-1\},\infty}([0,T]\times\Omega)}$的常数 $C_{T,k-1}[u]$ ，使得

$$\mathcal{E}[\ell,k-1;\theta]^2 \le C_{T,k-1}[u]\mathcal{E}_G[\ell+k-1,k-1;\theta]^2, \tag{4.24}$$

对所有 $\ell\in\mathbb{N}$ 均成立。我们证明

$$\mathcal{E}[\ell,k;\theta]^2 \le C_{T,k}[u]\mathcal{E}_G[\ell+k,k;\theta]^2 \tag{4.25}$$

对所有 $\ell\in\mathbb{N}$ 均成立。为此，我们固定 $\ell\in\mathbb{N}$ 并区分以下两种情况：

情况 1. ℓ **为偶数**。 在这种情况下，我们有

$$\begin{aligned}
&\partial_t^{(k+1)}(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) + \partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}\Lambda^\alpha(\psi - J_\epsilon\widehat{\psi})\\
&\quad = -\partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}\mathcal{R}_i - \partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi})).
\end{aligned} \tag{4.26}$$

我们指出， $\partial_t^{(k)}$ ， ∇ 和 $(-\Delta)^{\frac{\ell}{2}}$ 都是局部算子，并且彼此可交换。将后者乘以 $\partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi})$ 并在 Ω 上积分，得到

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\Lambda_D^{\ell}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 + \|\Lambda_D^{\ell+\frac{\alpha}{2}}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 \\
&\quad = -\int_\Omega \partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}\mathcal{R}_i\partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) \\
&\qquad -\int_\Omega \partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}))\partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}).
\end{aligned} \tag{4.27}$$

根据莱布尼茨定理，我们将 u 项分解如下，

$$\begin{aligned}
&-\int_\Omega \partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(u\cdot\nabla(\psi - J_\epsilon\widehat{\psi}))\partial_t^{(k)}(-\Delta)^{\frac{\ell}{2}}(\psi - J_\epsilon\widehat{\psi}) \\
&\quad = -\sum_{i=0}^{k}\binom{k}{i}\int_\Omega(-\Delta)^{\frac{\ell}{2}}\left[\partial_t^{(i)}u\cdot\nabla\partial_t^{(k-i)}(\psi - J_\epsilon\widehat{\psi})\right](-\Delta)^{\frac{\ell}{2}}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi}) \\
&\quad = -\sum_{i=1}^{k}\binom{k}{i}\int_\Omega(-\Delta)^{\frac{\ell}{2}}\left[\partial_t^{(i)}u\cdot\nabla\partial_t^{(k-i)}(\psi - J_\epsilon\widehat{\psi})\right](-\Delta)^{\frac{\ell}{2}}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi}) \\
&\qquad -\int_\Omega(-\Delta)^{\frac{\ell}{2}}\left[u\cdot\nabla\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})\right](-\Delta)^{\frac{\ell}{2}}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi}) \\
&\quad := \mathcal{U}_1 + \mathcal{U}_2.
\end{aligned} \tag{4.28}$$

利用索博列夫乘积估计以及 $\mathrm{D}(\Lambda_D^m)$ 到 H^m 的连续嵌入（对任意 $m\in\mathbb{N}$），我们估计

$$\begin{aligned}
\mathcal{U}_1 &\le C\sum_{i=1}^{k}\|\partial_t^{(i)}u\|_{W^{\ell,\infty}}\|\partial_t^{(k-i)}(\psi - J_\epsilon\widehat{\psi})\|_{H^{\ell+1}}\|\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})\|_{H^{\ell}} \\
&\le C\sum_{i=1}^{k}\|\partial_t^{(i)}u\|_{W^{\ell,\infty}}^2\|\Lambda_D^{\ell+1}\partial_t^{(k-i)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2 + \|\Lambda_D^{\ell}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2.
\end{aligned} \tag{4.29}$$

我们指出，

$$\begin{aligned}
&\int_0^T\sum_{i=1}^{k}\|\partial_t^{(i)}u\|_{W^{\ell,\infty}}^2\|\Lambda_D^{\ell+1}\partial_t^{(k-i)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2dt \\
&\le C\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}^2\int_0^T\sum_{i=1}^{k}\|\Lambda_D^{\ell+1}\partial_t^{(k-i)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2dt \\
&\le C\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}^2\mathcal{E}[\ell+1,k-1;\theta]^2 \\
&\le C\|u\|_{W^{\max\{\ell,k\},\infty}([0,T]\times\Omega)}^2\mathcal{E}_G[\ell+k,k-1;\theta]^2,
\end{aligned} \tag{4.30}$$

根据归纳假设。对于项 U_2, ，我们利用消去律

$$\int_\Omega u\cdot\nabla[(-\Delta)^{\frac{\ell}{2}}\partial_t^{k}(\psi - J_\epsilon\widehat{\psi})](-\Delta)^{\frac{\ell}{2}}\partial_t^{k}(\psi - J_\epsilon\widehat{\psi}) = 0, \tag{4.31}$$

将 U_2 表示为交换子 $[(-\Delta)^{\frac{\ell}{2}}, u\cdot\nabla]$，如下所示，

$$\mathcal{U}_2 = -\int_\Omega[(-\Delta)^{\frac{\ell}{2}}, u\cdot\nabla]\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})(-\Delta)^{\frac{\ell}{2}}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi}), \tag{4.32}$$

然后我们利用交换子估计将 U_2 控制为

$$\mathcal{U}_2 \le C\|u\|_{W^{\ell,\infty}}\|\Lambda_D^{\ell}\partial_t^{(k)}(\psi - J_\epsilon\widehat{\psi})\|_{L^2}^2. \tag{4.33}$$

综合所有这些估计并应用Gronwall不等式，可得

$$\mathcal{E}[\ell,k;\theta]^2 \le C_{T,k}[u]\mathcal{E}_G[\ell+k,k;\theta]^2, \tag{4.34}$$

对于某个正常数 $C_{T,\ k}[u]$ 该常数仅取决于 T，k，ℓ 以及 $W^{\max\{\ell,k\},\ \infty}$的 $[0,T]\times\Omega$ 范数。

情形 2. ℓ **为奇数**。这种情形与前面的情形类似。为避免重复，这里省略证明。□

5. 总结。本文为定义在具有谱狄利克雷边界条件的有界区域上的分数阶扩散方程，建立了物理信息神经网络的收敛理论。通过引入一种谱定义的光滑化策略，我们确保了边界相容性，并在标准索博列夫范数下推导出严格的误差估计。我们的分析以偏微分方程理论和泛函分析中的工具为基础，而非数值离散格式，并为物理信息神经网络的分析研究日益丰富的成果添砖加瓦，尤其是在经典方法面临结构性挑战的非局部问题中。

本文采用的光滑化策略允许用神经网络在任意时空索博列夫空间中逼近解。使用截断函数时，这种逼近则无法实现。事实上，将神经网络乘以截断函数 χ_ϵ 会使其在边界上消失，但由此需要在索博列夫空间中适当估计差值 $\Lambda_D^{\alpha}(\chi_\epsilon\psi-\psi)$。更准确地说，在任意 $\mathrm{D}(\Lambda_D^{j})$ 空间中，$\chi_\epsilon\psi$ 与 ψ 之间的距离都需要足够小。这在 L^2 中显然可以实现，但在 $\mathrm{D}(\Lambda^{j})$ 中则未必。例如，当 $j=1$ 时，有

$$\|\Lambda_D(\chi_\epsilon\psi-\psi)\|_{L^2} = \|\nabla((\chi_\epsilon-1)\psi)\|_{L^2}^2. \tag{5.1}$$

但是，当 ϵ 趋近于 0 时，χ_ϵ 的梯度会发散，因此，是否存在一个足够小的 ϵ，使得后一个范数足够小，并不清楚。相比之下，我们知道 $J_\epsilon\psi$ 在 $\mathrm{D}(\Lambda^{j})$ 中收敛于 ψ，因此优先使用 J_ϵ 而非 χ_ϵ。

在未来工作中，将这一框架推广到更广泛的偏微分方程类别会很有意义，这些方程涉及某些非线性因素以及不同类型的非局部算子。

致谢。L.C.得到中国国家重点研发计划项目（编号：2021YFA1003001）和国家自然科学基金项目（编号：12271537）的部分资助。R.H.得到美国海军研究办公室项目（编号：N00014-24-1-2432）、西蒙斯基金会（MP-TSM-00002783）以及美国国家科学基金项目（编号：DMS-2420988）的部分资助。

附录A。整体光滑解的存在性与唯一性。

THEOREM A.1. 设 $\alpha\in[0,2]$。设 $T>0$。设 u 为一个光滑的无散向量场，使得 $u\cdot n|_{\partial\Omega}=0$。设 f 为给定的光滑函数，使得对于任意 $k\in\mathbb{N}$，$(-\Delta)^k f|_{\partial\Omega}=0$。令 $\psi_0\in\bigcap_{k=0}^{\infty}\mathcal{D}(\Lambda_D^k)$。则对流-扩散方程

$$\partial_t\psi + u\cdot\nabla\psi + \Lambda_D^{\alpha}\psi = f, \tag{A.1}$$

配备齐次狄利克雷边界条件和初始数据 ψ_0，存在唯一的解 ψ，满足

$$\psi\in L^{\infty}(0,T;\mathcal{D}(\Lambda_D^k)), \tag{A.2}$$

对所有 $k \in \mathbb{N}$ 均成立。

证明。 对于 $\epsilon \in (0, 1)$，我们考虑正则化方程

$$\partial_t \psi^\epsilon + u \cdot \nabla \psi^\epsilon + \Lambda_D^\alpha \psi^\epsilon - \epsilon \Delta \psi^\epsilon = f, \tag{A.3}$$

其初始数据为

$$\psi^\epsilon|_{\partial\Omega} = 0, \tag{A.4}$$

以及边界条件为

$$\psi^\epsilon(x, 0) = \psi_0(x). \tag{A.5}$$

这些系统具有全局光滑解

$$\psi^\epsilon \in L^\infty(0, T; \mathcal{D}(\Lambda_D^k)), \tag{A.6}$$

对于任意 $k \in \mathbb{N}$，这一事实可由经典的 Galerkin 逼近格式以及通过紧性论证取极限得到（例如参见 [1]）。我们为这些正则化系统的解建立与 ϵ 无关的界，并推导出解的存在性

$$\psi \in L^\infty(0, T; \mathcal{D}(\Lambda_D^k)), \tag{A.7}$$

对于任意 $k \in \mathbb{N}$。事实上， ψ^ϵ 的 $\mathrm{D}(\Lambda_D^k)$ 范数按照以下方式演化：

$$\frac{1}{2}\frac{d}{dt}\|\Lambda_D^k \psi^\epsilon\|_{L^2}^2 + \|\Lambda_D^{k+\frac{\alpha}{2}} \psi^\epsilon\|_{L^2}^2 = -\int_\Omega \Lambda_D^k (u \cdot \nabla \psi^\epsilon) \Lambda_D^k \psi^\epsilon dx + \int_\Omega \Lambda_D^k f \Lambda_D^k \psi^\epsilon dx. \tag{A.8}$$

我们指出，对于所有 $k \in \mathbb{N}$， $u \cdot \nabla \psi^\epsilon \in \mathrm{D}(\Lambda_D^k)$，因为

$$u \cdot \nabla \psi^\epsilon = -\partial_t \psi^\epsilon - \Lambda_D^\alpha \psi^\epsilon + \epsilon \Delta \psi^\epsilon + f, \tag{A.9}$$

其中，右端对所有 $k \in \mathbb{N}$都属于 $\mathrm{D}(\Lambda_D^k)$。

为了估计 u 中的项，我们使用抵消律

$$\int_\Omega u \cdot \nabla \Lambda_D^k \psi^\epsilon \Lambda_D^k \psi^\epsilon dx = 0, \tag{A.10}$$

当 k 为偶数时，并且

$$\int_\Omega u \cdot \nabla \nabla \Lambda_D^{k-1} \cdot \nabla \Lambda_D^{k-1} \psi^\epsilon dx = 0, \tag{A.11}$$

当 k 为奇数时，并且我们得到

$$-\int_\Omega \Lambda_D^k (u \cdot \nabla \psi^\epsilon) \Lambda_D^k \psi^\epsilon dx = -\int_\Omega [\Lambda_D^k, u \cdot \nabla] \psi^\epsilon \Lambda_D^k \psi^\epsilon dx \le C \|u\|_{W^{k,\infty}} \|\Lambda_D^k \psi^\epsilon\|_{L^2}^2, \tag{A.12}$$

当 k 为偶数且

$$-\int_\Omega \Lambda_D^k (u \cdot \nabla \psi^\epsilon) \Lambda_D^k \psi^\epsilon dx = -\int_\Omega [\nabla \Lambda_D^{k-1}, u \cdot \nabla] \psi^\epsilon \nabla \Lambda_D^{k-1} \psi^\epsilon dx \le C \|u\|_{W^{k,\infty}} \|\Lambda_D^k \psi^\epsilon\|_{L^2}^2, \tag{A.13}$$

当 k 为奇数时。最后两个估计可通过展开交换子、应用 H¨older 不等式，并在奇数情形下使用恒等式 $-\nabla \cdot \nabla = -\Delta = \Lambda_D^2$ 得出。

由于 A.10 和 A.11 成立， u 满足无散性质，且边界假设 $u \cdot n|_{\partial\Omega} = 0$ 成立，因此有上述消去关系。对这一性质的需求，正是采用能够保持无散向量场输运的抛物型正则化方案的理由；相比之下，伽辽金近似会破坏上述结构。最后，我们应用Gronwall不等式，推得

$$\begin{aligned} &\sup_{0\le t\le T} \|\Lambda_D^k \psi^\epsilon(t)\|_{L^2}^2 \\ &\qquad \le \left(\|\Lambda_D^k \psi_0\|_{L^2}^2 + C\int_0^T \|\Lambda_D^{k-\frac{\alpha}{2}} f\|_{L^2}^2 dt \right) \exp\left\{ C\int_0^T \|u\|_{W^{k,\infty}} dt \right\}. \end{aligned} \tag{A.14}$$

我们不再赘述。 □

REMARKA.1. 我们指出，根据偏微分方程 *A.1*，定理 (A.1)中得到的解在空间上的光滑性意味着其在时间上也光滑。因此，唯一解 ψ 在 C^∞ 空间和时间上均光滑。

参考文献

[1] E. ABDO, *On the regularity of solutions to some active scalar equations on bounded domains*, 《SIAM 数学分析杂志》, 56（2024）, 第 7306–7326 页。[2] F. ABDO, L. CHAL, B. HU, AND X. YANG, *Error estimates of physics-informed neural networks for approximating boltzmann equation*, 即将发表于《IMA 数值分析杂志》; arXiv:2407.08383, （2024）。[3] E. ABDO, R. HU, AND Q. LIN, *Neural network solutions to the critical sqg equations via ap-proximating nonlocal periodic operators*, 《物理学 D：非线性现象》, 476（2025）, 第 134652 页。[4] F. ABDO AND Q. LIN, *On the dirichlet fractional laplacian and applications to the SQG equation on bounded domains* 2409 2024, arXiv 预印本 arXiv: .05209, （）。[5] C. BECK, M. HUTZENTHALER, A. JENTZEN, AND B. KUCKUCK, *An overview on deeplearning-based approximation methods for partial differential equations*, arXiv 预印本 arXiv:2012.12348, （2020）。[6] D. A. BENSON, S. W. WHEATCRAFT, AND M. M. MEERSCHAERT, *Application of a fractional advection-dispersion equation* 36 2000 1403 1412, 《水资源研究》, （）, 第 – 页。[7] A. BISWAS, J. TIAN, AND S. ULUSOY, *Error estimates for deep learning methods in fluid dynamics* 151 2022 753 777, 《数值数学》, （）, 第 – 页。[8] L. A. CAFFARELLI AND P. R. STINGA, *Fractional elliptic equations caccioppoli estimates and regularity*, 《亨利·庞加莱研究所年鉴 C：非线性分析》, 第 Elsevier 卷, 2016 年, 第 767–807 页。[9] Z. GAO, J. CHEN, AND M. LIU, *Least-squares ReLU neural network (LSNN) method for linear advection-reaction equation* 2021 110514, 《计算物理学杂志》, （）, 第 页。[10] A. CARTEA AND D. DEL CASTILLO-NEGRETE, *Fractional diffusion models of option prices in markets with jumps*, 《物理学 A：统计力学及其应用》, 374（2007）, 第 749–763 页。[11] W. CHEN, G. DIMARCO, AND L. PARESCHI, *Structure and asymptotic preserving deep neural surrogates for uncertainty quantification in multiscale kinetic equations*, arXiv 预印本 arXiv:2506.10636, （2025）。[12] P. CONSTANTIN, T. ELGINDI, M. IGNATOVA, AND V. VICOL, *On some electroconvection models*, 《非线性科学杂志》, 27（2017）, 第 197–211 页。[13] W. E AND B. YU, *The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems* 6 2018 1 12, 《数学与统计学通讯》, （）, 第 – 页。[14] G. GILBOA AND S. OSHER, *Nonlocal operators with applications to image processing*, 《多尺度建模与仿真》, 7（2009）, 第 1005–1028 页。[15] B. HAO, U. BRAGA-NETO, C. LIU, L. WANG, AND M. ZHONG, *Structure preserving pinn for solving time dependent pdes with periodic boundary* 2404 16189, 2024, arXiv 预印本 arXiv:。（）。[16] R. HU, Q. LIN, A. RAYDAN, AND S. TANG, *Higher-order error estimates for physics-informed neural networks approximating the primitive equations*, 《偏微分方程及其应用》, 4（2023）, 第 34 页。[17] S. JIN, Z. MA, AND K. WU, *Asymptotic-preserving neural networks for multiscale time-dependent linear transport equations* 94 2023 57, 《科学计算杂志》, （）, 第 页。

[18] S. Jin, Z. Ma and T.-a. Zhang *Asymptotic-preserving neural networks for multiscale vlasov–*, , 《科学计算杂志》, *poisson–fokker–planck system in the high-field regime* 99 （2024）, 第 61 页。[19] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. and Bhattacharya, A. Stuart, A. Anandkumar ,

Fourier neural operator for parametric partial differential equations, arXiv 预印本 arXiv:2010.08895 （2020）。[20] J. Lu and M. Wang, *Neural network based variational methods for solving quadratic porous medium equations in high dimensions* 2205 02927, 2022 ,

arXiv 预印本 arXiv:。（）。[21] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, *Learning nonlinear operatorsvia deeponet based on the universal approximation theorem of operators*, 《自然·机器智能》, 3（2021）, 第 218–229 页。[22] R. Metzler and J. Klafter, *The random walk's guide to anomalous diffusion: a fractional dynamics approach* 339 2000 1 77 , 《物理学报告》, （）, 第 – 页。

[23] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Physics-informed neural networks: A deep 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。[24] J. Sirignano and K. Spiliopoulos, *DGM: A deep learning algorithm for solving partial differential equations* 375 2018 1339 1364 , 《计算物理学杂志》, （）, 第 – 页。[25] J. Wan and L. Liu, *Error estimates of asymptotic-preserving neural networks in approximating stochastic linearized boltzmann equation* 2503 01643, 2025 ,

arXiv 预印本 arXiv:。（）。[26] Y. Zang, G. Bao, X. Ye, and H. Zhou, *Weak adversarial networks for high-dimensional partial differential equations* 2020 109409 , 《计算物理学杂志》, （）, 第 页。[27] Z. Zhu, Y. Huang, and L. Liu, *Physicssolver: Transformer-enhanced physics-informed neural networks for forward and forecasting problems in partial differential equations*, arXiv 预印本 arXiv:2502.19290, （2025）。