用于形式化不确定性传播与概率评估的非线性连续时间系统学习

Peter Amorese
Peter.Amorese@colorado.edu 科罗拉多大学博尔德分校 美国科罗拉多州博尔德

Morteza Lahijanian
Morteza.Lahijanian@colorado.edu
科罗拉多大学博尔德分校 美国科罗拉多州博尔德

arXiv:2602.05103v1 [eess.SY] 4 Feb 2026

摘要

非线性常微分方程（ODE）是对现实世界中的动力系统进行建模的有力工具。然而，将初始状态不确定性通过非线性动力学进行传播，尤其是在常微分方程未知且由数据学习得到时，仍是一项重大挑战。本文提出一种用于模型学习的全新连续体动力学视角，通过构造概率事件的泰勒级数近似，实现形式化不确定性传播。我们建立了该方法可靠性的充分条件，并证明了其渐近收敛性。实证结果表明，该框架十分有效，尤其是在预测罕见事件方面。

关键词

不确定性传播、非线性常微分方程、泰勒级数近似、连续体动力学、可达性

ACM 引用格式：
PeterAmorese 和 MortezaLahijanian。2026。用于形式化不确定性传播与概率评估的非线性连续时间系统学习。载于 .ACM，美国纽约州纽约市，共 11 页。https://doi.org/10.1145/ nnnnnnn.nnnnnnn

1 引言

对现实世界动力系统进行建模，是预测其未来行为的核心要素。具体而言，非线性常微分方程（ODE）能够描述许多连续空间、连续时间系统的动力学 [11, 25]。当常微分方程未知时，可以利用机器学习方法从数据中生成准确的估计。在这类情况下，模型的初始条件往往也存在不确定性，这可能导致截然不同的预测。然而，使用这类模型预测系统的未来行为，尤其是在初始条件存在不确定性的情况下，颇具挑战，因为非线性常微分方程通常不存在解析解。本文旨在探究模型选择与不确定性下预测之间的联系。具体而言，我们能否学习一种非线性常微分方程模型，使形式化的不确定性传播更易实现？我们旨在从连续动力学视角探讨不确定性传播与建模问题，以回答这一问题。

不确定性传播通常用于概率评估，即预测在未来某个关注时刻 τ，系统状态$\mathbf{x}$ 位于某个区域 R 内的概率。这项任务结合了两个数学上具有挑战性的步骤：(i) 将概率密度函数（pdf）通过非线性模型进行传播；(ii) 随后对所得概率密度函数进行积分 R。直接传播概率密度函数会面临与求解非线性常微分方程相同的解析难题，因此通常需要进行近似。此外，即使假设密度具有易于处理的表达式，对其进行积分 R 也可能难以实现，因此还需要另一种近似技术。

密度传播是一个经过充分研究的问题，由刘维尔偏微分方程 [5, 18]所支配。物理信息方法通过物理信息学习 [15, 17, 19, 21]，将刘维尔方程的近似解训练为神经网络或生成模型。然而，对于概率评估而言，概率密度函数的候选模型（例如神经网络）通常无法进行解析积分，因此难以直接用于概率评估。为了量化概率事件，研究人员曾结合高斯混合模型与动力学的线性近似来重建可积分的传播后密度[26], ，但这些方法难以为概率事件提供形式化上界。基于矩的多项式优化技术 [10, 24]能够为多项式常微分方程的概率事件提供形式化上界，但仅适用于有界支撑初始分布。基于离散化的可达性方法利用重叠的前向可达集来估计概率 [13]；然而，由于对可达集进行严格过近似，这类方法难以提供收敛保证。

基于采样的粒子模拟是通过非线性常微分方程进行不确定性传播的一种有效近似方法 [29]。与密度传播方法不同，基于采样的方法可以非常直接地近似传播和概率评估问题。可以对初始状态进行采样，并通过数值积分向前模拟每个样本。在 τ 时刻，最终落在 R 中的样本数与样本总数之比，可以估计所需的概率。尽管这类方法简单易用，却无法为所得概率估计提供正式的严格上界，因此只能依赖基于置信度的集中不等式。此外，在预测稀有事件 [7]，（即概率在 10^{-6} 量级及以下）时，这类方法的准确性会大幅下降。

特此授权免费制作本作品全部或部分内容的数字副本或纸质副本，供个人或课堂使用，但不得为营利或商业利益而制作或传播副本，且副本须附有本声明及首页完整引文。作者以外人员拥有的本作品组成部分的版权必须予以尊重。注明出处后，允许摘编。若要以其他方式复制、再版、发布至服务器或向列表重新分发，须事先获得具体许可和/或支付费用。请通过 per missions@acm.org 申请许可。*Conference'17*，美国华盛顿特区 © 2026 版权所有者/作者所有。出版权已授权给 ACM。ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM https://doi.org/10.1145/nnnnnnn.nnnnnnn

在建模方面，也就是学习常微分方程方面，SINDy [6] 是一种重要框架，可用于根据数据高效地恢复系统的动力学。深度学习方法也可以将常微分方程学习为表达能力强的神经网络，并可直接用于通过连续归一化流 [8]近似密度传播。虽然这些方法能够学习到精度很高的模型，但它们并未考虑正式的不确定性传播与评估，因此通常必须依赖基于采样的方法。近期研究 [2] 尝试解决联合建模与不确定性传播问题；然而，这些方法仅适用于离散时间系统。

本文通过引入一种基于连续体动力学的简化方法来处理概率密度函数传播与评估问题，并使学习得到的常微分方程架构满足传播流程的要求，从而应对概率密度函数传播所面临的挑战。通过巧妙的域变换，我们可以将复杂的状态预测与积分问题转化为对一个标量函数的近似，以此捕捉整个时间过程中的事件概率。我们构造了概率函数的泰勒级数近似，对模型的概率预测给出形式化上界。我们深入分析了学习得到的模型架构所受的限制，并提出了一种具有强渐近收敛保证的学习框架。

本文的主要贡献有四个方面：

i. 一种新颖的、基于连续体动力学的不确定性传播方法；ii. 对该方法的可靠性与建模约束进行严格分析；iii. 一种利用现成可达性工具来改进较长时间范围内预测结果的算法；以及 iv. 该方法的实验结果与分析。

1.1 符号说明

向量用粗体表示，例如 $\mathbf{x} = [x_0, \ldots, x_n]$。对于向量 $y \in \mathbb{R}^n$，我们用 i维表示其第 y_i 个分量。长度为 n 的全 1 向量记为 $\mathbf{1}_n$。令 $\mathrm{C}^\omega(U)$ 表示集合 U 上所有实值解析函数构成的集合。

2 问题表述

考虑由下式描述的动力系统

$$\dot{\mathbf{x}} = f(\mathbf{x}) \tag{1}$$

其中 $\mathbf{x} \in \mathbb{R}^n$，且 $f\colon \mathbb{R}^n \to \mathbb{R}^n$ 是向量场，并且可能是非线性的。我们假设 f 是全局利普希茨连续的，且每个 $f_i \in \mathrm{C}^\omega(\mathbb{R}^n)$。从初始状态 $\mathbf{x}_0$, 出发，系统沿着唯一的轨迹流函数 $\phi(\mathbf{x}_0, t)$ 演化，其中 $t \in \mathbb{R}_{\geq 0}$和 ϕ 满足常微分方程

$$\frac{\partial \phi}{\partial t}(\mathbf{x}_0, t) = f(\phi(\mathbf{x}_0, t)). \tag{2}$$

本文聚焦于 f未知，且初始状态 $\mathbf{x}_0$ 具有不确定性的情形，其由分布 $p_0(\mathbf{x}_0)$ 描述。具体而言，我们假设 $p_0(\mathbf{x}_0)$ 服从均值为 $p_0 \quad _0 = \quad \mu \Sigma \qquad \Sigma \qquad _0 \sim \quad _1$、（非退化）协方差矩阵为 Σ 的正态分布，即 $\mathbf{x}(\mathbf{x})\ \mathrm{N}(,)$，并且不失一般性地假设其为对角矩阵。

[1]利用（线性）马氏变换，总能将系统 (1) 映射到一个协方差空间，使初始状态的协方差为对角矩阵。

由于 f 未知，我们假设可以获得形式为 $\mathrm{D} = \{(\hat{}\mathbf{x}_i, f(\hat{}\mathbf{x}_i))\}_{i\in\mathbb{N}}$ 的数据集，从而能够学习得到 f 的近似函数。令 $\hat{} f$ 表示学习得到的向量场，$\hat{} \phi(\mathbf{x}\ t)$ 表示其对应的流函数解。

我们的总体目标是确定系统 (1) 在特定时间 $t = \tau$ 位于给定感兴趣区域 $R \subset \mathbb{R}^n$ 内的概率（上界），记为 $P(\phi(\mathbf{x}_0, \tau) \in R)$。这通常涉及两个步骤：学习模型，以及通过学习得到的模型传播不确定性。本文主要关注后者，即针对学习得到的 模型计算具有形式化保证的概率（上界）这一子问题，即 $P(\hat{}\phi(\mathbf{x}_0, \tau) \in R)$。$\hat{}$ 因此，通过确保学习得到的模型 f 具有通用近似性和渐近收敛性质，模型的预测也就继承了相同的收敛保证。形式上，我们可以如下定义收敛性质。

定义 1（收敛通用估计器）。设 $\mathrm{F}_\theta \subset \mathrm{C}^\omega(\mathbb{U}^n)$ 是一类参数化函数，$\mathrm{L}_\mathcal{D}\colon \mathrm{F}_\theta \to \mathbb{R}$ 是关于参数 θ 为凸的损失函数。令 $\hat{} f_\theta^\star \in \mathrm{F}_\theta$ 表示该损失函数的全局极小值，即

$$\hat{f}_\theta^\star = \arg\min_{\hat{f}_\theta \in \mathcal{F}_\theta} \mathcal{L}_\mathcal{D}(\hat{f}_\theta). \tag{3}$$

那么，对于任意闭超矩形 $U \subset \mathbb{U}^n$，如果$\hat{} f_\theta^\star$ 是收敛的通用估计器，则

$$\hat{f}_\theta^\star = \arg\min_{\hat{f}_\theta \in \mathcal{F}_\theta} \mathcal{L}_\mathcal{D}(\hat{f}_\theta) \tag{4a}$$

and

$$\lim_{|\theta|, |\mathcal{D}| \to \infty} \|\hat{f}_\theta^\star - f\|_{C^k} = 0. \tag{4b}$$

对于任意 $k > 0$，其中 $\|\cdot\|_{C^k}$表示 C^k-范数。

$\hat{}$ 给定一个学习得到的模型 f_θ，我们感兴趣的是传播$\mathbf{x}_0$ 中的不确定性，以预测状态在给定时间 $t = \tau$进入感兴趣区域 R 的概率，即

$$P(\hat{\phi}(\mathbf{x}_0, \tau) \in R) = \int_R p(\mathbf{x}_\tau) d\mathbf{x}_\tau \tag{5}$$

其中，$p(\mathbf{x}_\tau)$ 是系统在时间 $t = \tau$的状态分布，通过以下方式计-算得到：

$$\log p(\mathbf{x}_\tau) = \log p_0(\mathbf{x}_0) - \int_{t=0}^{t=\tau} \nabla \cdot \hat{f}_\theta(\mathbf{x}_t) dt. \tag{6}$$

$\hat{}\ \hat{}$ 由于 f_θ 是非线性的，使用式（5）和（6）预测 $P(\phi(\mathbf{x}_0, \tau) \in R)$ 通常会涉及两个（一般而言）难以求解的积分。

本文聚焦于通用模型与计算式（5）收敛的形式化上界的方法之间的耦合。形式-上，我们考虑如下所述的两部分问题。

问题 1. 考虑系统 (1)，其 f未知，且初始状态不确定 $\mathbf{x}_0 \sim \mathrm{N}(\mu, \Sigma)$. 给定数据 D 、感兴趣的超矩形区域 $R \subset \mathbb{R}^n$ 以及感兴趣的时间 τ，

$\hat{}$ (1) 学习收敛通用估计器模型 f_θ，以及 (2) 计算 $\bar{P}_\tau \in [0, 1]$，使得 $P(\hat{}\phi(\mathbf{x}_0, \tau) \in R) \leq \bar{P}_\tau$。

起初，模型学习问题1.1和不确定性传播问题 1.2 可能看起来是彼此分离的问题。在这项工作中，我们研究形式化不确定性传播问题，以

揭示 $\hat{}\, f_\theta$ 的一种结构，使得能够在不损害通用表征能力的前提下，对 $\bar{P}_\tau$ 进行可处理的计算。为便于 $\hat{}$ 展示，我们省略下标 θ，并假设 f 是一个参数化估计器。

3 不确定性传播

本节提出了一种计算概率界 $\bar{P}_\tau$ 的方法。根据该过程的要求，我们可以选择一个回归器 $\hat{}\, f$，从而实现对 $\bar{P}_\tau$ 的易处理计算。

3.1 控制质量视角

不确定性传播和积分问题也可以被视为跟踪一个控制质量的属性，即由一组固定的无限小物质元（例如状态）组成的集合；这些物质元在受 $\hat{}\, f$ [14]约束的情况下运动。非线性 $\hat{}\, f$ 会导致每个物质元的体积随时间发生变化，从而使控制质量的形状发生变形。

图 1 展示了控制质量相关的连续体动力学，以及密度随时间的演化。我们研究的是在时间 $\tau\hat{}$ 占据区域 R 的粒子所构成的控制质量 Ω_τ（最右侧图中的粉色方框）。由于 ϕ 是一个单参数群 [30],，因此可以直接定义为

$$\Omega_\tau(t) = \{\mathbf{x}_t \mid \hat{\phi}(\mathbf{x}_t, \tau - t) \in R\}. \tag{7}$$

确定 Ω_τ 的总质量等价于计算式 (5)。在图 1a 中，状态空间密度最初呈高斯分布，并根据 (6) 变形为复杂分布。相反，在 $t = 0$ 时，$\Omega_\tau(0)$ 以高度复杂的形状开始，因此在 $t = \tau$ 时，粒子占据了 R 的简单矩形形状。给定固定的 R，控制质量由时间点 τ 唯一确定，因为在 $t = \tau_1$ 占据 R 的物质元（或初始状态）集合，可能不同于在另一时间 $\tau_2 \neq \tau_1$ 占据 R 的物质元集合。

采用这一视角，计算 $P(\phi(\mathbf{x}_0, \tau) \in R)$ 有两种等价的方法：(i) 在 $t = \tau$ 处的一个简单矩形积分区域上，对复杂的变形密度进行积分：

$$\int_{R=\Omega_\tau(\tau)} p(\mathbf{x}_\tau) d\mathbf{x}_\tau \tag{8}$$

或者 (ii) 在一个复杂积分区域 $t = 0$ 上，对简单密度进行积分：

$$\int_{\Omega_\tau(0)} p_0(\mathbf{x}_0) d\mathbf{x}_0. \tag{9}$$

乍看之下，这两种方案似乎同样棘手。为阐明方案 (ii) 的优势，我们采用域变换使初始分布变为均匀分布，从而将式 (9) 中复杂的积分简化为计算 $\Omega_\tau(0)$ 的体积。

令 $\mathbb{U}^n = (0,1)^n$ 表示开单位超矩形。考虑关于 $p_0(\mathbf{x}_0)$ 的参数 μ 和 Σ 定义的微分同胚 ψ： $R^n \to \mathbb{U}^n$

$$\psi_i(\mathbf{x}) = c(x_i; \mu_i, \Sigma_{ii}) \tag{10}$$

其中，$c(\cdot; \mu_i, \Sigma_{ii})$ 是 $p_0(x_{0,i})$ 的累积分布函数。令 $\mathbf{u} = \psi(\mathbf{x})$ 表示 $\mathbb{U}^n$ 中得到的变换状态。由于 ψ 被构造为每个初始状态分量的累积分布函数，

$\mathbb{U}^n$ 中得到的初始分布是均匀分布，即 $p_0(\mathbf{u}_0) = \mathrm{U}((0,1)^n)$。因此，(9) 可简化为

$$\int_{\Omega_\tau(0)} p_0(\mathbf{x}_0) d\mathbf{x}_0 = \int_{\psi(\Omega_\tau(0))} 1 d\mathbf{u}_0 = \mathrm{Vol}\big(\psi(\Omega_\tau(0))\big). \tag{11}$$

此外，利用以下链式法则关系：

$$\begin{aligned} \dot{\mathbf{u}} &= \frac{\partial \psi}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial t} \\ &= p_0\big(\psi^{-1}(\mathbf{u})\big) \dot{\mathbf{x}}\big(\psi^{-1}(\mathbf{u})\big), \end{aligned} \tag{12}$$

任何状态空间向量场模型都可以等价地表示为 $\mathbb{U}^n$。然而，这样做会对$\mathbf{u}$空间模型施加边界条件。具体而言，

$$\dot{\mathbf{u}}_i = 0 \quad \text{if} \quad u_i = 0 \text{ or } u_i = 1. \tag{13}$$

这确保系统永远不会离开 $\mathbb{U}^n$（因此也永远不会离开状态空间），因为 $\mathbb{R}^n$ 与 $\mathbb{U}^n$之间必须始终存在双射。感兴趣的超矩形区域 R 也可以映射到 $\mathbb{U}^n$，即 $R_u = \psi(R)$，从而得到另一个超矩形区域，因为 ψ完全解耦且单调。

图 1b 展示了在变换空间 $\mathbb{U}^n$中的等效传播情形。在 $\mathbb{R}^n$ 和 $\mathbb{U}^n$ 中，Ω_τ 的概率质量始终守恒。然而，在 $\mathbb{U}^n$ 中，密度在 $t = 0$处均匀分布，因此仅知道 Ω_τ 的体积就足以计算其质量。因此，在 $\mathbb{U}^n$中进行传播，无需 (i) 对初始分布进行积分，(ii) 跟踪$\mathbf{u}$空间密度，以及 (iii) 跟踪 $\Omega_\tau(t)$的形状。式 (5) 中涉及的两个积分都可简化为跟踪$\mathbf{u}$空间控制质量 $\psi(\Omega_\tau(t))$的体积（一个标量）的演变，并沿时间反向追踪，直到 $t = 0$

.

因此，我们以下研究$\hat{}\mathbb{U}^n$中控制质量演化的过程。为简化记号，我们用 f、Ω_τ和 R 分别表示$\mathbf{u}$空间向量场、$\mathbf{u}$空间控制质量以及 R_u

.

3.2 体积函数泰勒展开 on

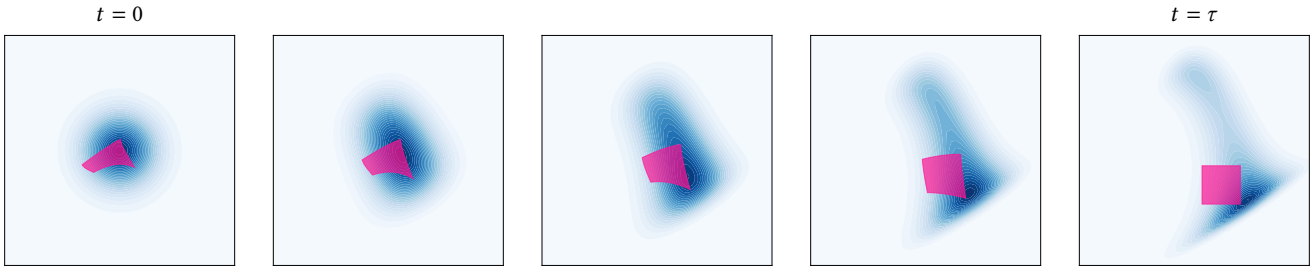
令函数 $V_\Omega(\cdot;\tau)$： $\mathbb{R}_{\geq 0} \to [0,1]$ 表示$\Omega_\tau(t)$随时间变化的体积。变量 t表示特定控制质量在时间何处，而参数 τ 表示控制质量何时占据 R，从而有效地确定正在跟踪哪个控制质量。例如，$V_\Omega(t_1;\tau)$和 $V_\Omega(t_2;\tau)$描述的是不同时刻的同一控制质量（初始条件），而 $V_\Omega(t;\tau_1)$和 $V_\Omega(t;\tau_2)$描述的是同一时刻的不同控制质量。利用前述关于 $\mathbb{U}^n$的性质，初始体积等于所需的概率测度

$$V_\Omega(t=0;\tau) = P(\hat{\phi}(\mathbf{u}_0, \tau) \in R). \tag{14}$$

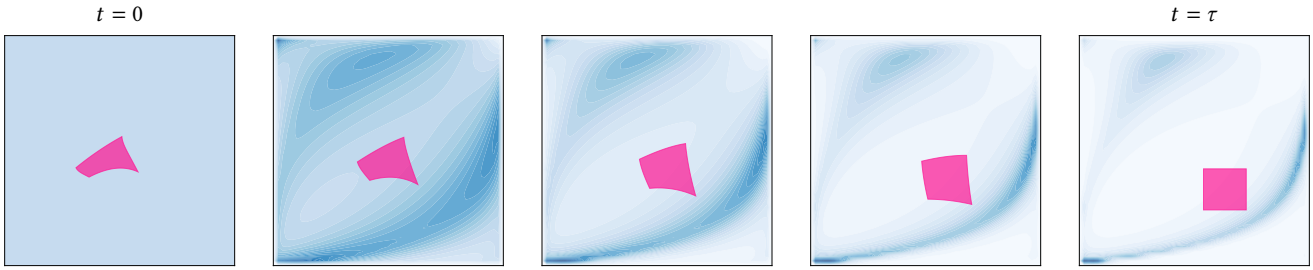
注意，$V_\Omega(\tau;\tau) = (R)$ $V_\Omega(0;\tau) \leq (;)$ Vol，且 $\bar{P}_\tau$ 1 2 计算一个 $\bar{P}_\tau$，使得 $(;)$ 满足问题 . .

非线性模型通常无法用闭式形式表示的体积函数。此外，$\Omega_\tau(t)$ 的形状只有在时间 $t = \tau$ 才能用简单方式描述。为应对这些挑战，我们围绕点 $t = \tau$ 对 $V_\Omega(t;\tau)$ 构建一个泰勒级数近似：

$$V_\Omega(t;\tau) \approx \tilde{V}_\Omega^m(t;\tau) = \sum_k^m \frac{d^k V_\Omega}{dt^k}\Big|_{t=\tau} \frac{(t-\tau)^k}{k!}. \tag{15}$$



(a) 状态空间（$\mathbb{R}^n$）演化



(b) 变换空间（$\mathbb{U}^n$）演化

ˆ图1：受非线性 f作用的概率密度函数演化示例。我们希望在时刻 $t = \tau$对 $p(\mathbf{x}_\tau)$（蓝色密度）在$\Omega_\tau(\tau) = R$（粉色区域）上进行积分。在时刻 $t = 0$，密度较为简单，即在 $\mathbb{R}^n$ 中为高斯分布，或在 $\mathbb{U}^n$中为均匀分布，但随着时间推移会变得复杂。在时刻 $t = \tau$， R是一个简单矩形，而随着时间回溯，其形状会变得更加复杂。由于概率质量是守恒量， Ω_τ 的体积和密度会发生波动，从而使其概率质量在整个时间内保持恒定。

泰勒级数仅利用 $V_\Omega(t;\tau)$ 在 $t = \tau$ 处的时间导数等点态信息，来近似其在其他时刻的值 $V_\Omega(t;\ \tau)$。通过围绕 $t = \tau$ 展开泰勒级数， Ω_τ 超矩形形状的简单性使得某些类别的 ˆ f 的展开系数可以精确计算。

为了计算系数，我们使用 $\Omega_\tau(t)$的连续体动力学。物质元 $\Omega_\tau(t)$上某一（累积的）随时间变化的标量量（例如体积）的时间导数由雷诺输运定理 [14],支配，该定理针对给定的感兴趣累积量 $\gamma(\mathbf{u}\ t)$指出，

$$\frac{d}{dt}\int_{\Omega_\tau(t)}\gamma(\mathbf{u},t)d\mathbf{u} = \int_{\Omega_\tau(t)}\frac{\partial\gamma}{\partial t} + \nabla\cdot\big(\gamma(\mathbf{u},t)\hat{f}(\mathbf{u})\big)d\mathbf{u}. \tag{16}$$

令 $\gamma(\mathbf{u}\ t) = 1$，即可得到体积函数的一阶时间导数表达式：

$$\frac{d}{dt}V_\Omega(t;\tau) = \int_{\Omega_\tau(t)}\nabla\cdot\hat{f}(\mathbf{u})d\mathbf{u}, \tag{17}$$

这是连续体动力学中一个众所周知的性质。利用这一结果，可以改用 $\gamma(\mathbf{u}\ t) = \nabla\cdot$ ˆ $f(\mathbf{u})$ 来求得体积函数的二阶时间导数：

$$\begin{aligned}\frac{d^2}{dt^2}V_\Omega(t;\tau) &= \frac{d}{dt}\int_{\Omega_\tau(t)}\nabla\cdot\hat{f}d\mathbf{u} \\ &= \int_{\Omega_\tau(t)}\nabla\cdot\big((\nabla\cdot\hat{f})\hat{f}\big)d\mathbf{u}.\end{aligned} \tag{18}$$

可以递归地应用相同的代换，推导出 k 阶时间导数 $V_\Omega(t)$ 的表达式，如以下引理所示。

引理 1. 设 $\Gamma_{k=0}(\mathbf{u}) = 1$ 和 $\Gamma_k(\mathbf{u})$ 递归定义如下：

$$\Gamma_k(\mathbf{u}) = \nabla\cdot\big(\Gamma_{k-1}\hat{f}\big). \tag{19}$$

那么，体积函数的 k阶时间导数为

$$\frac{d^k}{dt^k}V_\Omega(t;\tau) = \int_{\Omega_\tau(t)}\Gamma_k d\mathbf{u}. \tag{20}$$

证明。根据数学归纳法即可证明。对于 $k = 1$，式（17）成立。假设

$$\frac{d^k}{dt^k}V_\Omega(t;\tau) = \int_{\Omega_\tau(t)}\Gamma_k(\mathbf{u})d\mathbf{u}. \tag{21}$$

那么，

$$\begin{aligned}\frac{d^{k+1}}{dt^{k+1}}V_\Omega(t;\tau) &= \frac{d}{dt}\int_{\Omega_\tau(t)}\Gamma_k(\mathbf{u})d\mathbf{u} \\ &= \int_{\Omega_\tau(t)}\frac{\partial\Gamma_k}{\partial t} + \nabla\cdot\big(\Gamma_k(\mathbf{u})\hat{f}(\mathbf{u})\big)d\mathbf{u}\end{aligned} \tag{22}$$

并且$\frac{\partial\Gamma}{\partial t} = 0$，因为 $\Gamma_k(\mathbf{u})$ 不随时间变化。 □

引理 1 描述了一种递归方法，用于计算体积函数的时间导数。函数 Γ_k 与逐点的 k 阶体积变化率相关；对 Ω_τ（t）积分后，可得到控制质量体积的累积变化率。(20) 中的积分是在 Ω_τ（t）上进行的，

在展开时刻 $t = \tau$ 处求值后，得到超矩形积分区域 R。因此，如果能够解析地确定 Γ_k 的原函数，就可以精确计算体积函数导数。由此，泰勒展开近似仅 随时间而退化，不依赖任何空间近似。因此，预测质量不受 R 所在位置的影响；即使 R 位于分布尾部，也就是发生稀有事件（见第 6 节），该方法仍能给出准确的估计。

注释 1. 按照相同的递归过程，使用时变向量场模型 $f(\mathbf{u}\ t)$ 可以计算体积函数的时间导数ˆ。

然而，使用任意的 ˆf 构建泰勒级数展开（15），实际上可能无法收敛到真实的体积函数。事实上，当且仅当 $V_\Omega(t)$ 是解析函数时，这种收敛才存在。由于 ˆ f 隐式地 ˆ 生成 $V_\Omega(t)$，我们必须刻画哪些模型 f 能够生成解析体积函数。为此，我们可以研究算子（19）如何生成导数，并证明导数的增长不会过快（否则会导致泰勒级数发散）。然而，对于任意解析的 f，这一过程 ˆ 极具挑战性；不过，通过对 f 在复平面中的解析延拓 作出 ˆ 假设，可以得到次阶乘界，从而确保泰勒级数收敛。

定理 2。设 ˆ$g(w)$: $\mathbb{C}^n \to \mathbb{C}^n$ 是 ˆ $f(u)$ 在复域 $w \in \mathbb{C}^n$ 上的解析延拓，其中 $\mathbb{R}e(w) \in \mathbb{U}^n$。定义 $B($ r $_0$ = $\in$ $\mathbb{C}^n$ $|w_i - w_{0\,i}| < r\ \forall 1 \le i \le n$ g $, w)$ 为 -维复多圆盘 $B(, w)\ \{w\ s.t.,\}$。那么，若 ˆ(w) 在 $B(r, w_0)$ 上全纯，且 $r > 1$ 并且 $w_0 = \frac{1}{2}1_n + 0i$，则 $V_\Omega(t)$ 是解析函数，且收敛半径为无穷大。

证明。首先，我们回顾体积函数具有无穷收敛半径的解析性的充分条件。设 $r^k(t)\ V_\Omega\ t\ -\tilde{V}_\Omega\ t$ k ( ) ( ) 为 -次泰勒展开的余项。根据泰勒定理，

$$|r^k(t)| \le M_k \frac{(t-\tau)^{k+1}}{(k+1)!} \tag{23}$$

其中，对于所有 $\xi \in [t, \tau]$，$\left|\frac{d^{k+1}}{dt^{k+1}} V_\Omega(\xi)\right| \le M_k$。若 $V_\Omega(t)$ 是解析的，则该展开必定收敛，即 $\lim_{k\to\infty} |r^k(t)| = 0$。因此，只需证明

$$\lim_{k\to\infty} \frac{M_k}{(k+1)!} = 0, \tag{24}$$

即，M_k 的增长速度慢于阶乘。

接下来，为了得到关于 M_k 的次阶乘界，我们利用多变量全纯函数的柯西估计 [16]。具体而言，对于在 B 上全纯的复值函数 g，

$$\left|\frac{\partial g}{\partial w_i}(\mathbf{w})\right| \le |w_i - w_{0,i}|^{-1} \sup_{\mathbf{w}\in B} |g|. \tag{25}$$

我们的目标是使用 (25) 来界定连续 $\Gamma_k(\mathbf{w})$ 的增长，这些项是使用 ˆg 而非 ˆ f 生成的。由于算子 (19) 由乘法、加法和微分构成，每个 $\Gamma_k(\mathbf{w})$ 都由代数运算组成，因此必定在 B 上全纯。此外，任何在 B 上全纯的函数，在半径小于或等于 r 的多圆盘上必定有界。

代入 ˆg 后，算子 (19) 可等价地写为

$$\Gamma_k(\mathbf{w}) = \sum_{i=1}^{n} \frac{\partial \Gamma_{k-1}}{\partial w_i}(\mathbf{w})\hat{g}_i(\mathbf{w}) + \Gamma_{k-1}(\mathbf{w})\frac{\partial \hat{g}_i}{\partial w_i}(\mathbf{w}) \tag{26}$$

令 $A_k = \sup_{\mathbf{w}\in B(1,\mathbf{w}_0)} |\Gamma_k(\mathbf{w})|$，且 $\epsilon = \max_{1\le i\le n} |w_i - w_{0,i}|^{-1}$。然后，将 (25) 应用于 (26)，得到递推关系

$$\begin{aligned} A_k &\le \sum_{i=1}^{n} \epsilon A_{k-1}|\hat{g}_{u,i}(\mathbf{w})| + \epsilon A_{k-1}\left|\frac{\partial \hat{g}_i}{\partial w_i}(\mathbf{w})\right| \\ &= A_{k-1}\epsilon\Big(\sum_{i=1}^{n} |\hat{g}_{u,i}(\mathbf{w})| + \left|\frac{\partial \hat{g}_i}{\partial w_i}(\mathbf{w})\right|\Big) \end{aligned} \tag{27}$$

注意到 $s = \epsilon \sum_{i=1}^{n} |\hat{}g_i(\mathbf{w})| + \left|\frac{\partial \hat{}g_i}{\partial w_i}(\mathbf{w})\right|$ 与 k 无关，可得关于 A_k 的非递推上界：

$$A_k \le s^k. \tag{28}$$

根据式 (20)，要得到 $V_\Omega(t)$ 的 k阶时间导数的上界，可以用被积函数的上界乘以 Ω_τ 所能占据的最大积分区域（即整个 $\mathbb{U}^n$）来为积分给出上界，即

$$\begin{aligned} \int_{\Omega_\tau(t)} \Gamma_k d\mathbf{u} &\le \Big(\sup_{t\in[0,\infty)} \mathrm{Vol}\big(\Omega_\tau(t)\big)\Big) \sup_{\mathbf{w}\in B(1,\mathbf{w}_0)} |\Gamma_k(\mathbf{w})| \\ &\le \mathrm{Vol}\big(\mathbb{U}^n\big)A_k. \end{aligned} \tag{29}$$

因此，由于 $\mathrm{Vol}(\mathbb{U}^n) = 1$，$\frac{d^k}{dt^k} V_\Omega(t) \le s^k$，这表明对所有 t，增长均以指数有界（次阶乘）方式进行，足以保证 $V_\Omega(t)$ 是解析的，且收敛半径无限。□

定理 2 给出了保证˜体积函数具有解析性的充分条件，从而确保式 (15) 中的 $V_\Omega(t;\ \tau)$ 对所有时间都收敛到 $V_\Omega(t;\tau)$。特别地，只需确保所选回归器 f 的解析延拓在包含 $\mathbb{U}^n$ 的多圆盘上全纯。许多类别的代数函数都具有这一性质，例如多项式、三角函数等。

3.3 体积函数的界定

上一节介绍了一种方法，用于计算模型真实体积函数关于 $t = \tau$ 的截断泰勒展开。然而，截断近似无法得到严格界 $\bar{P}_\tau$。为计算 $\bar{P}_\tau$，我们采用泰勒定理，即对截断结果与真实函数之间的误差进行界定。

回顾一下，泰勒定理指出，阶泰勒展开的余项 $r^k(t) = V_\Omega(t)$ $\tilde{V}_\Omega\ t$ k 23 $[t_1, t_2]()$ 的上界为 ( )。令 R 为一个集合，其中包含范围 $t_1 \le t \le t_2$ 内所有 $\Omega_\tau(t)$ 的并集，即，

$$\mathcal{R}[t_1, t_2] \supseteq \{\mathbf{u}\ \text{s.t.}\ \mathbf{u} \in \Omega_\tau(t)\ \text{for some}\ t \in [t_1, t_2]\}. \tag{30}$$

直观地说，R 描述了 Ω_τ 在场中移动时所划出的“管状区域”的过近似。可以利用 $\mathrm{R}[t, \tau]$ 上逐点体积增长的“最坏情况”来计算一个直接的上界M_k。具体而言，例如$\int_{\Omega_\tau(t)} \Gamma_{k+1}(\mathbf{u})d\mathbf{u}$ 在任意可能的 $\Omega_\tau(t) \subseteq \mathrm{R}[t, \tau]$ 上的最大值，都受 $\Gamma_{k+1}(\mathbf{u})$ 的极值乘以 $R[t, \tau]$ 的体积所界定。以下定理将为 $t \le \tau$ 给出一个上界。

定理 3。 假设 $\hat{f}$ 满足定理 *2*。令

$$\delta_k \geq \sup_{\mathbf{u}\in\mathcal{R}[t,\tau]} \Big((-1)^{k+1}\Gamma_{k+1}(\mathbf{u})\Big). \tag{31}$$

于是，对于 $t \in [0,\tau]$,

$$r^k(t) \leq \delta_k \mathrm{Vol}\big(\mathcal{R}[t,\tau]\big)\frac{|t-\tau|^{k+1}}{(k+1)!}. \tag{32}$$

证明。我们的目标是在区间 $[t,\tau]$ 上求泰勒级数展开的上界。考虑定义为下式的映射 $t' = g(t)$:

$$g : t \mapsto \tau - t \tag{33}$$

该映射将展开重新以 0 为中心并反转时间，从而将展开函数关于其中心进行反射。我们将平移并反射后的展开记为 $V'_{\Omega}(t;\tau)$，使得 $V'_{\Omega}(t;\tau) = V_{\Omega}(g(t);\tau)$。根据泰勒定理可知，对于 $t' > 0$，

$$r'^k(t') \leq M_k^u \frac{(t')^{k+1}}{(k+1)!} \tag{34}$$

其中，对于所有 $\xi \geq 0$，有 $M_k^u \geq \left|\frac{d^{k+1}}{dt^{k+1}} V_{\Omega}(\xi)\right|$。$t'$ 阶导数和 t 阶导数之间的关系为

$$\frac{d^{k+1}}{(dt')^{k+1}}V'_{\Omega}(t') = (-1)^{k+1}\frac{d^{k+1}}{(dt)^{k+1}}V_{\Omega}(t). \tag{35}$$

使用 (20),

$$\begin{aligned}\sup_{\xi'\in[0,t']}\frac{d^{k+1}}{(dt')^{k+1}}V'_{\Omega}(\xi') &= \sup_{\xi\in[t,\tau]}(-1)^{k+1}\frac{d^{k+1}}{(dt)^{k+1}}V_{\Omega}(\xi)\\ &= \sup_{\xi\in[t,\tau]}\int_{\Omega_\tau(\xi)}(-1)^{k+1}\Gamma_k d\mathbf{u}.\end{aligned} \tag{36}$$

令 $\Omega_\tau^\star$ 为 $\sup^{\xi\in[t,\tau]}\mathrm{Vol}\big(\Omega_\tau(\xi)\big)$ 的最大化点。那么，

$$(36) \leq \mathrm{Vol}\big(\Omega_\tau^\star\big)\sup_{\mathbf{u}\in\Omega_\tau^\star}\Big((-1)^{k+1}\Gamma_{k+1}(\mathbf{u})\Big). \tag{37}$$

由此可得，$\mathrm{Vol}\big(\Omega_\tau^\star\big) \leq \mathrm{Vol}\big(\mathrm{R}[t,\tau]\big)$，且 $\sup^{\mathbf{u}\in\Omega_\tau^\star}(\cdot) \leq \sup^{\mathbf{u}\in R[t,\tau]}(\cdot)$。□

给定 Γ_{k+1} 在 $\mathrm{R}[t,\tau]$ 上的上界，可使用定理 3 计算 $\bar{P}_\tau$。这有助于选择 $\hat{f}$，因为已知要为任意连续、非线性、非单调函数生成形式化上界是一项颇具挑战的问题 [22]。选择属于某个函数类的 $\hat{f}$，使其生成 Γ_k，且存在一种可行的方法来计算紧上界，可以大幅降低 $\bar{P}_\tau$ 的保守性。然而，上界不能过于宽松，否则可能无法满足泰勒级数收敛所需的次阶乘增长条件（参见定理 2 的证明）。

推论 4. 假设 $\hat{f}$ 满足定理 *2*。令

$$e_k = \delta_k - \sup_{\mathbf{u}\in\mathcal{R}[t,\tau]}\Big((-1)^{k+1}\Gamma_{k+1}(\mathbf{u})\Big). \tag{38}$$

那么，若

$$\lim_{k\to\infty}\frac{e_k}{(k+1)!} = 0, \tag{39}$$

余项界(32)从上方收敛至零。

证明。该结论可直接由定理 2 得出，但需注意，上界计算过程的保守性增长速度不得快于阶乘。□

推论 4 明确阐述了上界计算过程与上界收敛之间的关系。基本而言，重要的是验证对于高阶展开，上界不能变得过于保守。

在 $\delta_k > 0$的情况下，定理 3 中的上界可以进一步改进。定理 3 使用 $\mathrm{Vol}(\mathrm{R}[t,\tau])$ 作为在整个区间内 Ω_τ可能达到的最大体积的安全（但可能偏保守）上界。认识到 $V_\Omega(t;\tau)$本身就在估计同一数量，可以改进这一估计。利用定理 3 在区间 $[t,\tau]$ 上得到上界 $V_{max} \geq V_\Omega(t;\tau)$，即可替代 $\mathrm{Vol}(\mathrm{R}[t,\tau])$，从而得到更紧的余项上界。此过程可以任意次重复，生成一个上界收敛序列。形式化地，令 $\{V_{max,i}\}_{i\geq 0}$ 为一个无限序列，使得$V_{max,0} = \mathrm{Vol}(\mathrm{R}[t,\tau])$并且

$$V_{max,i+1} = \sup_{\xi\in[t,\tau]}\Big(\tilde{V}_\Omega(\xi;\tau) + \delta_k V_{max,i}\frac{|\xi-\tau|^{k+1}}{(k+1)!}\Big). \tag{40}$$

~~在满足关于 $V_\Omega(t;\tau)$ 的某些条件时，我们证明 $V_{max,i}$是一个收敛的几何级数，且具有已知的闭式极限值。~~

定理 5. 设对所有 $\xi \in [t,\tau]$, $\kappa \geq \tilde{V}_\Omega(\xi)$成立。那么，对于 ~~$0 \leq \delta_k \leq \frac{(k+1)!}{|t-\tau|^{k+1}}$~~，

$$\sup_{\xi\in[t,\tau]} V_\Omega(t;\tau) \leq \kappa\Big(1 - \delta_k\frac{|t-\tau|^{k+1}}{(k+1)!}\Big)^{-1} \tag{41}$$

此外，

$$V_\Omega(t;\tau) \leq \tilde{V}_\Omega(t,\tau) + \kappa\Big(\frac{(k+1)!}{|t-\tau|^{k+1}} - \delta_k\Big)^{-1}. \tag{42}$$

证明。根据引理3中的式（32），我们可以将第 i个余项上界函数定义为

$$\bar{r}_i^k(t) = \delta_k V_{max,i}\frac{|t-\tau|^{k+1}}{(k+1)!}. \tag{43}$$

对于 $\delta_k \geq 0$，每个 $\bar{r}_i^k(t)$ 在 $[t,\tau]$ 上单调递减，因此在 t 处取得最大值。因此，（40）可简化为

$$\begin{aligned}V_{max,i+1} &\leq sup_{\xi\in[t,\tau]}\Big(\kappa + \delta_k V_{max,i}\frac{|\xi-\tau|^{k+1}}{(k+1)!}\Big)\\ &= \kappa + \delta_k V_{max,i}\frac{|t-\tau|^{k+1}}{(k+1)!}.\end{aligned} \tag{44}$$

令 $\alpha = \delta_k\frac{|t-\tau|^{k+1}}{(k+1)!}$。于是 $V_{max,i+1} = \kappa + \alpha V_{max,i}$，以非递归形式生成级数

$$V_{max,i} = \sum_{j=0}\kappa\alpha^j \tag{45}$$

对于 $\alpha \in [0,1)$，当 $i \to \infty$ 时，它收敛到 (41) 的右侧。将 (32) 中的 $\mathrm{Vol}(\mathrm{R}[t,\tau])$替换为 (41) 中收敛的上界，即可得到上界 (42)。□

虽然定理 5 中的上界似乎总是优于定理 3，但它需要在计算 κ时增加少量额外开销，即需要求出一个一元多项式[2]的上界。此外，如果估计~的展开 V_Ω 在 $[t,\tau]$上大幅减小，κ 可能会足够大，使得定理 5 比定理

[2]计算一元多项式上界的方法有很多。我们的实验采用了伯恩斯坦基转换（见 (48)）。

3 更保守。然而在实践中，余项界的增长速度通常远快于˜ ˜，而且 V_Ω 在 $\tau - t$较小时通常变化不大，因此定理 5 的保守性较低。

4 选择回归器

现在我们聚焦于问题1.1中的模型学习部分。为了选择参数化回归器 f，我们首先回顾一下，为实现渐近普适收敛和不确定性传播，f 必须具备哪些性质。令每个 f_i 都属于参数化函数类 $\mathrm{F}_\theta \subset \mathrm{C}^\omega(\mathbb{U}^n)$。那么，$f$必须满足以下性质：

（P1）每个 f_i 都是收敛通用估计器（定义 1）；（P2）可以精确施加边界条件（13）；（P3）每个 f_i 都满足定理 2；（P4） Γ_k 对所有 $k > 0$ 都有解析反导函数；（P5） Γ_k 可以同时取上界和下界，使得式（39）成立。

每个性质都对 F_θ施加了独特的限制。在本文中，我们研究伯恩斯坦多项式回归器，并证明其满足条件（P1）至（P5）。

一个（多元）伯恩斯坦多项式 $b(\mathbf{u}) \in \mathrm{C}^\omega(\mathbb{U}^n)$，其次数为 $\mathbf{d} = (d_1, \ldots, d_n)$，定义为

$$b(\mathbf{u}) = \sum_{\mathbf{0} \le \mathbf{j} \le \mathbf{d}} \theta_{\mathbf{j}} \prod_{i=1}^{n} \beta_i(u_i; \mathbf{j}, \mathbf{d}), \tag{46}$$

其中，$\mathbf{j} = (j_1, \ldots, j_n)$ 是一个多重指标，并且

$$\beta_i(u_i; \mathbf{j}, \mathbf{d}) = \binom{j_i}{d_i} u_i^{d_i} (1 - u_i)^{d_i - j_i} \tag{47}$$

它们是伯恩斯坦基函数。伯恩斯坦多项式构成一个多项式基，即每个（最大）次数为 $\mathbf{d}$ 的多项式都可以表示为伯恩斯坦多项式，反之亦然。每个 ˆ 分量 f_i 都作为一个独立的多元伯恩斯坦多项式进行学习。下面，我们将展示伯恩斯坦多项式回归器如何满足上述各项条件。

条件（*P1*）（收敛通用估计器）。魏尔斯特拉斯逼近定理 [23] 指出，（伯恩斯坦）多项式是通用逼近器。事实上，众所周知，伯恩斯坦多项式相对于 C^k-范数 [27], 具有强收敛性，这正是定义 1 所要求的。在线性回归中使用伯恩斯坦多项式基作为多项式特征，可以确保损失景观是凸的。因此，随着特征数量增加，即 $\mathbf{d}$ 增大，全局极小值（按照最小二乘法意义）既可计算，又能收敛到真实系统。

条件（*P2*）（边界条件）。回想一下，要使 ˆ f 成为有效的 ˆ$\mathbf{u}$-空间向量场，f 沿 $[0, 1]^n$ 边界的垂直分量必须为零。伯恩斯坦多项式提供了一种简单的方法来施加此类边界条件。具体而言，对于多重指标 $\mathbf{j} = (j_1, \ldots, j_n)$，若 $\theta_{\mathbf{j}} = 0$ 在 $j_i = 0$ 或 $j_i = d_i$ 时成立，则 ˆ $f_i(\mathbf{u}) = 0$ 在 $u_i = 0$ 或 $u_i = 1$时成立。

条件 *(P3)*（解析体积函数）。多项式已知是整函数，即对所有 $\mathbb{C}^n$，其在复平面上的解析延拓都是全纯的。因此，定理 2 中的条件很容易满足，伯恩斯坦多项式回归器会生成一个解析体积函数。

条件 *(P4)*（闭式反导数）。多项式总是有易于计算的原函数。用伯恩斯坦基表示时，原函数可计算为系数的线性组合，或通过不完全贝塔函数 [20] 计算。

条件（*P5*）（界收敛）。伯恩斯坦多项式提供了一种高效且紧致的界定方法。具体而言，任何伯恩斯坦多项式 $b(\mathbf{u})$ 均受其系数界定，即

$$\min_{\mathbf{0} \le \mathbf{j} \le \mathbf{d}} \theta_{\mathbf{j}} \le \min_{\mathbf{u} \in \mathbb{U}^n} b(\mathbf{u}) \le \max_{\mathbf{u} \in \mathbb{U}^n} b(\mathbf{u}) \le \max_{\mathbf{0} \le \mathbf{j} \le \mathbf{d}} \theta_{\mathbf{j}}. \tag{48}$$

此外，[12] 概述了生成任意紧上界的方法，从而满足推论 4。不过在实践中，通常不需要此类界定方法，式（48）已足够。

除了满足上述性质外，代数运算（例如引理 1 中使用的乘法、微分和积分）具有相对良好的数值稳定性，并且可以在伯恩斯坦基中进行 [20, 22]。请注意，上述标准并非仅限于伯恩斯坦多项式，而是概述了针对给定回归器评估所提方法有效性的维度。研究其他回归器及其对这些标准的符合程度，是未来值得探索的方向。

伯恩斯坦多项式的理论性质使所提出的联合建模与传播过程具有很强的渐近收敛性，这直接源于条件 (P1)–(P5) 得到了满足。

ˆ**定理 6.** 假设每个 f_i 都是满足边界条件 (13)的$\mathbf{d}$ 次伯恩斯坦多项式估计器。令 $\bar{P}_\tau^k(\hat{} f)$ 表示ˆ对于给定模型 f 和 k 次展开，使用定理 *3* 计算得到的上界。那么，

$$\lim_{\mathbf{d}, |\mathcal{D}|, k \to \infty} |\bar{P}_\tau^k(\hat{f}) - P(\phi(\mathbf{x}_0, \tau) \in R)| = 0, \tag{49}$$

也就是说，$\bar{P}_\tau^k(\hat{} f)$ 收敛于真实系统的概率。

定理 6 阐明了使用伯恩斯坦多项式回归器时，所提出的不确定性传播方法的收敛性质。

5 后向集合传播

第 3 节和第 4 节介绍了一种为模型预测生成渐近收敛上界的方法。然而在实践中，即使底层函数的收敛半径为无穷大，截断的泰勒近似通常也只能在展开点附近的局部邻域内提供准确估计。为解决这一问题，我们可以利用可达集传播 [1, 4, 9]，以维持对 $\Omega_\tau(t)$ 真实形状的一种粗略且在数学上易于处理的过近似。可达集传播会生成一条流管，即一系列转移区域，并保证这些区域在有限个时间区间内包含 $\Omega_\tau(t)$。参见图 1，我们可以计算一条从 $t = \tau$ 开始的流管，并对 $\Omega_\tau(t)$ 随时间向后推移至 $t = 0$ 的过程进行过近似。

形式上，流管是一个由管状区域构成的序列 $S = \{\mathrm{R}_l[t_l, t_{l+1}]\}$，其中 $l \in \{0, \ldots, L\}$，且每个 R_l都满足 (30)、$t_{l=0} = \tau$ 和 $t'_L = 0$。此外，流管理论还提供特定时刻 t 的逐点集合过近似，记作 $\mathrm{R}[t] \supseteq$

$\Omega_\tau(t)$。我们假设每个 R_l 都是超矩形。由于计算每个 R_l的体积很简单，因此可以利用流管本身计算一个非常保守的 $\bar{P}_\tau$。为简单起见，令 $\bar{V}_\Omega$表示近似扩张加上有界单项式，使其成为 V_Ω 的上界。

我们将方框流管的计算实用性与泰勒体积函数展开的局部精度结合起来。对于每个流管盒，我们可以构造一个以 t_l 为中心的泰勒展开 $\bar{V}^l_\Omega(t;\, t_l)$，用于估算区间 $[t_l, t_{l+1}]$ 上的体积。每个泰勒展开都是通过对超矩形过近似 $\mathrm{R}_l[t_l]$ 进行积分而得到的。仅凭这一点，与直接使用每个 R_l 的体积相比，并不能有意义地降低保守性。不过，我们可以利用每个 $\bar{V}^{l+1}_\Omega(t;\, t_l)$ 的计算系数，来"驯服"前一个展开 $\bar{V}^l_\Omega(t;\, t_l)$ 中发散的导数。

考虑第一和第二个流管区域的展开（不失一般性）：$\bar{V}^0_\Omega(t;\, t_0 = \tau)$ 是通过对 $\mathrm{R}_0[t_0] = R$ 积分构造的，而 $\bar{V}^1_\Omega(t;\, t_1)$ 是通过对 $\mathrm{R}_1[t_1]$ 积分构造的。在时刻 $t = t_1$，前一个展开的导数为所有 k时间导数提供上界，即

$$\left.\frac{d^k V_\Omega}{dt^k}\right|_{t=t_1} \le \left.\frac{d^k \bar{V}^0_\Omega}{dt^k}\right|_{t=t_1}. \tag{50}$$

另一方面，我们可以对 Γ_k 在过近似矩形区域 $R_1[t_1]$ 上的积分进行计算，从而构造一个替代的（上界）估计。为说明这一点，考虑分解式 $\Gamma_k = \Gamma_k^+ + \Gamma_k^-$，其中 $\Gamma_k^+ \ge 0$、$\Gamma_k^- \le 0$，使得 $\Gamma_k^+ \ge \Gamma_k$ 和 $\Gamma_k^- \le \Gamma_k$。由于 Γ_k 是伯恩斯坦多项式，因此可以将 Γ_k 中所有负系数设为零，以构造 Γ_k^+。我们的目标是对 $\int_{\Omega_\tau(t_1)}\Gamma_k(\mathbf{u})d\mathbf{u}$ 求上界；然而，在 t_1 时（或任何未来时刻 $t_l > t_0$），我们并不知道 $\Omega_\tau(t_1)$ 的确切位置，只知道它必定包含在 $R_1[t_1]$ 中。因此，对于任何可能的 $\Omega_\tau(t_1) \subseteq R_1[t_1]$，$\int_{\Omega_\tau(t_1)}\Gamma_k(\mathbf{u})d\mathbf{u}$ 的上界如下

$$\begin{aligned}\left.\frac{d^k V_\Omega}{dt^k}\right|_{t=t_1} &= \int_{\Omega_\tau(t_1)} \Gamma_k(\mathbf{u})d\mathbf{u} = \int_{\Omega_\tau(t_1)} \Gamma_k^+(\mathbf{u}) + \Gamma_k^-(\mathbf{u})d\mathbf{u} \\ &\le \int_{\Omega_\tau(t_1)} \Gamma_k^+(\mathbf{u})d\mathbf{u} \\ &\le \int_{R_1[t_1]} \Gamma_k^+(\mathbf{u})d\mathbf{u}. \end{aligned}\tag{51}$$

(50) 中的上界在先前的局部泰勒展开较为准确时会更紧，而 (51) 在可达集近似更为准确时会更紧（即先前的泰勒级数正在发散）。这两个上界都是真实导数的可靠过近似，因此对 $\tilde{V}^0_\Omega(t;\, t_1)$ 展开的每个系数选取较小的那个。该过程对流管中的每个集合进行迭代，如下述算法所述。

算法 1 概述了创建分段式"驯化"体积函数展开的过程。该算法的输入为流管 S 和展开阶数。对于第一个流管盒，即 R，使用精确积分创建展开（第 4 行）。对于后续的每个流管盒，使用（50）计算前一个展开的导数估计（第 6 行），并通过在当前流管盒上积分，使用（51）计算"新的"导数估计（第 7 行）。选择较不保守的导数（第 8 行）。所选导数用于

算法 1流管驯化体积函数展开

1: **输入：** 流管 S，展开阶数 m
2: **对于** $l = 1$ **到** $L - 1$ **执行**
3: **如果** $l = 0$ **则**
4: $v_{k,\ curr} \leftarrow \int_{\mathcal{R}}\Gamma_k(\mathbf{u})d\mathbf{u}\ \forall k \in \{1, \ldots, m\}$
5: **else**
6: $v_{k,prev} \leftarrow \left.\frac{d}{dk}\bar{V}^{l-1}_\Omega(t;\, t_{l-1})\right|_{t=t_l}\ \forall k \in \{1, \ldots, m\}$
7: $v_{k,\ new} \leftarrow \int_{\mathcal{R}[t_l]}\Gamma_k^+(\mathbf{u})\ d\mathbf{u}\ \forall k \in \{1, \ldots, m\}$
8: $v_{k,curr} \leftarrow$ 最小值$(v_{k,new}, v_{k,prev})\ \forall k \in \{1, \ldots, m\}$
9: **结束如果**
10: 在$\mathrm{R}[t_l, t_{l+1}]$上创建有界单项式
11: 创建 $\bar{V}^l_\Omega(t;\, t_l)$，系数为 $v_{k,\ curr}$，加上单调上界-mial
12: **结束循环**
13: **输出：** $\{\ \bar{V}^l_\Omega(t;\, t_l)\}_{l\in\{1,\ldots,L\}}$

（15）的系数，并在过渡区域上计算一个有界单项式（使用定理 3 或 Theorem5 计算），从而形成下一个展开（第 10–11 行）。

6 实验结果

本节评估所提出的建模与传播方法的实际效用。

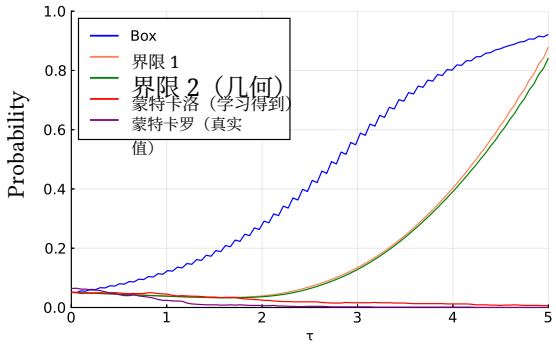
实验设置。 我们在两个系统上进行了实验：一个二维范德波尔系统和一个四维倒立摆。范德波尔系统由以下方程描述

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_2(1 - x_1^2) - x_1\end{aligned}\tag{52}$$

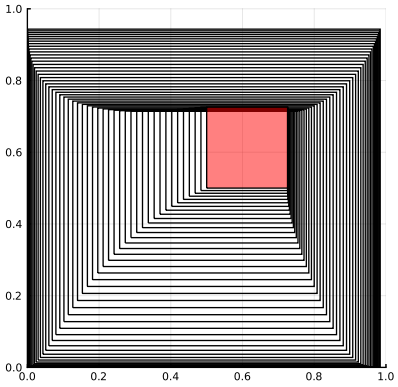
其初始分布为 $\mathbf{N}(\mathbf{0}, 0.5^2 I_{2\times 2})$。倒立摆系统由$\mathbf{x} = [y, v_y, \theta, \omega]$描述

$$\begin{aligned}\dot{y} &= v_y \\ \dot{v}_y &= \frac{m_p \sin(\theta)(l\omega^2 + g\cos(\theta))}{m_c + m_p \sin^2(\theta)} \\ \dot{\theta} &= \omega \\ \dot{\omega} &= \frac{m_p l\omega^2 \cos(\theta)\sin(\theta) - (m_c + m_p)g\sin(\theta)}{l(m_c + m_p \sin^2(\theta))}\end{aligned}\tag{53}$$

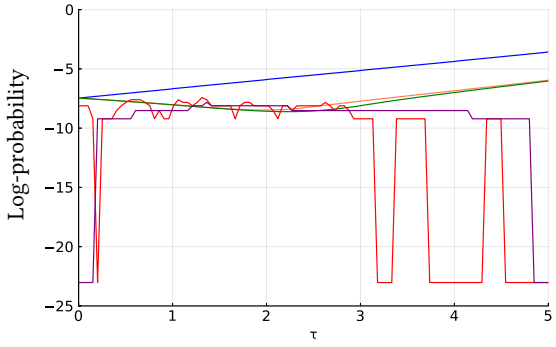
其中 $g = 9.81$，$l = 1.0$，$m_p = 0.1$和 $m_c = 1.0$，初始分布为 $\mathbf{N}([0.0, 0.0, 0.1, 0.5],\ \mathrm{diag}(0.5_2,\ 0.1_2,\ 0.2_2,\ 0.4_2))$。针对学习得到的模型和底层真实系统（真实值），将三种界定方法与采用 10k 个样本的蒙特卡洛前向欧拉积分仿真结果进行比较。第一种界定方法（"方框"）为每个流管盒逐步构建泰勒展开。然后，我们使用算法 1 中的驯化展开过程，对比两种界定方法，即定理 3（"界限 1"）和定理 5（"界 2"）。对于范德波尔系统，模型被回归为 7 阶伯恩斯坦多项式，并选择了展开阶数 $m = 5$。对于倒立摆系统，模型为 6 阶，且 $m = 4$。



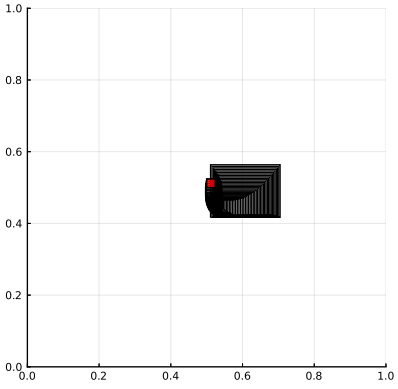
(a.1) 概率 vs. τ



(a.2) 流管（$\mathbb{U}^n$）（R 以红色显示）



(b.1) 对数概率 vs. τ



(b.2) 流管（$\mathbb{U}^n$）

图2：二维范德波尔：概率与流管的演变（上方，a.），以及稀有事件对应项（下方，b.）。

6.1 传播结果

图 2a图1展示了常规大小的 R的结果。盒方法给出的 $\bar{P}_\tau$最为保守，因为它大致反映了流管本身的保守性，这一点可从图 2a.2中看出。直到 $\tau = 2.0$为止，驯化方法给出的 $\bar{P}$都非常紧。超过$\tau = 2.0$后，流管变得更加保守，“驯化”效果也随之减弱。这意味着，在驯化展开导数时，流管所传递的信息不再那么有用，因此展开结果会变得更加保守，最终趋近于盒方法。

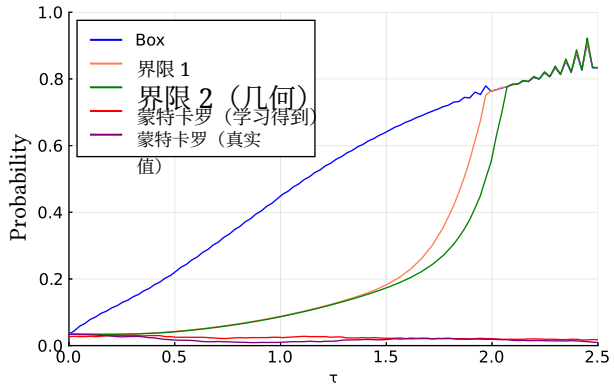
图 2b.1展示了极小 R（即稀有事件）的结果。请注意，图 2b.1中的概率采用对数尺度表示。由于该事件极为罕见，最终落入 R的前向仿真寥寥无几，因此即使使用 10k 个样本，概率估计也会非常不准确。相比之下，驯化展开方法能够在 $\tau = 3.0$之前给出非常紧致且信息丰富的上界。这项实验凸显了所提方法的优势：与基于采样的仿真不同，目标事件的稀有程度不会影响其预测性能。

四维倒立摆系统的结果如图 3 所示。其行为与二维范德波尔系统非常相似。具体而言，有一段时间，驯化方法给出的上界非常紧；随后，随着流管变得高度保守，驯化方法的准确性降低。几何界能够在发散前更长时间地保持更紧的估计，这凸显了使用定理 5 而非定理 3 的优势。稀有事件的结果（图 3b）与范德波尔实验一致：与蒙特卡罗仿真相比，在相当长的一段时间内都能得到非常准确的上界。

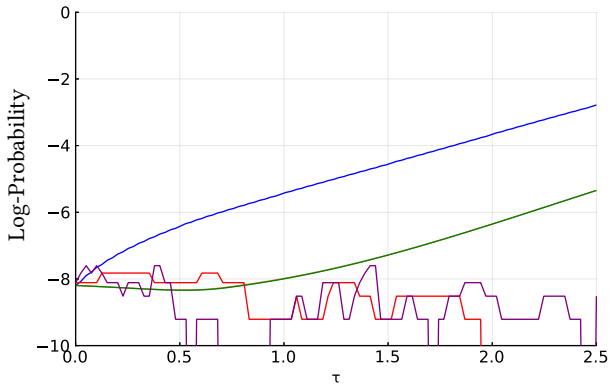
注 2。 当前先进的可达集工具 *[1, 4, 9]* 会采用更高保真度的近似来刻画区域的形状，例如区间多面体，或使用集合边界传播方法 *[28]*。这显著减轻了“包裹”效应，并使流管的保守性大幅降低，即使后处理中使用的是超矩形过近似。我们实验中使用的可达集实现采用伯恩斯坦多项式界，而非传统的区间算术界，但未采用高保真度的形状近似。将伯恩斯坦界嵌入当前先进的可达集工具中，可以显著延长准确预测的时间范围。

6.2 展开阶数消融研究

图 4 展示了范德波尔系统在不同展开阶数 k 下的消融研究。正如预期，增大 m 会得到更紧的上界。然而，当 m 较大时，例如 $m = 7$，上界单项式



(a) 概率与 τ的关系



(b) 对数概率与 τ 的关系（稀有事件）

图 3：四维倒立摆

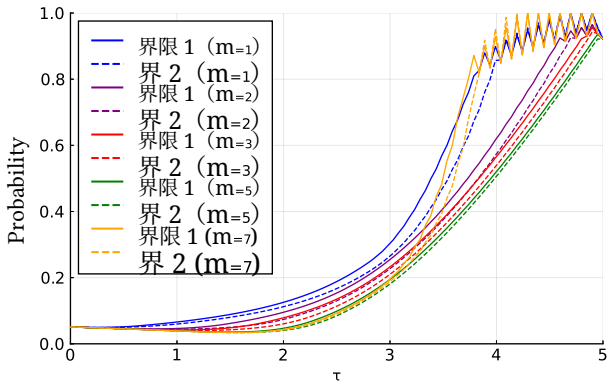


图 4：展开阶数 m 消融研究

（例如，(32)）的次数会更高，因此随着时间推移，更有可能发散。这与泰勒级数近似的结论一致：更高的展开阶数可以改进局部邻域估计，但可能导致在这些邻域之外更急剧地发散。

此外，由于范德波尔系统在状态幅值较大时会表现出混沌行为，占据状态空间大部分区域的流管区域可能具有急剧递减的体积函数，但为了保证包含所有状态，其体积仍必须增大。这解释了概率估计超过 0.8 时观察到的振荡行为。

7 讨论

所提出的方法利用一种特殊的域变换，无需追踪 $\Omega_\tau(t)$形状的高保真估计，而只需精确追踪体积。将这一思路与粗略的形状近似（通过可达集实现）相结合，可以在相当长的时间内得到非常紧的概率界。该方法针对学习得到的模型生成形式化证书，从而加深对模型预测的理解并增强对其的信任。此外，所提出的方法还能提升采用更高保真度的形状近似时，性能会进一步提升（见备注 2）。由于我们的方法会对流管进行“后处理”，因此可以轻松整合可达集工具和方法的未来进展。

所提出的方法可能适用于罕见事件的不确定性量化。仅用于近似（不提供形式化上界）时，使用帕德近似式 [3] 可能会得到显著˜更准确的 V_Ω结果，因为有理函数不一定像高阶多项式那样发散。此外，用于改善 $\Omega_\tau(t)$预测形状的离散化方法也是一个有前景的未来研究方向。

局限性。 虽然第 4 节所述条件并未假设使用伯恩斯坦多项式回归器，但满足所有条件的回归器类别可能很少。由于模型是从数据中学习得到的，因此总会存在统计学习误差；然而，本研究并未通过有限样本保证对该误差进行形式化量化。对于需要端到端 安全证书的应用，我们认为这是一个有价值的未来研究方向。此外，伯恩斯坦多项式中的系数数量随维度呈指数增长。 Γ_k 所需的内存为 $O((kd)^n)$，因此所提出的方法在高维系统或大阶数展开中会占用大量内存。

8 结论

本文针对受初始状态不确定性影响的非线性连续时间系统，研究了联合建模与不确定性传播/评估问题。所提出的方法将连续体动力学与泰勒级数近似理论相结合，从而针对学习得到的模型给出具有形式化保证的概率界，并渐近收敛到真实系统的概率。理论结果为模型选择提供了一份严谨的检查清单，以确保数学保证成立。所选用的伯恩斯坦多项式回归器与所提出的传播方法相结合，在生成基于模型的证书方面展现出良好的前景，尤其适用于稀有事件。

参考文献

[1] Althoff, M . CORA 2015简介。载于应用研讨会论文集

连续与混合系统的验证(2015)，第120–151页。[2] Amorese, P.与 Lahijanian, M. 使用Bernstein归一化流进行精确信念传播的随机动力学通用学习。载于*AAAI*人工智能会议论文集 (2026)，第40卷。[3] Baker Jr, G. A.与 Gammel, J. L. Padé逼近。数学分析与应用期刊*2*，1（1961），21–30。[4] Bogomolov, S.、Forets, M.、Frehse, G.、Potomkin, K.与Schilling, C. JuliaReach：基于集合的可达性工具箱。载于第*22*届*ACM*混合系统国际会议：计算与控制论文集 (2019年)，第39–44页。[5] Brockett, R. W. 刘维尔方程的最优控制。*AMS IP*高等数学研究*39* (2007)，23。[6] Brunton, S. L.、Proctor, J. L.与Kutz, J. N. 通过非线性动力系统稀疏辨识从数据中发现控制方程。美国国家科学院院刊*113*，15（2016年），3932–3937。[7] Chan, J. C.与Kroese, D. P. 基于条件蒙特卡洛的稀有事件概率估计。运筹学年鉴*189*，1（2011），43–61。[8] Chen, R. T.、Rubanova, Y.、Bettencourt, J.与Duvenaud, D. K. 神经常微分方程。神经信息处理系统进展*31* (2018)。[9] Chen, X.、Ábrahám, E.与 Sankaranarayanan, S. Flow*：非线性混合系统分析器。载于计算机辅助验证国际会议 (2013)，Springer，第258–263页。[10] Covella, F.与Fantuzzi, G. 非线性动力学的不确定性传播：一种多项式优化方法。*arXiv*预印本 *arXiv:2209.07432* (2022年)。[11] Ehrendorfer, M. 刘维尔方程及其对预报技巧预测的潜在用途。第一部分：理论。每月天气评论*122*，4（1994），703–713。[12] Garloff, J. 多元多项式值域的收敛上界。载于区间数学国际研讨会 (1985年)，Springer，第37–56页。[13] Gray, A.、Forets, M.、Schilling, C.、Ferson, S.与Benet, L. 非线性常微分方程中不精确概率的可验证传播。国际近似推理期刊*164* (2024)，109044。[14] Gurtin, M. E. 连续介质力学导论，第158卷。Academic Press，1982。[15] He, J.、Liao, Q.与Wan, X. 随机动力系统的自适应深度密度逼近。科学计算期刊*102*，3（2025），57。

[16] Hormander, L. 多变量复分析导论. Elsevier, 1973.[17] Kong, C.-W., Laurenti, L., McMahon, J., and Lahijanian, M. 物理信息神经网络用于福克-普朗克偏微分方程的误差上界。载于第四十一届人工智能不确定性会议（*UAI*）论文集（2025年7月）。[18] Li, J., and Chen, J. 概率守恒原理与广义密度演化方程。*Structural Safety 30*, 1（2008年），65–77。[19] Liu, S., Li, W., Zha, H., and Zhou, H. 神经参数化福克–普朗克方程。*SIAM*数值分析杂志 *60*, 3（2022年），1385–1449。[20] Lorentz, G. G. 伯恩斯坦多项式. American Mathematical Soc., 2012年。[21] Meng, Y., Sun, D., Qiu, Z., Waez, M. T. B., and Fan, C. 自主系统可达状态密度分布的学习。载于机器人学习会议（2022年），PMLR，第124–136页。[22] Murty, K. G., and Kabadi, S. N. 二次规划和非线性规划中的若干NP完全问题。技术报告，1985年。[23] Stone, M. H. 广义魏尔斯特拉斯逼近定理。*Mathematics Magazine 21*, 5（1948年），237–254。[24] Streif, S., Henrion, D., and Findeisen, R. 利用线性矩阵不等式进行概率和基于集合的模型失效判定与估计。*IFAC*会议论文集 *47*, 3（2014年），4110–4115。[25] Sun, Z.-J., Luo, Y.-Z., Di Lizia, P., and Zazzera, F. B. 利用微分代数和高斯混合模型进行非线性轨道不确定性传播。《中国科学：物理学 力学 天文学》*62*, 3（2019年），34511。[26] Terejanu, G., Singla, P., Singh, T., and Scott, P. D. 利用高斯混合模型进行非线性动力系统的不确定性传播。《制导、控制与动力学杂志》*31*, 6（2008年），1623–1633。[27] Veretennikov, A. Y., and Veretennikova, E. V. 多元伯恩斯坦多项式的偏导数。《西伯利亚数学进展》*26*, 4（2016年），294–305。[28] Xue, B., Easwaran, A., Cho, N.-J., and Fränzle, M. 基于边界分析的非线性系统可达-避障验证。*IEEE*自动控制汇刊 *62*, 7（2016年），3518–3523。[29] Yang, C., and Kumar, M. 蒙特卡洛方法在非线性动力系统初始不确定性预测中的有效性。*Automatica 87*（2018年），301–309。[30] Zeidler, E. 应用泛函分析：基本原理及其应用，第109卷。Springer Science & Business Media，2012年。