arXiv:2301.03080v1 [math.NA] 8 Jan 2023

Perron-Frob enius 随机动力系统的算子滤波器

刘宁欣* 姜立坚†

摘要

滤波问题源于对一个二次函数进行序贯最小化，该函数表示模型与数据之间的折中。本文利用随机过程中的 Perron-Frobenius 算子，开发了一种 Perron-Frobenius 算子滤波器。所提出的方法属于贝叶斯滤波，适用于非线性随机动力系统的非高斯分布。滤波递推过程可由 Perron-Frobenius 算子与似然算子的复合来表征。这建立了 Perron-Frobenius 算子与贝叶斯滤波之间的重要联系。我们通过乌拉姆方法近似 Perron-Frobenius 算子，以数值方式实现递推过程。由此，后验测度可表示为乌拉姆方法中示性函数的凸组合。为获得 Perron-Frobenius 算子滤波器的低秩近似，我们利用离散化 Perron-Frobenius 算子的特征函数对后验测度进行谱分解。本文还进行了收敛性分析，结果表明，Perron-Frobenius 算子滤波器的收敛速率高于使用狄拉克测度表示后验的粒子滤波器。本文探讨了所提出方法在随机动力系统数据同化中的应用，并通过若干数值算例展示 Perron-Frobenius 算子滤波器相较于粒子滤波器和扩展卡尔曼滤波器的优势。

关键词：Perron-Frobenius 算子、贝叶斯滤波、随机动力系统、粒子滤波器

1 引言

近年来，基于算子的方法已被广泛用于分析动力系统。该方法的两个主要候选算子是 Perron-Frobenius 算子及其对偶算子 Koopman 算子。人们已开发出许多数据驱动方法，用于对这些算子进行数值近似。这两个算子旨在从不同角度近似动力系统的行为。Koopman 算子用于研究观测的演化，而 Perron-Frobenius 算子（PFO）则刻画密度的转移。因此，PFO 处理的是系统状态概率密度函数形式的不确定性

*同济大学数学科学学院，中国上海 200092。（nxliu@tongji.edu.cn）。
†同济大学数学科学学院，中国上海 200092。（ljjiang@tongji.edu.cn）。

。在实践中，它确定了测度空间上由给定可测变换保持不变的一个绝对连续概率测度。

Perron-Frobenius 算子已广泛用于刻画源自许多不同领域的动力系统的全局渐近行为，例如流体动力学 [1], 分子动力学 [2], 气象学和大气科学 [3], ，并借助类似于 [4] 中的工具箱估计不变集或亚稳集。研究 Perron-Frobenius 算子的不变密度 [5] 并设计高效的数值方法，具有重要意义。这样便可以将遍历定理应用于确定性动力系统的统计性质。

由于 PFO 能够传输马尔可夫过程的密度，因此有必要对其进行近似，以便对马尔可夫过程的转移概率进行数值建模。人们提出了许多不同的数值方法 [6], ，例如乌拉姆方法和 Petrov-Galerkin 方法，用于近似 Perron-Frobenius 算子。由于 PFO 作用于无限维空间，将其投影到由适当基函数张成的有限维子空间上是很自然的。该投影通常通过 Galerkin 方法以弱近似的方式实现。这一方法最初由 Ulam 提出 [7], ，他指出，可以根据区域的有限划分，在由示性函数构成的有限维子空间 L^1 上研究离散 Perron-Frobenius 算子。许多文献讨论了乌拉姆方法的收敛性分析 [8, 9]。

随机过程中的经典滤波问题已在 [10, 11] 中得到研究。该论文考虑了离散时间和连续时间随机过程所对应的滤波问题模型，这些随机过程由随机微分方程的解定义，能够模拟现实世界中的大多数随机动力系统。滤波方法已广泛应用于地球物理领域，例如海洋学 [12], 、石油采收 [13], 、大气科学和天气预报 [14]。值得注意的是，作为一种滤波方法，卡尔曼滤波器 [15] 因在线性高斯模型中的低维工程应用而广为人知，也已在许多其他领域得到发展和应用 [16–18]。对于非线性问题，经典滤波器（如三维变分 [19], 、扩展卡尔曼滤波器 [10] 和集合卡尔曼滤波器 [20], ）通常采用高斯假设。它们常用于观测噪声较小且空间维度较高的情形。然而，这些扩展方法依赖于高斯假设。作为一种序贯蒙特卡洛方法，粒子滤波器 [21] 能够有效处理非线性、非高斯滤波问题。可以证明，当粒子数量趋于无穷大时，它能够估计真实的后验滤波结果。

尽管粒子滤波器（PF）能够处理非线性非高斯滤波问题，但在实际应用中仍有一些局限。首先，粒子滤波器在低维系统中表现良好，但当系统规模非常大时，可能会出现退化 [22] 。这意味着随着系统维数趋于无穷大，与样本集合相关联的最大权重会趋于 1。为避免退化，所需粒子数量必须随系统维数呈指数级增长。这正是维数灾难的一种表现。为规避维数灾难并准确估计高维概率密度函数（PDF），人们采用了重采样、添加抖动和局地化等方法。粒子滤波器在地球物理数据同化应用中的表现也不佳，因为这些应用中的数据对粒子位置有很强的约束 [23]。这会影响滤波性能。

此外，粒子滤波器需要已知转移概率密度函数的先验知识；对于复杂的密度函数，还需要使用接受-拒绝法和马尔可夫链蒙特卡罗等高效采样方法。在高维空间中，采样尤其具有挑战性。为克服这些困难，我们提出了一种 Perron-Frobenius 算子滤波器（PFOF），它不使用粒子或任何采样方法。该方法不需要先验概率密度信息，而是需要数据信息。该方法适用于非线性非高斯模型。

本文提出 PFOF，用于处理非线性随机滤波问题。该方法通过 Perron-Frobenius 算子传递滤波分布。对于滤波问题，滤波分布的更新包含两个步骤：预测和分析。在预测阶段，密度通过由底层系统的马尔可夫链给出的转移密度函数进行传输。在分析阶段，根据观测给出的似然函数，利用贝叶斯法则校正密度。因此，滤波分布的更新可以表示为 PFO 与似然函数的复合。在模拟过程中，使用乌拉姆方法对 PFO 进行离散化，并将其投影到由示性函数构成的有限维空间中。因此，滤波密度也被投影到该子空间，并表示为示性函数的线性凸组合。随后，滤波分布的递归可通过与基函数相关联的权重向量的线性映射来表示，该映射由 PFO 和似然函数确定。对于高维问题，我们提出一种采用谱分解的低秩 PFOF（lr-PFOF）。为此，我们首先使用离散化 PFO 的特征函数来表示密度函数的谱分解。然后对该分解进行截断，并使用对应于前几个主导特征值的特征函数。这可以提高在线同化效率。PFO 的思想还扩展到了连续时间滤波问题。在这些问题中，Zakai方程刻画了滤波密度的演变。我们利用 Perron-Frobenius 算子的近似来计算 Zaikai 方程，并获得连续时间滤波问题的后验密度函数。

我们将所提出的方法与粒子滤波器进行比较。对于 PFOF，我们给出了近似后验测度与真实后验之间全变差距离的误差估计。该估计表明，PFOF 的收敛速率为 $O(\frac{1}{N})$，快于使用相同数量 N 个基函数的粒子滤波器。数值结果表明，PFOF 的精度也优于扩展卡尔曼滤波器。

本文其余部分安排如下。第 2 节中，我们用 Perron-Frobenius 算子表述贝叶斯滤波器。第 3 节中，我们给出采用近似 Perron-Frobenius 算子的滤波经验测度递归公式。随后，我们推导 PFOF 和 lr-PFOF，并分析误差估计。我们还将 PFOF 扩展到连续时间滤波问题。第 4 节中，我们将其与粒子滤波器进行全面比较。第 5 节中，我们给出随机滤波问题的一些数值结果。第 6 节对全文进行总结。

2 预备知识

本节回顾 Perron-Frobenius 算子 [24] （PFO）和贝叶斯滤波器的相关背景。Perron-Frobenius 算子将分布映射到状态空间上，并描述动力系统的随机行为。我们介绍贝叶斯滤波器的框架，并将其概括为一个基于 PFO 的递归公式。

2.1 Perron-Frobenius 算子

设 X 为一个度量空间，$\mathcal{B}$ 为 X上的 Borel-σ-代数，且 $\Phi: X \to X$ 为一个非奇异变换。令 $\mathcal{M}$ 表示 (X, $\mathcal{B}$) 上所有有限测度构成的空间，且 μ 为一个有限测度。相空间定义在测度空间 (X, $\mathcal{B}$, μ) 上。Perron-Frobenius 算子 $\mathcal{P}: \mathcal{M} \to \mathcal{M}$ 是一个线性且无限维的算子，定义如下：

$$\mathcal{P}\mu(A) = \mu(\Phi^{-1}(A)), \quad \forall A \in \mathcal{B}. \tag{2.1}$$

PFO 是一个线性、正且非扩张的算子，因此是一个马尔可夫算子。我们也可以追踪函数空间中分布的作用。在本文中，我们将$L^1(X) := L^1(X, \mathcal{B}, \mu)$记作。令 $f \in L^1(X)$ 为取值于 X 的随机变量 x 的概率密度函数（PDF）。由于 Φ 相对于 μ是非奇异的，因此存在一个 $g \in L^1(X)$，满足 $\int_{\Phi^{-1}(A)} f d\mu = \int_A g d\mu$ 对所有 $A \in \mathcal{B}$均成立。于是，g 是刻画 $\Phi(x)$ 分布的函数。映射 $f \mapsto g$由线性算子 $\mathcal{P}: L^1(X) \to L^1(X)$ 唯一确定：

$$\int_A \mathcal{P}f \, d\mu = \int_{\Phi^{-1}(A)} f \, d\mu, \quad \forall A \in \mathcal{B}. \tag{2.2}$$

算子 $\mathcal{P}$ 称为 Perron-Frobenius 算子。根据定义 (2.1) 和 (2.2)，我们建立了概率密度函数与和 PFO 相关联的测度之间的联系。当 f 是相对于绝对连续概率测度 $\mu \in \mathcal{M}(X)$的概率密度函数时，g 是相对于绝对连续概率测度 $\mu \circ \Phi^{-1}$的另一个概率密度函数。此外，当 $\mathcal{P}\mu = \mu$ 成立时，测度 $\mu \in \mathcal{M}(X)$是 $\mathcal{P}$ 的不变测度。

设 $\Psi: \mathbb{R}_+ \times X \to X$ 是确定性连续时间动力系统的一个非奇异流映射。那么，对于每个 $\tau \in \mathbb{R}^+$，$\Psi_\tau: X \to X$ 都是非奇异的。传递算子$\mathcal{P}_\tau: L^1(X) \to L^1(X)$依赖于时间，其定义与 (2.2) 类似，因此

$$\int_A \mathcal{P}_\tau f \, d\mu = \int_{\Psi_\tau^{-1}(A)} f \, d\mu.$$

$\{\mathcal{P}_\tau\}_{\tau \geq 0}$构成 Perron-Frobenius 算子的一个半群。我们注意到，根据希勒-吉田定理，$\{\mathcal{P}_\tau\}_{\tau \geq 0}$具有无穷小生成元 $\mathscr{A}_{PF}$。

我们来考察随机动力系统中的 Perron-Frobenius 算子，以及与随机动力学方程（SDE）所诱导的随机解半流相关联的 PFO 无穷小生成元。设 $b: X \to X$ 和 $\sigma: X \to X$ 是光滑的时不变函数。假设随机过程 x_t 是以下时齐次随机微分方程的解：

$$dx_t = b(x_t)dt + \sigma(x_t)dW_t, \quad x(t_0) \sim \rho_0, \tag{2.3}$$

其中 W_t 是标准布朗运动。在这种情况下，随机过程 x_t 的分布可以由定义在 $L^1(X)$ 上的 Perron-Frobenius 算子半群 $\{\mathcal{P}_\tau\}_{\tau>0}$ 来描述。$\{\mathcal{P}_\tau\}_{\tau>0}$ 的生成元是定义在 X 上的二阶微分算子。由该生成元定义的偏微分方程描述了 x_t的概率密度演化。

假设 Φ 是随机动力系统 (2.3) 的映射，且 $\Phi(x)$ 是概率空间 $(X, \mathcal{B}, \mu)$ 上的一个 X 值随机变量。给定由 Φ 诱导的随机转移函数 $p_\tau: X \times \mathcal{B} \to [0,1]$，我们考虑由根据转移函数 $p_\tau(x,\cdot)$ 定义的线性算子变换后的概率测度 μ。随机佩龙-弗罗贝尼乌斯算子 [25] $\mathcal{P}_\tau: \mathcal{M} \to \mathcal{M}$ 定义为

$$\mathcal{P}_\tau \mu(A) = \int_X p_\tau(x, A)\, d\mu(x), \quad \forall A \in \mathcal{B}. \tag{2.4}$$

如果对于所有 $x \in X$，$p_\tau(x,\cdot)$ 都关于 μ 绝对连续，则存在一个非负转移密度函数 $q_\tau: X \times X \to \mathbb{R}$，满足 $q_\tau(x,\cdot) \in L^1(X)$，且

$$P(x_{t+\tau} \in A | x_t = x) = \int_A q_\tau(x,y) d\mu(y), \quad \forall A \in \mathcal{B}.$$

转移密度函数是马尔可夫链转移矩阵的无限维对应物。现在我们定义与转移密度相关的随机 PFO。若 $f \in L^1(X)$ 是概率密度函数，则算子 Perron-Frobenius 半群$\mathcal{P}_\tau: L^1(X) \to L^1(X)$ $\tau > 0$ 定义为

$$\mathcal{P}_\tau f(y) = \int_X q_\tau(x,y) f(x)\, d\mu(x).$$

此处定义的 PFO $\mathcal{P}_\tau$ 将 x_t 的概率密度函数随时间演化。设 ρ为概率密度。$\mathcal{P}_\tau$ 的无穷小生成元 $\mathscr{A}_{PF}$ 为

$$\mathscr{A}_{PF}\rho = -\nabla \cdot (b\rho) + \frac{1}{2}\nabla \cdot \nabla \cdot (\sigma\sigma^T \rho).$$

我们假设 $\widetilde{\rho}: [0,\infty) \times X \to [0,\infty)$ 是式 (2.3) 中解x_t 的概率密度函数，而 ρ_0 是初始条件 x_0的密度函数。那么，$\widetilde{\rho}$ 满足福克-普朗克方程:

$$\begin{cases} \dfrac{\partial \widetilde{\rho}}{\partial t} = -\nabla \cdot (b\widetilde{\rho}) + \dfrac{1}{2}\nabla \cdot \nabla \cdot (\sigma\sigma^T \widetilde{\rho}), \quad (t,x) \in (0,\infty) \times X, \\ \widetilde{\rho}(0,x) = \rho_0(x). \end{cases}$$

如果相空间 X 是紧致的，且 $b \in C^3(X,X)$，则该方程有唯一解，其表达式为

$$\widetilde{\rho}(t,x) = \mathcal{P}_t \rho_0(x).$$

2.2 贝叶斯滤波器

本节从贝叶斯法则的角度介绍离散时间贝叶斯滤波器的框架。在滤波问题中，状态模型和观测模型

相结合，用于估计后验分布，即给定观测的状态条件分布。考虑由流 $\Psi \in C(\mathbb{R}^n, \mathbb{R}^n)$ 所控制的动力学模型，其中含噪观测 $y = \{y_j\}_{j\in\mathbb{Z}^+}$ 取决于函数$h(x)$: $\mathbb{R}^n \to \mathbb{R}^p$:

$$\begin{cases} x_{j+1} = \Psi(x_j) + \xi_j,\ j \in \mathbb{N},\ x_0 \sim \rho_0, \\ y_{j+1} = h(x_{j+1}) + \eta_{j+1},\ j \in \mathbb{N}, \end{cases} \tag{2.5}$$

其中，$\xi := \{\xi_j\}_{j\in\mathbb{N}}$是一个服从 $\xi_j \sim N(0, \Sigma)$的独立同分布序列，而 $\eta := \{\eta_j\}_{j\in\mathbb{Z}^+}$是一个服从 $\eta_j \sim N(0, R)$的独立同分布序列。令 $Y_j = \{y_l\}_{l=1}^j$ 表示截至时间 t_j的数据。滤波问题旨在确定随机变量 $x_j|Y_j$ 的后验概率密度函数 $p(x_j|Y_j)$，并随着数据增加而逐步更新概率密度函数。贝叶斯滤波包含两个步骤：预测和分析。它推导出 $p(x_j|Y_j)$对应的 $p(x_{j+1}|Y_{j+1})$。预测关注映射 $p(x_j|Y_j) \mapsto p(x_{j+1}|Y_j)$，而分析则利用贝叶斯公式推导映射$p(x_{j+1}|Y_j) \mapsto p(x_{j+1}|Y_{j+1})$。

预测

$$\begin{aligned} p(x_{j+1}|Y_j) &= \int_{\mathbb{R}^n} p(x_{j+1}|Y_j, x_j)p(x_j|Y_j)dx_j \\ &= \int_{\mathbb{R}^n} p(x_{j+1}|x_j)p(x_j|Y_j)dx_j. \end{aligned} \tag{2.6}$$

注意，$p(x_{j+1}|Y_j, x_j) = p(x_{j+1}|x_j)$，因为 Y_j 为确定 x_{j+1}提供了间接信息。由于 $p(x_{j+1}|x_j)$由底层模型 (2.5) 指定，而

$$p(x_{j+1}|x_j) \propto \exp(-\frac{1}{2}|\Sigma^{-\frac{1}{2}}(x_{j+1} - \Psi(x_j))|^2), \tag{2.7}$$

预测给出了从 $p(x_j|Y_j)$ 到 $p(x_{j+1}|Y_j)$ 的映射。令 $\widehat{\mu}_j$为与密度 $p(x_j|Y_{j-1})$ 对应的先验概率测度，μ_j 为与密度 $p(x_j|Y_j)$ 对应的后验概率测度。(2.5) 中的随机过程 $\{x_j, j \in \mathbb{N}\}$ 是一个转移核为 $p(\cdot,\cdot)$ 的马尔可夫链，该转移核由 $p(x_j, x_{j+1}) = p(x_{j+1}|x_j)$ 确定。于是我们可以将 (2.6) 重写为

$$\widehat{\mu}_{j+1}(\cdot) = (\mathcal{P}\mu_j)(\cdot) := \int_{\mathbb{R}^n} p(x_j, \cdot)\mu_j(dx_j) = \int_{\mathbb{R}^n} p(x_j, \cdot)d\mu(x_j)^1. \tag{2.8}$$

特别地，算子 $\mathcal{P}$ 与 (2.4) 中定义的 Perron-Frobenius 算子一致。分析

$$\begin{aligned} p(x_{j+1}|Y_{j+1}) &= p(x_{j+1}|Y_j, y_{j+1}) \\ &= \frac{p(y_{j+1}|x_{j+1}, Y_j)p(x_{j+1}|Y_j)}{p(y_{j+1}|Y_j)} \\ &= \frac{p(y_{j+1}|x_{j+1})p(x_{j+1}|Y_j)}{p(y_{j+1}|Y_j)}. \end{aligned} \tag{2.9}$$

注意，$p(y_{j+1}|x_{j+1}, Y_j) = p(y_{j+1}|x_{j+1})$，第二个等式中使用了贝叶斯公式。似然函数 $p(y_{j+1}|x_{j+1})$ 由观测模型确定：$p(y_{j+1}|x_{j+1}) \propto$

[1]参见 [26], 若测度空间 (X, $\mathcal{B}$, μ) 上的函数 $f \in L^1(X)$ 被称为 μ 可积，则有$\int f d\mu = \int f(x)\mu(dx) = \int f(x)d\mu(x)$。

$\exp(-\frac{1}{2}|R^{-\frac{1}{2}}(y_{j+1}-h(x_{j+1}))|^2)$。令

$$g_j(x_{j+1}) := \exp(-\frac{1}{2}|R^{-\frac{1}{2}}(y_{j+1}-h(x_{j+1}))|^2). \tag{2.10}$$

分析给出了从 $p(x_{j+1}|Y_j)$到 $p(x_{j+1}|Y_{j+1})$的映射，因此我们可以将测度 $\mu_{j+1}(\cdot)$的更新表示为

$$\mu_{j+1}(\cdot) = (L_j\widehat{\mu}_{j+1})(\cdot) := \frac{g_j(x_{j+1})\widehat{\mu}_{j+1}(\cdot)}{\int_{\mathbb{R}^n} g_j(x_{j+1})\widehat{\mu}_{j+1}(\cdot)}, \tag{2.11}$$

其中，似然算子 L_j 定义为

$$(L_j\mu)(dx) = \frac{g_j(x)\mu(dx)}{\int_{\mathbb{R}^n} g_j(x)\mu(dx)}. \tag{2.12}$$

一般而言，预测和分析给出了从 μ_j 到 μ_{j+1}的映射。预测通过Perron-Frobenius 算子 $\mathcal{P}$ 将 μ_j 映射到 $\widehat{\mu}_{j+1}$ ，而分析通过似然算子 L_j将 $\widehat{\mu}_{j+1}$ 映射到μ_{j+1} 。随后，我们使用公式（2.8）和（2.11）表示 μ_{j+1} ，并将贝叶斯滤波总结如下:

$$\mu_{j+1} = L_j\mathcal{P}\mu_j. \tag{2.13}$$

假设 μ_0 是已知的初始概率测度。注意， $\mathcal{P}$ 不依赖于 j，因为预测步骤在每次j都由相同的马尔可夫过程控制。然而， L_j 依赖于 j ，因为每次 j都会使用不同的观测来计算似然。这样， μ_j 的演化过程经过线性算子 $\mathcal{P}$和非线性算子 L_j。通过对（2.13）进行数值迭代，可以实现对 μ_j 的近似。

3 基于 Perron-Frobenius 算子的贝叶斯滤波器

值得注意的是，Perron-Frobenius 算子会根据动力系统的流随时间传递概率密度函数。我们将这一思想扩展到滤波问题，用以表示后验概率密度函数（即滤波分布）的转移。因此，我们提出了一种滤波方法：Perron-Frobenius 算子滤波器（PFOF）。在该方法中，密度函数被投影到由示性函数张成的近似子空间上。密度函数的变化由权重向量近似，并通过 PFO 和似然算子传递。此外，我们还提出了低秩 Perron-Frobenius 算子滤波器（lr-PFOF），这是对 PFOF 的一种改进。

3.1 Perron-Frobenius 算子滤波器

迭代 (2.13) 有助于设计一种滤波方法。根据定义 (2.8)，迭代中的算子 $\mathcal{P}$ 是与模型的流 Ψ 相对应的 Perron-Frobenius 算子

(2.5)。基于这一思路，我们提出了一种 Perron-Frobenius 算子滤波器，它利用乌拉姆方法 [7] 来近似迭代中的算子 $\mathcal{P}$。在 PFOF 中，由于状态模型的离散时间步长保持不变，我们简单地使用 $\mathcal{P}$代替 $\mathcal{P}_\tau$。这样，PFOF 的滤波分布迭代变为

$$\mu_{j+1}^N = L_j \mathcal{P}^N \mu_j^N, \quad \mu_0^N = \mu_0, \tag{3.14}$$

其中，通过乌拉姆方法计算得到的 $\mathcal{P}^N$ 是 $\mathcal{P}$的近似。乌拉姆方法是一种用于离散化 Perron-Frobenius 算子的 Galerkin 投影方法。我们首先对相空间进行离散化。设 $B = \{\mathbb{B}_1, \cdots, \mathbb{B}_N\} \subset \mathcal{B}$为有限个测度盒，且构成相空间 X的不相交划分。示性函数是分段常数函数，定义如下：

$$\mathbb{1}_{\mathbb{B}_i}(x) = \begin{cases} 1, & \text{if x} \in \mathbb{B}_\text{i}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

Ulam 提出将一族示性函数$\{\mathbb{1}_{\mathbb{B}_1}, \cdots, \mathbb{1}_{\mathbb{B}_N}\}$所张成的空间用作 PFO 的逼近空间。我们定义投影空间 V_N：$=\text{span}\{\mathbb{1}_{\mathbb{B}_1}, \cdots, \mathbb{1}_{\mathbb{B}_N}\}$。 $V_N \in L^1(X)$ 被视为 $L^1(X)$ 的近似子空间。对于$L^1(X)$中的每个 $\rho \geq 0$，我们定义算子 π_N： $L^1(X) \to V_N$ 如下：

$$\pi_N \rho = \sum_{i=1}^N \omega^{(i)} \mathbb{1}_{\mathbb{B}_i}, \quad \text{where} \quad \omega^{(i)} := \frac{\int_{\mathbb{B}_i} \rho \, d\mu}{\mu(\mathbb{B}_i)}. \tag{3.16}$$

于是， π_N 是到 V_N上的投影。借助该投影，我们将离散化的 PFO$\mathcal{P}^N$ 定义为：

$$\mathcal{P}^N = \pi_N \circ \mathcal{P}.$$

我们可以用矩阵 $P^N = (P_{ij}^N) \in \mathbb{R}^{N \times N}$表示线性映射 $\mathcal{P}^N|_{V_N^1}$: $V_N^1 \to V_N^1$，其中 V_N^1: $= \left\{ f \in V_N : \int |f| d\mu = 1 \right\}$。矩阵的元素 $P_{ij}^N = {}_\mu{}^1{}_{\underline{(\mathbb{B}_i)}} \int_{\mathbb{B}_i} \mathcal{P} \mathbb{1}_{\mathbb{B}_j} d\mu$表示在流映射 Ψ作用下，从盒 $\mathbb{B}_i$到盒 $\mathbb{B}_j$的转移概率。我们可以证明

$$P_{ij}^N = \frac{\mu(\mathbb{B}_i \cap \Psi^{-1}(\mathbb{B}_j))}{\mu(\mathbb{B}_i)}. \tag{3.17}$$

对于 Ψ，马尔可夫链是 PFO 的离散化 $\mathcal{P}^N$，其转移矩阵为 P^N。因此，乌拉姆方法既可以用算子$\mathcal{P}^N$来描述，也可以用矩阵 P^N来描述。通过投影 π_N，密度 ρ可以表示为向量 $\mathbf{W} = [\omega^{(1)}, \cdots, \omega^{(N)}]$，其中 $\omega^{(i)}$是基函数 $\mathbb{1}_{\mathbb{B}_i}$的权重。由于矩阵元素 P_{ij}^N表示从 $\mathbb{B}_i$ 到 $\mathbb{B}_j$的转移概率，因此可以通过蒙特卡洛方法对其进行估计，从而得到乌拉姆方法的一种数值实现 [6]。我们在每个 $\mathbb{B}_i$ 中随机选取大量点 $\{x_i^l\}_{l=1}^n$，并统计其中落在盒 $\mathbb{B}_j$中的点 $\Psi(x_i^l)$ 的数量。然后按以下方式计算 P_{ij}^N：

$$P_{ij}^N \approx P_{n,ij}^N = \frac{1}{n} \sum_{l=1}^n \mathbb{1}_{\mathbb{B}_j}(\Psi(x_i^l)). \tag{3.18}$$

蒙特卡洛方法用于近似计算积分（3.17）。乌拉姆方法的收敛性取决于区域的划分方式和点的数量。基于示性函数基，我们将 PFOF 中关于测度 μ_j^N的经验密度记为

$$\rho_j^N(x) = \sum_{i=1}^{N} \omega_j^{(i)} \mathbb{1}_{\mathbb{B}_i}(x), \tag{3.19}$$

其中， $\mathbb{1}_{\mathbb{B}_i}(\cdot)$ 是由 (3.15) 定义的示性函数， j 表示时间序列的索引。这样，密度 ρ_j^N 便可表示为权重向量 $\mathbf{W}_j = [\omega_j^{(1)}, \cdots, \omega_j^{(N)}]$。假设 $\mathbf{W}_j = [\omega_j^{(1)}, \cdots, \omega_j^{(N)}]$ 和 $\mathbf{W}_{j+1} = [\omega_{j+1}^{(1)}, \cdots, \omega_{j+1}^{(N)}]$ 分别是 $\pi_N \rho_j$和 $\pi_N \rho_{j+1} := \pi_N \mathcal{P} \rho_j$的权重。当区域被均匀划分时，密度函数的演化可表示为权重的马尔可夫转移方程:

$$\mathbf{W}_{j+1} = \mathbf{W}_j P^N, \tag{3.20}$$

其中， P^N是离散化的 PFO 的矩阵形式。我们考虑 $\mathcal{P}\rho_j$的投影，即:

$$\pi_N \mathcal{P} \rho_j = \sum_{i=1}^{N} \omega_{j+1}^{(i)} \mathbb{1}_{\mathbb{B}_i}. \tag{3.21}$$

此外，注意到

$$\pi_N \mathcal{P} \rho_j = \pi_N \mathcal{P} \sum_{i=1}^{N} \omega_j^{(i)} \mathbb{1}_{\mathbb{B}_i} = \sum_{i=1}^{N} \omega_j^{(i)} \pi_N(\mathcal{P} \mathbb{1}_{\mathbb{B}_i}).$$

我们将 $\pi_N(\mathcal{P}\mathbb{1}_{\mathbb{B}_i})$ 记为 $= \sum_{k=1}^{N} c_{ik} \mathbb{1}_{\mathbb{B}_k}$，其中

$$\begin{aligned} c_{ik} &= \frac{\int_X \mathcal{P}(\mathbb{1}_{\mathbb{B}_i}) \mathbb{1}_{\mathbb{B}_k} dx}{\mu(\mathbb{B}_k)} = \frac{\int_{\mathbb{B}_k} \mathcal{P}(\mathbb{1}_{\mathbb{B}_i}) dx}{\mu(\mathbb{B}_k)} \\ &= \frac{\int_{\Psi^{-1}(\mathbb{B}_k)} \mathbb{1}_{\mathbb{B}_i} dx}{\mu(\mathbb{B}_k)} = \frac{\mu(\mathbb{B}_i \cap \Psi^{-1}(\mathbb{B}_k))}{\mu(\mathbb{B}_k)}. \end{aligned}$$

于是我们得到

$$\begin{aligned} \pi_N \mathcal{P} \rho_j &= \sum_{i=1}^{N} \omega_j^{(i)} \sum_{k=1}^{N} c_{ik} \mathbb{1}_{\mathbb{B}_k} = \sum_{k=1}^{N} \sum_{i=1}^{N} \omega_j^{(i)} c_{ik} \mathbb{1}_{\mathbb{B}_k} \\ &= \sum_{k=1}^{N} \sum_{i=1}^{N} \frac{\mu(\mathbb{B}_i)}{\mu(\mathbb{B}_k)} \omega_j^{(i)} P_{ik}^N \mathbb{1}_{\mathbb{B}_k}. \end{aligned}$$

与 (3.21) 比较，可得

$$\omega_{j+1}^{(k)} = \sum_{i=1}^{N} \frac{\mu(\mathbb{B}_i)}{\mu(\mathbb{B}_k)} \omega_j^{(i)} P_{ik}^N. \tag{3.22}$$

因此，若我们对 X 进行均匀划分，即 $\mu(\mathbb{B}_i) = \mu(\mathbb{B}_j) \quad \forall i, j \in N$，便可得到结果 (3.20)。根据这一表达式，我们设计了以下预测步骤和分析步骤，以近似后验分布 $p(x_{j+1}|Y_{j+1})$。

预测 在此步骤中，我们给出一组盒子 $\{\mathbb{B}_1, \cdots, \mathbb{B}_N\} \subset \mathcal{B}$，它们构成了 X 的均匀划分，并将每个盒子的质量点记为 $x^{(i)}, i = 1, \cdots, N$。令 $\widehat{\mathbf{W}}_j = [\widehat{\omega}_j^{(1)}, \cdots, \widehat{\omega}_j^{(N)}]$ 为先验权重向量，$\mathbf{W}_j = [\omega_j^{(1)}, \cdots, \omega_j^{(N)}]$ 为后验权重向量。在公式 (2.8) 中，我们注意到先验密度 $p(x_{j+1}|Y_j)$ 是通过线性算子 $\mathcal{P}$ 计算得到的。为将公式 $\widehat{\mu}_{j+1} = \mathcal{P}\mu_j$ 离散化，我们在密度函数的权重之间建立映射，

$$\widehat{\mathbf{W}}_{j+1} = \mathbf{W}_j P^N.$$

该公式包含底层系统 (2.5) 的先验信息，因为公式中的 PFO 是由系统的转移核 p 定义的。利用基函数$\mathbb{1}_{\mathbb{B}_i}(\cdot)$，经验先验测度可表示为

$$\widehat{\mu}_{j+1}^N = \sum_{i=1}^N \widehat{\omega}_{j+1}^{(i)} \mathbb{1}_{\mathbb{B}_i}(dx).$$

分析 在此步骤中，我们推导后验测度 μ_{j+1}^N。为此，我们将贝叶斯公式 (2.9) 应用于权重，并按如下方式更新权重：

$$\omega_{j+1}^{(i)} = \widetilde{\omega}_{j+1}^{(i)} / (\sum_{n=1}^N \widetilde{\omega}_{j+1}^{(n)}), \quad \widetilde{\omega}_{j+1}^{(i)} = g_j(x^{(i)}) \widehat{\omega}_{j+1}^{(i)}, \tag{3.23}$$

其中，由 (2.10) 给出的 $g_j(x)$ 与之前一样表示似然函数。然后，由示性函数测度近似得到的 μ_{j+1}^N 为

$$\mu_{j+1}^N = \sum_{i=1}^N \omega_{j+1}^{(i)} \mathbb{1}_{\mathbb{B}_i}(dx).$$

请注意，我们选择每个盒 $\mathbb{B}_i$ 的质量点 $x^{(i)}$ 来计算 $g_j(x)$，即似然函数。这是对 $\mathbb{B}_i$ 中各点的似然函数进行近似的合理选择。在预测步骤和分析步骤中，它们只通过 $\{\widehat{\omega}_{j+1}^{(i)}\}_{i=1}^N$ 将权重 $\{\omega_j^{(i)}\}_{i=1}^N$ 演化为 $\{\omega_{j+1}^{(i)}\}_{i=1}^N$，并提供从 μ_j^N 到 μ_{j+1}^N 的变换。完整流程总结于算法 2，称为 Perron-Frobenius 算子滤波器。该算法分为两个阶段：离线阶段通过乌拉姆方法计算 P^N，在线阶段则更新后验测度的近似。此外，标准乌拉姆方法在高维动力系统中会因维数灾难而效率低下。对于这种情况，我们可以改用稀疏乌拉姆方法 [27]。它构造一个最优近似子空间，并且在要求达到特定精度时，比标准乌拉姆方法所需的计算量更少。

算法1 Perron-Frobenius 算子滤波器

离线：通过乌拉姆方法计算 P^N 在线：1: 设置 $j=0$ 和 $\mu_0^N(dx_0)=\mu_0(dx_0)$，计算 $\omega_0^{(i)}=\frac{\int_{\mathbb{B}_i}\mu_0 dx_0}{\mu\ \mathbb{B}_i}$ ()2: 令 W $_j=[\omega_j^{(1)},\cdots,\omega_j^{(N)}]$，计算 $\widehat{\ }$W$_{j+1}=$W$_jP^N$3: 定义 $\widehat{\mu}_{j+1}^N=\sum_{i=1}^N\widehat{\omega}_{j+1}^{(i)}\mathbb{1}_{\mathbb{B}_i}(x)$4: 根据 (3.23) 表示 $\omega_{j+1}^{(i)}$，定义 $\mu_{j+1}^N=\sum_{i=1}^N\omega_{j+1}^{(i)}\mathbb{1}_{\mathbb{B}_i}(x)$5: j+1→j6: 转到步骤 2

3.2 误差估计分析

本节分析 Perron-Frobenius 算子滤波器的误差估计，以探究决定算法收敛性的因素。定义两个概率测度 μ 和 ν 之间的全变差距离 $d(\cdot,\cdot)$如下：

$$d(\mu,\nu)=\frac{1}{2}\sup_{|f|_\infty\le 1}|\mathbb{E}^\mu(f)-\mathbb{E}^\nu(f)|,$$

其中，对于 $f\in L^1(X)$，有 $\mathbb{E}^\mu(f)=\int_X f(x)\mu(dx)$，且 $|f|_\infty=\sup_x|f(x)|$。距离 $d(\cdot,\cdot)$也可以用分别对应测度 μ 和 ν的两个概率密度函数 ρ 和 ρ'之差的 L^1 范数来刻画，即，

$$d(\mu,\nu)=\frac{1}{2}\int_X|\rho(x)-\rho'(x)|dx. \tag{3.24}$$

该距离诱导出一个度量。为估计误差，我们回顾迭代式（3.14），并注意到概率的近似误差来自算子 $\mathcal{P}^N$。为此，我们需要以下引理。

引理 3.1。（[23]中的定理 4.8）设 $\mathcal{P}$ 是式（2.8）中定义的 Perron-Frobenius 算子。设 μ 和 ν 是任意两个概率测度。则

$$d(\mathcal{P}\mu,\mathcal{P}\nu)\le d(\mu,\nu).$$

引理 3.2。（[23]中的引理 4.9）令 g_j 为由（2.10）定义的似然函数，L_j 为由（2.12）定义的算子。假设存在 $\kappa\in(0,1]$，使得对所有 $x\in X$ 和 $j\in\mathbb{N}$,

$$\kappa\le g_j(x)\le\kappa^{-1}. \tag{3.25}$$

则有

$$d(L_j\mu,L_j\nu)\le 2\kappa^{-2}d(\mu,\nu).$$

引理 3.3。（[28] 中的定理 2.4.1）设 $\mathcal{C}_a$ 为定义如下的离散利普希茨锥

$$\mathcal{C}_a=\{\phi:\frac{\phi(x)}{\phi(y)}\le e^{a|x-y|},\forall x,y\in\mathcal{R}\}.$$

对于每个 $N>0$，由 (3.16) 给出的 π_N 表示将 L^1 投影到 V_N 上。则对于任意函数 $f\in\mathcal{C}_a$,

$$\|f-\pi_N f\|_{L^1}\leq(e^{a/N}-1)\|f\|_{L^1}.$$

根据上述三个引理，我们分析近似测度 μ_J^N 与真实测度 μ_J 之间的全变差距离，并给出以下定理。

定理 3.4. 若 $g_j(x)$ 满足条件 (3.25)，且测度 μ_j^N的概率密度 ρ_j^N满足 $\mathcal{P}\rho_j^N\in\mathcal{C}_a$, $\forall j\in\mathbb{N}$, 则

$$d(\mu_J^N,\mu_J)\leq\sum_{j=1}^{J}(2\kappa^{-2})^j\frac{e^a-1}{2N}.$$

证明. 根据公式 (2.13) 和 (3.14)，我们对距离 $d(\mu_{j+1}^N,\mu_{j+1})$ 应用三角不等式，得到

$$\begin{aligned}d(\mu_{j+1}^N,\mu_{j+1})&=d(L_j\mathcal{P}^N\mu_j^N,L_j\mathcal{P}\mu_j)\\&\leq d(L_j\mathcal{P}^N\mu_j^N,L_j\mathcal{P}\mu_j^N)+d(L_j\mathcal{P}\mu_j^N,L_j\mathcal{P}\mu_j).\end{aligned}$$

根据引理 3.2 和引理 3.1，可得

$$\begin{aligned}d(\mu_{j+1}^N,\mu_{j+1})&\leq 2k^{-2}\big[d(\mathcal{P}^N\mu_j^N,\mathcal{P}\mu_j^N)+d(\mathcal{P}\mu_j^N,\mathcal{P}\mu_j)\big]\\&\leq 2k^{-2}\big[d(\mathcal{P}^N\mu_j^N,\mathcal{P}\mu_j^N)+d(\mu_j^N,\mu_j)\big],\end{aligned}\tag{3.26}$$

考虑 $d(\mathcal{P}^N\mu_j^N,\mathcal{P}\mu_j^N)$。设 ρ'_{j+1}是与测度 $\mathcal{P}\mu_j^N$相关联的密度函数。设 ρ_j^N和 ρ_{j+1}^N分别为 μ_j^N和 μ_{j+1}^N的密度函数。根据式 (3.24) 中全变差距离的定义，我们有

$$\begin{aligned}d(\mathcal{P}^N\mu_j^N,\mathcal{P}\mu_j^N)&=\frac{1}{2}\int_X|\rho_{j+1}^N(x)-\rho'_{j+1}(x)|dx\\&=\frac{1}{2}\int_X|\mathcal{P}^N\rho_j^N(x)-\mathcal{P}\rho_j^N(x)|dx\\&=\frac{1}{2}\int_X|\pi_N\circ\mathcal{P}\rho_j^N(x)-\mathcal{P}\rho_j^N(x)|dx,\end{aligned}$$

其中，最后一个等式用到了方程 $\mathcal{P}^N=\pi_N\circ\mathcal{P}$。由于 $\mathcal{P}\rho_j^N(x)\in\mathcal{C}_a$，我们使用引理 3.3，得到

$$\begin{aligned}d(\mathcal{P}^N\mu_j^N,\mathcal{P}\mu_j^N)&\leq\frac{1}{2}(e^{a/N}-1)\\&\leq\frac{1}{2N}(e^a-1).\end{aligned}\tag{3.27}$$

根据 $\mu_0^N=\mu_0$，我们将 (3.27) 与 (3.26) 结合，并重复迭代以完成证明。

□

定理 3.4 估计了 PFOF 算法的在线误差。由于 Perron-Frobenius 算子在离线阶段通过式 (3.18) 给出的矩阵形式 P_n^N进行数值近似，我们将分析该近似产生的离线误差。P_n^N中的每个系数都是通过使用一组采样点 $\{x_i^l\}_{l=1}^n$对式 (3.17)进行蒙特卡洛近似计算得到的。我们证明矩阵 P_n^N依概率收敛到矩阵 P^N。

命题 3.5. 若矩阵 P_n^N由式 (3.18)定义，且 P^N由式 (3.17)定义，则有如下依分布收敛：

$$\sqrt{n}((P_n^N)_{ij} - (P^N)_{ij}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma_{ij}^{n,N}), \tag{3.28}$$

其中

$$(\sigma_{ij}^{n,N})^2 = \int_X (\mathbb{1}_{\Psi_\tau^{-1}(\mathbb{B}_j)} \cdot \mathbb{1}_{\mathbb{B}_i})^2 d\mu - (\int_X \mathbb{1}_{\Psi_\tau^{-1}(\mathbb{B}_j)} \cdot \mathbb{1}_{\mathbb{B}_i} d\mu)^2, \tag{3.29}$$

且 $\mathcal{N}(0, \sigma_{ij}^{n,N})$ 是均值为 0、标准差为 $\sigma^{n,N}_{,i,j}$ 的正态分布 .

证明。注意， P_n^N的各项为

$$P_{n,ij}^N = \frac{1}{n} \sum_{l=1}^{n} \mathbb{1}_{\mathbb{B}_j}(\Psi_\tau(x_i^l)),$$

这是下式的蒙特卡洛近似

$$P_{ij}^N = \frac{\int_X \mathbb{1}_{\Psi_\tau^{-1}(\mathbb{B}_j)} \cdot \mathbb{1}_{\mathbb{B}_i} d\mu}{\int_X \mathbb{1}_{\mathbb{B}_i} d\mu},$$

采样点 x_i^l 从盒 $\mathbb{B}_i$中独立且均匀地抽取。分母 $\int_X \mathbb{1}_{\mathbb{B}_i} d\mu$ 对各项 P_{ij}^N进行归一化，使得 P^N 成为右随机矩阵，每行的元素之和为 1。收敛结果 (3.28) 直接由蒙特卡洛积分的收敛性 [29] 得出。

□

命题 3.5 表明，存在一个常数 $C^N(\Psi_\tau, \alpha)$，由标准差 $\sigma^{n,N}_{,ij}$和给定的置信率 $\alpha \in [0, 1)$ 决定，使得当 m足够大时，以下估计以至少 α 的概率成立：

$$\| (P_n^N)_{ij} - (P^N)_{ij} \|_\infty \leq C^N(\Psi_\tau, \alpha) n^{-\frac{1}{2}}. \tag{3.30}$$

结果表明， P_n^N 收敛到 P^N 的速度为 $\mathcal{O}(n^{-\frac{1}{2}})$。

3.3 低秩 Perron-Frobenius 算子滤波器

在 PFOF 中，我们注意到块的数量会随着维数呈指数级增长，导致基函数数量迅速增加。因此，我们提出一种由 Perron-Frobenius 算子的特征函数构成的低秩近似，用于表示密度。这种方法可以有效减少所需基函数的数量。设 ρ 仍为由 Ψ 控制的动力系统的概率密度函数。那么，它可以表示为 $\mathcal{P}$ 的独立特征函数 φ_i 的线性组合。因此

$$\rho(x, t) = \sum_{i=1}^{\infty} a_i \varphi_i(x), \quad a_i \in \mathbb{C}.$$

假设 λ_i 是 $\mathcal{P}$ 的特征函数 φ_i 对应的特征值，那么

$$\mathcal{P}\rho(x,t)=\sum_{i=1}^{\infty}\lambda_i a_i\varphi_i(x).$$

实际上，离散化的 PFO 的特征函数可由以下命题确定。

命题 3.6. 设 $B=\{\mathbb{B}_1,\cdots,\mathbb{B}_N\}\subset\mathcal{B}$是相空间$X$的均匀划分。如果 ξ 是对应于特征值 λ 的 P^N 的左特征向量，那么 λ 也是限制算子 $\pi_N\mathcal{P}$的特征值，其特征函数为 φ，$\xi^T\mathrm{U}$，其中 $\mathrm{U}=[\mathbb{1}_{\mathbb{B}_1}(x)\ \cdots,\mathbb{1}_{\mathbb{B}_N}(x)]^T$。

证明。设 $\varphi=\sum_i\xi^{(i)}\mathbb{1}_{\mathbb{B}_i}$。由式 (3.21) 和式 (3.22),

$$\pi_N\mathcal{P}\varphi=\sum_{j}^{N}\sum_{i}^{N}\xi^{(i)}P_{ij}^{N}\mathbb{1}_{\mathbb{B}_j}.$$

由于 $\xi P^N=\lambda P^N$，即 $\lambda\xi^{(j)}=\sum_i\xi^{(i)}P_{ij}^N$, $\forall j\in\mathbb{N}$，我们得到

$$\pi_N\mathcal{P}\varphi=\sum_{j}^{N}\lambda\xi^{(j)}\mathbb{1}_{\mathbb{B}_j}=\lambda\varphi.$$

因此，λ 也是限制算子 $\pi_N\mathcal{P}$ 的特征值，其特征函数为 φ。 □

为了获得密度函数 $\rho_j:=\rho(x,t_j)$ 的谱展开，我们定义矩阵 $\boldsymbol{\varphi}=[\varphi_1,\cdots,\varphi_N]^T$，其中 $\{\varphi_i\}_{i=1}^N$ 是对应于特征值 $\{\lambda_i\}_{i=1}^N$ 的特征函数，且 $|\lambda_1|=1\geq|\lambda_2|\geq\cdots\geq|\lambda_N|\geq 0$。令 $\rho_0=\mathrm{W}_0\mathrm{U}$，并且

$$\Xi=\begin{bmatrix}\xi_1^T\\ \xi_2^T\\ \vdots\\ \xi_N^T\end{bmatrix}.$$

于是，特征函数记为 $\boldsymbol{\varphi}=\Xi\mathrm{U}$，而 ρ_1 的密度函数为

$$\rho_1=\pi_N\mathcal{P}\rho_0=\pi_N\mathcal{P}\mathbf{W}_0\mathbf{U}=\pi_N\mathcal{P}\mathbf{W}_0\Xi^{-1}\boldsymbol{\varphi}=\Lambda\mathbf{W}_0\Xi^{-1}\boldsymbol{\varphi}=\sum_{i=1}^{N}\lambda_i\varphi_i v_i,\tag{3.31}$$

其中，Λ 是 $\pi_N\mathcal{P}$ 的对角特征值矩阵，而 v_i 是矩阵$V=\mathrm{W}_0\Xi^{-1}$的列向量。使用前 r 个主要特征值及其对应的特征函数来近似密度函数。若公式（3.31）截断至 $r<N$，则 ρ_1 的低秩模型为

$$\rho_1=\sum_{i=1}^{r}\lambda_i\varphi_i v_i.$$

这样，时刻 t_j 的密度函数低秩模型为

$$\rho_j = \sum_{i=1}^{r} \lambda_i^j \varphi_i v_i, \quad j \in \mathbb{N}.$$

我们将 Perron-Frobenius 算子的低秩近似记为 $\rho_j = \widetilde{\mathcal{P}}\rho_{j-1}$,$\sum_{i=1}^{r} \lambda_i \varphi_i v_{j-1,\ i}$，其中 $v_{j-1,\ i}$ 是矩阵 $V_{j-1} = \mathsf{W}_{j-1}\Xi^{-1}$的列向量。我们在贝叶斯滤波器中应用$\sim\mathcal{P}$，得到低秩 Perron-Frobenius 算子滤波器（lr-PFOF），其概率测度满足递归公式

$$\mu_{j+1}^N = L_j \widetilde{\mathcal{P}} \mu_j^N, \quad \mu_0^N = \mu_0.$$

为描述接下来的预测和分析步骤，我们首先计算弱近似 P^N，并得到 P^N的特征值 Λ和左特征向量 Ξ。

预测 在此步骤中，我们对先验密度 $p(x_{j+1}|Y_j)$进行模型分解。首先，得到满足 $p(x_j|Y_j) = \mathsf{W}j\mathsf{U}$ 的 W_j，该矩阵来自前一分析步骤。接下来，计算矩阵 $V_j = \mathsf{W}_j\Xi^{-1}$ 并

$$\widehat{\rho}_{j+1} = p(x_{j+1}|Y_j) = \sum_{i=1}^{r} \lambda_i \varphi_i v_{j,i}.$$

分析 在这一步中，我们通过贝叶斯公式推导后验密度 $p(x_{j+1}|Y_{j+1})$。将 $p(x_{j+1}|Y_j)$ 乘以似然函数 g_j，得到

$$\rho_{j+1} = p(x_{j+1}|Y_{j+1}) \propto \sum_{i=1}^{r} \lambda_i \varphi_i v_{j,i} g_j.$$

为了对 ρ_{j+1}进行归一化，我们重写 $\widehat{\rho}_{j+1}$。由于 $\varphi_i = \xi_i^T\mathsf{U} = \sum_{k=1}^{N} \xi_i^{(k)} \mathbb{1}_{\mathbb{B}_k}$，得到

$$\widehat{\rho}_{j+1} = \sum_{i=1}^{r} \lambda_i \sum_{k=1}^{N} \xi_i^{(k)} \mathbb{1}_{\mathbb{B}_k} v_{j,i} = \sum_{k=1}^{N} \big(\sum_{i=1}^{r} \lambda_i \xi_i^{(k)} v_{j,i} \big) \mathbb{1}_{\mathbb{B}_k}.$$

然后，我们乘以 $g_j(x)$，并对 $\mathbb{1}_{\mathbb{B}_k}$ 的权重进行归一化，使得

$$\omega_{j+1}^{(k)} = \widetilde{\omega}_{j+1}^{(k)} / (\sum_{n=1}^{N} \widetilde{\omega}_{j+1}^{(n)}), \quad \widetilde{\omega}_{j+1}^{(k)} = \sum_{i=1}^{r} \lambda_i \xi_i^{(k)} v_{j,i} g_j(x^{(k)}), \tag{3.32}$$

其中，$x^{(k)}$ 仍然是每个盒 $\mathbb{B}_k$ 的质量点。后验密度变为

$$\rho_{j+1} = \sum_{k=1}^{N} \omega_{j+1}^{(k)} \mathbb{1}_{\mathbb{B}_k} = \mathbf{W}_{j+1}\mathbf{U}.$$

注释 3.1。请注意，在矩阵 P^N 的特征分解中可能会出现复特征值和特征向量。当系统的平稳分布 $\widetilde{\pi}$ 满足细致平衡时，[30] 中提出了一种对称化方法来解决这一问题。由于

算法 2 低秩 Perron-Frobenius 算子滤波器

离线:
计算 P^N 及其特征值 Λ 和左特征向量 Ξ。给出特征函数 $\boldsymbol{\varphi} = \Xi \mathsf{U}$。

在线:
1: 设置 $j = 0$ 和 $\rho_0 = \mathsf{W}_0 \mathsf{U}$，计算 $\omega_0^{(i)} = \frac{\int_{\mathbb{B}_i} \mu_0 dx_0}{\mu\ \mathbb{B}_i}$ ()2: 记 $V_j = \mathsf{W}_j \Xi^{-1}$，计算 $\widehat{\rho}_{j+1} = \sum_{i=1}^r \lambda_i \varphi_i v_{j,i}$3: 根据 (2.10) 定义 g_j，给出 $\rho_{j+1} \propto \sum_{i=1}^r \lambda_i \varphi_i v_{j,i} g_j$
4: 根据 (3.32) 对权重进行归一化，得到 W_{j+1}，令 $\rho_{j+1} = \sum_{k=1}^N \omega_{j+1}^{(k)} \mathbb{1}_{\mathbb{B}_k}$。
5: j+1→j6: 转到步骤 2

$\rho_0, \rho_1, \cdots$ 可以视为一个转移矩阵为 P^N 的马尔可夫链，我们假设 P^N 相对于 $\widetilde{\pi}$ 满足细致平衡，即，

$$\widetilde{\pi}_i P_{ij}^N = \widetilde{\pi}_j P_{ji}^N, \quad \forall i, j \in N.$$

那么， P^N 可以通过相似变换进行对称化

$$S = \widetilde{\Lambda} P^N \widetilde{\Lambda}^{-1}, \quad \text{where } \widetilde{\Lambda} = \begin{bmatrix} \sqrt{\widetilde{\pi}_1} & & & \\ & \sqrt{\widetilde{\pi}_2} & & \\ & & \ddots & \\ & & & \sqrt{\widetilde{\pi}_N} \end{bmatrix}.$$

这里， S 是一个对称矩阵，这一点可以通过细致平衡方程轻松验证。已知 S 具有一组完整的实特征值 $\alpha_j \in \mathbb{R}$ 和一组正交特征向量 w_j。因此， P^N 具有相同的特征值 α_j 以及实左特征向量

$$\psi_j = \widetilde{\Lambda} w_j.$$

3.4 扩展至连续时间滤波问题 s

在本小节中，我们考虑一个连续时间滤波问题，其中状态模型和观测由以下随机微分方程给出:

$$\frac{dx}{dt} = f(x) + \sqrt{\Sigma_c} \frac{dW_t}{dt}, \quad x(t_0) \sim \mathcal{N}(m_0, C_0), \tag{3.33}$$

$$\frac{dz}{dt} = h(x) + \sqrt{R_c} \frac{dW_t}{dt}, \quad z(0) = 0. \tag{3.34}$$

其中， Σ_c 是模型误差的协方差， R_c 是观测误差的协方差。假设连续时间问题所支配的后验测度 μ_t 具有勒贝格密度 $\rho(\cdot, t)$: $\mathbb{R}^n \mapsto \mathbb{R}^+$，其中t为固定值。令 $\rho(x,t) = r(x,t) / \int_{\mathbb{R}^n} r(x,t) dx$，其中 r 是未归一化密度。对于正定对称矩阵 $A \in \mathbb{R}^{p \times p}$，我们在空间 $L^2([0,T]; \mathbb{R}^p)$上定义加权内积 $\langle \cdot, \cdot \rangle_A = \langle A^{-\frac{1}{2}} \cdot, A^{-\frac{1}{2}} \cdot \rangle$。由此得到的

范数 $|\cdot|_A = |A^{-\frac{1}{2}}\cdot|$. 在连续滤波问题中，我们关注的是随着时间 t 推移，随机变量 $x(t)$ $|\{z(s)\}_{s\in[0,t]}$ 的分布。Zakai随机偏微分方程（SPDE）是一种著名的方程，其解刻画了后验分布 [35]的未归一化密度。Zakai方程的形式为

$$\frac{\partial r}{\partial t} = \mathscr{A}_{PF} r + r\left\langle h, \frac{dz}{dt}\right\rangle_{R_c}. \tag{3.35}$$

偏微分算子 $\mathscr{A}_{PF}$ 生成连续的 Perron-Frobenius 半群$\{\mathcal{P}_t, t\geq 0\}$。令 $\{\mathcal{Q}_s^t, 0\leq s\leq t\}$ 为与以下随机微分方程相关联的随机半群 [31]

$$\frac{dr'}{dt} = r'\left\langle h, \frac{dz}{dt}\right\rangle_{R_c}. \tag{3.36}$$

然后，可以通过以下类Trotter乘积公式近似 Zakai方程 (3.35)

$$r_{j+1} = Q_{t_j}^{t_{j+1}} \mathcal{P}_\tau r_j, \tag{3.37}$$

其中 $\tau = t_{j+1} - t_j$, $\forall j\in\mathbb{N}$。对于固定的 τ， $\mathcal{P}_\tau$仍记为 $\mathcal{P}$。根据参考文献 [31]，$Q_{t_j}^{t_{j+1}}$ 描述了方程 (3.36) 的解，即,

$$Q_{t_j}^{t_{j+1}} r(x) = \exp\left(\langle h(x), z_{j+1} - z_j\rangle_{R_c} - \frac{\tau}{2}|h(x)|_{R_c}^2\right) r(x).$$

利用离散格式 (3.37)，我们使用 Perron-Frobenius 算子求解 Zakai方程，而不是使用 Fokker-Planck算子 $\mathscr{A}_{PF}$。因此，我们通过乌拉姆方法离散化 $\mathcal{P}$，并将密度函数投影到 V_N上。令 P^N为 $\mathcal{P}$的离散化。令 $\mathbf{W}_j$和 $\mathbf{W}_{j+1}$为关于 $\pi_N r_j$和 $\pi_N r_{j+1}$的权重向量。记 $g_j^c(x) = \exp\left(\langle h(x), z_{j+1} - z_j\rangle_{R_c} - \frac{\tau}{2}|h(x)|_{R_c}^2\right)$ 和

$$\mathbf{G}_j = \begin{bmatrix} g_j^c(x^{(1)}) \\ g_j^c(x^{(2)}) \\ \vdots \\ g_j^c(x^{(N)}) \end{bmatrix},$$

其中， $x^{(i)}$ 是 $\mathbb{B}_i$的质量点。于是，密度函数的转移转化为权重映射,

$$\mathbf{W}_{j+1} = \mathbf{G}_j \odot \left(\mathbf{W}_j P^N\right).$$

这里， $\odot$ 表示Hadamard积。在这种情况下，PFO被扩展到连续时间滤波问题，以估计后验密度函数。

4 与粒子滤波器的比较

粒子滤波器（PF） [32, 33] 是一种重要的滤波方法，在粒子数量趋于无穷大时，可逐步逼近真实的后验滤波分布 $p(x_j|Y_j)$。在

实践中，我们通过粒子的位置与狄拉克函数对应权重的组合来近似概率密度。粒子滤波器通过调整权重并确定粒子的狄拉克测度来运行。它能够处理非高斯和非线性模型。本节将比较 PFOF 与 PF 的计算精度及其差异。

因此，我们将 μ_j^N定义为 $\mathbb{R}^N$ 上的后验经验测度，用以近似真实后验概率测度 μ_j；并将 $\widehat{\mu}_j^N$定义为 $\mathbb{R}^N$ 上的先验概率测度 $\widehat{\mu}_j$的近似。令

$$\mu_j \approx \mu_j^N := \sum_{n=1}^{N} \omega_j^{(n)} \delta_{x_j^{(n)}}, \quad \widehat{\mu}_{j+1} \approx \widehat{\mu}_{j+1}^N := \sum_{n=1}^{N} \widehat{\omega}_{j+1}^{(n)} \delta_{\widehat{x}_{j+1}^{(n)}},$$

其中， $x_j^{(n)}$ 和 $\widehat{x}_{j+1}^{(n)}$ 是粒子位置， $\omega_j^{(n)} > 0$和 $\widehat{\omega}_{j+1}^{(n)} > 0$ 是相应的权重，满足 $\sum_{n=1}^{N} \omega_j^{(n)} = 1$和 $\sum_{n=1}^{N} \widehat{\omega}_{j+1}^{(n)} = 1.$。经验分布完全由粒子位置和权重确定。粒子滤波器的目标是计算更新量 $\{x_j^{(n)}, \omega_j^{(n)}\} \to \{\widehat{x}_{j+1}^{(n)}, \widehat{\omega}_{j+1}^{(n)}\}$和 $\{\widehat{x}_{j+1}^{(n)}, \widehat{\omega}_{j+1}^{(n)}\} \to \{x_{j+1}^{(n)}, \omega_{j+1}^{(n)}\}$，分别定义预测步骤和分析步骤。预测步骤中使用蒙特卡洛采样确定粒子位置，分析步骤中使用贝叶斯规则更新权重。

预测 在此步骤中，预测阶段由马尔可夫链 $\{\Psi(x_j)\}_{j\in\mathbb{N}}$ 近似，其转移核为 $p(x_j, x_{j+1}) = p(x_{j+1}|x_j)$。我们从以 $x_j^{(n)}$ 为起点的核 p 中抽取 $\widehat{x}_{j+1}^{(n)}$，即 $\widehat{x}_{j+1}^{(n)} \sim p(x_j^{(n)}, \cdot)$。我们保持权重不变，使得 $\widehat{\omega}_{j+1}^{(n)} = \omega_j^{(n)}$，并得到先验概率测度

$$\widehat{\mu}_{j+1}^N = \sum_{n=1}^{N} \omega_j^{(n)} \delta_{\widehat{x}_{j+1}^{(n)}}.$$

分析 在此步骤中，我们应用贝叶斯公式来近似后验概率测度。为此，我们固定粒子的位置并更新权重。根据式 (2.10) 中 $g_j(x)$ 的定义，我们得到经验后验分布

$$\mu_{j+1}^N = \sum_{n=1}^{N} \omega_{j+1}^{(n)} \delta_{\widehat{x}_{j+1}^{(n)}},$$

其中

$$\omega_{j+1}^{(n)} = \widetilde{\omega}_{j+1}^{(n)} / (\sum_{n=1}^{N} \widetilde{\omega}_{j+1}^{(n)}), \quad \widetilde{\omega}_{j+1}^{(n)} = g_j(\widehat{x}_{j+1}^{(n)}) \omega_j^n. \tag{4.38}$$

(4.38) 中的第一个方程是一个归一化过程。序贯重要性重采样（SIR）粒子滤波器是一种基本的粒子滤波器，如算法1所示。该算法引入了一个重采样步骤。这样，当初始测度 μ_0 不是狄拉克函数的组合时，我们也能处理这种情况。对于某些粒子权重接近 1 的情形，我们也能进行处理。该算法表明，每个粒子都按照

底层模型移动，并根据似然重新赋予权重。通过贝叶斯滤波的迭代，我们将粒子滤波器改写为如下近似形式

$$\mu_{j+1}^N = L_j S^N \mathcal{P} \mu_j^N, \quad \mu_0^N = \mu_0, \tag{4.39}$$

其中算子 S^N 定义如下:

$$(S^N \mu)(dx) = \frac{1}{N} \sum_{n=1}^{N} \delta_{x^{(n)}}(dx), \quad x^{(n)} \sim \mu \quad \text{i.i.d..}$$

算法 3 序贯重要性重采样粒子滤波器

1: Set $j = 0$ and $\mu_0^N(dx_0) = \mu_0(dx_0)$
2: Draw $x_j^{(n)} \sim \mu_j^N, n = 1, \cdots, N$
3: Set $\omega_j^{(n)} = 1/N, n = 1, \cdots, N$, redefine $\mu_j^N := \sum_{n=1}^N \omega_j^{(n)} \delta_{x_j^{(n)}}$
4: Draw $\widehat{x}_{j+1}^{(n)} \sim p(x_j^{(n)}, \cdot)$
5: Define $\omega_{j+1}^{(n)}$ by (3.23) and $\mu_{j+1}^N = \sum_{n=1}^N \omega_{j+1}^{(n)} \delta_{\widehat{x}_{j+1}^{(n)}}$
6: j+1→ j
7: Go to step 2

根据 (4.39)，我们发现，概率测度的随机性是由采样算子 S^N 引起的，而粒子滤波器的收敛取决于粒子数量。随着粒子数量趋于无穷，粒子滤波器确实能够恢复真实的后验分布 [34]。下面的定理给出了粒子滤波器的收敛结果。

定理 4.1。（ [23]中的定理 4.5）设 m 为粒子数量， μ_j^m为 SIR粒子滤波器中的近似测度。假设 $\kappa \in (0, 1]$ 为引理 3.2 中定义的常数），则 μ_J^m与 μ_J 之间的全变差距离可估计为

$$d(\mu_J^m, \mu_J) \leq \sum_{j=1}^{J} (2\kappa^{-2})^j \frac{1}{\sqrt{m}}. \tag{4.40}$$

在定理 3.4 和定理 4.1 中令 J 固定。我们发现，粒子滤波器的收敛速率取决于粒子数量 m。类似地，PFOF 的收敛由乌拉姆方法中使用的块数 N 决定。当 $N = m$时，即两种方法中的基函数数量相同时，PFOF 的收敛速率为 $\mathcal{O}(\frac{1}{N})$，而 SIR粒子滤波器的收敛速率为 $\mathcal{O}(\frac{1}{\sqrt{m}})$。分析表明，PFOF 比粒子滤波器收敛得更快。

在粒子滤波器中，从高维且复杂的转移核中采样很难实现。PFOF 避免了采样，转而采用数据驱动近似，这种方法需要短期路径模拟，而不需要转移密度的形式。粒子退化也是一个重要问题。随着有效粒子数量逐渐减少，粒子滤波器的效率会越来越低。

众所周知，粒子滤波器会因退化问题而不适用于高维模型。因此，要准确估计后验概率密度函数，需要大量粒子，其数量

会随着系统规模呈指数级增长。除了重采样，添加抖动和局地化也是解决这一问题的有效改进方法。随着划分尺度的扩大，PFOF 在高维情况下也会遭遇“维数灾难”。稀疏乌拉姆方法是规避这一问题的一种解决方案。低秩 Perron-Frobenius 算子滤波器能够提高滤波问题的处理效率。

5 数值结果

本节通过一些数值示例，展示所提出的 PFOF 在滤波问题中的应用。滤波问题中的系统动力学未知，且已给出一些观测。我们采用 PFOF 和 lr-PFOF 对随机滤波问题的后验概率密度函数进行估计。在第 5.1 节中，我们考虑一个 Ornstein-Uhlenbeck（O-U）过程，以识别系统的高斯概率密度函数，并利用已知观测估计其后验概率密度函数。在第 5.2 节中，我们考虑一个由 Beneš 随机微分方程支配的非线性滤波问题，并估计非高斯后验概率密度函数。在第 5.3 节中，我们考虑一个连续时间滤波问题，即带有观测的经典混沌系统 Lorenz’63模型，以对状态的后验密度进行建模。我们将所提出的 PFOF/lr-PFOF 与粒子滤波器和扩展卡尔曼滤波器（ExKF）进行比较。数值结果表明，PFOF 的后验概率密度函数估计优于 PF，且其状态估计比 ExKF 更准确。

5.1 奥恩斯坦-乌伦贝克过程

考虑 一种奥恩斯坦-乌伦贝克过程，它是一个一维线性动力系统 ，

$$dx_t = -\lambda x_t dt + dW_t, \quad x(t_0) \sim \mathcal{N}(m_0, C_0),$$

其中 $\lambda > 0$ 和 W_t 是标准布朗运动。现在我们考虑由奥恩斯坦-乌伦贝克过程的离散化以及对状态的离散观测构成的状态空间模型，如下所示，

$$\begin{cases} x(t_{k+1}) = \exp(-\lambda \Delta t_k) x(t_k) + q_k, \quad q_k \sim \mathcal{N}(0, \Sigma_k), \\ y(t_k) = H x(t_k) + r_k, \quad r_k \sim \mathcal{N}(0, R), \end{cases} \tag{5.41}$$

其中 $\Sigma_k = \exp(-2\lambda\Delta t_k)$、 $H = I$ 和 $R = \sigma^2$。参数由 $\lambda = 1/2$、$m_0 = 2$、 $C_0 = 0.1$ 和 $\sigma = 1$给出。为应用 PFOF，我们使用乌拉姆方法，以时间步长 $\tau = 0.1$ 计算 Perron-Frobenius 算子 P_τ，并得到 $\mathcal{P}_\tau$的近似形式 $P_\tau^N \in \mathbb{R}^{N\times N}$。我们取 x_t的相空间为 $[-6, 6]$，并将其划分为 $N = 100$ 个网格；每个区间 $[z_k, z_{k+1}]$、 $k = 0$、$\cdots$、$N - 1$定义一个盒 $\mathbb{B}_k$。我们在每个 $\mathbb{B}_k$上定义示性函数$\mathbb{1}_{\mathbb{B}_k}(x)$，并在盒中随机选取 $n = 100$ 个样本点来计算 P_τ^N。给定初始高斯分布 $\mathcal{N}(2, 0.1)$，我们将 μ_0 改写为向量 W_0，其表示 μ_0^N 的系数。 P_τ^N 作用于权重向量，以估计 $\mathbb{B}_k$上 x_t的概率值，即 $\mathbb{P}(x_t \in \mathbb{B}_k)$、 $t = q\tau$、 $q = 0$、1、$2\cdots$。由此，我们得到 x_t 在 t处的离散概率密度函数 (PDF)。不同时间的模拟概率密度函数显示在图 5.1左列。从图中可以看出，PFO 估计的概率密度函数接近

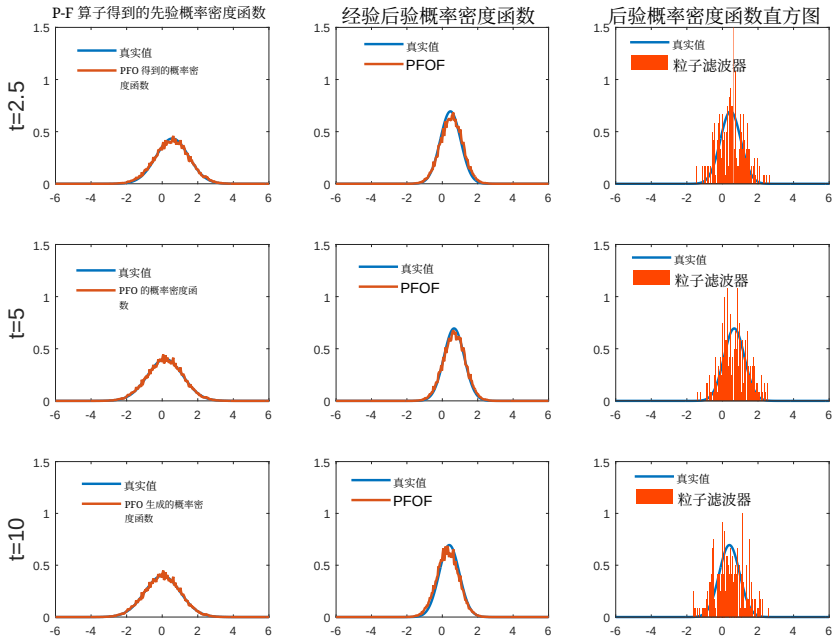

图5.1： 不同时刻由 PFO 估计的先验概率密度函数（左列）、由 PFOF 估计的后验概率密度函数（中列）以及由粒子滤波器估计的后验概率密度函数（右列）。

真实值。通过这种方式，无需求解福克-普朗克方程即可计算概率密度函数，而概率密度函数的估计实际上就是模型 (5.41) 中的先验密度。

接下来，我们计算状态空间模型（5.41）的后验概率密度。我们设定$N = 500$ 和 $n = 100$。使用算法1估计后验概率密度函数，结果显示在图5.1的中列。从图5.1可以看出，PFOF估计的经验后验概率密度函数与高斯后验密度接近。为了与PFOF进行比较，我们也使用粒子滤波器求解滤波问题。在粒子滤波器中，随机抽取500个粒子以生成狄拉克测度并构造经验测度。因此，这两种方法中的基函数数量相同。图5.1清楚地表明，PFOF计算出的经验概率密度函数比PF计算出的更加准确。数值结果支持定理 3.4和定理 4.1。

5.2 Beneš-Daum 滤波器

在本小节中，我们将 PFOF 应用于一个非线性滤波问题，其状态空间模型由 Beneš 随机差分方程定义，

$$dx_t = \tanh(x_t)dt + dW_t, \tag{5.42}$$

初始条件为 $x_0 = 0$。参见 [36], , 方程（5.42）的概率密度函数为

$$p(x(t)) = \frac{1}{\sqrt{2\pi t}} \frac{\cosh(x(t))}{\cosh(x_0)} \exp\Big(-\frac{t}{2}\Big) \exp\Big(-\frac{1}{2t}(x(t) - x_0)\Big).$$

我们取相空间 $[-15, 15]$ 并将其均匀划分为 100 ($N = 100$) 个网格 $[z_k, z_{k+1}]$, k =0, ..., $N-1$, 每个网格对应一个盒 $\mathbb{B}_k$。使用乌拉姆方法近似 PFO。时间步长设为 $\tau = 0.5$，随机样本点的数量为 $m = 400$。图 5.2 展示了 x_t 在 $t = 1$、$t = 2.5$ 和 $t = 5$ 时的预测概率密度函数。概率密度函数分别通过离散化的 PFO 矩阵 $P^N \in \mathbb{R}^{100\times100}$ 和截断为 $r = 30$的 PFO 低秩近似来估计。可以看到，$t = 2.5$ 和 $t = 5$ 时的概率密度函数有两个众数，且 PFO 能较好地近似这两个众数。

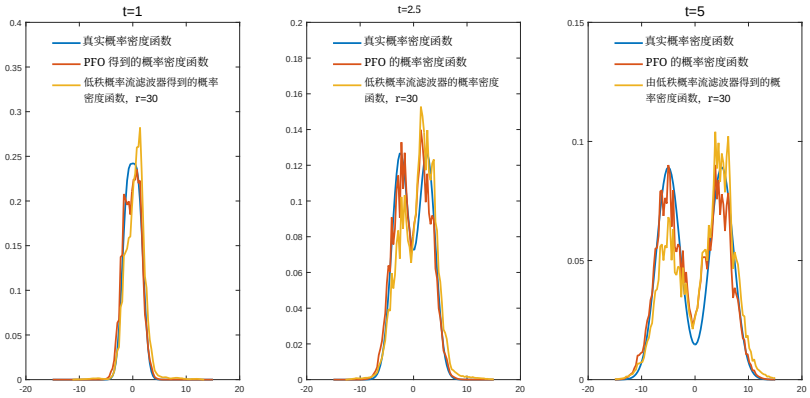

图 5.2：不同时间由 PFO 和低秩模型估计的概率密度函数。

首先，我们要计算模型 (5.42) 在给定观测条件下的真实后验滤波分布。在此示例中，观测模型满足

$$p(y_k|x(t_k)) = \mathcal{N}(y_k|x(t_k), \sigma^2). \tag{5.43}$$

根据 [36]（第 10.5 章），Beneš SDE 的转移密度为

$$p(x(t_k)|x(t_{k-1})) = \frac{1}{\sqrt{2\pi\Delta t_{k-1}}} \frac{\cosh(x(t_k))}{\cosh(x(t_{k-1}))} \exp(-\frac{1}{2}\Delta t_{k-1}) \times \exp\left(-\frac{1}{2\Delta t_{k-1}}(x(t_k) - x(t_{k-1}))^2\right),$$

其中 $\Delta t_{k-1} = t_k - t_{k-1}$。如果假设时刻 t_{k-1} 的滤波解具有以下形式

$$p(x(t_{k-1})|y_{1:k-1}) \propto \cosh(x(t_{k-1})) \exp\left(-\frac{1}{2P_{k-1}}(x(t_{k-1}) - m_{k-1})^2\right)$$

对于给定的 m_{k-1} 和 P_{k-1}。然后我们使用查普曼-科尔莫戈罗夫方程，并给出先验密度

$$p\big(x(t_k)|y_{1:k-1}\big) \propto \cosh\big(x(t_k)\big) \exp\left(-\frac{1}{2P_k^-}(x(t_k) - m_k^-)^2\right),$$

其中

$$m_k^- = m_{k-1},$$

$$P_k^- = P_{k-1} + \Delta t_{k-1}.$$

m_k^- 和 P_k^- 是表示先验密度函数的充分统计量。根据贝叶斯公式， $x(t_k)$的后验密度为

$$p\big(x(t_k)|y_{1:k}\big) \propto \cosh\big(x(t_k)\big)\exp\bigg(-\frac{1}{2P_k}\big(x(t_k)-m_k\big)^2\bigg), \tag{5.44}$$

后验密度中参数 m_k 和 P_k 的方程满足

$$m_k = m_k^- + \bigg(\frac{P_k^-}{P_k^- + \sigma^2}\bigg)(y_k - m_k^-),$$

$$P_k^- = P_{k-1} + \Delta t_{k-1}.$$

因此，参考后验分布由式（5.44）定义。

为了将 PFOF 应用于非线性滤波问题，我们对相位区间 $[-15, 15]$ 进行更细的划分，得到 400 个盒子。此外，我们在乌拉姆方法中选取足够多的样本点，以尽可能减小蒙特卡洛误差。观测 y_k 通过模拟底层模型（5.42）并按照（5.43）添加噪声而人工生成，其中 $\sigma = 1$。可观测区间为 $[0, 5]$ ，时间步长为 $\Delta t_k = 0.1$。滤波过程的初始分布选为 $m_0 = 0, P_0 = 2$。特别地，我们还使用粒子滤波器进行比较。在预测过程中，由于转移概率密度函数较为复杂，我们无法直接抽取样本点。因此，我们采用接受-拒绝法解决这一问题。我们首先在图 5.3 中展示 PFOF 和 lr-PFOF（r=40）估计的后验均值结果，并同时给出真实值和观测值。后验均值通过对 PFOF/lr-PFOF 的后验分布取平均得到；如图所示，该均值接近真实值。

PFOF、lr-PFOF 和粒子滤波器估计得到的后验密度，以及真实值，均显示在图 5.4 中。lr-PFOF 的截断参数分别设为 $r = 10$、 $r = 20$ 和 $r = 40$。随着截断基函数数量增加，lr-PFOF 的估计精度逐渐提高；当 $r = 40 < N = 400$时，其精度几乎与 PFOF 相同。尽管 PFOF 和粒子滤波器使用的基函数数量相同，但这两种方法之间存在明显差异。结果表明，在非高斯非线性滤波问题中，PFOF 的精度高于粒子滤波器。这进一步证实了第 3 节中的理论分析。如表 1 所示，PFOF 和 lr-PFOF 所用的 CPU 时间都少于 SIR粒子滤波器。实际上，粒子滤波器的 CPU 时间主要耗费在接受-拒绝采样上。从表中可以看出，与 PFOF 相比，lr-PFOF 能够减少在线计算时间。

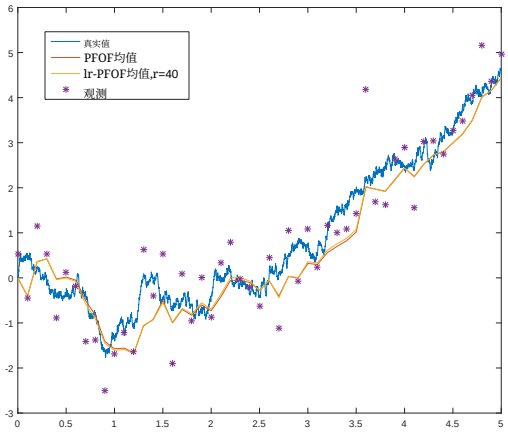

图 5.3：PFOF 和 lr-PFOF 估计的均值。

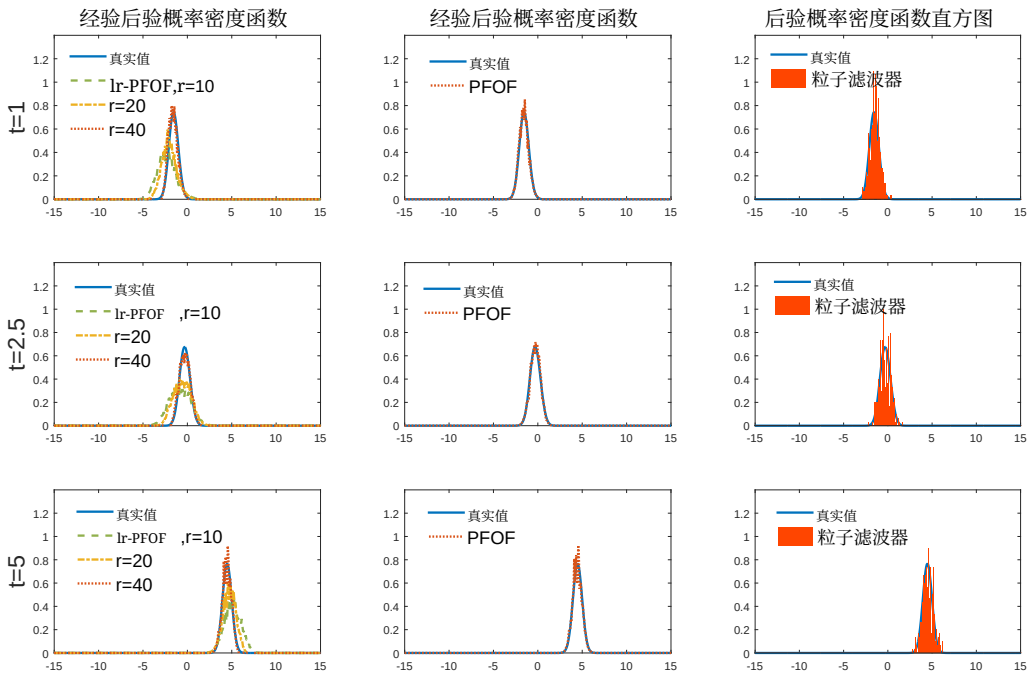

图 5.4：不同时间下，lr-PFOF（左列）、PFOF（中列）和粒子滤波器（右列）得到的后验概率密度函数。

表 1：不同方法计算后验概率密度函数所需的 CPU 时间（秒）

方法	PFOF	lr-PFOF（r=10）	lr-PFOF（r=20）	lr-PFOF（r=40）	粒子滤波器
离线	0.1599	0.2536	0.2649	0.2689	6906.9486
在线	0.0673	0.0150	0.0431	0.0641	

5.3 Lorenz'63模型

Lorenz 于 1963 年提出了一个用于描述大气对流的数学模型。Lorenz'63模型是最简单的、能够表现出对初始条件的敏感性和混沌的连续时间系统，也是数据同化中常用的示例。对于某些参数和初始条件，该系统可能呈现混沌行为。该模型由三个耦合非线性常微分方程组成，其解为 $v = (v_1, v_2, v_3) \in \mathbb{R}^3$。我们考虑带有加性白噪声的 Lorenz'63模型，

$$\begin{cases} \dfrac{dv_1}{dt} = a(v_2 - v_1) + \sigma_1 \dfrac{dW_1}{dt} \\ \dfrac{dv_2}{dt} = -av_1 - v_2 - v_1 v_3 + \sigma_2 \dfrac{dW_2}{dt} \\ \dfrac{dv_3}{dt} = v_1 v_2 - bv_3 - b(r + a) + \sigma_3 \dfrac{dW_3}{dt} \\ v(0) \sim \mathcal{N}(m_0, C_0), \end{cases}$$

其中 W_j 为假设相互独立的布朗运动。我们采用经典参数值 $(a, b, r) = (10, 8/3, 28)$，并设定 $\sigma_1 = \sigma_2 = \sigma_3 = 2$。初始均值 m_0 为 (0, 0, 0)，协方差矩阵为单位矩阵 $I_3 \in \mathbb{R}^{3\times3}$。我们给出连续观测 $z(t)$，其由一个随机微分方程控制

$$\begin{cases} \dfrac{dz}{dt} = h(v) + \gamma \dfrac{dW_z}{dt} \\ z(0) = 0, \end{cases}$$

其中 $\gamma = 0.2$。本例旨在探究 PFOF 在连续时间滤波问题中的表现。我们比较了基于 Perron-Frobenius 算子和连续时间扩展卡尔曼滤波器的同化结果。两种方法估计的后验均值如图5.5和图5.7所示。这两幅图分别对应不同的观测 $h(v) = Hv$，前者由 $H = [0, 1, 0]$ 确定，后者由 $H = [0, 0, 1]$确定。特别是，我们发现 Lorenz 模型中观测的选择影响很大，尤其是对扩展卡尔曼滤波器而言。扩展卡尔曼滤波器的稳定性很大程度上取决于观测。观测不足可能使滤波器偏离真实值并导致显著的模型误差，而且一旦出现偏差，就很容易导致数值不稳定性。然而，PFOF 的重建结果受观测模型的影响小得多，因此该方法比扩展卡尔曼滤波器表现出更强的鲁棒性。

图5.6展示了分别将 v_2 或 v_3 作为不同观测时均方误差的结果。对于扩展卡尔曼滤波，我们发现，当观测第三个分量 v_3 时，其均值估计误差较大。为了更直观地展示结果，我们在图5.8中将PFOF和扩展卡尔曼滤波得到的均值轨迹与真实值进行了比较。我们发现，PFOF的均值轨迹比扩展卡尔曼滤波的均值轨迹更接近真实值。

当 v_3 作为观测时，图 5.9展示了一维和二维边缘概率分布。该图旨在直观地描述单个值的分布以及不同分量之间的相关性。如图所示，

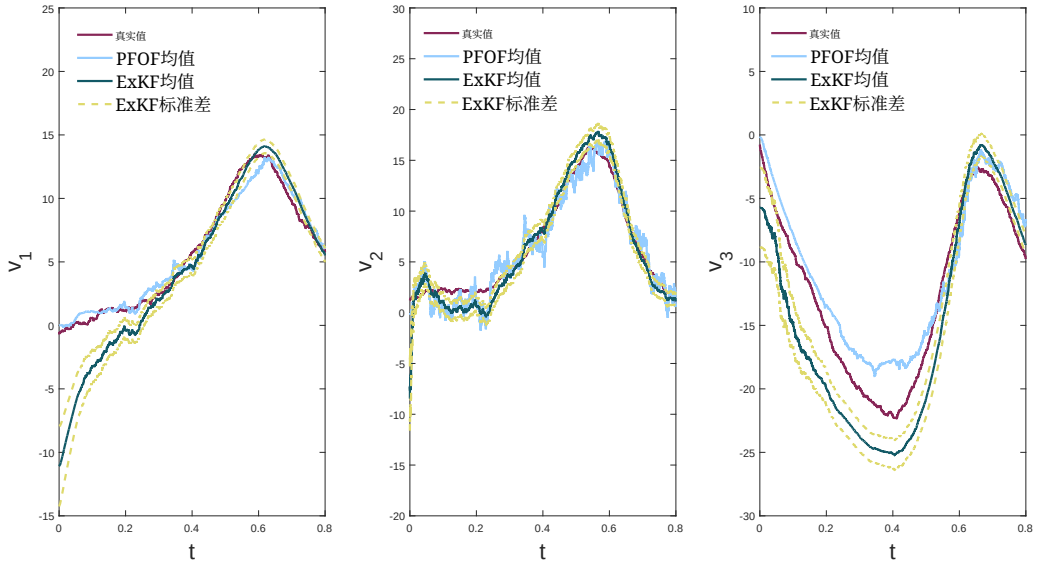

图5.5： 连续观测条件下，Lorenz'63模型中扩展卡尔曼滤波和PFOF对各分量的后验均值。观测分量为 v_2 。

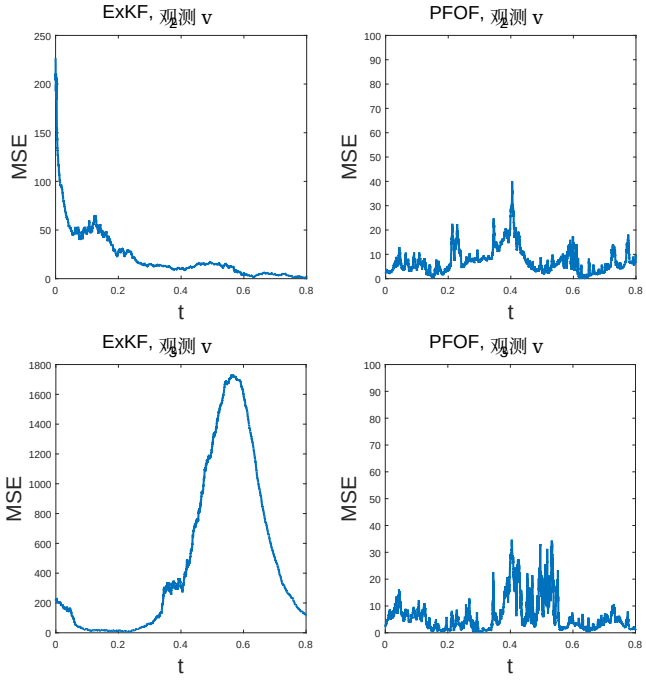

图5.6： 滤波器的均方误差 $\|v(t) - m(t)\|_2^2$ 。

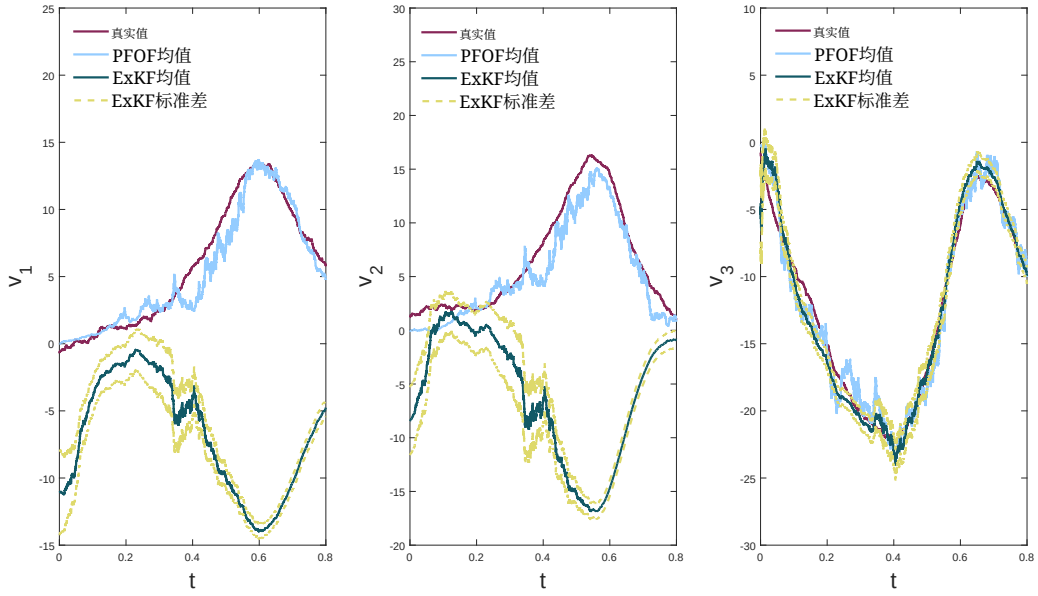

图5.7： 连续观测下，Lorenz'63模型中 ExKF 和 PFOF 对各分量的后验均值。观测分量为 v_3 。

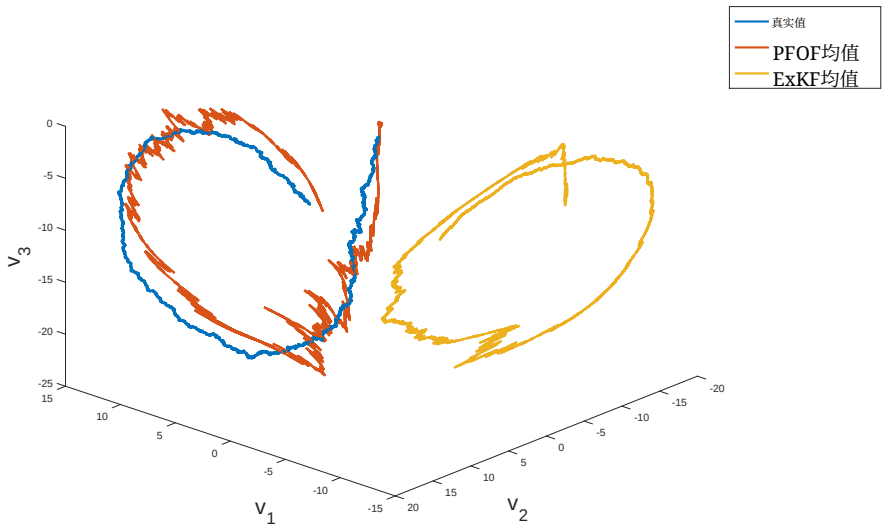

图5.8： PFOF 和 ExKF 的均值轨迹。

观测分量的一维边缘分布比另外两个分量更接近高斯分布。这一现象表明，当某个分量用作观测时，其均值估计会比其他未观测分量更准确。

上述结果表明，在这个混沌非线性系统中，PFOF 的状态估计精度高于扩展卡尔曼滤波。此外，前者还能估计概率密度函数，从而以概率意义获取更多状态信息。

6 结论

本文提出了一种新的滤波方法，在 Perron-Frobenius 算子框架下估计状态的滤波分布。我们针对离散和连续随机动力系统建立了滤波问题，并将 Perron-Frobenius 算子应用于后验概率密度函数的传播。通过乌拉姆方法实现了 PFO 的有限维近似，该方法构造了一个由示性函数张成的伽辽金投影空间。利用乌拉姆方法，对后验概率密度函数进行离散化，并将其表示为基函数的权重。于是，概率密度函数的演化转化为权重向量的迭代过程，该过程由 PFO 和似然函数共同作用。我们将这一过程称为佩龙-弗罗贝尼乌斯算子滤波器。因此，经验概率密度函数由示性函数的凸组合确定。我们给出了所提方法的误差估计，并证明其精度高于粒子滤波器。此外，我们提出了一种低秩 Perron-Frobenius 算子滤波器，通过谱分解近似密度函数。该分解通过对离散化的 PFO 进行特征值分解实现。最后，我们将所提方法应用于三个随机滤波问题，包括线性离散系统、非线性离散系统和非线性连续混沌系统。数值结果表明，与粒子滤波器和扩展卡尔曼滤波相比，所提方法具有更高的精度和更好的鲁棒性。

致谢：L. Jiang 感谢国家自然科学基金项目 12271408 和中央高校基本科研业务费的支持。

参考文献

[1] G. Froyland, R. Stuart and E. Sebille, 全球海洋表面的连通性如何？, 《混沌：非线性科学跨学科期刊》, 24 (2014), 033126。[2] C. Schütte and M. Sarich, 《分子动力学中的亚稳态与马尔可夫状态模型》, 美国数学学会, 2013。[3] A. Tantet, F. Burgt and H. Dijkstra, 利用转移算子预警大气阻塞事件的早期指标, 混沌, 25 (2015), 036406。

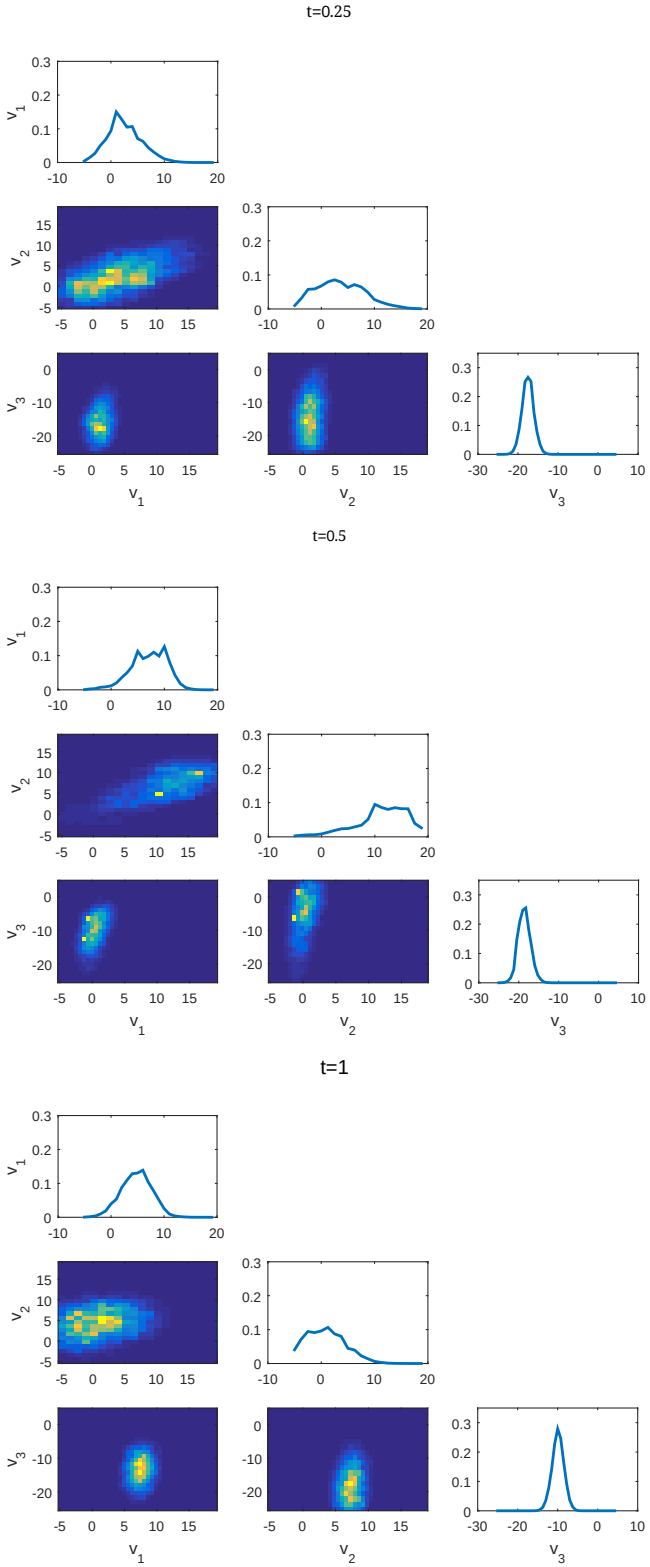

图 5.9: v 的一维和二维后验边缘概率密度函数。

[4] M. Dellnitz, G. Froyland, and O. Junge，GAIO 导向的动力系统数值方法背后的算法，载于《遍历理论、分析与动力系统的高效模拟》，施普林格，2001 年，第 145-174 页。[5] K. Krzyżewski and W. Szlenk，关于膨胀可微映射的不变测度，《数学研究》，33（1969），第 83-92 页。[6] S. Klus, P. Koltai and C. Schütte，关于 Perron-Frobenius 算子和 Koopman 算子的数值逼近，《计算动力学杂志》，3（2016），第 51-79 页。[7] S. Ulam，《数学问题集》，Interscience 出版社，1960 年。[8] C. Bose and R. Murray，乌拉姆方法的精确逼近速率，《离散与连续动力系统》，7（2001），第 219-235 页。[9] J. Ding and A. Zhou，弗罗贝尼乌斯-佩龙算子的有限逼近：多维变换中乌拉姆猜想的一个解，《物理学 D：非线性现象》，92（1996），第 61-68 页。[10] A. Jazwinski，《随机过程与滤波理论》，Dover 出版社，2007 年。
[11] P. Maybeck，《随机模型、估计与控制》，学术出版社，1979 年。
[12] G. Evensen，《数据同化：集合卡尔曼滤波器》，施普林格，2006 年。
[13] D. Oliver, A. Reynolds and N. Liu，《石油储层表征与历史拟合的反演理论》，剑桥大学出版社，2008 年。[14] E. Kalnay，《大气建模、数据同化与可预测性》，剑桥大学出版社，2003 年。[15] R. Kalman，线性滤波与预测问题的新方法，《基础工程杂志》，82（1960），第 35-45 页。[16] Y. Ba and L. Jiang，一种用于数据同化的两阶段变量分离卡尔曼滤波器，《计算物理学杂志》，434（2021），110244。[17] Y. Ba, L. Jiang and N. Ou，一种基于多尺度模型降阶的两阶段集合卡尔曼滤波器，用于时间分数阶扩散波方程的反问题，《计算物理学杂志》，374（2018），第 300-330 页。
[18] L. Jiang and N. Liu，用卡尔曼滤波器校正含噪动态模态分解，《计算物理学杂志》，461（2022），111175。[19] A. Lorenc，数值天气预报的分析方法，《英国皇家气象学会季刊》，112（2000），第 1177-1194 页。[20] D. Kelly, K. Law and A. Stuart，离散时间和连续时间集合卡尔曼滤波器的适定性与精度，《非线性》，27（2014），第 25-79 页。

[21] A. Doucet and A. Johansen，粒子滤波与平滑教程：15 年之后，《牛津非线性滤波手册》，牛津大学出版社，纽约，2011 年，第 656-704 页。[22] C. Snyder, T. Bengtsson, P. Bickel and J. Anderson，高维粒子滤波面临的障碍，《月刊天气评论》，136（2008），第 4629-4640 页。[23] A. Stuart and K. Zygalakis，《数据同化：数学导论》，施普林格，2015 年。[24] P. Koltai，动力系统全局长期行为的高效逼近方法：理论、算法与示例，柏林 Logos 出版社，2011 年。
[25] D. Goswami, E. Thackray and D. Paley，受约束的乌拉姆动态模态分解：确定性与随机系统中 Perron-Frobenius 算子的逼近，IEEE《控制系统快报》，2（2018），第 809-814 页。[26] R. Schilling，《测度、积分与鞅》，剑桥大学出版社，2017 年。
[27] O. Junge and P. Koltai，利用稀疏哈尔张量基离散化弗罗贝尼乌斯–佩龙算子：稀疏乌拉姆方法，《SIAM 数值分析杂志》，47（2009），第 3464-3485 页。
[28] R. Murray，《不变密度的离散逼近》，博士论文，剑桥大学，1997 年。
[29] H. Niederreiter and J. Spanier，《蒙特卡罗方法与拟蒙特卡罗方法》，施普林格，1999 年。[30] E. Weinan, L. Tiejun, E. Vanden-eijnden，《应用随机分析》，美国数学会，2019 年。[31] P. Florchinger, F. Gland，对观测于相关噪声中的扩散过程的 Zakai方程进行时间离散化，《随机学：概率与随机过程国际期刊》，35（1991），第 233-256 页。[32] J. Carpenter, P. Clifford and P. Fearnhead，非线性问题的改进粒子滤波器，《IEE 雷达、声纳与导航学报》，146（1999），第 2-7 页。
[33] M. Bolic, P. Djuric and S. Hong，分布式粒子滤波器的重采样算法与架构，IEEE《信号处理汇刊》，53（2005），第 2442-2450 页。
[34] D. Crisan and A. Doucet，面向实践者的粒子滤波方法收敛结果综述，IEEE《信号处理汇刊》，50（2002），第 736-746 页。[35] A. Bain and D. Crisan，《随机滤波基础》，施普林格，2009 年。[36] S. Särkkä and A. Solin，《应用随机微分方程》，剑桥大学出版社，2019 年。