arXiv:2304.05392v1 [math.OC] 7 Apr 2023

用于随机反应扩散系统状态估计的分块粒子滤波器

若泽 奥古斯托 F. 马加良艾斯 * 奥塔西利奥B. L. 内图 * Francesco Corona *

*化学工程学院，阿尔托大学，芬兰。（电子邮件：{augusto.magalhaes,otacilio.neto, francesco.corona}@aalto.fi）

摘要： 本文考虑环境波动下反应扩散系统状态演化的微分描述。我们关注的是在仅有部分观测可用时，如何估计系统状态。为描述观测与状态之间的关系，我们将乘性噪声驱动动力学与观测模型相结合。具体而言，我们确保观测受到加性噪声形式的误差影响。我们重点研究别洛乌索夫–扎博廷斯基化学反应的状态估计。我们模拟在准二维物理域中进行的反应，例如培养皿表面上的反应。我们的目标是根据带噪声的光谱观测重建形成的化学图案。为此，我们考虑空间域上的有限差分表示，其中节点根据观测点选取。我们采用分块粒子滤波器（一种序贯蒙特卡罗方法）近似求解这一状态估计问题，该方法能够应对这种状态空间表示所伴随的高维性。

1. 引言

自“图灵斑图”概念提出以来，反应扩散模型在理论生物学中发挥了重要作用（图灵，1952年）。这类表示已被用于解释若干自然产生的图案形成现象：胚胎中的发育结构（Kondo 和Miura，2010年；Green和 Sharpe，2015年）皮肤和色素沉着图案（Kondo，2009年；Nakamasu 等，2009年），以及肿瘤细胞生长（Ferreira 等，2002年；Hogea等，2008年）。此外，反应扩散模型也被用于描述野火蔓延动力学（Asensio和Ferragut，2002年），以及合成计算机图形学中的纹理（Witkin和 Kass，1991年）。而且，如果能够实时确定反应扩散系统的状态，就可以利用反馈理论出于多种目的控制图案形成过程（Ardizzone等，2013年）。在实际生物化学应用中采用这些模型一直颇具挑战：这类模型必须通过间接方法（例如光谱分析）依赖于对真实系统的部分观测，而且其典型实现是高维随机系统。因此，状态估计任务必然会带来沉重的计算负担，即使该任务确实可计算。

粒子滤波器（PF）广泛用于一般非线性随机动力系统的状态估计（Chopin 和 Papaspiliopoulos，2020年）。每当获得新的观测值，由带有重要性权重的一组粒子便会进行一次抽样过程（Kitagawa 和 Gersch，1996年）。由于这种粒子表示具有递归性质，粒子集合不可避免地会出现退化问题（Musso 等，2001年）。缓解这一现象的一些技术包括增加粒子间的多样性（Gilks 和 Berzuini，2001年），或对其路径中固定滞后区间内的一段进行重采样（Doucet 等，2006年）。然而，即使采用了这些技术，经典粒子滤波器应用于高维状态空间模型时，仍会在仅经过几次迭代后出现退化（Van Leeuwen，2003年）。另一种方法是利用许多高维系统的相关性衰减特性，将状态空间划分为低维子区域，并据此进行全状态估计。这类方法称为局部粒子滤波器，包括多重粒子滤波器（MPF，Djuric 等，2007年）以及较新的分块粒子滤波器（BPF，Rebeschini 和 van Handel，2015年）。

本文旨在利用含噪光谱观测，估计一般反应—扩散系统中化学物质的浓度。我们聚焦于由时空动力学的有限差分近似所产生的状态空间表示。针对这一高维问题，我们采用基于空间域不相交划分的块粒子滤波器（BPF）。我们以重构 Belousov-Zhabotinskii 型反应—扩散系统中涌现的化学图案为任务，展示了该估计器的性能，该系统即奥列戈纳托模型（Field and Noyes，1974年）。根据我们的结果，对于相对较小的集合集合规模，朴素 块粒子滤波器无法重构状态。然而，我们表明，在特定的抽样过程选择下，块粒子滤波器能够提供准确的状态估计。

本文结构安排如下：第 2 节概述反应-扩散系统及其有限差分表示方法。第 3 节讨论适用于一般和高维状态空间的粒子滤波器。第 4 节展示将状态估计器应用于基准反应扩散模型的仿真结果。最后，第 5 节总结本文并提出若干未来研究方向。

*“©2023 年作者。本文已获 IFAC 接收，将依据知识共享许可协议 CC-BY-NC-ND 发表”

2. 反应-扩散系统

我们考虑分布在某个有界空间 $U \subseteq \mathbb{R}^3$ 中的一组化学物种 $\mathcal{S} = \{\mathcal{S}_1, \ldots, \mathcal{S}_{N_\mathcal{S}}\}$，它们按照如下形式的化学反应网络相互作用

$$\sum_{n_s=1}^{N_\mathcal{S}} \underline{S}_{n_r,n_s} \mathcal{S}_{n_s} \xrightarrow{\kappa_{n_r}} \sum_{n_s=1}^{N_\mathcal{S}} \overline{S}_{n_r,n_s} \mathcal{S}_{n_s}, \quad n_r = 1, \ldots, N_R, \tag{1}$$

速率常数为 $\kappa \in \mathbb{R}_{\geq 0}^{N_R}$，反应物和生成物的化学计量系数矩阵分别为 $\underline{S} \in \mathbb{N}^{N_R \times N_\mathcal{S}}$ 和 $\overline{S} \in \mathbb{N}^{N_R \times N_\mathcal{S}}$。从系统分析的角度来看，该过程的动力学可用浓度 $z(u,\ t) = ([\mathcal{S}_1](u,\ t)\ \ldots\ldots,\ [\mathcal{S}_{N_\mathcal{S}}](u,\ t)) \in \mathbb{R}_{\geq 0}^{N_\mathcal{S}}$ 在各空间坐标 $u \in U$ 和时刻 $t \in \mathbb{R}_{\geq 0}$ 处描述。随后，可将浓度的演化描述为

$$\partial^t z(u,t) = \left[(\overline{S} - \underline{S})^\mathsf{T} \nu\big(z(u,t)\big) + D_z \nabla^2 z(u,t)\right] \partial t, \tag{2}$$

给定反应速率 $\nu(\cdot) = (\nu_1(\cdot)\ \ldots\ldots,\ \nu_{N_R}(\cdot)) \in \mathbb{R}^{N_R}$、扩散系数矩阵 $D_z \in \mathbb{R}^{N_\mathcal{S} \times N_\mathcal{S}}$ 以及拉普拉斯算子 $\nabla^2 z(u,\ t) = \sum_{i=1}^3 \partial z^2 / \partial u_i^2$。对于网络方程（1）中的第 n_r 个反应，函数 $\nu_{n_r}(n_r = 1,\ \ldots\ldots,\ N_R)$ 遵循质量作用定律（Murray，2002年），

$$\nu_{n_r}\big(z(u,t)\big) = \kappa_{n_r} \prod_{n_s=1}^{N_\mathcal{S}} z_{n_s}(u,t)^{\underline{S}_{n_r,n_s}}. \tag{3}$$

最后，我们考虑所有边界点 $u \in \partial U$ 上的齐次诺伊曼边界条件 $(\partial z / \partial u)\,|_{u=u} = 0$。

假设每种物质的浓度只能通过其混合物间接观测。具体而言，我们假设光谱测量过程会发射

$$y(u,t|\lambda) = \sum_{n_s=1}^{N_\mathcal{S}} \phi_{\mathcal{S}_{n_s}}(\lambda) z_{n_s}(u,t), \tag{4}$$

给定波长 $\lambda \in \mathbb{R}_{\geq 0}$ 和函数 $\phi_{\mathcal{S}_{n_s}}:\ \mathbb{R}_{\geq 0} \to \mathbb{R}$，用于描述与第 n_s 种物质相关的特征响应光谱（$n_s = 1,\ \ldots\ldots,\ N_\mathcal{S}$）。这对应于光吸光度与物质浓度呈线性关系，如比尔定律所示（Nadler 和 Coifman，2005年）。下文中，对于特定波长 $\Lambda = (\lambda_1,\ \ldots\ldots,\ \lambda_{N_\Lambda})$，我们考虑响应 $\Phi_{\mathcal{S}_{n_s}}(\Lambda) = (\phi_{\mathcal{S}_{n_s}}(\lambda_1)\ldots\ldots,\ \phi_{\mathcal{S}_{n_s}}(\lambda_{N_\Lambda}))$，因此测量过程可表示为 $y(u,\ t) = H_z z(u,\ t)$，其中 $H_z = [\Phi_{\mathcal{S}_1} \cdots \Phi_{\mathcal{S}_{N_\mathcal{S}}}]$。

带有测量过程 Eq. (4) 的反应扩散系统 Eq. (1) 可表示为状态空间模型

$$\partial^t z(u,t) = f_{\theta_x}^z(z(u,t), \nabla^2 z(u,t)) \partial t; \tag{5a}$$

$$y(u,t) = H_{\theta_y}^z z(u,t), \tag{5b}$$

状态为 $z:\ U \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{N_\mathcal{S}}$，测量值为 $y:\ U \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{N_\Lambda}$。函数 $f^z(\cdot)$ 和矩阵 $H^z \in \mathbb{R}^{N_\Lambda \times N_\mathcal{S}}$ 分别由从公式 (1)–(4) 中收集的固定参数 θ_x 和 θ_y 确定。

一般而言，要获得状态方程 Eq. (5a) 的精确解并不现实。为对其进行数值积分，我们将空间 U 离散为一个晶格图，以表示间距均匀的坐标。随后，可通过有限差分法（Strauss，2007年）将状态动力学近似为一组常微分方程。下文将概述这一方法。

2.1 有限差分近似

我们将仅限于准二维空间 $U = \{\underline{u} \in \mathbb{R}^2 :\ 0 \leq u_{i=1,\ 2} \leq \overline{U}\}$ 其尺寸有限 $U > 0$：这些空间对应于正方形表面上的化学混合物（例如培养皿）。给定某个长度 $\Delta u > 0$，我们将其离散化 U 到正方晶格 $V = \{v \in \mathbb{N}^2 :\ 1 \leq v_{i=1,\ 2} \leq \overline{V}\}$ 该晶格由间距均匀的网格点组成，大小为 $V = (U/\Delta u)$。随后，将空间离散化的动力学近似为

$$dz^{(v)} = \left[(\overline{S} - \underline{S})^\mathsf{T} \nu\big(z^{(v)}\big) + D_z \widetilde{\nabla}^2 z^{(v)}\right] dt, \quad \forall v \in V, \tag{6}$$

其中 $z^{(v)}(t) = z\big((v_1, v_2)\Delta u, t\big)$，以及拉普拉斯算子 $\nabla^2 z$ 在 $v = (v_1, v_2)$ 处的一阶近似，

$$\widetilde{\nabla}^2 z^{(v)} = \sum_{v' \in \mathcal{N}(v)} \frac{1}{\Delta u^2} \left(z^{(v')} - z^{(v)}\right), \tag{7}$$

其中 $\mathcal{N}(v) = \{v' \in \mathbb{N}^2 :\ |v_1 - v_1'| + |v_2 - v_2'| = 1\}$，$v$ 的1 步邻域。输出方程变为 $y^{(v)}(t) = H_z z^{(v)}(t)$，其中 $y^{(v)}(t) = y((v_1,\ v_2)\Delta u,\ t)$。因此，该测量过程对应于一种典型的成像光谱学，其像素 被分配给 V 上的网格点。

最后，反应扩散模型公式（5）通过传统的状态空间表示进行近似

$$dx(t) = f_{\theta_x}(x(t))dt; \tag{8a}$$

$$y(t) = H_{\theta_y} x(t), \tag{8b}$$

以及状态向量 $x(t) = (z^{(v_1,v_2)}(t))_{v_1,v_2=1}^{V} \in \mathbb{R}^{N_x}$ 和输出向量 $y(t) = (y^{(v_1,v_2)}(t))_{v_1,v_2=1}^{V} \in \mathbb{R}^{N_y}$，其通过在正方晶格的每个点收集 $(z^{(v)}, y^{(v)})$ 得到 V。输出矩阵为 $H_{\theta_y} = [I_{V^2} \otimes \Phi_{\mathcal{S}_1} \cdots I_{V^2} \otimes \Phi_{\mathcal{S}_{N_\mathcal{S}}}]$，其中 $M_1 \otimes M_2$ 表示矩阵 M_1 和 M_2 之间的克罗内克积。该模型因此对应于 $N_x = N_\mathcal{S} V^2$ 状态变量和 $N_y = N_\Lambda V^2$ 输出变量：当 V 较大时，状态空间是高维的（即空间被精细离散化，且 $\Delta u \ll U$）。

3. 预备知识：非线性滤波

我们考虑一个自治系统的随机微分方程（SDE）表示，

$$dX(t) = f_{\theta_x}(X(t))dt + g_{\theta_b}(X(t))dB(t), \tag{9a}$$

$$Y(t) = h_{\theta_y}(X(t)) + e(t), \tag{9b}$$

其中状态方程式 (9a) 描述状态过程 X 的时间演化：$[0, \infty) \times \Omega \to \mathbb{X}$ 给定其当前值、随机驱动过程 $B = (B(t))_{t \geq 0}$ 以及初始状态 $X(0) = \xi(\omega^*)\ \omega^* \in \Omega$。输出方程式 (9b) 描述当前状态如何作为输出向量 $y(t) \in \mathbb{Y}$ 发出，此前受到测量噪声 $e(t) \in \mathbb{R}^{N_y}$ 的干扰。我们考虑 Polish 空间 $\mathbb{X} = \mathbb{R}^{N_x}$ 和 $\mathbb{Y} = \mathbb{R}^{N_y}$。参数向量 $\theta = (\theta_x, \theta_b, \theta_y)$ 决定函数 $f_{\theta_x}(\cdot)$、$g_{\theta_b}(\cdot)$ 和 $h_{\theta_y}(\cdot)$。为便于说明，以下省略对 θ 的依赖。对于所有 $0 \leq t' < t$，$B(t) - B(t') \sim \mathcal{N}(0, (t - t')I_{N_x})$ 以及 $e(t) \sim \mathcal{N}(0, \Sigma_y)$，其协方差 Σ_y 已知。最后，我们将范围限制在线性方程 $Y(t) = HX(t) + e(t)$ 上，其中给定矩阵 $H \in \mathbb{R}^{N_y \times N_x}$。

我们感兴趣的是根据一系列离散时间测量值 $(Y_1, \ldots, Y_k)$ 估计状态 $(X(t))_{t \geq 0}$，其中 $Y_k = Y(k\Delta t)$ 对应的采样时间为 $\Delta t > 0$。假设

在驱动噪声测量之间使用零阶保持器（即，对于所有 $t \in [t_{k-1}, t_k)$，$g(X(t)) = g(X(t_{k-1}))$），离散时间状态动力学表示为

$$X_k = \underbrace{X_{k-1} + \int_{t_{k-1}}^{t_k} f(X(t))dt}_{F(X_{k-1})} + g(X_{k-1})\Delta B_k, \qquad (10)$$

其中，$X_k = X(k\Delta t)$、$\Delta B_k = \mathcal{N}(0,\ \Delta t \cdot I_{N_x})$以及 $t_k = k\Delta t$。离散时间输出方程为 $Y_k = h(X_k) + e_k$，其中测量噪声为 $e_k \sim \mathcal{N}(0,\ \Sigma_y)$。因此，估计问题可通过随机过程$\pi = (\pi_k)_{k\in\mathbb{N}}$形式化，其中 $\pi_k = \mathbb{P}(X_k \in \cdot | Y_1, \ldots\ldots, Y_k)$是滤波分布，用于概括基于历史观测（$Y_1, \ldots\ldots, Y_k$）对 X_k的不确定性。假设分布 π_k相对于勒贝格测度具有密度 p_k。

我们考虑粒子滤波器（PF，Chopin 和 Papaspiliopoulos，2020年）方法，通过递归采样N_p 个基于观测 Y_k 和先前$\pi_{k-1}^{N_p}$的 X_k 候选样本，计算蒙特卡洛近似 $\pi^{N_p} = (\pi_k^{N_p})_{k\in\mathbb{N}}$。针对高维空间，我们将该方法专门应用于一类局部化估计器，即块粒子滤波器（BPF，Rebeschini 和 van Handel，2015年）。下文将概述这些方法。

3.1 经典粒子滤波器

我们假设 X 是一个马尔可夫链：对于所有 $k \in \mathbb{N}$ 和任意博雷尔集 $A \in \mathcal{B}(\mathbb{R}^{N_x})$，

$$\mathbb{P}(X_k \in A | X_{0:k-1}) = \mathbb{P}(X_k \in A | X_{k-1}). \qquad (11)$$

我们考虑转移核 P_k：$\mathbb{R}^{N_x} \times \mathcal{B}(\mathbb{R}^{N_x}) \to \mathbb{R}_{\geq 0}$，使得对于所有 $k \in \mathbb{N}$、$x \in \mathbb{R}^{N_x}$ 以及任意 $A \in \mathcal{B}(\mathbb{R}^{N_x})$，

$$\mathrm{P}_k(x, \mathrm{A}) = \mathbb{P}(X_k \in \mathrm{A} | X_{k-1} = x). \qquad (12)$$

因此，过滤器 π_k 可以递归计算为

$$\pi_{k-1} \xrightarrow{\text{prediction}} \tilde{\pi}_k = \mathrm{P}_k \pi_{k-1} \xrightarrow{\text{correction}} \pi_k = \mathrm{C}_k \tilde{\pi}_k,$$

其中函数 P_k 和 C_k 满足

$$(\mathrm{P}_k \pi_{k-1})(A) := \int \mathrm{P}_k(x, A)\pi_{k-1}(dx), \qquad (13)$$

$$(\mathrm{C}_k \tilde{\pi}_k)(A) := \frac{\int_A \mathrm{C}_k(x)\tilde{\pi}_k(dx)}{\tilde{\pi}_k(\mathrm{C}_k)}, \qquad (14)$$

对于非负函数C_k，满足 $\tilde{\pi}_k(\mathsf{C}_k) > 0$。此处，$\tilde{\pi}_k(\mathsf{C}) := \int \mathsf{C}(x)\tilde{\pi}_k(dx)$。一般情况下，递推关系$\pi_k = \mathsf{C}_k\mathsf{P}_k\pi_{k-1}$ 没有平凡解。因此，估计问题可以通过序贯蒙特卡洛方法求解：滤波器 π_k 在给定采样算子 $\mathsf{S}_k^{N_p}\pi := N_p^{-1}\sum_{n=1}^{N_p}\delta_{x^n}$ 以及样本 $x^1, \ldots, x^{N_p} \overset{i.i.d}{\sim} \pi$的情况下，由经验分布 $\mathsf{C}_k\mathsf{S}_k^{N_p}\tilde{\pi}_k$ 近似。

令 $\{X_k^{(n)}\}_{n=1}^{N_p}$ 为 N_p 相互独立的随机过程，且都与 Y_k 独立，并满足式 (9a)。各对 $(X_k^{(n)}, Y_k)$与 (X_k, Y_k) 同分布，其中 $n = 1, \ldots, N_p$。现在，令$\pi^{N_p} = (\pi_k^{N_p})_{k\in\mathbb{N}}$ 为经验分布序列

$$\pi_k^{N_p} \triangleq \frac{1}{N_p}\sum_{n=1}^{N_p} w_k^{(n)}\delta_{X_k^{(n)}}, \qquad (15)$$

其中，我们为每个 $X_k^{(n)}$ 赋予一个（归一化的）权重 $w_k^{(n)}$。于是，我们几乎必然以稍慢一些的速率得到 $\pi^{N_p} \to \pi$。低于 $1/\sqrt{N_p}$ (Bain 和 Crisan，2008年)。由于无法获得滤波器π_k，我们从具有以下形式密度的重要性分布 ρ_k 中采样独立同分布的 $X_k^{(1)}, \ldots, X^{(N_p)}{}_k$

$$q(x_{0:k}|y_{1:k}) = q(x_0)\prod_{k'=1}^{k} q(x_{k'}|y_{0:k'}, x_{0:k'-1}), \qquad (16)$$

其中 x_k 和 y_k 分别表示 X 和 Y 在时间t_k的一个实现。在此设置下，归一化权重$\widetilde{w}^{(n)}(x_{0:k}) = w^{(n)}(x_{0:k-1})\ \widetilde{w}_k^{(n)}$($n = 1, \ldots, N_p$)可以根据以下公式递归计算

$$\widetilde{w}_k^{(n)} = \frac{p(y_k|y_{1:k-1}, x_{0:k}^{(n)})p(x_k|x_{k-1}^{(n)})}{q(x_k|y_{1:k}, x_{0:k-1}^{(n)})} := l_k(x_k^{(n)}, y_k), \quad (17)$$

每当新的观测 Y_k 可用时。它们通过 $w_k = [\sum_{n'=1}^{N_p}\widetilde{w}^{(n')}{}_k\ ]^{-1}\widetilde{w}_k^{(n)}$ 进行归一化。

$q(x_{0:k}|y_{1:k})$ 的选择是任意的：它只需具有一个包含 $p(x_{0:k}|y_{0:k})$ 支持集的支持集。通常，我们从动力学方程 (9a) 的转移分布中采样粒子，即 $q(x_k|y_{1:k},\ x_{0:k-1}) = p(x_k|x_{k-1})$。随后根据似然计算重要性权重，

$$\widetilde{w}_k^{(n)} \propto w_{k-1}^{(n)} p(y_k|x_k^{(n)}). \qquad (18)$$

这种提议密度的选择虽属惯例，但并不太有信息量。实践中，可能只有少数选定的粒子会被赋予有意义的权重。在极端情况下，这种情况称为退化：$w_k^{(n)} \approx 1$ 对单个 n 而言，而 $w^{(n')}{}_k \approx 0$ 对所有 $n' \neq n$ 而言。此时，似然的近似效果较差，且方差 $\mathrm{Var}_\rho(\widetilde{w})$ 较大。

最优采样分布 通过选择一种重要性分布，可解决退化问题；该分布使（未归一化）权重的方差最小化，且计算时不仅以 $x_{0:k}$ 和 $y_{0:k-1}$ 为条件，也以 y_k 为条件。命题 *1*。（Doucet 等，2000）。使 $\mathrm{Var}_\rho(\widetilde{w})$ 最小化的（最优重要性）分布，其密度为

$$q(x_k|y_{1:k}, x_{0:k-1}) = \frac{p(y_k|y_{1:k-1}, x_{0:k})p(x_k|x_{k-1})}{p(y_k|y_{1:k-1}, x_{0:k-1})}, \quad (19)$$

其中，相应的未归一化权重满足

$$\widetilde{w}_k \propto w_{k-1}\underbrace{\int p(y_k|x_{k-1:k})p(x_k|x_{k-1})\,dx_k}_{p(y_k|x_{k-1})}. \qquad (20)$$

对于过程 $X_k = F(X_{k-1}) + g(X_{k-1})\Delta B_k$及测量模型 $Y_k = HX_k + e_k$（给定 $e_k \sim \mathcal{N}(0, \Sigma_y)$），最优重要性分布具有解析形式

$$X_k|Y_k, X_{k-1} \sim \mathcal{N}(M_k^{\mathrm{opt}}, \Sigma_k^{\mathrm{opt}}) \qquad (21)$$

均值为 M_k^{opt} 且协方差为 Σ_k^{opt}，由

$$\Sigma_B = (g(x_{k-1})g(x_{k-1})^{\mathsf{T}})\Delta t, \qquad (22)$$

$$\Sigma_k^{\mathrm{opt}} = (\Sigma_B^{-1} + H^{\mathsf{T}}\Sigma_y^{-1}H)^{-1}, \qquad (23)$$

$$M_k^{\mathrm{opt}} = \Sigma_k^{\mathrm{opt}}(\Sigma_B^{-1}F(x_{k-1})) + H^{\mathsf{T}}\Sigma_y^{-1}y_k. \qquad (24)$$

Moreover, the likelihood distribution satisfies

$$Y_k|X_{k-1} \sim \mathcal{N}(HF(x_{k-1}),\ \Sigma_B + H\Sigma_y H^{\mathsf{T}}). \qquad (25)$$

对于正系统，一种常见的建模假设是，在给定 $\Sigma_x^{1/2} \succ 0$ 的情况下，$g(X_{k-1}) = \Sigma_B^{1/2} = \mathrm{diag}(X_{k-1})\Sigma_x^{1/2}$：这保证了方程 (9a) 存在唯一的正解 X（Yang 等人，2020年）。因此，当 H 具有某种特殊结构（如第 2 节所述）时，可以高效地计算高维系统中方程 (21) 和 (25) 的方差。

3.2 分块粒子滤波

一般而言，无论选择何种重要性分布，经典粒子滤波器应用于高维系统时都会面临退化问题（Snyder 等人，2015年）。一种可扩展的解决方案是设计空间局部化滤波器：假设空间位置处的动力学和观测仅取决于与其邻域相关联的状态变量（Rebeschini 和 van Handel，2015年）。

将每个时间 t_k 下的这对（X_k, Y_k）视为一个随机场（X_k, Y_k）$_{v\in V}$，该随机场由一个（有限）无向图$\mathcal{G} = (V, E)$索引。X_k 和 Y_k 的空间 $\mathbb{X}$ 和 $\mathbb{Y}$ 以及滤波器 π_k 可以表示为乘积形式

$$\mathbb{X} = \prod_{v\in V}\mathbb{X}^v, \quad \mathbb{Y} = \prod_{v\in V}\mathbb{Y}^v, \quad \pi_k = \bigotimes_{v\in V}\pi_k^v. \tag{26}$$

此处，π_k^v 是状态空间 $\mathbb{X}^v$上的条件分布，而 $\bigotimes_{v\in V}$是关于顶点集 V的（测度）乘积。然后，将式 (12) 中的转移核 P_k 和式 (17) 中的似然函数l 分解为其局部化形式 $\mathrm{P}_k(x, A) = \prod_{v\in V}\mathrm{P}_k^v(x, A^{(v)})$ $l_k(x, y) = \prod_{v\in V} l_k^v(x^{(v)}, y^{(v)})$.

我们聚焦于图 $\mathcal{G}$ 表示正方晶格（如第 2.1 节所示）且具有顶点集的情形

$$V = \{v \in \mathbb{N}^2 : 1 \le v_{i=1,2} \le \overline{V}\} \quad (\overline{V} \in \mathbb{N}). \tag{27}$$

考虑互不重叠的块集合

$$\mathcal{K} = \{(v_0 + \{1, \dots, \overline{V}_b\}^2) \cap V : v_0 \in \overline{V}_b\mathbb{N}^2\}, \tag{28}$$

其大小为 $\overline{V}_b < \overline{V}$， $\mathcal{G}$ 的顶点可划分为

$$V = \bigcup_{V_b\in\mathcal{K}} V_b, \quad V_b \cap V_{b'} = \emptyset \text{ for } V_b \ne V_{b'}, \quad V_b, V_{b'} \in \mathcal{K}.$$

在这种分解下，我们将 $X^{(V_b,n)} = \otimes_{v\in V_b} X^{(v,n)}$和 $Y^{(V_b)} = \otimes_{v\in V_b} Y^{(v)}$ 定义为与块 $V_b \in \mathcal{K}$ 相关联的（n 个粒子）状态变量和输出变量。

令分块算子为 $\mathrm{B}\pi_k := \bigotimes_{V_b\in\mathcal{K}} \mathrm{B}^{V_b}\pi_k$，其中 $\mathrm{B}^{V_b}\pi_k$ 表示 π_k 在 $\prod_{v\in V_b}\mathbb{X}^v$ 上的边缘分布。在这种情况下，我们通过以下方式递归计算滤波器 π_k：

$$\pi_{k-1} \xrightarrow{\text{prediction}} \tilde{\pi}_k = \mathrm{P}_k\pi_{k-1} \xrightarrow{\text{correction}} \pi_k = \mathrm{C}_k\mathrm{B}\tilde{\pi}_k.$$

随后，分块粒子滤波算法按照算法 1[†] 实现。在这种方法中，粒子滤波器的收敛性质取决于各个块的基数 $|V_b| = V_b^2$ （$V_b \in \mathcal{K}$），而不是原始晶格的基数 $|V| = V^2$。

—

4. 案例研究：奥列戈纳托系统 TEM

在本节中，我们介绍使用 BPF（第 3.2 节）从基准反应扩散系统中重建浓度所获得的结果。具体而言，我们关注从双稳态、振荡性的 Belousov-Zhabotinskii（BZ）反应系统（Zhabotinsky，1991年）中估计状态的任务。奥列戈纳托模型（Field 和Noyes，1974年）由 BZ 反应动力学的最简单现实模型构成，其网络

[†] 在其标准形式中，算法1假设 P_k^v 和 l_k^v 分别从具有密度 $p(\tilde{x}_k|x_{(n)\ k-1})$ 和 $p(y_k|\tilde{x}_{(n)\ k})$的相应边缘分布中获得。使用最优重要性分布时，边缘分布取自方程（21）和（25）。

算法1： 分块粒子滤波器

使用理想的分布初始化 $\pi_0^{N_p}$； **对于** $k = 1\ 2\ .,.\ ,K$ **执行**

Resample $X_{k-1}^{(1)}, \dots, X_{k-1}^{(N_p)} \overset{i.i.d.}{\sim} \pi_{k-1}^{N_p}$;

Sample $\widetilde{X}_k^{(v,1)}, \dots, \widetilde{X}_k^{(v,N_p)} \overset{i.i.d.}{\sim} \mathrm{P}_k^v(x_{k-1}^{(n)}, \cdot),\ \forall v \in V$;

Compute $w_k^{(V_b,n)} = \dfrac{\prod_{v\in V_b} l^v\left(\tilde{x}_k^{(v,n)}, y_k^{(v)}\right)}{\sum_{n'=1}^{N_p}\prod_{v\in V_b} l^v\left(\tilde{x}_k^{(v,n')}, y_k^{(v)}\right)}$ for every $V_b \in \mathcal{K}$ and $n = 1, \dots, N_p$;

Let $\pi_k^{N_p} = \bigotimes_{V_b\in\mathcal{K}}\sum_{n=1}^{N_p} w_k^{(V_b,n)}\delta_{\tilde{x}_k^{(V_b,n)}}$.

end

$$\begin{aligned} 2\mathcal{S}_1 &\xrightarrow{\kappa_1} \mathcal{S}_4 + \mathcal{S}_5 & \mathcal{S}_1 + \mathcal{S}_3 &\xrightarrow{\kappa_2} 2\mathcal{S}_5 \\ \mathcal{S}_1 + \mathcal{S}_4 &\xrightarrow{\kappa_3} 2\mathcal{S}_1 + 2\mathcal{S}_2 & \mathcal{S}_3 + \mathcal{S}_4 &\xrightarrow{\kappa_4} \mathcal{S}_1 + \mathcal{S}_5 \\ \mathcal{S}_2 + \mathcal{S}_6 &\xrightarrow{\kappa_5} 0.5\sigma\mathcal{S}_3, \end{aligned} \tag{29}$$

浓度为 $z = \{[\mathcal{S}_1], \dots, [\mathcal{S}_6]\}$。通常，物种（$\mathcal{S}_4$， $\mathcal{S}_5$， $\mathcal{S}_6$）的密度较高，因此假设浓度$z_{4,\ 5,\ 6}$（s， t）在几个振荡周期的时间尺度内保持恒定。此外，当 z_3 缓慢变化时，系统的动力学仅由 z_1 和 z_2 的演化概括，其中 f_z 以缩放形式[‡] 表示。

我们考虑定义在准二维空间U 上的系统，其大小为 $U = 2$ 空间单位。我们将 U 离散化为$\Delta u = 0.02$ 个尺寸单位，并将动力学的有限差分近似（第 2.1 节）定义在晶格 V上，其中每个维度包含 $V = U/\Delta u = 100$ 个网格点。随后，使用 $\Delta t = 0.01$对动力学进行时间离散化。假设该过程受到驱动噪声的作用，其系数为 $g(X_{k-1}) = \sigma_x\mathrm{diag}(X_{k-1})$，给定 $\sigma_x = 10^{-2}$。测量过程定义为 $H = [I_{V^2} \otimes \Phi_{\mathcal{S}_1}\ I_{V^2} \otimes \Phi_{\mathcal{S}_2}]_2$，其光谱 $(\Phi_{\mathcal{S}_1}, \Phi_{\mathcal{S}_2})$ 通过响应函数$\phi_{S_1}(\lambda) = \exp\left(-\frac{(\lambda-10)}{30}_2\right)$ 和 $\phi_{S_2}(\lambda) = \exp\left(-\frac{(\lambda-40)^2}{30}\right)$，在 10 个等间距的波长 $\lambda \in [0, 50)$ 处采集。测量噪声为 $e_k \sim \mathcal{N}(0, \sigma_y^2 I_{N_y})$，~~其中~~ $\sigma_y^2 = 10^{-5}$。最后，我们假设初始条件为 $X_0 \sim \delta_{x_0}$，其中δ 表示狄拉克 delta 分布，而 x_0 是奥列戈纳托模型动力学的非平凡稳态解。

对于估计任务，我们通过将晶格 V 划分为大小为 $V_b = 5$ 的块来设计块粒子滤波器（总共包含$(V/V_b)^2 = 400$ 个块）。对于每个块，我们考虑由 $N_p = 128$ 个粒子组成的集合，并假设每个粒子都按照与仿真系统相同的动力学演化。我们分析使用传统的($p(x_k|x_{k-1})$)重要性密度或最优重要性密度（公式 19）获取粒子和重要性权重时的滤波结果。

4.1 估计结果

我们采用上述配置的块粒子滤波器，利用生成的光谱数据估计奥列戈纳托模型中的浓度 $z = (z_1,\ z_2)$。不同时间 t_k 的结果如图 1 所示。

[‡] 根据其质量作用动力学，奥列戈纳托模型的动力学 $f_z(\cdot)$可写成无量纲形式，定义为 $[f_z]_1 = \epsilon^{-1}\left(z_1\ 1 - z_1\ -\ \frac{\sigma z_2\ z_1 - q}{z_1 + q}\right) + (\ )$ $D_{z_1}\nabla^2 z_1$ $[f_z]_2 = z_1 - z_2\ + D_{z_2}\nabla^2 z_2$ 1986 $(\ )$和$(\ )$（Keener 和 Tyson，）。无量纲常数为$(\epsilon, \sigma, q) = (0.08, 0.95, 0.0075)$。至于扩散系数，$(D_{z_1}, D_{z_2}) = (5 \times 10^{-4}, 5 \times 10^{-6})$。

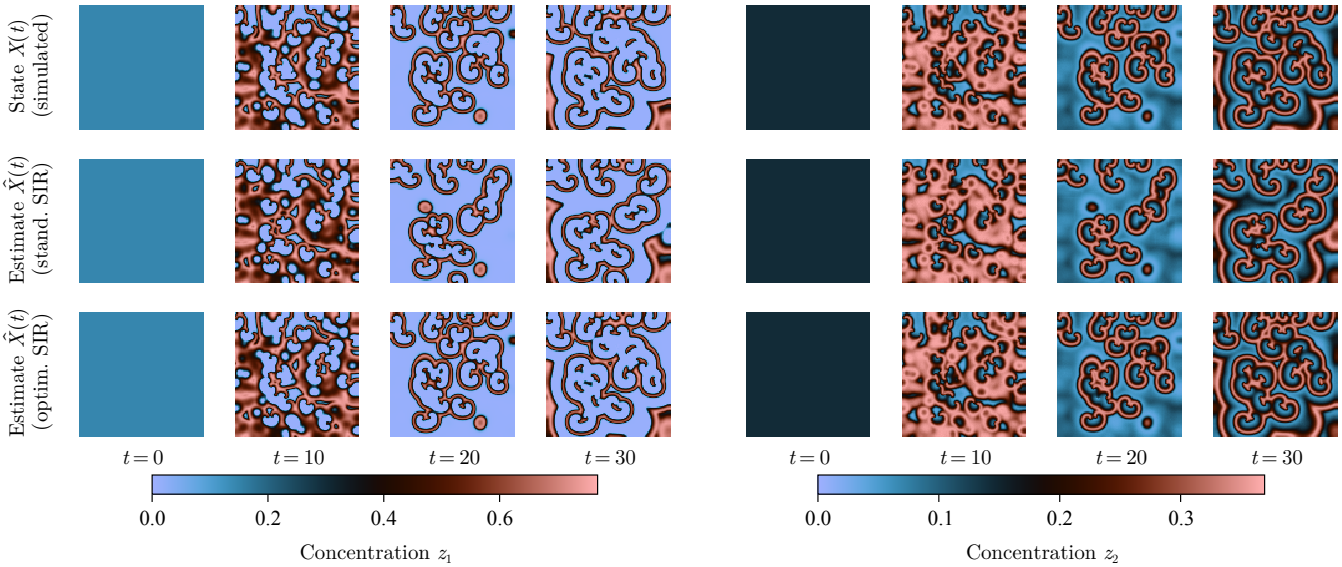


图 1. 奥列戈纳托模型在不同时间 t_k 的 $[X_k]_1$ （前 4 列）和 $[X_k]_2$ （后 4 列）快照。$\hat{}$ 第一行展示了用于生成观测 y_k 的 X_k 一次抽样。我们使用带有标准（第二行）和最优（第三行）重要性采样分布选择及重采样步骤（SIR）的块粒子滤波器估计了 X_k 。

近似条件分布的均值 π_k，即$\hat{}X^{(V_b)}{}_k := \sum_{n=1}^{N_p} w^{(V_b,n)}{}_k X_k^{(V_b,n)}$ ，以及实际状态 X_k。从系统的演化过程可以观察到，经过几次迭代后形成了振荡的“螺旋图案”。预计系统会在激活剂占主导的状态（物种 $\mathcal{S}_1$ 的浓度较高）和抑制剂占主导的状态（物种$\mathcal{S}_2$ 的浓度较高）之间来回切换。起初，我们观察到整个空间域内的浓度都在振荡。在这一瞬态期间（$t < 20$ 个时间单位），新出现的图案不断改变其形状和位置。之后，可以观察到图案在固定位置附近出现（如 $t \geq 20$ 个时间单位时所见）。结果表明，采用标准重要性分布的块粒子滤波器无法准确估计状态：粒子近似在瞬态区域内逐渐偏离，最终形成的图案与实际系统中的图案不同。相比之下，采用最优重要性分布的滤波器能够准确估计系统的图案形成过程。

接下来，我们通过比较两种近似的度量指标，考察图 1 中的不匹配情况。对于每个块 $V_b \in \mathcal{K}$，我们使用均方根误差（RMSE）来量化估计精度，

$$\mathrm{RMSE}_k^{(V_b)} = \sqrt{\|\hat{y}_k^{(V_b)} - y_k^{(V_b)}\|_2^2}, \tag{30}$$

其中预测输出为 $\hat{}y^{(V_b)}{}_k := \sum_{n=1}^{N_p} w_k^{(V_b,n)}$ （$Hx^{(V_b,n)}{}_k$ ），观测值为 y_k，而 $w^{(V_b,n)}{}_k$ 和 $x^{(V_b,n)}{}_k$ 分别表示块 V_b 中第 n 个粒子 X_k 的归一化权重和实现值。将各块的精度结果相加后，得到图 2 所示的结果。

结果（图 2）表明，由于系统具有双稳态，估计精度会发生振荡；波峰对应激活剂或抑制剂占主导状态的时期，波谷则对应瞬态时期。此外，这一指标进一步表明，只有在滤波器

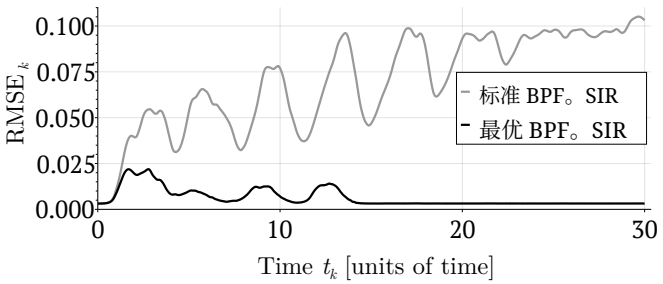


图 2. 观测值 y_k 与预测值 $\hat{}y_k$之间的 RMSE_k 。

采用最优重要性分布时，才能获得准确的估计。在这种情况下，估计误差在 $t \leq 15$ 期间略有增加，随后降至接近零。相反，使用标准重要性分布时，估计精度会迅速下降。

此外，我们通过对状态空间进行蒙特卡洛积分来计算边际似然函数 $p(y_{1:K})$，以比较两种滤波器的性能。具体而言，对于每个区块 $V_b \in \mathcal{K}$ ，我们近似计算其密度

$$p(y_{1:K}^{(V_b)}) \approx \prod_{k=1}^{K} \left(\frac{1}{N_p} \sum_{n=1}^{N_p} w_k^{(V_b,n)} \delta_{X_k^{(V_b,n)}} \right). \tag{31}$$

为避免数值问题，我们在对数尺度上近似边际似然。最后，将各区块计算出的值相加，近似得到整个输出空间上的 $\log(p(y_{1:K}))$，如图 3 所示。

这里，边缘化似然是在不假设状态具有特定实现的情况下，根据公式 (9)中的模型得到观测 $y_{1:K}$ 的概率。在这一定义下，只要使用足够多的粒子以避免退化，无论采用哪种重要性采样分布进行滤波，对 $p(y_{1:k})$的近似都应得到相似的结果。然而，采用标准选择的滤波器所获得的近似结果，与采用其最优对应方法所获得的结果并不一致。相对较小的概率密度的

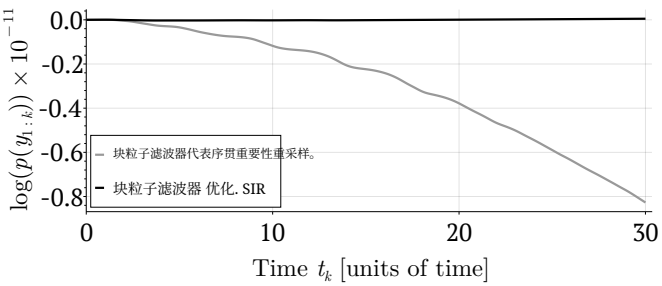


图 3. 估计的对数边际似然 $\log(p(y_{1:k}))$。

图 3 中观测值的集合是路径退化的直接结果。

5. 结论性评述

在这项工作中，我们重点介绍了对滤波方程（13）和（14）解的序贯蒙特卡洛近似。在反应-扩散系统的背景下，我们考察了块粒子滤波器在采用标准重要性采样分布和最优重要性采样分布时如何处理状态的高维性。

BZ反应系统的特征是长期不可预测性，这源于其对初始条件和过程噪声的极端敏感性。根据这些参数，系统可能会呈现图1所示的振荡模式，也可能不会。在此，我们展示了在系统噪声方差取中等值的情况下重构浓度的情景。我们想指出的是，正如Snyder 等人（2015年）所讨论的，随着这种噪声变小，最优提议分布所带来的益处将变得微乎其微。

块粒子滤波器在设计上存在一些固有局限。例如，分块算子引入的偏差在空间上并不均匀。在本研究的扩展版本中，我们计划：(i) 评估这种空间非均匀性，(ii) 研究块粒子滤波器的一些扩展方法如何处理该偏差，以及 (iii) 通过一项更广泛的实验研究，与其他高维滤波器进行比较。

参考文献

Ardizzone, V., Lewandowski, P., Luk, M.H., Tse, Y.C., Kwong, N.H., L¨ucke, A., Abbarchi, M., Baudin, E., Galopin, E., Bloch, J., 等（2013）。相干量子流体中的图灵斑图形成与控制。《科学报告》，3(1)。Asensio, M.I. 和 Ferragut, L.（2002）。关于一种包含辐射作用的野火模型。《国际工程数值方法杂志》，54(1)。Bain, A. 和 Crisan, D.（2008）。《随机滤波基础》。《随机建模与应用概率》。施普林格纽约。Chopin, N. 和 Papaspiliopoulos, O.（2020）。《序贯蒙特卡罗导论》。施普林格。Djuric, P. M.、Lu, T. 和 Bugallo, M.F.（2007）。多粒子滤波。载于*IEEE*声学、语音与信号处理国际会议，第3卷。Doucet, A.、Godsill, S.J. 和 Andrieu, C.（2000）。关于贝叶斯滤波的序贯蒙特卡罗抽样方法。《统计与计算》，10。Doucet, A.、Briers, M. 和 S´en´ecal, S.（2006）。序贯蒙特卡罗的高效分块抽样策略方法。计算与图形统计学杂志，15(3)。Ferreira, S.C.、Martins, M.L. 和 Vilela, M.J.（2002）。无血管肿瘤生长的反应扩散模型。《物理评论*E*》，65。Field, R.J. 和 Noyes, R.M.（1974）。化学系统中的振荡。IV. 真实化学反应模型中的极限环行为。《化学物理杂志》，60。Gilks, W.R. 和 Berzuini, C.（2001）。追踪移动目标：动态贝叶斯模型的蒙特卡洛推断。英国皇家统计学会会刊*B*辑，63，127–146。Green, J.B.A. 和 Sharpe, J.（2015）。位置��息与反应扩散：发育生物学中的两大理念相结合。《发育》，142(7)。Hogea, C.、Davatzikos, C. 和 Biros, G.（2008）。一种基于图像的参数估计问题：具有质量效应的反应扩散胶质瘤生长模型。《数学生物学杂志》，56(6)。Keener, J.P. 和 Tyson, J.J.（1986）。别洛乌索夫-扎博廷斯基反应中的螺旋波。《物理学*D*辑：非线性现象》，21(2)。Kitagawa, G. 和 Gersch, W.（1996）。时间序列的平滑先验分析。统计学讲义。施普林格纽约。Kondo, S.（2009）。动物如何形成皮肤图案：鱼类色素图案作为活体图灵波。《系统生物学》。Kondo, S. 和 Miura, T.（2010）。作为理解生物图案形成框架的反应扩散模型。《科学》，329(5999)。Murray, J.D.（2002）。《数学生物学：*I.* 导论》。施普林格。Musso, C.、Oudjane, N. 和 Gland, F.L.（2001）。改进正则化粒子滤波器。载于《实践中的序贯蒙特卡洛方法》。Nadler, B. 和 Coifman, R.R.（2005）。偏最小二乘法、比尔定律与净分析物信号：统计建模与分析。《化学计量学杂志》，19(1)，45–54。Nakamasu, A.、Takahashi, G.、Kanbe, A. 和 Kondo, S.（2009）。斑马鱼色素细胞之间促成图灵斑图形成的相互作用。《美国国家科学院院刊》，106(21)。Rebeschini, P. 和 van Handel, R.（2015）。局部粒子滤波器能否克服维数灾难?《应用概率年刊》，25(5)。Snyder, C.、Bengtsson, T. 和 Morzfeld, M.（2015）。采用最优提议分布的粒子滤波器性能界限。《月度天气评论》，143，4750–4761。Strauss, W.A.（2007）。《偏微分方程：导论》。约翰·威立父子出版公司，第2版。Turing, A.M.（1952）。形态发生的化学基础。《伦敦皇家学会哲学汇刊*B*辑》，237(641)。Van Leeuwen, P.J.（2003）。一种适用于大规模应用的方差最小化滤波器。《月度天气评论》，131。Witkin, A. 和 Kass, M.（1991）。反应扩散纹理。《*SIGGRAPH*计算机图形学》，25(4)。Yang, Y.、Jiang, D.、O'Regan, D. 和 Alsaedi, A.（2020）。随机别洛乌索夫-扎博廷斯基化学反应模型的动力学。《数学》，8。Zhabotinsky, A.M.（1991）。化学振荡与波的历史。《混沌》，1。