这是已被《英国皇家气象学会季刊》接受发表的文章的同行评审版本，最终发表版本可在以下网址获取：https://doi.org/10.1002/qj.70177

arXiv:2603.03926v1 [physics.ao-ph] 4 Mar 2026

一种结构局地化集合卡尔曼滤波方法

Boujemaa Ait-El-Fquih与Ibrahim Hoteit

阿卜杜拉国王科技大学（KAUST），Thuwal，沙特阿拉伯

摘要

当前最先进的集合卡尔曼滤波（EnKF）算法需要引入局地化技术，以应对其误差协方差矩阵的秩亏以及由此产生的伪相关。局地化技术大多是特设的，依据状态变量与观测变量之间的某种距离，并需要繁重的人工调参。本文提出一种新的集合滤波方法，其本身即具有局地化特性，从而无需任何辅助局部化技术。该方法并不对集合显式施加局地化，而是先对连续的分析概率密度函数（pdf）进行局地化，然后再进行集合采样。分析概率密度函数的局地化通过变分贝叶斯优化实现，将其近似为对应于状态向量各个小分区的独立边缘概率密度函数之积。随后，按照随机集合卡尔曼滤波和确定性集合变换卡尔曼滤波（ETKF）流程对这些边缘分布进行采样，所用集合的规模大于分区大小。所得滤波器的预报步骤与标准EnKF和ETKF对应方法相同，但分析步骤不同：它们依据其他分区的集合均值，迭代调整每个分区的EnKF和ETKF更新。为展示所提滤波器的潜力，在不同情景下使用Lorenz-96模型开展了数值实验。新滤波器的性能与已完成局地化调参的EnKF和ETKF相当，无论是在计算负担还是估计精度方面皆是如此。

关键词： 数据同化；EnKF；ETKF；局地化；变分贝叶斯。

1 引言

贝叶斯滤波方法，也称为序贯数据同化方法，遵循概率框架，将底层的状态估计问题分解为一系列交替进行的预报步骤和分析步骤（Künsch，2001年；Hoteit 等人，2018年）。预报步骤通过将先前的分析概率密度函数与状态模型相结合，计算给定过去观测时当前状态的预报概率密度函数（pdf）。随后，在分析步中，利用新到达的观测更新预报概率密度函数，从而得到给定截至当前时刻所有观测的状态分析概率密度函数。理论上，在特定的线性高斯状态空间系统情形下，可以使用著名的卡尔曼滤波（KF）精确计算此类预报和分析概率密度函数（Kalman，1960；Anderson and Moore，1979）。然而在实践中，为避免计算和存储大型误差协方差矩阵的成本过高，卡尔曼滤波的应用仅限于相对较小的系统（例如，Ait-El-Fquih 和 Hoteit，

通讯作者：boujemaa.aitelfquih@gmail.com，ibrahim.hoteit@kaust.edu.sa

2022b）。因此，卡尔曼滤波无法应用于现实中的大规模地球物理应用，而这类应用还涉及非线性系统。

对于非线性但规模较小的系统，可以使用粒子滤波器（PF）获得预报分布和分析分布的良好蒙特卡洛近似。粒子滤波器是一种广泛使用的数值滤波方法，能够提供具有数学依据的方案，并具备渐近收敛特性且易于实现（例如，Gordon 等人，1993年；Doucet 等人，2000年）。由于状态维数与为有效采样状态空间而需生成的粒子数量之间存在指数关系，PF 的适用性仅限于小型系统，这就是所谓的维数灾难问题（Crisan 和 Doucet，2002；Snyder 等人，2008年；Hoteit 等人，2008年）。为同时应对非线性和高维问题，人们提出了一些 PF 的改进变体（例如，Ades 和 van Leeuwen，2013；Beskos 等人，2014年；Septier 和 Peters，2015年；Ait-El-Fquih 和 Hoteit，2016；Poterjoy，2016；Potthast 等人，2019年；Mauro 等人，2022）。与此同时，一类不依赖重要性加权的贝叶斯推断方法——粒子流滤波器——已展现出应用于高维地球物理问题的能力（例如，Pulido 和 van Leeuwen，2019年；Hu 和 van Leeuwen，2021；Hu 等人，2024年）。然而，尽管与原始 PF 相比，这些方法对维数灾难具有更好的稳健性，但在业务地球物理应用中，它们尚未达到实际应用标准。

迄今为止，凭借非侵入式的公式构造、易于实现、出色的稳健性和有效性，以及合理的计算需求，集合卡尔曼滤波器（EnKF）仍是许多地球物理应用中的黄金标准（例如参见 Reichle 等人，2002年；Edwards 等人，2015年；Anderson 等人，2009年；Hoteit 等人，2015年，2018及其中的参考文献）。EnKF 通过随机实现构成的集合（通常称为集合成员），对状态预报分布和分析分布进行基于高斯分布的蒙特卡洛近似（Evensen，1994年，2006；Hoteit 等人，2015年）。在一个 EnKF 循环中，预报步骤利用动力状态模型向前积分（分析）集合，从而得到新状态的预报集合；分析步则根据输入观测和观测误差模型，使用类似卡尔曼滤波的更新方法，从所得预报集合中计算同一状态的分析集合（Hoteit 等人，2018）。

在其原始形式中（下文称为随机集合卡尔曼滤波（SEnKF）），该方法会先对观测进行随机扰动，再对每个集合成员应用卡尔曼滤波更新（例如，Burgers 等，1998年；Houtekamer和Mitchell，1998年；Hoteit 等，2015年）。扰动观测可确保所抽样分析集合的协方差渐近匹配理论卡尔曼滤波器的协方差（Burgers 等，1998年）。然而，尽管随机 SEnKF 在许多应用中表现出色，但当集合规模小于观测噪声的秩时，它会低估协方差（Hoteit 等，2015年）。这一现象可以通过以下方式缓解：i）例如，在 SEnKF 中采用改进的抽样方案，如 Evensen（2004年、2009年）所述，或采用 Hoteit 等（2015年）提出的所谓二阶重采样方案；也可以通过 ii）使用确定性（平方根）集合卡尔曼滤波器形式来完全避免这一问题。事实上，确定性集合卡尔曼滤波器无需扰动观测，而是基于实际观测，按照与卡尔曼滤波器完全相同的方式，通过更新均值和预报集合协方差的平方根形式来抽样生成分析集合（例如，Pham，2001年；Bishop 等人，2001年；Hoteit 等，2002；Tippett 等人，2003年；Hunt 等，2007；Hoteit 等，2015年）。文献中已有多种确定性集合卡尔曼滤波器，其主要区别在于分析协方差平方根的形式及其匹配方式。最受欢迎且提供公开代码的包括集合变换卡尔曼滤波器（ETKF）（Bishop 等人，2001年；Wang 等人，2004年；Hunt 等，2007）、奇异演化插值卡尔曼滤波器（Pham，2001年；Hoteit 等，2002）以及集合调整卡尔曼滤波器（Anderson，2001年；Anderson 等人，2009年）。

在实际应用中，所有随机性和确定性集合卡尔曼滤波器都表现良好；若在实现过程中不考虑不可避免的抽样误差和系统误差，这种表现便无从实现（Whitaker 等人，2008年；Houtekamer和Mitchell，2005年）。抽样—

误差源于使用较小的集合（相对于较大的状态维数而言），以限制计算成本。系统误差包括由其他缺陷引起的不确定性，具体而言，有关状态误差统计和观测模型误差统计的认知不足、模型的非线性特征与推导滤波器时所作高斯假设之间的不匹配等（例如，Ait-El-Fquih 和 Hoteit，2020及其中的参考文献）。

忽略采样误差和系统误差会限制滤波器预报（背景）集合协方差的代表性；这会导致秩亏、对角方差被低估以及虚假的非对角项（交叉方差）。这显著限制了滤波器拟合观测并产生有意义的状态估计的能力（Whitaker和Hamill，2002年；Furrer和Bengtsson，2007年；T. M. Hamill等人，2009年；Houtekamer和Zhang，2016；Hoteit 等人，2018年）。在已提出的诸多用于考虑这些误差并减轻其影响的方法中，将协方差膨胀（Anderson，2001年）和局地化（Houtekamer和Mitchell，1998年；Hamill等人，2001年）技术纳入集合卡尔曼滤波器仍然最为流行。膨胀通过人为增大集合离散度来抵消预报方差被低估的问题，方式包括乘性因子（例如Anderson和Anderson，1999年；Anderson，2007b；Miyoshi，2011）或加性因子（例如Mitchell和Houtekamer，2000年；Whitaker和Hamill，2012年）等。局地化使交叉协方差项随空间距离衰减，从而有助于缓解与秩亏和交叉协方差高估相关的问题（Houtekamer和Mitchell，1998年；Hamill等人，2001年）。

大多数局地化技术都基于状态变量与观测的位置之间的某种距离，其中最常见的是局地分析（Houtekamer和Mitchell，1998年）和协方差局地化（Hamill等，2001年）技术。局地分析仅使用位于各状态变量邻域内的观测来更新该状态变量，而协方差局地化则使用截尾矩阵对预报协方差矩阵进行舒尔积运算（Gaspari和Cohn，1999年）。其他基于物理距离的局地化方法也已被提出，例如 Houtekamer和Mitchell（2001年）、Zhang和Oliver（2010年）、Fertig等（2007年）以及Bishop和Hodyss（2007年）提出的方法。此外，研究者还开展了进一步工作，以应对部分或全部状态变量和/或观测变量无法与物理位置相关联的情况（例如，Bocquet，2016；Bishop和Hodyss，2007年；Anderson，2007a；De La Chevrotière和Harlim，2017；Luo等，2018年）。

所有 EnKF 局地化方案都明确作用于集合（离散分布），即先推导连续分析概率密度函数的集合采样器（即 EnKF），再将局地化纳入其中。本文反其道而行之：先对连续分析概率密度函数进行局地化，再推导所得局部化概率密度函数的集合采样器。分析概率密度函数的局地化，是通过（函数式）变分贝叶斯（VB）优化，使用一组可分离（独立）的边缘概率密度函数的乘积来近似它，即依据最小化 Kullback–Leibler 散度（KLD）的准则（Smidl和Quinn，2006；Blei 等人，2017；Urozayev等人，2022）。这相当于消除状态向量各分区之间的后验相依性。随后，我们按照类似 SEnKF 和 ETKF 的分析步骤对所得边缘分析密度进行采样，但也可以采用其他抽样方案。如果集合规模大于分区大小，所提出的集合方案无需结合任何其他局地化方案即可发挥作用。

所提出的分区集合平方根卡尔曼滤波（pSEnKF）和 ETKF（pETKF）采用与标准 SEnKF 和 ETKF 相同的预报步骤，但其分析步骤不同，会以迭代方式逐个分区进行。具体而言，对所有分区应用经典 SEnKF（分别对应经典 ETKF）更新后，再根据其他分区的集合均值，迭代调整／修正每个分区得到的集合（分别对应均值集合）。这种调整在一定程度上弥补了状态分区之间被施加的后验独立性。

该 c 采用变分贝叶斯方法进行状态分区的概念已在 th 中采用 e

线性高斯框架，首先见于（Ait-El-Fquih 和 Hoteit，2015年），其中提出了对卡尔曼滤波进行逐元素均值场（对角后验协方差）近似；随后见于（Ait-El-Fquih 和 Hoteit，2022b），其中将逐元素近似放宽为允许按分区进行均值场（块对角后验协方差）近似。本文将这项先前的工作扩展到了更一般的非线性集合类框架。

本文其余部分的组织如下。第 2节阐述贝叶斯滤波问题，并回顾标准 SEnKF 和 ETKF。第 3 节介绍连续分析概率密度函数的局地化方法。第 4 节重点讨论其类似 SEnKF 和 ETKF 的采样方法，并对由此产生的新型局部化滤波器进行全面讨论。第5 节展示并分析采用 Lorenz-96模型开展的数值实验结果，最后在第 6节进行总结并展开一般性讨论。

2 经典集合卡尔曼滤波：SEnKF 和 ETKF

2.1 问题 formulation

令 $\mathbf{x} = \{\mathbf{x}_n\}_{n=0}^N$ 和 $\mathbf{y} = \{\mathbf{y}_n\}_{n=0}^N$分别表示离散时间的（未知）状态过程和观测过程，其中 $\mathbf{x}_n \in \mathbb{R}^{d_x}$ 和 $\mathbf{y}_n \in \mathbb{R}^{d_y}$。在数据同化中，这些过程通常根据如下形式的状态空间系统进行描述：

$$\left\{ \begin{array}{rcl} \mathbf{x}_n & = & \mathcal{M}_{n-1}\left(\mathbf{x}_{n-1}\right) + \mathbf{u}_{n-1}, \\ \mathbf{y}_n & = & \mathbf{H}_n \mathbf{x}_n + \mathbf{v}_n; \end{array} \right. \tag{1}$$

$\mathcal{M}_{n-1}$（.）是一个将系统的状态从时间 t_{n-1} 积分到 t_n的（可能为非线性的）动力算子，而 $\mathbf{H}_n$ 是时间 t_n处的观测算子；为简便起见，此处假设其为线性的。非线性 $\mathbf{H}_n$ 的情形可以按照 EnKFs 中的通常方法处理（例如 Liu等人，2015年）。状态噪声过程 $\mathbf{u} = \{\mathbf{u}_n\}_n$与观测噪声过程 $\mathbf{v} = \{\mathbf{v}_n\}_n$被假设为相互独立、联合独立，并且与初始状态 $\mathbf{x}_0$相互独立。进一步假设噪声 $\mathbf{u}_n$ 和 $\mathbf{v}_n$ 均为均值为零的高斯噪声，其协方差分别为 $\mathbf{Q}_n$ 和 $\mathbf{R}_n$。

滤波问题是指根据可用的观测值，对状态进行（在线）推断 $\mathbf{x}_n$ ，即 $\mathcal{Y}_n = \{\mathbf{y}_0、\cdots、\mathbf{y}_n\}$，旨在求得分析概率密度函数 p_n（$\mathbf{x}_n$），也就是在给定 $\mathbf{x}_n$ 条件下的 $\mathcal{Y}_n$概率密度函数。当需要点估计时，后验均值（PM），即 p_n（$\mathbf{x}_n$）的均值，通常被选用，因为根据包括均方误差（MSE）最小化在内的若干准则，它是最优解（van Trees，1968年）。

上述独立性假设使得可以基于状态转移概率密度函数递归计算 $p_n(\mathbf{x}_n)$，

$$p(\mathbf{x}_n|\mathbf{x}_{n-1}) = \mathcal{N}_{\mathbf{x}_n}(\mathcal{M}_{n-1}(\mathbf{x}_{n-1}), \mathbf{Q}_{n-1}), \tag{2}$$

以及似然概率密度函数[1]，

$$p(\mathbf{y}_n|\mathbf{x}_n) = \mathcal{N}_{\mathbf{y}_n}(\mathbf{H}_n\mathbf{x}_n, \mathbf{R}_n); \tag{3}$$

$\mathcal{N}_{\mathbf{x}}(\mathbf{m},\mathbf{C})$ 表示高斯概率密度函数[2]，其自变量为 $\mathbf{x}$，参数为 $(\mathbf{m}, \mathbf{C})$（例如，Künsch，2001年；Ait-El-Fquih 和 Desbouvries，2006）。一个 $(n-1,\ n)$ 同化循环（递归）通常按以下方式进行[3]：预报步骤使用转移概率密度函数得到预报概率密度函数，$p_{n-1}(\mathbf{x}_n)$，

[1]值得回顾的是，当 $\mathbf{y}_n$ 为固定值时，“似然”一词对应于 $p(\mathbf{y}_n|\mathbf{x}_n)$，即条件概率密度函数 $p(\mathbf{y}_n|\mathbf{x}_n)$ 在给定点 $\mathbf{y}_n$ 处的值。然而，本文沿用常见做法，对此作了灵活处理：无论 $\mathbf{y}_n$ 是否固定，均将 $p(\mathbf{y}_n|\mathbf{x}_n)$称为“似然”。[2]高斯形式 (2) 和 (3) 分别是对噪声 $\mathbf{u}_{n-1}$ 和 $\mathbf{v}_n$作出高斯假设后直接得到的结果。[3]Ait-El-Fquih 和 Hoteit（2022a）及其参考文献中介绍了另一种滤波方法，该方法涉及一步超前平滑概率密度函数。

由前一步的分析概率密度函数得到，如下:

$$p_{n-1}(\mathbf{x}_n) = \int p(\mathbf{x}_n|\mathbf{x}_{n-1})p_{n-1}(\mathbf{x}_{n-1})d\mathbf{x}_{n-1}, \tag{4}$$

随后是一个分析步，该步骤利用似然，按照贝叶斯定理得到当前状态的分析概率密度函数，

$$p_n(\mathbf{x}_n) \propto p(\mathbf{y}_n|\mathbf{x}_n)p_{n-1}(\mathbf{x}_n). \tag{5}$$

2.2 集合高斯滤波：SEnKF 与 ETKF

SEnKF 和 ETKF 是通用滤波算法（4）-（5）的基于高斯分布的蒙特卡洛实现，其引入旨在避免模型的线性化，并高效处理大维系统（Hoteit 等人，2018年）。已证明，即使以较小的集合实现，它们也能提供稳健的 PM 估计（Hoteit 等人，2015年）。

两种滤波器都包含相同的预报步骤，作为式（4）的蒙特卡罗实现，生成状态预报集合 $\{\mathbf{f}_n^{(m)}\}_{m=1}^M$，即在用状态模型传播先前的分析集合之后$\{\mathbf{a}_{n-1}^{(m)}\}_{m=1}^M$，即

$$\mathbf{f}_n^{(m)} = \mathcal{M}_{n-1}(\mathbf{a}_{n-1}^{(m)}) + \mathbf{u}_{n-1}^{(m)}, \quad m = 1, \cdots, M; \tag{6}$$

$\mathbf{u}_{n-1}^{(m)}$ 是高斯模型误差概率密度函数 $\mathcal{N}(0, \mathbf{Q}_{n-1})$ 的一个样本。

给定预报集合和新的观测 $\mathbf{y}_n$，在高斯[4] $p(\mathbf{x}_n,\mathbf{y}_n|\mathcal{Y}_{n-1})$假设下，推导 SEnKF 和 ETKF 分析步骤，从而将问题限制为类似卡尔曼的线性更新。根据通用更新步骤（5），SEnKF 和 ETKF 分别被归类为同一分析密度的随机性和确定性采样器，$p_n(\mathbf{x}_n)$。

2.2.1 （随机性）SEnKF分析步骤

SEnKF分析步骤通过添加随机观测误差 $\mathbf{v}_n^{(m)} \sim \mathcal{N}(0, \mathbf{R}_n)$ 对观测 $\mathbf{y}_n^{(m)}$进行扰动，即 $\mathbf{y}_n^{(m)} = \mathbf{y}_n + \mathbf{v}_n^{(m)}$，然后根据 KF更新步骤更新预报成员，以获得分析集合成员:

$$\mathbf{a}_n^{(m)} = \mathbf{f}_n^{(m)} + \mathbf{K}_n(\mathbf{y}_n^{(m)} - \mathbf{H}_n\mathbf{f}_n^{(m)}), \quad m = 1, \cdots, M. \tag{7}$$

卡尔曼增益 $\mathbf{K}_n$ 是 $\mathrm{cov}[\mathbf{x}_n, \mathbf{z}_n|\mathcal{Y}_{n-1}] \times (\mathrm{cov}[\mathbf{z}_n|\mathcal{Y}_{n-1}] + \mathrm{cov}[\mathbf{v}_n])^{-1}$ 的蒙特卡罗近似，其中 $\mathbf{z}_n = \mathbf{H}_n\mathbf{x}_n$，可表示为:

$$\mathbf{K}_n = \mathbf{S}_{\mathbf{f}_n}\widetilde{\mathbf{S}}_{\mathbf{f}_n}^T \left[\widetilde{\mathbf{S}}_{\mathbf{f}_n}\widetilde{\mathbf{S}}_{\mathbf{f}_n}^T + \mathbf{R}_n\right]^{-1}, \tag{8}$$

其中 $\mathbf{S}_{\mathbf{f}_n}$ 表示预报扰动矩阵（即预报样本~误差协方差的平方根），且 $\mathbf{S}_{\mathbf{f}_n} = \mathbf{H}_n\mathbf{S}_{\mathbf{f}_n}$。

需要强调随机更新机制中的一个理论区别。虽然上述标准形式涉及对观测向量进行扰动（Burgers 等人，1998年），但从贝叶斯观点严格推导随机滤波器后可以发现，随机扰动在结构上应施加于 $\mathbf{H}_n\mathbf{f}_n^{(m)}$，而非观测（van Leeuwen，2020；Ait-El-Fquih 和 Hoteit，2022a），因此式（7）中应包含 $-\mathbf{v}_n^{(m)}$，而不是 $+\mathbf{v}_n^{(m)}$。不过，鉴于 $+\mathbf{v}_n^{(m)}$和 $-\mathbf{v}_n^{(m)}$都是同一个对称高斯分布 $\mathcal{N}(0, \mathbf{R}_n)$ 的样本，我们在本文中也依照惯例，允许一定的灵活性，改用 $+\mathbf{v}_n^{(m)}$。

[4]请注意，由于在系统（1）中，观测模型是线性的，且噪声为高斯噪声，因此只需假设 e 预报 概率密度函数 $p_{n-1}(\mathbf{x}_n)$ 为高斯分布，因为这意味着 $p(\mathbf{x}_n, \mathbf{y}_n|\mathcal{Y}_{n-1})$ 为高斯分布。

2.2.2 （确定性）ETKF分析步骤

与SEnKF分析步骤不同，ETKF分析步骤是确定性的，因为它直接利用观测，基于卡尔曼滤波更新步骤更新预报成员的均值 $\mathbf{f}_n$及其扰动矩阵 $\mathbf{S}_{\mathbf{f}_n}$。随后，利用得到的分析均值 $\hat{\mathbf{a}}_n$和分析扰动矩阵 $\mathbf{S}_{\mathbf{a}_n}$ （即分析样本误差协方差的平方根），通过匹配KF解前两阶矩的过程，确定性地抽样得到分析集合（例如，参见 Bishop 等人，2001年；Tippett 等人，2003年；Hoteit 等人，2015年；Raboudi等人，2019年）。形式上：

- $\hat{\mathbf{a}}_n$ 由基于 $\mathbf{y}_n$ 对 $\mathbf{f}_n$ 进行的卡尔曼滤波更新得到，

$$\hat{\mathbf{a}}_n = \hat{\mathbf{f}}_n + \mathbf{K}_n(\mathbf{y}_n - \mathbf{H}_n\hat{\mathbf{f}}_n); \tag{9}$$

- $\mathbf{S}_{\mathbf{a}_n}$ 通过对 $\mathbf{S}_{\mathbf{f}_n}$ 按照协方差的卡尔曼滤波更新进行变换而得到，

$$\mathbf{S}_{\mathbf{a}_n} = \mathbf{S}_{\mathbf{f}_n}\mathbf{T}_n\mathbf{\Omega}_n; \tag{10}$$

  $\mathbf{T}_n$ 是 $\mathbb{I}_M + \tilde{\mathbf{S}}_{\mathbf{f}_n}^T\mathbf{R}_n^{-1}\tilde{\mathbf{S}}_{\mathbf{f}_n}$的一个平方根，且$\mathbf{\Omega}_n$ 是保证二阶矩匹配的中心化矩阵（即确保 $\mathbf{S}_{\mathbf{a}_n}\mathbf{S}_{\mathbf{a}_n}^T$ 与卡尔曼滤波分析协方差的理论表达式相匹配）。$\mathbf{T}_n$ 与$\mathbf{\Omega}_n$ 的表达式例如可参见 Bishop 等人（2001年）；Pham （2001年）；Wang 等人（2004年）。

- 分析成员最终按如下方式采样，

$$\mathbf{a}_n^{(m)} = \hat{\mathbf{a}}_n + \sqrt{M-1}(\mathbf{S}_{\mathbf{a}_n})_m; \quad m = 1, \cdots, M, \tag{11}$$

  其中 $(\mathbf{S}_{\mathbf{a}_n})_m$ 表示 $\mathbf{S}_{\mathbf{a}_n}$ 的第 m 列。

3 理论分析概率密度函数的局地化

设状态向量 $\mathbf{x}_n$ 被划分为 K 个维度相同的分区：$\mathbf{x}_n^1$、$\mathbf{x}_n^2$、 $\cdots$、$\mathbf{x}_n^K$，维度为 $d_p = d_x/K$。在更一般的情形下，若分区 $\{\mathbf{x}_n^k\}_k$ 的维度各不相同$\{d_{p^k}\}_k$，只需将 Kd_p 替换为 $\sum_{k=1}^K d_{p^k}$，所提出的方案仍然有效。通过消除其 K 个边缘分布之间的依赖关系，可以对分析概率密度函数 $p_n(\mathbf{x}_n)$ 进行局地化。具体而言，可以用可分离边缘概率密度函数的乘积 [5], 来近似 $p_n(\mathbf{x}_n)$。

$$\begin{aligned} p_n(\mathbf{x}_n) &\approx \pi_n(\mathbf{x}_n) \\ &= \pi_n(\mathbf{x}_n^1)\pi_n(\mathbf{x}_n^2)\cdots\pi_n(\mathbf{x}_n^K). \end{aligned} \tag{12}$$

我们采用变分贝叶斯方法，并寻求最佳近似 $\pi_n(\mathbf{x}_n)$，以 Kullback–Leibler 散度最小化，或等价地，以负自由能（NFE）最大化为准则（例如 Fraysse 和 Rodet，2014年，及其中的参考文献），即：

$$\pi_n(\mathbf{x}_n) = \arg\max_{q\in\Omega} \mathbb{E}_{q(\mathbf{x}_n)}\left[\ln \frac{p(\mathbf{x}_n, \mathbf{y}_n|\mathcal{Y}_{n-1})}{q(\mathbf{x}_n)}\right], \tag{13}$$

其中，Ω 表示满足以下条件的可分离概率密度函数空间 $q(\mathbf{x}_n) = \prod_{k=1}^K q(\mathbf{x}_n^k)$，且 $\mathbb{E}_{q(\mathbf{x}_n)}[.]$ 表示关于 $q(\mathbf{x}_n)$的期望值。对最大化问题（13）应用变分法，可直接得到解 $\pi_n(\mathbf{x}_n)$，其边缘分布满足：

$$\pi_n(\mathbf{x}_n^k) \propto \mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k) \times \pi_{n-1}(\mathbf{x}_n^k), \tag{14}$$

[5] The 相同的记号 $\pi(.)$ 将用于完整向量 $\mathbf{x}_n$ 的概率密度函数，以及分区 $\mathbf{x}_n^k$ 的概率密度函数

其中，

$$\mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k) \propto \exp\left(\mathbb{E}_{\pi_n(\mathbf{x}_n^{k-})}[\ln p(\mathbf{y}_n|\mathbf{x}_n)]\right), \tag{15}$$

$$\pi_{n-1}(\mathbf{x}_n^k) \propto \exp\left(\mathbb{E}_{\pi_n(\mathbf{x}_n^{k-})}[\ln p_{n-1}(\mathbf{x}_n)]\right), \tag{16}$$

其中 $k=1,\cdots,K$ 和 $\mathbf{x}_n^{k-}$ 表示 $\mathbf{x}_n^k$ 在 $\mathbf{x}_n$（Smidl和Quinn，2008年）.

公式（14）—（16）表明，所提出的方法通过对对数似然 $\ln p(\mathbf{y}_n|\mathbf{x}_n)$ 和对数预报 $\ln p_{n-1}(\mathbf{x}_n)$ 关于 $\pi_n(\mathbf{x}_n^{k-})$ 进行“平均”，在 $\pi_n(\mathbf{x}_n^k)$ 与 $\pi_n(\mathbf{x}_n^{k-}) = \prod_{j\neq k}\pi_n(\mathbf{x}_n^j)$ 之间施加独立性。换言之，对于每个 $k=1,\cdots,K$，全状态 $\mathbf{x}_n$ 的分区 $\mathbf{x}_n^k$ 并非通过粗略地忽略其他分区 $\mathbf{x}_n^{k-}$ 来局地化，而是借助上述“平均”将它们“冻结”。另一方面，可以注意到，公式（14）的形式类似于贝叶斯更新步骤，利用似然 $\mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k)$ 将先验 $\pi_{n-1}(\mathbf{x}_n^k)$ 与后验 $\pi_n(\mathbf{x}_n^k)$ 联系起来。

真实分析概率密度 $p_n(\mathbf{x}_n^k)$（即 $p_n(\mathbf{x}_n)$ 的边缘分布）的表达式与其变分贝叶斯估计 $\pi_n(\mathbf{x}_n^k)$ 的表达式之间的联系，可以通过将 $p_n(\mathbf{x}_n^k)$ 改写为与 $\pi_n(\mathbf{x}_n^k)$（式 (14)–(16)）类似的函数形式来建立。具体做法是对式 (5) 两边取对数，关于 $p_n(\mathbf{x}_n^{k-})$ 求期望，最后再取指数。由此得到:

$$p_n(\mathbf{x}_n^k) \propto \frac{\kappa_{\mathbf{y}_n}(\mathbf{x}_n^k)\times\tilde{p}_{n-1}(\mathbf{x}_n^k)}{\exp\left(\mathbb{E}_{p_n(\mathbf{x}_n^{k-})}[\ln p(\mathbf{x}_n^{k-}|\mathbf{x}_n^k,\mathcal{Y}_n)]\right)}, \tag{17}$$

其中，

$$\kappa_{\mathbf{y}_n}(\mathbf{x}_n^k) \propto \exp\left(\mathbb{E}_{p_n(\mathbf{x}_n^{k-})}[\ln p(\mathbf{y}_n|\mathbf{x}_n)]\right), \tag{18}$$

$$\tilde{p}_{n-1}(\mathbf{x}_n^k) \propto \exp\left(\mathbb{E}_{p_n(\mathbf{x}_n^{k-})}[\ln p_{n-1}(\mathbf{x}_n)]\right). \tag{19}$$

现在，鉴于式 (18)–(19) 与式 (15)–(16) 类似，$p_n(\mathbf{x}_n^k)$ 的表达式 (17) 比 $\pi_n(\mathbf{x}_n^k)$ 的表达式 (14) 多出一项，该项取决于 $\mathbf{x}_n^k$，$T_{\mathbf{y}_n}(\mathbf{x}_n^k) = \exp\left(\mathbb{E}_{p_n}(\mathbf{x}_n^{k-})[\ln p(\mathbf{x}_n^{k-}|\mathbf{x}_n^k,\mathcal{Y}_n)]\right)$。在分区本质上以 $\mathcal{Y}_n$ 为条件相互独立的特殊情况下，这个额外项会消失（因为它与 $\mathbf{x}_n^k$ 无关），此时表达式 (14) 和 (17) 重合（即 $\mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k) = \kappa_{\mathbf{y}_n}(\mathbf{x}_n^k)$，$\pi_{n-1}(\mathbf{x}_n^k) = \tilde{p}_{n-1}(\mathbf{x}_n^k)$ 和 $\pi_n(\mathbf{x}_n^k) = p_n(\mathbf{x}_n^k)$）。这种情形对应于不同分区之间不存在连接的状态空间系统（就好像有 K 个独立系统，每个分区对应一个系统）。这是一个非常严格的条件，在现实应用中通常并不成立。

在标准系统中（即状态分区彼此相连的系统），变分贝叶斯方法通过将项 $T_{\mathbf{y}_n}(\mathbf{x}_n^k)$ 近似为一个与 $\mathbf{x}_n^k$ 无关的常量（从而使表达式（14）和（17）相同），来消除该项；随后将其纳入贝叶斯公式（14）的归一化项中。可以证明，这种近似是相对于 $p_n(\mathbf{x}_n^k)$ 进行“平均”，即:

$$\begin{aligned}\pi_n(\mathbf{x}_n^k) &\overset{(14)}{\propto} \frac{\mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k)\times\pi_{n-1}(\mathbf{x}_n^k)}{\int_{\mathbb{R}^{d_p}}\mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k)\times\pi_{n-1}(\mathbf{x}_n^k)d\mathbf{x}_n^k},\\ &\propto \frac{\mathcal{L}_{\mathbf{y}_n}(\mathbf{x}_n^k)\times\pi_{n-1}(\mathbf{x}_n^k)}{\mathbb{E}_{p_n(\mathbf{x}_n^k)}[\tilde{T}_{\mathbf{y}_n}(\mathbf{x}_n^k)]};\end{aligned} \tag{20}$$

$\tilde{T}_{\mathbf{y}_n}(\mathbf{x}_n^k) = \exp\left(\mathbb{E}_{\pi_n(\mathbf{x}_n^{k-})}[\ln p(\mathbf{x}_n^{k-}|\mathbf{x}_n^k,\mathcal{Y}_n)]\right)$。这意味着，为了对分析概率密度函数进行可分离的 KLD 估计，变分贝叶斯方法通过对 $\mathbf{x}_n^k$ 进行“平均”，将标准（真实的）贝叶斯方法中的 K 项 $T_{\mathbf{y}_n}(\mathbf{x}_n^k)$“冻结”。

最后，值得注意的是，在高斯框架下，变分贝叶斯方法并不强制真实后验概率密度函数具有块对角结构，而是试图找到与真实后验概率密度函数最接近的（按准则 (13) 衡量）具有块对角协方差的高斯概率密度函数。因此，即使采用了在高斯框架下意味着协方差矩阵为块对角结构的可分离近似 (12)，真实后验概率密度函数的协方差也不一定是块对角结构。

3.1 局部分析概率密度函数的计算 $\pi_n(\mathbf{x}_n^k)$

如上所述，尽管变分贝叶斯边缘分析概率密度函数 $\pi_n(\mathbf{x}_n^k)$ 按构造彼此独立，但它们各自仍与一种函数形式相关，该函数形式涉及对其他概率密度函数取期望 $\pi_n(\mathbf{x}_n^{k-})$。这种函数关系弥补了状态分区之间强制施加的后验独立性，这是变分贝叶斯方法的一项关键优势。然而，它也有一个缺点，即无法精确求解 K 解（14），因此需要进行近似。

一种自然的做法是进行（类似坐标下降的）迭代（例如，Bertsekas，1995）。也就是说，在第 i 次迭代中，依次更新每个 K 边缘分布 $\pi_n^i(\mathbf{x}_n^k)$，并使用最新更新的 $\pi_n(\mathbf{x}_n^{k-})$。在这方面，第 14 次更新 i可写为:

$$\pi_n^i(\mathbf{x}_n^k) \propto \mathcal{L}_{\mathbf{y}_n}^i(\mathbf{x}_n^k) \times \pi_{n-1}^i(\mathbf{x}_n^k), \tag{21}$$

其中，$\mathcal{L}_{\mathbf{y}_n}^i(\mathbf{x}_n^k)$ 和 $\pi_{n-1}^i(\mathbf{x}_n^k)$ 分别由式 (15) 和 (16) 给出，在期望算子中使用 $\pi_n^{(i,i-1)}(\mathbf{x}_n^{k-}) = \prod_{j=1}^{k-1}\pi_n^i(\mathbf{x}_n^j)\times\prod_{j=k+1}^{K}\pi_n^{i-1}(\mathbf{x}_n^j)$（回忆一下，$\pi_n(\mathbf{x}_n^{k-}) = \prod_{j=1}^{k-1}\pi_n(\mathbf{x}_n^j)\times\prod_{j=k+1}^{K}\pi_n(\mathbf{x}_n^j)$）。

3.2 通用局部化滤波算法

这里，我们给出系统 (1) 的通用局部化滤波算法的预报步和分析步。以下记 $\mathbf{H}_n^k$ 和 $\mathbf{H}_n^{k-}$ 为 $\mathbf{H}_n$ 关于位置 k 和 $k-$ 的分区，即满足 $\mathbf{H}_n\mathbf{x}_n = \mathbf{H}_n^k\mathbf{x}_n^k + \mathbf{H}_n^{k-}\mathbf{x}_n^{k-}$。另记 $(\hat{\mathbf{x}}^{n|n-1}, \mathbf{P}^{n|n-1})$ 为预报概率密度函数 $p_{n-1}(\mathbf{x}_n)$ 的前两阶矩，$(\mathbf{m}_{n|j}^{k\,i}$，$\mathbf{C}_{n|j}^k)$ 为 $\pi_j^i(\mathbf{x}_n^k)$ 的前两阶矩，其中 $j = n-1, n$（我们在附录中证明，协方差 $\mathbf{C}_{n|j}^k$ 在各次迭代中保持不变）。

•预报步骤。$p_{n-1}(\mathbf{x}_n)$ 由 $\pi_{n-1}(\mathbf{x}_{n-1})$ 计算得到，遵循标准预报步骤，即遵循式（4），并使用 $\pi_{n-1}(\mathbf{x}_{n-1})$而非 $p_{n-1}(\mathbf{x}_{n-1})$。

• 分析步骤。所得预报概率密度函数 $p_{n-1}(\mathbf{x}_n)$ 根据新获取的观测 $\mathbf{y}_n$ 进行更新，并遵循迭代过程以得到局部分析概率密度函数 $\pi_n(\mathbf{x}_n)$。从初始化 $\pi_n^0(\mathbf{x}_n) = p_{n-1}(\mathbf{x}_n)$ 出发，第 i 次迭代依次计算边缘分布 $\pi_n^i(\mathbf{x}_n^k)$，从 $k=1$ 到 K，分三步进行:

1. 第一步计算变分贝叶斯似然 $\mathcal{L}_{\mathbf{y}_n}^i(\mathbf{x}_n^k)$，依据式（15），并在期望算子中使用 $\pi_n^{(i,i-1)}(\mathbf{x}_n^{k-})$。由此得到,

$$\mathcal{L}_{\mathbf{y}_n}^i(\mathbf{x}_n^k) = \mathcal{N}_{\mathbf{y}_n}(\mathbf{H}_n^k\mathbf{x}_n^k + \mathbf{H}_n^{k-}\mathbf{m}_{n|n}^{k-,i}, \mathbf{R}_n); \tag{22}$$

$$\mathbf{m}_{n|n}^{k-,i} = (\mathbf{m}_{n|n}^{1,i}, \cdots, \mathbf{m}_{n|n}^{k-1,i}, \mathbf{m}_{n|n}^{k+1,i-1}, \cdots, \mathbf{m}_{n|n}^{K,i-1}).$$

2. 第二步根据方程（$p_{n-1}(\mathbf{x}_n)$16），在期望算子中使用$\pi_n^{(i,i-1)}(\mathbf{x}_n^{k-})$，对预报 $\pi_{n-1}^i(\mathbf{x}_n^k)$进行“调整”，从而得到 $p_{n-1}(\mathbf{x}_n)$。假设 $p_{n-1}(\mathbf{x}_n)$ 是

高斯分布，则可得，

$$\pi_{n-1}^{i}(\mathbf{x}_n^k) = \mathcal{N}_{\mathbf{x}_n^k}(\mathbf{m}_{n|n-1}^{k,i}, \mathbf{C}_{n|n-1}^{k}); \tag{23}$$

$$\mathbf{m}_{n|n-1}^{k,i} = \hat{\mathbf{x}}_{n|n-1}^{k} + \mathbf{G}_n^k(\mathbf{m}_{n|n}^{k-,i} - \hat{\mathbf{x}}_{n|n-1}^{k-}), \tag{24}$$

$$\mathbf{C}_{n|n-1}^{k} = \mathbf{P}_{n|n-1}^{k} - \mathbf{G}_n^k\mathbf{P}_{n|n-1}^{k-,k}, \tag{25}$$

其中 $\mathrm{G}_n^k = \mathrm{P}_{n|n-1}^{k\,k-}(\mathrm{P}_{n|n-1}^{k-})^{+}$，其中 $(.)^{+}$ 表示伪逆，而协方差 $\mathrm{P}_{n|n-1}^{k}$ 和 $\mathrm{P}_{n|n-1}^{k-}$ 以及交叉协方差 $\mathrm{P}_{n|n-1}^{k\,k-}$，和 $\mathrm{P}_{n|n-1}^{k-\,k}$，根据位置 k 和 k^{-}从 $\mathrm{P}_{n|n-1}$中提取。

3. 第三步根据贝叶斯更新（22），将高斯变分贝叶斯似然（23）-（25）与高斯变分贝叶斯预报结合，得到高斯变分贝叶斯分析概率密度函数：

$$\pi_{n}^{i}(\mathbf{x}_n^k) = \mathcal{N}_{\mathbf{x}_n^k}(\mathbf{m}_{n|n}^{k,i}, \mathbf{C}_{n|n}^{k}); \tag{26}$$

$$\mathbf{m}_{n|n}^{k,i} = \mathbf{m}_{n|n-1}^{k,i} + \mathbf{L}_n^k(\mathbf{y}_n - \mathbf{H}_n^k\mathbf{m}_{n|n-1}^{k,i} - \mathbf{H}_n^{k-}\mathbf{m}_{n|n}^{k-,i}), \tag{27}$$

$$\mathbf{C}_{n|n}^{k} = \mathbf{C}_{n|n-1}^{k} - \mathbf{L}_n^k\mathbf{H}_n^k(\mathbf{C}_{n|n-1}^{k})^{T}, \tag{28}$$

其中 $\mathrm{L}_n^k = \mathrm{C}_{n|n-1}^{k}(\mathrm{H}_n^k)^{T}[\mathrm{H}_n^k\mathrm{C}_{n|n-1}^{k}(\mathrm{H}_n^k)^{T} + \mathrm{R}_n]^{-1}$。

收敛时（最后一次迭代，$i = I$），对于所有 k，$\pi_n(\mathrm{x}_n^k) = \pi_n^I(\mathrm{x}_n^k)$。

4 实际实现：分块集合平方根集合卡尔曼滤波（pSEnKF）和 -ETKF（pETKF）

这里，我们重点介绍第 3.2 节所述通用算法的类 SEnKF 和类 ETKF 集合实现。为简单起见，我们沿用第 2 节中用于集合及其相应矩（样本均值、样本协方差和扰动矩阵）的记号。

滤波器 pSEnKF 和 pETKF 的预报步骤与其标准对应形式相同（式（6））。基于通用类 KF更新（（24）-（25），（27）-（28））推导其分析步骤时，可能会遇到与增益矩阵 $\mathbf{G}_n^k$ 有关的计算难题。事实上，在集合框架中，这些矩阵可基于预报扰动矩阵计算如下：

$$\begin{aligned}\mathbf{G}_n^k &\approx \mathbf{S}_{\mathbf{f}_n^k}\mathbf{S}_{\mathbf{f}_n^{k-}}^{T}(\mathbf{S}_{\mathbf{f}_n^{k-}}\mathbf{S}_{\mathbf{f}_n^{k-}}^{T})^{+},\\ &= \mathbf{S}_{\mathbf{f}_n^k}(\mathbf{S}_{\mathbf{f}_n^{k-}})^{+}.\end{aligned} \tag{29}$$

在这一局地化框架下，需要预先计算 K $(M \times (d_x - d_p))$ 阶伪逆，随后进行 K 次乘积（29），此外还需要在所有迭代期间存储 K 计算得到的$(d_p \times (d_x - d_p))$ 阶 $\mathbf{G}_n^k$ 矩阵，因为每次迭代的更新（24）都需要用到这些矩阵。处理大维系统时，这很快就会变得繁琐。本初步研究通过直接忽略式（$\mathbf{G}_n^k$24）-（25）中涉及 G 的项来解决这一问题，从而将其简化为：

$$\mathbf{m}_{n|n-1}^{k,i} = \hat{\mathbf{x}}_{n|n-1}^{k} \quad \text{and} \quad \mathbf{C}_{n|n-1}^{k} = \mathbf{P}_{n|n-1}^{k}, \tag{30}$$

从而将变分贝叶斯分析更新（27）-（28）简化为：

$$\begin{aligned}\mathbf{m}_{n|n}^{k,i} &= \hat{\mathbf{x}}_{n|n-1}^{k} + \mathbf{L}_n^k(\mathbf{y}_n - \mathbf{H}_n^k\hat{\mathbf{x}}_{n|n-1}^{k} - \mathbf{H}_n^{k-}\mathbf{m}_{n|n}^{k-,i}),\\ &= \underbrace{\hat{\mathbf{x}}_{n|n-1}^{k} + \mathbf{L}_n^k(\mathbf{y}_n - \mathbf{H}_n^k\hat{\mathbf{x}}_{n|n-1}^{k})}_{\Theta_n^k\,(\text{classical KF})} \underbrace{- \mathbf{L}_n^k\mathbf{H}_n^{k-}\mathbf{m}_{n|n}^{k-,i}}_{\text{adjustment}},\end{aligned} \tag{31}$$

$$\mathbf{C}_{n|n}^{k} = \mathbf{P}_{n|n-1}^{k} - \mathbf{L}_n^k\mathbf{H}_n^k(\mathbf{P}_{n|n-1}^{k})^{T}. \tag{32}$$

在 $p_{n-1}(\mathbf{x}_n)$服从高斯假设的情况下，忽略$\mathbf{G}_n^k$ 在方程（24）-（25中的项，相当于忽略$\mathbf{x}_n^k$与$\mathbf{x}_n^{k-}$在给定 $\mathcal{Y}_{n-1}$条件下的依赖关系，也就是说，假设：

$$p(\mathbf{x}_n^k|\mathbf{x}_n^{k-},\mathcal{Y}_{n-1}) = p_{n-1}(\mathbf{x}_n^k). \tag{33}$$

注意到这一假设会导致忽略不同分区之间的预报交叉协方差，因此可以将其解释为对预报分布的一种隐式局地化。

式（31）表明，在给定迭代中，某个状态分区的 VB分析均值 $\mathbf{x}_n^k$，实际上就是经典KF分析均值 Θ_n^k，再加上其余分区的 VB分析均值最新值的线性组合所形成的偏移 $\mathbf{x}_n^{k-}$。此外，Θ_n^k 只需计算一次，因为它不依赖于迭代；这意味着 VB KF 分析均值 31）的迭代更新 $\mathbf{m}_{n|n}^k$，只是根据最新 Θ_n^k的线性组合，对经典KF分析均值 $\mathbf{m}_{n|n}^{k-}$ 进行迭代调整。

4.1 pSEnKF分析步骤

pSEnKF分析步骤是迭代变分贝叶斯分析步骤（31）的一种集合形式。也就是说，对于每个第 k 个分区，首先执行经典集合平方根卡尔曼滤波更新，以获得集合$\{\boldsymbol{\theta}_n^{k,(m)}\}_{m=1}^M$，然后通过加上最新样本均值 $\hat{\mathrm{a}}_n^{k-}$ 的线性组合所形成的偏移，对其进行迭代调整，从而得到所需的变分贝叶斯分析集合 $\{\mathrm{a}_n^{k\,(m)}\}_{m=1}^M$。

类比于式（ Θ_n^k 中31）（以及式（7）-（8））的表达式，经典 SEnKF 更新对于 $k=1,\ \cdots,\ K$,

$$\tilde{\mathbf{L}}_n^k = \mathbf{S}_{\mathbf{f}_n^k}\tilde{\mathbf{S}}_{\mathbf{f}_n^k}^T[\tilde{\mathbf{S}}_{\mathbf{f}_n^k}\tilde{\mathbf{S}}_{\mathbf{f}_n^k}^T + \mathbf{R}_n]^{-1}, \tag{34}$$

$$\boldsymbol{\theta}_n^{k,(m)} = \mathbf{f}_n^{k,(m)} + \tilde{\mathbf{L}}_n^k(\mathbf{y}_n^{(m)} - \mathbf{H}_n^k\mathbf{f}_n^{k,(m)}), \tag{35}$$

其中 $\tilde{\mathrm{S}}_{\mathrm{f}_n^k} = \mathrm{H}_n^k\mathrm{S}_{\mathrm{f}_n^k}$，$\mathrm{y}_n^{(m)} = \mathrm{y}_n + \mathrm{v}_n^{(m)}$ 和 $\mathrm{v}_n^{(m)} \sim \mathcal{N}(0, \mathrm{R}_n)$。随后开始迭代。在第 i 次迭代中，通过对所得的 $\{\boldsymbol{\theta}_n^{k,(m)}\}_{m=1}^M$ 进行调整，并平移 $-\tilde{\mathrm{L}}_n^k\mathrm{H}_n^{k-}\hat{\mathrm{a}}_n^{k-\,i}$，得到变分贝叶斯分析集合 $\{\mathrm{a}_n^{k\,(m)\,i}\}_{m=1}^M$。其中，样本均值 $\hat{\mathrm{a}}_n^{k-\,i}$ 表示为 $\hat{\mathrm{a}}_n^{k-\,i} = (\hat{\mathrm{a}}_n^{1\,i}, \cdots, \hat{\mathrm{a}}_n^{k-1\,i}, \hat{\mathrm{a}}_n^{k+1\,i-1}, \cdots, \hat{\mathrm{a}}_n^{K\,i-1})$。具体而言，在第 i 次迭代中，对于 $k=1,\cdots,K$，有：

$$\mathbf{a}_n^{k,(m),i} = \boldsymbol{\theta}_n^{k,(m)} - \tilde{\mathbf{L}}_n^k\mathbf{H}_n^{k-}\hat{\mathbf{a}}_n^{k-,i};\ \ m=1,\cdots,M. \tag{36}$$

在收敛时（最后一次迭代 I），对所有 k，$\{\mathrm{a}_n^{k\,(m)}\}_{m=1}^M = \{\mathrm{a}_n^{k\,(m)\,I}\}_{m=1}^M$，和 $\hat{\mathrm{a}}_n^k = \hat{\mathrm{a}}_n^{k\,I}$。例如，若超过预先设定的最大迭代次数 I，或基于两个连续的 $\hat{\mathrm{a}}_n = (\hat{\mathrm{a}}_n^1, \cdots, \hat{\mathrm{a}}_n^K)$ 之间差异（即基于 $\hat{\mathrm{a}}_n^i - \hat{\mathrm{a}}_n^{i-1}$）计算的误差小于给定阈值，则认为已达到收敛。

pSEnKF 的$(n-1,n)$ 同化循环如算法 1所示。计算量将在第 4.3 节进一步讨论。

4.2 pETKF分析步

pETKF 分析步基于类似卡尔曼滤波的方程（31）-（32），遵循 ETKF 的原理（第 2.2.2 节），即先按分区计算“矩” $\hat{\mathbf{a}}_n$ 和 $\mathbf{S}_{\mathbf{a}_n}$，然后用它们对成员 $\mathbf{a}_n^{(m)}$进行采样。

分析均值，$\hat{\mathbf{a}}_n$，按照迭代KF调整过程计算（31）。也就是说，对于每个 k分区，首先执行一次经典KF更新，以得到分析均值$\hat{\boldsymbol{\theta}}_n^k$，然后基于最新$\hat{\mathbf{a}}_n^{k-}$的线性组合对其进行迭代调整，从而得到所需的VB分析均值$\hat{\mathbf{a}}_n^k$。

类比于 Θ_n^k 式（31）中的表达式，经典卡尔曼滤波更新对于 $k=1$、$\cdots$、K,

$$\hat{\boldsymbol{\theta}}_n^k = \hat{\mathbf{f}}_n^k + \tilde{\mathbf{L}}_n^k(\mathbf{y}_n - \mathbf{H}_n^k\hat{\mathbf{f}}_n^k). \tag{37}$$

算法1： pSEnKF

- 预报：根据式（ $\{\mathbf{f}_n^{(m)}\}_{m=1}^{M}$ ）对集合进行采样，然后在计算 $\hat{\mathbf{f}}_n$和 $\mathbf{S}_{\mathbf{f}_n}$之前对其进行膨胀。

- 分析：

1) 经典集合卡尔曼滤波：对于 $k=1,\ \cdots,\ K$，使用公式（34）计算 $\tilde{\mathbf{L}}_n^k$，然后使用公式（35）计算 $\{\boldsymbol{\theta}_n^{k,(m)}\}_{m=1}^{M}$。

2) 迭代调整：

▷初始化 $\hat{\mathbf{a}}_n = \hat{\mathbf{f}}_n$ 以及 $\mathbf{h}_n = \mathbf{H}_n\hat{\mathbf{f}}_n$.

▷**当** 未收敛时 **执行**：对于 $k=1,\ \cdots,\ K$,

$$\begin{aligned}
\mathbf{h}_n &= \mathbf{h}_n - \mathbf{H}_n^k\hat{\mathbf{a}}_n^k \\
\mathbf{a}_n^{k,(m)} &= \boldsymbol{\theta}_n^{k,(m)} - \tilde{\mathbf{L}}_n^k\mathbf{h}_n; \quad m=1,\cdots,M \\
\hat{\mathbf{a}}_n^k &= \frac{1}{M}\sum_{m=1}^{M}\mathbf{a}_n^{k,(m)} \\
\mathbf{h}_n &= \mathbf{h}_n + \mathbf{H}_n^k\hat{\mathbf{a}}_n^k
\end{aligned}$$

End

在第 i 次迭代中，通过对已获得的 $\hat{\boldsymbol{\theta}}_n^k$ 进行调整，并平移 $-\tilde{\mathbf{L}}_n^k\mathbf{H}_n^{k-}\hat{\mathbf{a}}_n^{k-,i}$，计算得到 VB分析均值 $\hat{\mathbf{a}}_n^{k,i}$，：

$$\hat{\mathbf{a}}_n^{k,i} = \hat{\boldsymbol{\theta}}_n^k - \tilde{\mathbf{L}}_n^k\mathbf{H}_n^{k-}\hat{\mathbf{a}}_n^{k-,i}. \tag{38}$$

迭代循环按照 pSEnKF 中讨论的相同停止准则终止。假设经过 I 次迭代后达到收敛，则对所有 k，取 $\hat{\mathbf{a}}_n^k = \hat{\mathbf{a}}_n^{k,I}$，。

样本分析协方差的平方根按照与式（32）类似的方式，根据式（10）计算如下：

$$\mathbf{S}_{\mathbf{a}_n^k} = \mathbf{S}_{\mathbf{f}_n^k}\mathbf{T}_n^k\boldsymbol{\Omega}_n, \tag{39}$$

其中 变换矩阵 $\mathbf{T}_n^k$ 在此由 $\mathbb{I}_M + \tilde{\mathbf{S}}^{T}_{\mathbf{f}_n^k}\mathbf{R}_n^{-1}\tilde{\mathbf{S}}_{\mathbf{f}_n^k}$ 的平方根给出 .

由此得到的 $\{\hat{\mathbf{a}}_n^k\}_{k=1}^{K}$ 和 $\{\mathbf{S}_{\mathbf{a}_n^k}\}_{k=1}^{K}$ 随后按列拼接，分别构成 $\hat{\mathbf{a}}_n$ 和 $\mathbf{S}_{\mathbf{a}_n}$，并据此按照式（$\mathbf{a}_n^{(m)}$）对分析成员进行采样。11)。

pETKF 的一个 $(n-1,n)$ 同化循环总结于算法 2，其计算量将在第 4.3 节讨论。

4.3 讨论

第 3.2 节提出的通用滤波方案采用了与标准通用滤波算法（式（4））相同的预报步骤，但其分析步有所不同：它遵循变分贝叶斯方法，按分区对状态 $\mathbf{x}_n$ 拆分为低维子状态 $\mathbf{x}_n^k$、 $k=1$、 $\cdots$、K。该分析步包含一个迭代过程；每次迭代 i 由 K 个连续的贝叶斯更新步骤构成，分别将 K 状态分区 $\mathbf{x}_n^k$ 与观测 $\mathbf{y}_n$关联起来，以计算变分贝叶斯分析概率密度 $\pi_n^i(\mathbf{x}_n^k)$。

在观测模型为线性的特定情况下，此类 K 贝叶斯更新可归结为 K 类KF更新，用于计算概率密度函数 $\pi_n^i(\mathbf{x}_n^k)$的均值和协方差。 $\pi_n^i(\mathbf{x}_n^k)$的协方差不随迭代变化，因此可在迭代循环之外，使用应用于模型 $\mathbf{y}_n = \mathbf{H}_n^k\mathbf{x}_n^k + \mathbf{v}_n$ 的经典KF更新一次性计算得到。每个 $\pi_n^i(\mathbf{x}_n^k)$的均值会随迭代变化，表现为对经典KF分析均值的迭代调整（即通过将KF应用于模型 $\mathbf{y}_n = \mathbf{H}_n^k\mathbf{x}_n^k + \mathbf{v}_n$ 获得），并加上其余分区所对应均值的最新值的线性

算法 2： pETKF

- 预报：对集合进行采样 $\{\mathbf{f}_{n}^{(m)}\}_{m=1}^{M}$ 使用式（6），然后在计算 $\mathbf{f}_n$ 以及 $\mathbf{S}_{\mathbf{f}_n}$。
- 分析：

  1) 均值：

  ▷ 经典卡尔曼滤波：对于 $k=1,\cdots,K$，计算 $\tilde{\mathbf{L}}_n^k$ 使用公式。(34)，然后 $\hat{\boldsymbol{\theta}}_n^k$ 使用公式 (37)。

  ▷ 迭代调整：

  ◇初始化 $\hat{\mathbf{a}}_n = \hat{\mathbf{f}}_n$ 和 $\mathbf{h}_n = \mathbf{H}_n\hat{\mathbf{f}}_n$.

  ◇**当** 尚未收敛 **时**：对于 $k=1,\cdots,K$,

$$\begin{aligned}\mathbf{h}_n &= \mathbf{h}_n - \mathbf{H}_n^k\hat{\mathbf{a}}_n^k\\ \hat{\mathbf{a}}_n^k &= \hat{\boldsymbol{\theta}}_n^k - \tilde{\mathbf{L}}_n^k\mathbf{h}_n\\ \mathbf{h}_n &= \mathbf{h}_n + \mathbf{H}_n^k\hat{\mathbf{a}}_n^k\end{aligned}$$

  End

  2) 协方差平方根：对于 $k=1、\cdots、K$，计算 $\mathbf{S}_{\mathbf{a}_n^k}$ 由 $\mathbf{S}_{\mathbf{f}_n^k}$ 使用式（39）。

  3）成员：对集合 $\{\mathbf{a}_n^{(m)}\}_{m=1}^{M}$ 进行采样（11），基于所得的 $\hat{\mathbf{a}}_n$ 和 $\mathbf{S}_{\mathbf{a}_n}$ 按照式。

组合，$\mathbf{x}_n^{k-}$。

所提出的pSEnKF和pETKF算法是前述通用滤波算法的集合变体，其基于预报概率密度函数的高斯假设以及独立性假设（33）。它们与标准 SEnKF和ETKF滤波器共享相同的预报步骤，但涉及不同的分析步骤，这些分析步骤分别是前述迭代通用分析步的随机性与确定性集合形式。更具体而言，分区集合平方根卡尔曼滤波分析步骤由基于样本分析均值线性组合的迭代 SEnKF 调整过程构成（即对标准 SEnKF分析所生成集合的调整），而pETKF分析步则在（确定性地）生成分析集合之前，精确采用上述通用迭代 KF 调整过程。

将状态向量划分为分区可视为一种内在局部化，从而无需（外部）辅助局地化，尤其是当集合规模 M大于分区大小 d_p时。事实上，与将局地化应用于离散分布（即（预报或分析）集合）的SEnKF和ETKF不同，所提出的方法首先基于变分贝叶斯方法对（连续）分析密度施加局地化，然后由pSEnKF和pETKF算法对所得局部化概率密度函数进行采样，且不为它们配备任何（外部）局地化，条件是 $M > d_p$。

还值得注意的是，前文（第3.1节）提到的函数关系，将每个 VB 分析概率密度函数 $\pi_n^i(\mathbf{x}_n^k)$ 与关于 $\pi_n^{(i,i-1)}(\mathbf{x}_n^{k^-})$ 的期望联系起来；在 pSEnKF 和 pETKF 分析步骤中，该关系以校正项 $-\tilde{\mathbf{L}}_n^k\mathbf{H}_n^k\hat{\mathbf{a}}_n^{k-\,i,}$ 的形式出现（即式（36）和（38））。这意味着，与人们可能预期的相反，使用 VB 方法对（连续）分析概率密度函数进行分区，并不会使分析 pSEnKF 和 pETKF 集合在各分区之间完全独立，因为它们仍通过样本均值相互关联，即 $\{\mathbf{a}_n^{k\,(m)\,i}\}_m$，，与 $\hat{\mathbf{a}}_n^{k-\,i,}$ 相关联。更准确地说，这种耦合发生在均值估计层面，但不会重新引入完整的跨分区（分析）协方差。

因此，当真实后验分布在分区边界两侧表现出强依赖关系时（即边界将强耦合的相邻部分分开时），分析更新可能通过迭代均值校正，在集合均值中较好地保留这些耦合关系；然而，分析集合中的跨边界相关性不会得到保留，从而造成物理失衡。尽管如此，这种现象与标准局地分析方案中观察到的失衡类似（例如 Hunt 等人，2007年），但并不会损害滤波器的稳定性。事实上，在随后的预报步骤中，模型动力学会发挥平滑器的作用，重新建立跨分区相关性，从而恢复物理平衡，然后进入下一个同化循环。

所提出的迭代变分贝叶斯算法不一定能收敛到理论Kullback–Leibler 散度的局部最小值，这主要是因为这些算法基于高斯分布和后验独立性假设（例如 Ait-El-Fquih 等人，2023；Smidl和Quinn，2006；Sato，2001；Chappell 等人，2009年；Blei 等人，2017；Ait-El-Fquih 等人，2020）。不过，按构造，这些算法保证能够收敛到“估计的”Kullback–Leibler 散度的局部最小值，即基于集合近似得到的局部最小值（例如 Ait-El-Fquih 等人，2020）。

要刻画类变分贝叶斯近似的准确性并不容易，因为这取决于多个因素，包括分区数量、 K 以及这些分区之间后验依赖关系的强弱（例如 Jordan 等人，1999年；Jaakkola 和 Jordan，2000年；Jaakkola，2001年；Smidl和Quinn，2006）。相较之下，我们可以直观地探讨何时变分贝叶斯近似的准确性可能较为理想，以及何时可能失效。显然，当状态分区之间的（后验）依赖关系较弱时，变分贝叶斯估计 $\pi_n(\mathbf{x}_n) = \prod_{k=1}^{K} \pi_n(\mathbf{x}_n^k)$的 $p_n(\mathbf{x}_n)$ 应当是合理的。相反，当这些分区之间的依赖关系变得很强时，对条件独立性的“假设”也会变得很强，这可能会损害变分贝叶斯估计的效果。

变分贝叶斯估计理论上应随着分区数量 K的减少而改善，或者等价地，随着维数 d_p的增加而改善。然而在实践中，由于 pSEnKF 和 pETKF 算法（不使用辅助局地化时）只有在 d_p 不超过集合规模 M时才有用，因此经验法则是：对于给定的 M，选择尽可能大的 d_p，但不超过 M。

所提出的 pSEnKF 和 pETKF 在实际应用中的算法与各自的标准对应算法相似，唯一额外的组成部分是迭代调整（参见算法1和2）。迭代调整步骤的计算成本随状态维数线性增长， d_x，约为 $\mathcal{C}^{\text{extra}}_{\text{pSEnKF}} = [2d_y + (6d_y + 2M)I]\, d_x$ 次浮点运算[6]（对于 pSEnKF），以及 $\mathcal{C}^{\text{extra}}_{\text{pETKF}} = [2d_y + 6d_y I]\, d_x$ 次浮点运算（对于 pETKF）。迭代次数 I通常远小于 d_x 和 d_y，这是变分贝叶斯型优化算法的一项关键特性（例如，Sato，2001年）。因此，在大规模地球物理应用中，与 SEnKF 和 ETKF 分析步骤（$\sim \mathcal{O}(d_y^2 d_x)$）（在线性观测模型中）以及预报步骤的成本相比，和的计算成本可以忽略不计。

最后值得指出的是，所提出的迭代方案属于 VB 坐标下降过程（具体而言，是对负自由能进行循环坐标上升），并采用解析的类卡尔曼更新，而非数值线性求解器（例如 Krylov 子空间方法）。因此，标准的物理预条件器并不适用于这些概率更新。不过，在高维或非高斯场景中，加速技术可能会有所帮助，例如使用预条件共轭梯度法求解大规模创新协方差求逆问题，或使用自然梯度方法（Martens，2020）进行基于优化的非高斯更新。

[6]浮点运算，参见例如 Ait-El-Fquih 和 Hoteit （2015年）。

5 数值实验

我们采用非线性洛伦兹-96（L96）模型（Lorenz和Emanuel，1998年）开展数值实验，以评估所提出的 pSEnKF 和 pETKF 的行为，并通过局地化将其性能与标准 SEnKF 和 ETKF 进行比较。L96模型基于微分方程模拟大气量随时间的演变：

$$\frac{dx(j,t)}{dt} = x(j-1,t)[x(j+1,t) - x(j-2,t)] - x(j,t) + F, \tag{40}$$

其中，$F = 8$ 表示外部强迫项，$x(j,t)$表示在时间 t状态的第 j个元素，其中 $j = 1$、$\cdots$、$d_x = 40$；边界条件为周期性的（即 $x(-1,t) = x(d_x - 1,t)$，$x(0,t) = x(d_x,t)$，以及 $x(1,t) = x(d_x + 1,t)$）。

模型（40）采用四阶龙格-库塔离散化方案进行数值积分，时间步长为 $\delta_m = 0.05$，相当于实际时间中的六小时。假设模型是完美的，并向前积分以模拟一条包含114600 个时间步长的轨迹，从一个初始状态开始，其中各元素（状态变量）设为 F，但第20个元素设为（$1 + 0.001$）F。所得轨迹的前100000 个时间步长作为自旋期被舍弃，剩余的14600 个时间步长（即模型时间中的10 年）被视为参考状态轨迹，之后将根据从这些参考状态中提取的观测，使用滤波器对其进行估计。观测是在状态中加入均值为零、协方差为$\mathbf{R} = r \times \mathbb{I}_{d_x}$的高斯噪声后提取的，其中 $r = 0.6$对应于信噪比（SNR）为15 分贝，即

$$r = \frac{\sum_{n=0}^{T-1} \|\mathbf{H}_n \mathbf{x}_n\|^2}{T d_y} \times 10^{-\mathrm{SNR}/10};$$

T 为模型时间步长的总数（例如 Ait-El-Fquih 和 Hoteit，2015年）.

我们首先展示一项实验的结果，该实验旨在说明所提出的 pSEnKF 和 pETKF 在具有 $M = 20$ 个成员、采用完整观测情景（即观测所有状态变量），且每四个模型时间步 assimilate 一次观测的设置下的适用性（即观测时间步长为 $\delta_o = 4\,\delta_m$）。在这两种滤波器中，初始预报集合由一个高斯分布生成，该分布以模型轨迹的均值为中心，协方差为 $3\mathbb{I}_{d_x}$。它们的（迭代）分析步骤将状态划分为 $K = 4$ 个分区（每个分区的维数为 $d_p = 10$），并假设当全状态分析估计的相对平方失配范数 $\mathrm{RSMN}^i = \|\hat{}\mathbf{a}_n^i - \hat{}\mathbf{a}_n^{i-1}\|^2 / \|\hat{}\mathbf{a}_n^{i-1}\|^2$低于 10^{-10} 时即达到收敛（这等同于使用 NFE（44）-（45）监测收敛，因为在所提出的滤波器中，迭代过程中只有分析均值会变化，而协方差保持不变）。此外，两种滤波器都采用了膨胀，膨胀因子为 1.1，该值通过反复试验调优选定。

为减少统计涨落，每项实验都使用不同的观测噪声和初始集合独立重复进行 50 次，并取这 50 次运行的平均值作为最终结果。

图 1 展示了在同化循环RSMN中的迭代变化过程 $n = 0$，$N/3$，$2N/3$和 N，其中 N 为同化循环总数。可以看出，两种算法均仅经过2次迭代便收敛，这与VB类滤波算法众所周知的快速收敛特性一致（例如，Sato，2001年）。

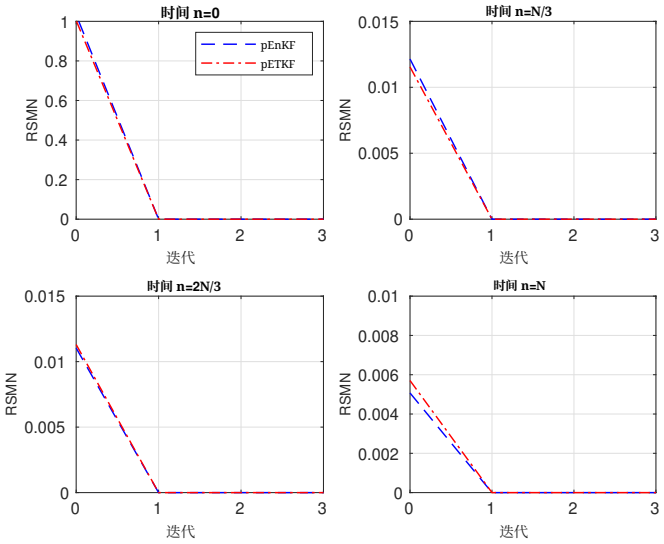


图 1：每经过 N/3 个同化循环后，状态分析估计的 RSMN 随迭代次数的演变，其中 N 为同化循环的总数。

图 2绘出了最后100 个同化循环中参考状态的前四个分量（以便更清晰地展示），以及两种所提出滤波器得到的分析估计。可以看出，两种算法都能很好地跟踪参考状态。全状态分析估计的均方误差也展示于图 3中，表明（确定性）pETKF的精度略高，因为与（随机性）pSEnKF相比，它对较小集合（此处为 $M = 20$）的支持更好。

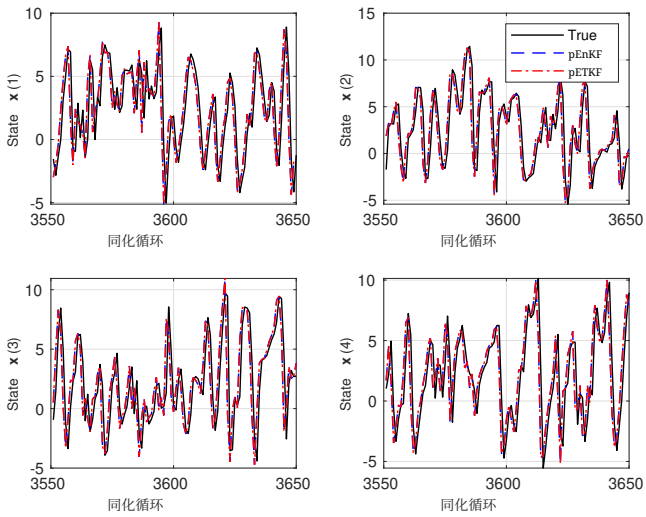


图2：在最后100个同化循环中，使用 pSEnKF 和 pETKF 跟踪前4个状态变量。

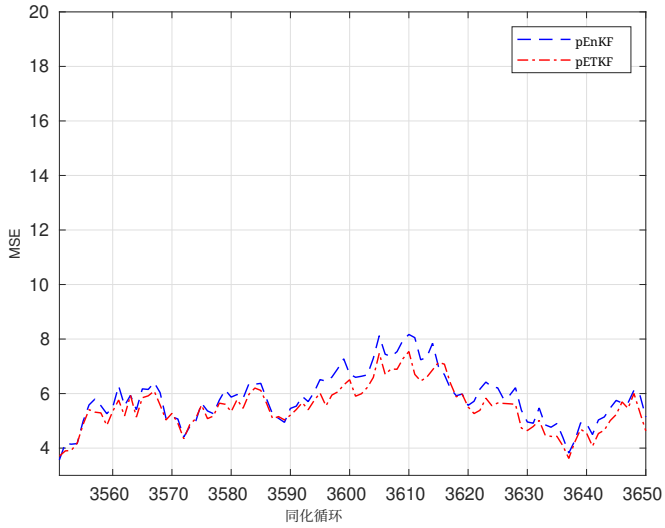


图3：在最后100个同化循环中，pSEnKF 和 pETKF 给出的分析状态估计的均方误差随时间的演变。

第二组实验旨在考察所提出滤波器的行为，并在两种不同的观测情景下，使用经经验调优的膨胀和局地化，将其性能与 SEnKF 和 ETKF 进行比较：完整观测（即观测所有状态变量），以及空间稀疏观测，即同化一半（观测奇数索引变量）或四分之一（每隔三个变量观测一个）状态变量。SEnKF 使用基于Gaspari和Cohn （1999年）(Hamill 等人，2001年)给出的五阶相关函数进行协方差局地化，而 ETKF 使用局地分析方法（Houtekamer和Mitchell，1998年）。在每种情景下，均针对不同的数据噪声水平（信噪比 $= 10$，分别为 15 分贝，对应于 $r = 1.87$，0.6）和不同的集合规模（$M = 10$，20，30）进行一系列敏感性实验。通过计算整个同化时段内参考状态与其分析估计之间偏差的均方根误差（RMSE），评估滤波器的性能。

5.1 全观测情景

当同化所有模型变量的观测时，四种滤波器所得的平均分析均方根误差如图4所示。图版(a)和(b)分别表示信噪比 $= 15$ 分贝和10分贝的情形，图版(c)为相应的比值，即信噪比 $= 10$ 分贝下所得均方根误差相对于信噪比 $= 15$ 分贝情形的比值。图4中pSEnKF和pETKF的均方根误差（下文图5、图6和图8同此）是在使用不同分区大小取值进行试错测试后得到的，即对 $M = 10$ 取 $d_p = 1$、2、5，对 $M = 20$取$d_p = 1$、2、5、10，对 $M = 30$取 $d_p = 1$、2、5、10、20。此处 d_p的“最佳”取值，即与图4中均方根误差相对应的那些取值，列于表1。

从表1可以看出， d_p 达到最低均方根误差的取值会随着集合规模增大而增大， M。根据第4.3节的讨论，这是由于这些滤波器基于串联近似（变分贝叶斯密度近似会随着 d_p增大而改善）与集合近似（会随着 M增大而改善）的结合。这种结合表明，要在增大 M 时获得最低均方根误差，必须同时增大 d_p，但 d_p 仍需小于 M （以避免需要进行局地化）。另一方面，表 1 还表明，对于固定的 M，当状态分区对应于（在测试值中）最大的 d_p 值或其前一个值时，可以达到最低均方根误差。以 pSEnKF 和信噪比 $= 10$ 分贝为例：

当集合包含 $M = 20$ 个成员时，在 $d_p = 10$ （测试值中最大的取值： $\{1, 2, 5, 10\}$）时可达到最低均方根误差；而当集合包含 $M = 30$ 个成员时，在 $d_p = 10$ （测试值中第二大的取值： $\{1, 2, 5, 10, 20\}$）时可达到最低均方根误差。不过，需要注意的是，与 10 相比，增大集合规模可使 $d_p = 20$ 获得更好的结果。对此，使用更大规模的集合（$M = 40$）进行的额外实验显示， $d_p = 10$ 的均方根误差为 $= 0.61628$ ，而 $d_p = 20$ 的均方根误差为 0.60297。

图4 中报告的均方根误差表明，正如预期，所有滤波器在集合较大且观测噪声较小的情况下表现更好。此外，这些滤波器总体上表现出相当的性能，所给出的均方根误差彼此较为接近，下文将进一步讨论。

如图版（a）和（b）所示，当使用较小的集合时（即 $M = 10$ 个成员），确定性集合变换卡尔曼滤波器和pETKF滤波器均优于各自的随机性对应滤波器，而与观测中的噪声水平无关（Hoteit 等人，2015年）。随着集合增大，这种差异变得不那么显著。

如图版 (c) 所示，当观测的信噪比从 SNR $= 15$ dB 降至 SNR $= 10$ dB 时，所有滤波器的误差均约增大一倍。

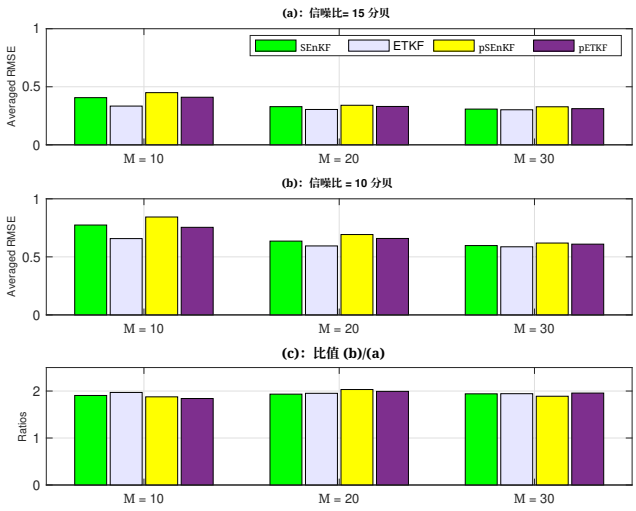


图4：SEnKF、ETKF、pSEnKF 和 pETKF 在不同集合规模下，对信噪比为 $= 15$ 分贝（图(a)）和 10 分贝（图(b)）的观测进行处理后得到的状态分析估计的平均均方根误差。图(c)报告了图(b)中的均方根误差与图(a)中对应均方根误差的比值。

表1： d_p 对应图4中所绘误差的数值。

	信噪比 $= 15$ 分贝			信噪比 $= 10$ 分贝		
	$M = 10$	$M = 20$	$M = 30$	$M = 10$	$M = 20$	$M = 30$
pSEnKF	5	10	20	5	10	10
pETKF	5	10	20	5	10	20

5.2 空间稀疏观测情景

我们先考察同化一半观测的情形，并在图中绘制所得的平均分析 REMS。5。通过试错测试得到对应这些（最小）误差的分区大小 d_p，列于表2。

表 2 中的结果与表 1中的结果接近。值得注意的是， d_p 使均方根误差最低的取值会随着集合规模的增大而增加， M，且对于每个 M，能使均方根误差最低的状态分区对应于测试值中最大的两个取值之一。 d_p 。

图 5 中的均方根误差大于图 4中的误差，这符合预期，因为本实验中同化的观测较少。随着集合规模增大以及观测噪声减小，均方根误差总体上也会降低。

图 5 还表明，这四种滤波器的整体性能相当，所得均方根误差彼此大致接近。具体而言，当使用 $M = 10$ 个成员时，所提出的 pSEnKF 和 pETKF 略逊于其标准版本（在 信噪比 $= 10$ 分贝时尤为明显）。随着集合规模增大，这种差异会减小（$M = 20$ 和 30 个成员）。

总体而言，在四分之一观测情景中得到了类似结果，如图6 和表3所示，但均方根误差对观测中更多噪声不太敏感，这一点除外。

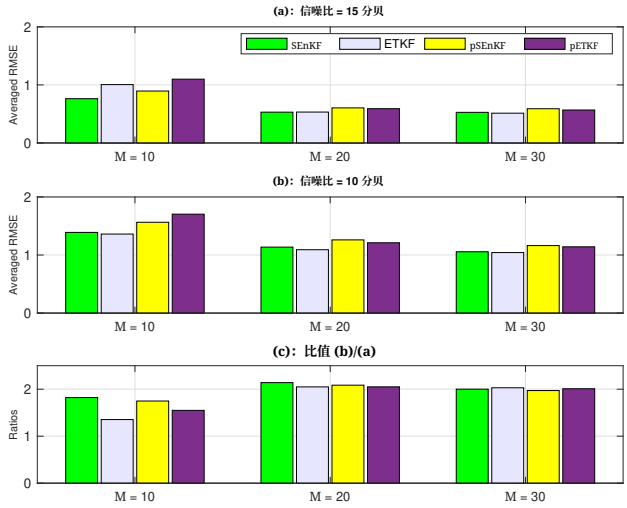


图5：与图 4相同，但仅观测奇数索引状态变量。

表2：与图 d_p 5中绘制的误差相对应的值。。

	信噪比 $= 15$ 分贝			信噪比 $= 10$ 分贝		
	$M = 10$	$M = 20$	$M = 30$	$M = 10$	$M = 20$	$M = 30$
pSEnKF	5	10	10	5	10	10
pETKF	2	5	20	2	10	20

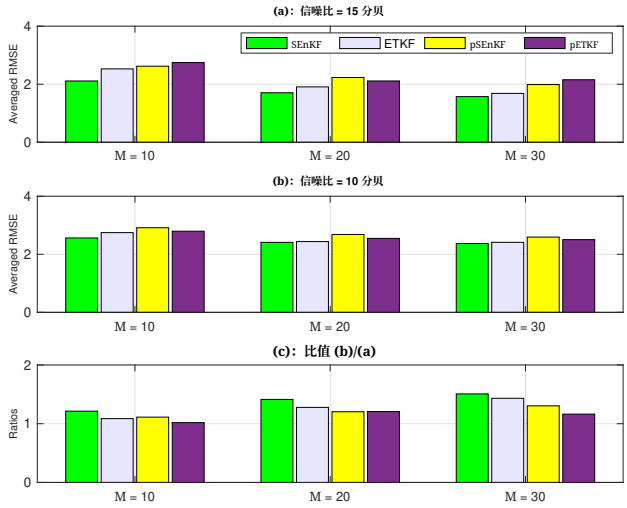


图6：与图4相同，但每隔四个状态变量进行一次观测。

表3：与图 d_p 中绘制的误差相对应的数值。6。

	信噪比 = 15 分贝			信噪比 = 10 分贝		
	$M = 10$	$M = 20$	$M = 30$	$M = 10$	$M = 20$	$M = 30$
pSEnKF	5	10	10	5	5	10
pETKF	2	5	10	5	5	10

我们进一步通过将分析集合离散度与（时间序列）分析均方根误差进行比较，评估所提出分区方法的统计一致性。为此，我们使用pSEnKF，取 $d_p = 10$ （即四个分区）和 $M = 30$ ，配置为“信噪比 $=$ 15分贝，半观测情景”（图5中的图(a)）。目标是确定分区更新是否会导致欠离散或过离散集合，并识别分区边界附近不确定性中潜在的伪影/不平衡。为此，我们根据观测状况（已观测与未观测）以及与分区边界的接近程度，将状态索引/元素划分为四个子集。若分量位于边界两个网格点以内，则视为“近”，其余（内部）分量则归类为“远”。

这四个子集分别对应： {已观测且靠近}、 {未观测且靠近}、 {已观测且远离}，以及 {未观测且远离}。对于每个同化循环，均使用 pSEnKF 分别计算这些子集的分析集合离散度和均方根误差。

结果见图 7。如预期，未观测变量（图版 b 和 d）相比已观测变量（图版 a 和 c）表现出更大的误差和集合离散度。集合离散度与均方根误差（RMSE）也明显一致，在所有子集中都紧密跟随其变化。特别是，即使对于未观测变量，滤波器在分区边界处（图版 a 和 b）也没有出现欠离散。换言之，分区并未在集合离散度和误差中引入明显的边界伪影。这表明，分析步中的迭代调整与预报步骤中模型动力学所提供的混合作用相结合，能够有效地在各分区间传播信息，并在全局范围内保持连贯的集合离散度，从而避免分区边界处出现人为的不连续。

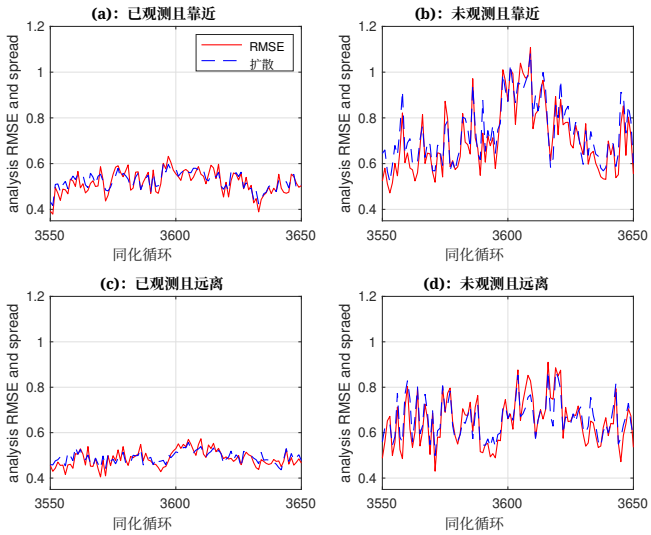


图 7：在信噪比为 100 分贝的半观测情景下，pSEnKF 在最近 $M = 30$ 个集合成员和 $d_p = 10$ 条件下给出的分析均方根误差与集合离散度的时间演变 $= 15$ （图 5 图(a)中的一个情景）。结果基于状态变量的可观测性及其与分区边界的距离，将其划分为四个子集，并分别取平均。

5.3 更具挑战性的情景

在迄今展示的情形中，信噪比 $= 10$ 分贝， $M = 10$ 和 $\delta_o = 4\delta_m$ 的四分之一观测情景（即图 6）最具挑战性（下文将此情形称为情形 0）。在此，我们考察滤波器在更具挑战性的情景下的稳健性：通过提高观测的时间稀疏性，并在滤波算法中使用错误模型，进一步加剧系统误差。以下列出四种情景，并按复杂程度排序：

- 情形1：情形 0，将 $4\delta_m$替换为 $\delta_o = 8\delta_m$ ；
- 情形2：情形1，在滤波过程中，状态模型中引入了偏差 $0.3F$ ；
- 情形3：情形2，在滤波过程中，观测模型中引入了 0.3Hx$_0$ 的偏差；
- 情形4：情形3，在滤波过程中将（真实）观测噪声协方差 R 减小 30%（即在同化时使用 0.7R 作为指定的观测噪声协方差，而 R 是根据真实状态生成观测时使用的真实观测噪声协方差）。

图 8 绘出了所得的平均分析均方根误差。如比较情形1与情形 0 所示，所有滤波器都明显受到观测频率降低的影响。在情形1中，所提出的滤波器仍逊于标准 SEnKF，与情形 0 的情况类似，但差距较小。然而，与 ETKF 相比，pSEnKF 的表现更加接近，而 pETKF 的精度甚至略有提高。

将情形2与情形1进行比较后发现，状态模型中的偏差对标准滤波器的影响略大于对所提出滤波器的影响。这可能反映了由于使用了有误差的模型，标准滤波器中的预报背景场退化程度略高。还可以注意到，在情形2中，pSEnKF 和 pETKF 的表现更接近 SEnKF，且明显优于 ETKF。

当观测模型中也引入偏差时（情形3），所有滤波器都受到显著影响，从情形2转至情形3时，其均方根误差显著增大。在情形3中，还发现这两种所提出的滤波器显著优于 ETKF，并略优于 SEnKF。

从情形3和情形4可以看出，所提出的 pSEnKF 和 pETKF 滤波器，以及表现稍逊的标准滤波器，似乎都能稳健应对对错误观测的置信度提高（将 R 缩减 30% 的影响或多或少可以忽略不计）。此外，从情形3转到情形4时，标准滤波器与所提出滤波器的均方根误差之间的差距有所扩大。所提出的滤波器在这类挑战性情形下表现更优，可能源于其设计：它们使用与标准滤波器相同的观测信息，但将其应用于更小的状态分区，并进行迭代更新。这种局地化的迭代更新方案降低了参数误差的影响，并增强了在挑战性条件下的稳定性。

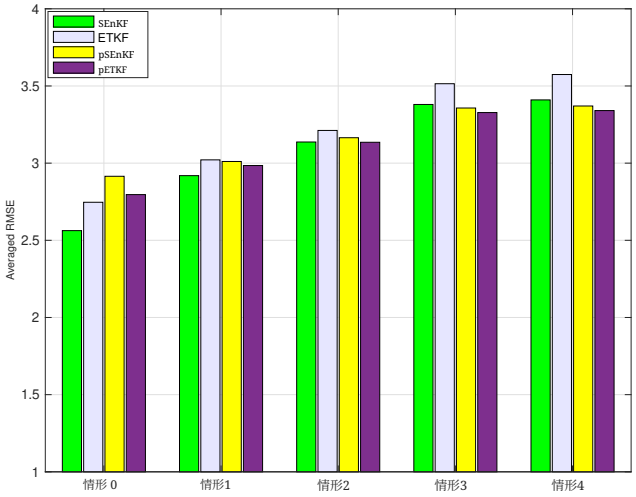


图8：在不同的挑战性情形下，SEnKF、ETKF、pSEnKF 和 pETKF 得出的状态分析估计的平均均方根误差。

此外，值得通过与基准进行比较来评估同化的有效性，例如气候变率水平 σ_{clim}（可用于了解真实状态（自然运行）的自然波动有多大），以及观测的时间序列均方根误差 rmse_{obs}。它们分别定义如下：

$$\begin{aligned} \sigma_{\text{clim}} &= \frac{1}{\sqrt{d_x}}\|\boldsymbol{\delta}\|, \\ \text{rmse}_{\text{obs}}(n) &= \frac{1}{\sqrt{T}}\|\mathbf{y}_n - \mathbf{H}_n\mathbf{x}_n\|, \end{aligned}$$

其中，大小为 $(d_x \times 1)$ 的向量 $\boldsymbol{\delta}$ 中第 i 个元素定义为 $\delta^i = \sqrt{\frac{1}{T}\sum_{n=0}^{T-1}(x_n^i - \overline{x}^i)^2}$，$\mathbf{x}_n$ 和 $\mathbf{y}_n$，分别表示同化时刻 n 的参考状态和观测；$\overline{x}^i$ 表示 $\{x_n^i\}_n$ 的（时间）平均值。

图 9绘出了这四种滤波器在案例 0和 4中最后 100 个同化循环期间的这些量以及状态分析估计的时间 均方根误差。根据图(a)中的结果，尽管在这种具有挑战性的情景下，这些滤波器（包括标准滤波器）似乎都无法很好地拟合数据（因为其分析 均方根误差误差超过 均方根误差$_{\text{obs}}$），但相对于气候平均值而言，它们仍明显有用（因为其分析 均方根误差误差

明显小于 σ_{clim}）。在图(a)中，还可以看出图 8所示的分析误差排序，即总体而言，SEnKF 优于 ETKF，其次是 pETKF 和 pSEnKF。

当滤波器在更长的同化循环和扰动参数下进一步运行时（图版 (b)），与图版 (a) 中的结果相比，分析误差有所增加，但总体上仍低于 σ_{clim}，这表明即使在如此具有挑战性的情景下，同化结果仍然具有意义。

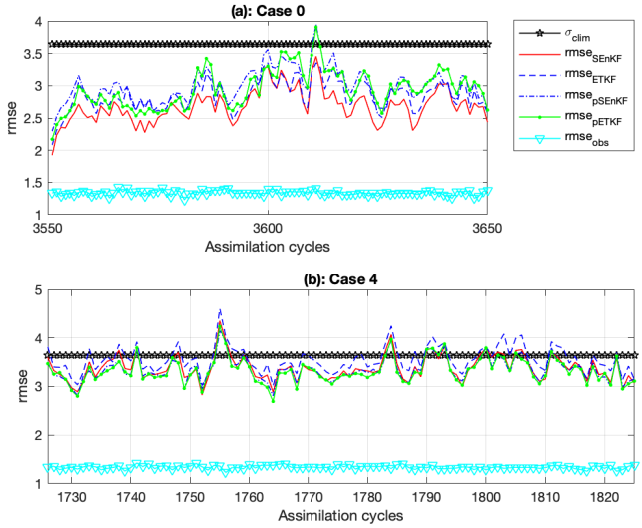


图 9：在最后 100 个同化循环内，四种滤波器在案例 0 和 4 中提供的观测与状态分析估计的气候变率水平和时间序列均方根误差。

5.4 非均匀状态分区案例

如前文所述，只需在算法 1 和 2 d_p 中将 K 分区 $\mathbf{x}_n^k$ 具有不同大小 d_{p^k} 替换为 d_{p^k}，并将 Kd_p 替换为 $\sum_{k=1}^K d_{p^k}$，即可将所提出的方案直接调整到各分区大小不同的情形。这里考虑两种配置：

- 配置 1：状态被划分为 $K=3$ 个维度为 $[d_{p^1}, d_{p^2}, d_{p^3}]$ 的分区，其维度可以等于 $\mathcal{D}_3^1=[15,13,12]$、$\mathcal{D}_3^2=[17,11,12]$、$\mathcal{D}_3^3=[16,15,9]$ 或 $\mathcal{D}_3^4=[10,20,10]$。
- 配置 2：状态被划分为 $K=4$个分区，维度为 $[d_{p^1}, d_{p^2}, d_{p^3}, d_{p^4}]$，可以等于 $\mathcal{D}_4^1=[9,2,15,14]$、$\mathcal{D}_4^2=[6,15,9,10]$、$\mathcal{D}_4^3=[10,10,10,10]$ 或 $\mathcal{D}_4^4=[10,8,13,9]$。

对于每种情况，使用 $M=30$ 个成员运行所提出的滤波器，采用非均匀的半观测场景，其中变量 $\{y_n^i; i=1$、4、5、8、10、12、13、14、17、20、24、25、27、28、31、32、34、35、37、$40\}$ 被观测，信噪比为 $=15$ 分贝，且 $\delta_o=4\delta_m$。标准滤波器也使用经过“最优”调优的膨胀和局地化运行。

图 10 绘出了采用所提出的滤波器与标准滤波器得到的平均状态分析均方根误差之间的差异百分比，定义如下：

$$\frac{\text{RMSE}_{\text{proposed filter}}-\text{RMSE}_{\text{standard filter}}}{\text{RMSE}_{\text{standard filter}}}\times 100.$$

总体而言，两种配置下所提出的滤波器性能均与标准滤波器相近。

从图(a)可以看出，对于所有三部分状态分区，pETKF 的均方根误差均不超过两种标准滤波器所得到均方根误差的 5.2%。它甚至

在采用 $\mathcal{D}_3^1$ 或 $\mathcal{D}_3^2$ 时，表现得与 ETKF 一样好。pSEnKF 的表现不如 pETKF；采用 $\mathcal{D}_3^1$ 或 $\mathcal{D}_3^2$ 时表现最佳，此时其均方根误差分别比 ETKF 高 3.79% 和 3.9%。

在四部分分区配置（子图 (b)）中，百分比误差与图(a)一样不超过 10%，但总体上高于图(a)中的误差。这表明，在此观测情景下，将状态划分为三个分区而非四个通常更有利。

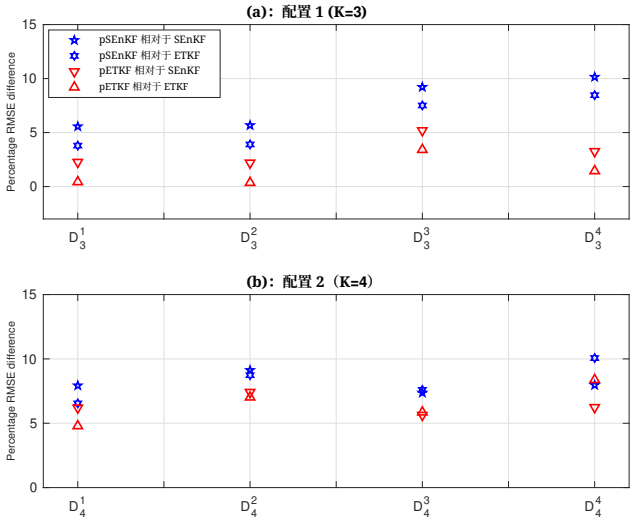


图 10：标准 SEnKF 和 ETKF 所提供的平均状态分析均方根误差与所提出的 pSEnKF 和 pETKF 所提供的平均状态分析均方根误差之间的差异（%）；后者按照配置 1（图(a)）和配置 2（子图 (b)）对状态进行非均匀划分，均使用$M = 30$ 个成员，处于信噪比 $= 15$ 分贝且 $\delta_o = 4\delta_m$ 的非均匀半观测情景中。

5.5 通过同化单个观测进行说明

为说明所提出的方法如何处理（采样和系统性）噪声，开展了单观测实验，仅观测第20个状态变量，信噪比为信噪比 $= 15$分贝。图11 绘制了最近100 个同化循环中分析增量（分析均值减去预报均值）的时间演变，该结果由 pSEnKF 使用 $M = 30$个成员、 $\delta_o = 4\delta_m$ 并将状态向量划分为大小为 $d_p = 8$ 的分区给出（pETKF 得到了类似结果，但为限制论文篇幅而未在此列出）。由图11可见，分析增量局地化于观测位置附近，表明局地化按预期发挥作用。

图 11 还表明，在分析循环期间，来自观测的信息保留在其所在的状态分区中（即分区 $\mathbf{x}_n^3 = (x_n^{17}, \cdots, x_n^{24})$），且未传播到其他分区。这是由于 i) 移除了预报交叉协方差（即令 $\mathbf{G}_n^k = \mathbf{0}$），再加上 ii) 单观测情形下 $\mathbf{H}\,(= [0, \cdots, 0, 1, 0, \cdots, 0])$ 的特定结构。事实上，由于 $\mathbf{G}_n^k = \mathbf{0}$，状态分区的更新采用分析方程（31），而非方程（27）；并且由于 $\mathbf{H}$ 的稀疏结构，与除包含该观测的分区（即 $\mathbf{x}_n^3$）之外的所有分区相关联的方程（31）中的卡尔曼增益 $\mathbf{L}_n^k$ 均变为零。因此，不包含该观测的分区中的分析估计与预报估计相同（即没有分析增量）。

出于类似的原因，采用经过调优的协方差局地化（具有一个“最优”长度尺度 $\ell_s = 2$）的标准 SEnKF 也被发现表现出类似的行为（参见图 12）。更确

切地说，较强的局地化（$\ell_s = 2$）极大地限制了预报交叉协方差；再加上 H 的特定结构，观测信息的传播也仅限于邻近区域。从这一点来看，这两个滤波器的均方根误差相近：pSEnKF 为 3.6315，SEnKF 为 3.6245。

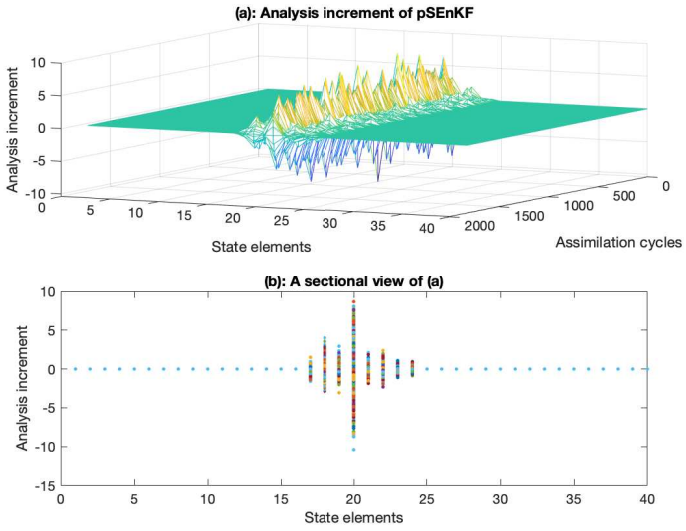


图11：a）pSEnKF 使用 $M = 30$ 个成员和“最优”膨胀（因子为 $\alpha = 1.1$），在最近 100 个同化循环中给出的所有状态元素分析增量的时间演变。（b）是（a）的局部视图。仅观测并同化第 20 个状态元素。

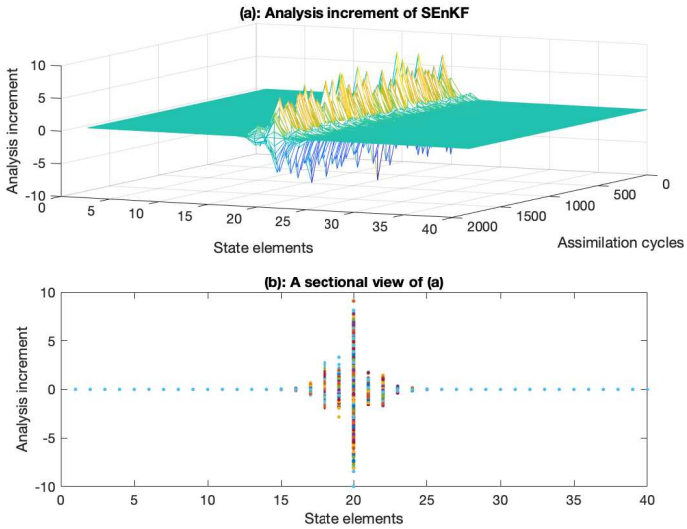


图12：a）SEnKF 使用 $M = 30$ 个成员、“最优”膨胀（因子为 $\alpha = 1.1$）和局地化（长度尺度为 $\ell_s = 2$），在最近 100 个同化循环中给出的所有状态元素分析增量的时间演变。（b）是（a）的局部视图。仅观测并同化第 20 个状态元素。

最后，为了考察在此单观测情景下增大分区大小的影响，我们采用 $d_p = 20$ 替代8，重复了图11 中的实验。此处未展示的结果表明，增量仍局限于包含观测分量的分区内（不存在

跨分区长程效应），而分区内空间范围自然扩大。与 $d_p = 8$ 情形（图11）相比，均方根误差略有增大；但随着 M增大，均方根误差会降低（均方根误差 $= 3.6383$ 和3.6287 分别对应 $M = 40$ 和45），这与 d_p 相对于 M变大时受到抽样限制的情况一致。

总体而言，这证实了在集合规模足够大的情况下，较大的分区最有益；同时也表明，d_p 能够有效控制离散的（逐分区）局地化尺度，但仍受常规集合规模采样约束的限制：可靠的分区内协方差估计要求 M 相对于 d_p 足够大（实际中为 $M > d_p$，理想情况下为 $M \gg d_p$）。

6 结论

本文提出了一种集合卡尔曼滤波（EnKF）方法，该方法本身具有局地化特性，无需任何辅助局地化技术。其原理是先对连续的分析概率密度函数（pdf）进行局地化，再在局地化后的概率密度函数上应用类高斯集合卡尔曼采样过程。分析概率密度函数的局地化通过变分贝叶斯优化实现：将其近似为一组相互独立的边缘概率密度函数的乘积，这些边缘概率密度函数分别对应状态向量的小分区。随后，按照随机集合卡尔曼滤波（SEnKF）和确定性集合变换卡尔曼滤波（ETKF）过程对这些边缘分布进行采样，但也可以采用其他采样过程。所提出方法的理论依据是：选择维数小于集合规模的分区，应能在不引入任何外部局地化方案的情况下提升性能，同时减轻欠采样和系统误差的影响。

从算法角度来看，这些滤波器采用与标准 SEnKF 和 ETKF 相同的预报步骤，但分析步骤有所不同：它们通过迭代调整各个分区的标准 SEnKF 和 ETKF 更新，并加上其他分区集合均值的线性组合偏移量。我们在各种具有挑战性的设置和情景下，使用 Lorenz-96模型开展了数值实验。结果表明，所提出的滤波器适用于地球物理应用；无论在计算负担还是估计精度方面，其整体表现均可与已调优局地化的 SEnKF 和 ETKF 相媲美，有时甚至更胜一筹。

未来工作将考虑通过在方程（24）-（25）中引入计算上合理的增益项近似，并通过采用例如（Ranganath 等人，2016）的方法，减轻状态分区之间被强加的后验独立性，从而提升新滤波器的性能。我们还将考虑引入膨胀因子的在线估计，例如（Ait-El-Fquih 和 Hoteit，2020）中的方法，旨在构建更加自主的集合滤波方案，避免对该因子进行反复试错调优。未来工作的其他重要方向包括：将所提出的内在局地化方法应用于其他确定性集合滤波器，例如 SEIK（Hoteit 等人，2002年）、LETKF（Hunt 等人，2007年），并将所有新滤波器应用于现实的海洋和大气数据同化问题。

附录

通用局部化滤波器的分析步

我们展示如何从21推导出22–28。高斯表达式3表示 $p(\mathbf{y}_n|\mathbf{x}_n)$，因此有：

$$p(\mathbf{y}_n|\mathbf{x}_n) \propto \exp\left(-\frac{1}{2}\|\mathbf{y}_n - \mathbf{H}_n^{k-}\mathbf{x}_n^{k-} - \mathbf{H}_n^{k}\mathbf{x}_n^{k}\|^2_{\mathbf{R}_n^{-1}}\right). \tag{41}$$

对ln 函数作用于41后，再对 $\pi_n^{(i,i-1)}(\mathbf{x}_n^{k^-}) = \prod_{j=1}^{k-1}\pi_n^i(\mathbf{x}_n^j)\times\prod_{j=k+1}^{K}\pi_n^{i-1}(\mathbf{x}_n^j)$求期望，然后应用 exp 函数，可得：

$$\begin{aligned}\mathcal{L}^i_{\mathbf{y}_n}(\mathbf{x}_n^k) &\propto \exp\left(\mathbb{E}_{\pi_n^{(i,i-1)}(\mathbf{x}_n^{k^-})}[\ln p(\mathbf{y}_n|\mathbf{x}_n)]\right),\\ &\propto \exp\left(-\frac{1}{2}\|\mathbf{y}_n-\mathbf{H}_n^{k^-}\hat{\mathbf{m}}_{n|n}^{k-,i}-\mathbf{H}_n^k\mathbf{x}_n^k\|^2_{\mathbf{R}_n^{-1}}\right),\\ &= (22).\end{aligned}$$

我们现在推导 $\pi^i_{n-1}(\mathbf{x}_n^k)$(23)-(25)。首先证明，

$$\begin{aligned}\pi^i_{n-1}(\mathbf{x}_n^k) &\propto \exp\left(\mathbb{E}_{\pi_n^{(i,i-1)}(\mathbf{x}_n^{k^-})}[\ln p_{n-1}(\mathbf{x}_n)]\right),\\ &\propto \exp\left(\mathbb{E}_{\pi_n^{(i,i-1)}(\mathbf{x}_n^{k^-})}[\ln p(\mathbf{x}_n^k|\mathbf{x}_n^{k^-},\mathcal{Y}_{n-1})]\right).\end{aligned} \tag{42}$$

另一方面，关于 $p_{n-1}(\mathbf{x}_n)$ 的高斯假设（即 $\mathcal{N}_{\mathbf{x}_n}(\hat{\mathbf{x}}_{n|n-1},\mathbf{P}_{n|n-1})$ 意味着条件分布 $p(\mathbf{x}_n^k|\mathbf{x}_n^{k^-},\mathcal{Y}_{n-1})$ 也是高斯分布，其均值为 $\boldsymbol{\mu}_n(\mathbf{x}_n^{k^-}) = \hat{\mathbf{x}}^k_{n|n-1}+\mathbf{G}_n^k(\mathbf{x}_n^{k^-}-\hat{\mathbf{x}}^{k^-}_{n|n-1})$，协方差为 $\boldsymbol{\Sigma}_n = \mathbf{P}^k_{n|n-1}-\mathbf{G}_n^k\mathbf{P}^{k^-\,k}_{n|n-1}$，其中 $\mathbf{G}_n^k = \mathbf{P}^{k\,k^-}_{n|n-1}(\mathbf{P}^{k^-}_{n|n-1})^{-1}$、$\mathbf{P}^k_{n|n-1}$、$\mathbf{P}^{k^-}_{n|n-1}$、$\mathbf{P}^{k\,k^-}_{n|n-1}$ 和 $\mathbf{P}^{k^-\,k}_{n|n-1}$ 是根据位置 k 和 k^- 从 $\mathbf{P}_{n|n-1}$ 中提取的（例如，Ait-El-Fquih 和 Desbouvries，2011，附录，命题 9）。在集合框架中，假定规模较大的 $(d_x-d_p)\times(d_x-d_p)$ 维协方差 $\mathbf{P}^{k^-}_{n|n-1}$ 是根据少量预报成员 M 近似得到的，这意味着它们是奇异的，因而不可逆。在这种情况下，$\mathbf{G}_n^k$ 表达式中的逆矩阵 $(\mathbf{P}^{k^-}_{n|n-1})^{-1}$ 被伪逆 $(\mathbf{P}^{k^-}_{n|n-1})^+$ 取代（例如，Anderson 和 Moore，1979，第 2.3 节）。

现在，按照与 $\mathcal{L}^i_{\mathbf{y}_n}(\mathbf{x}_n^k)$上文类似的推导，我们对ln 函数$p(\mathbf{x}_n^k|\mathbf{x}_n^{k^-},\mathcal{Y}_{n-1}) = \mathcal{N}_{\mathbf{x}_n^k}(\boldsymbol{\mu}_n,\boldsymbol{\Sigma}_n)$取关于 $\pi_n^{(i,i-1)}(\mathbf{x}_n^{k^-})$的期望，然后应用exp函数，即可直接得到（23）—（25）。

仍需推导 $\pi^i_n(\mathbf{x}_n^k)$（26）—（28）。由于 $\mathbf{x}_n^k$ 的 VB 概率密度函数 $\pi^i_{n-1}(\mathbf{x}_n^k)$（23）是高斯分布，而在给定 $\mathbf{x}_n^k$ 的条件下，$\mathbf{y}_n$ 的 VB 概率密度函数 $\mathcal{L}^i_{\mathbf{y}_n}(\mathbf{x}_n^k)$（22）也是高斯分布，且其均值关于 $\mathbf{x}_n^k$ 呈线性关系，因此相应的联合概率密度 $\kappa^i(\mathbf{x}_n^k,\mathbf{y}_n) = \pi^i_{n-1}(\mathbf{x}_n^k)\mathcal{L}^i_{\mathbf{y}_n}(\mathbf{x}_n^k)$ 也是高斯分布，其参数为（例如，Ait-El-Fquih 和 Desbouvries，2011，附录，命题 8）：

$$\kappa^i(\mathbf{x}_n^k,\mathbf{y}_n)=\mathcal{N}_{(\mathbf{x}_n^k,\mathbf{y}_n)}\left(\begin{bmatrix}\mathbf{m}^{k,i}_{n|n-1}\\ \mathbf{H}_n^k\mathbf{m}^{k,i}_{n|n-1}+\mathbf{H}_n^{k^-}\mathbf{m}^{k-,i}_{n|n}\end{bmatrix},\begin{bmatrix}\mathbf{C}^k_{n|n-1} & \mathbf{C}^k_{n|n-1}(\mathbf{H}_n^k)^T\\ \mathbf{H}_n^k\mathbf{C}^k_{n|n-1} & \mathbf{H}_n^k\mathbf{C}^k_{n|n-1}(\mathbf{H}_n^k)^T+\mathbf{R}_n\end{bmatrix}\right) \tag{43}$$

所关注的 VB 概率密度函数 $\pi^i_n(\mathbf{x}_n^k) = \kappa^i(\mathbf{x}_n^k|\mathbf{y}_n)$ 随后通过对公式（43）进行条件化，并使用（Ait-El-Fquih 和 Desbouvries，2011，附录）中的命题 9获得。这样即可立即得到（26）-（28）。

迭代过程中 NFE 的实际计算

根据公式（13），同化时刻 t_n 和迭代 i 处的 NFE 定义为：

$$\begin{aligned}\mathrm{NFE}^i &= \mathbb{E}_{\pi^i_n(\mathbf{x}_n)}[\ln p(\mathbf{x}_n,\mathbf{y}_n|\mathcal{Y}_{n-1})] - \overbrace{\mathbb{E}_{\pi^i_n(\mathbf{x}_n)}[\ln \pi^i_n(\mathbf{x}_n)]}^{\phi^i_3}\\ &= \underbrace{\mathbb{E}_{\pi^i_n(\mathbf{x}_n)}[\ln p(\mathbf{y}_n|\mathbf{x}_n)]}_{\phi^i_1}+\underbrace{\mathbb{E}_{\pi^i_n(\mathbf{x}_n)}[\ln p_{n-1}(\mathbf{x}_n)]}_{\phi^i_2}-\phi^i_3.\end{aligned} \tag{44}$$

将高斯分布 $p(\mathbf{y}_n|\mathbf{x}_n)=\mathcal{N}_{\mathbf{y}_n}(\mathbf{H}_n\mathbf{x}_n,\mathbf{R}_n)$、$p_{n-1}(\mathbf{x}_n)\approx\mathcal{N}_{\mathbf{x}_n}(\hat{\mathbf{f}}_n,\mathbf{P}_{\mathbf{f}_n})$ 和 $\pi^i_n(\mathbf{x}_n)\approx\mathcal{N}_{\mathbf{x}_n}(\hat{\mathbf{a}}^i_n,\mathbf{P}_{\mathbf{a}_n})$ 代入上述 ϕ^i 项，其中 $\mathbf{P}_{\mathbf{f}_n}$ 和 $\mathbf{P}_{\mathbf{a}_n}$ 是基于

集合扰动矩阵 $\{\mathrm{S}_{\mathbf{f}_n^k}\}_k$ 和 $\{\mathrm{S}_{\mathbf{a}_n^k}\}_k$ 构造的，可得：

$$\begin{cases} \phi_1^i & \approx & -\frac{1}{2}\left[\ln((2\pi)^{d_y}|\mathbf{R}_n|) + \mathrm{Tr}(\mathbf{P}_{\mathbf{a}_n}\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n) + ||\mathbf{y}_n - \mathbf{H}_n\hat{\mathbf{a}}_n^i||^2_{\mathbf{R}_n^{-1}}\right], \\ \phi_2^i & \approx & -\frac{1}{2}\left[\ln((2\pi)^{d_x}|\mathbf{P}_{\mathbf{f}_n}|) + \mathrm{Tr}(\mathbf{P}_{\mathbf{a}_n}\mathbf{P}_{\mathbf{f}_n}^{-1}) + ||\hat{\mathbf{a}}_n^i - \hat{\mathbf{f}}_n||^2_{\mathbf{P}_{\mathbf{f}_n}^{-1}}\right], \\ \phi_3^i & \approx & -\frac{1}{2}\left[\ln((2\pi)^{d_x}|\mathbf{P}_{\mathbf{a}_n}|) + d_x\right]. \end{cases} \tag{45}$$

公式（45）代入公式（44，得到NFE^{i}的基于集合的近似表达式。

致谢

作者感谢匿名审稿人提出宝贵且具有建设性的意见。

数据可用性声明

不适用数据共享——未生成新数据。

作者贡献

Boujemaa Ait-El-Fquih：概念化；调查；撰写初稿；方法论；撰写——审阅与编辑。

Ibrahim Hoteit：概念化；经费获取；资源；撰写——审阅与编辑。

开放研究者与贡献者身份标识

Boujemaa Ait-El-Fquih https://orcid.org/0000-0003-1119-7585

Ibrahim Hoteit https://orcid.org/0000-0002-3751-4393

参考文献

M. Ades 和 P. van Leeuwen. 等效权重粒子滤波器的探索。《英国皇家气象学会季刊》, 139:820–40, 2013. B. Ait-El-Fquih 和 F. Desbouvries. 三元马尔可夫链中的卡尔曼滤波。*IEEE*信号处理汇刊, 54(8):2957–63, 2006年8月. B. Ait-El-Fquih 和 F. Desbouvries. 奇异状态空间系统中的固定区间卡尔曼平滑算法。信号处理系统杂志, 65(3):469–78, 2011年12月. B. Ait-El-Fquih 和 I. Hoteit. 大维线性高斯状态空间模型中的快速类卡尔曼滤波。*IEEE*信号处理汇刊, 63(21):5853–67, 2015年. B. Ait-El-Fquih 和 I. Hoteit. 面向大维系统的变分贝叶斯多粒子滤波方案。*IEEE*信号处理汇刊, 64(20):5409–22, 2016. B. Ait-El-Fquih 和 I. Hoteit. 一种基于粒子滤波器的集合卡尔曼滤波器自适应膨胀方案。《英国皇家气象学会季刊》, 147:922–37, 2020. B. Ait-El-Fquih 和 I. Hoteit. 大气、海洋与水文应用的数据同化, 第4卷, 第3章：采用一步超前平滑实现高效数据同化的滤波。施普林格, 2022a. doi: 10.1007/978-3-030-77722-7_3.

B. Ait-El-Fquih 和 I. Hoteit. 面向大维系统快速卡尔曼滤波的并行与循环迭代变分贝叶斯。*IEEE*信号处理汇刊, 70:5871–84, 2022b. B. Ait-El-Fquih, J.-F. Giovannelli, N. Paul, A. Girard, 和 I. Hoteit. 地下水层点状污染源的参数贝叶斯估计。信号处理, 168:107339, 2020. B. Ait-El-Fquih, Aneesh C. Subramanian, 和 I. Hoteit. 随机参数化系统集合滤波的变分贝叶斯方法。《英国皇家气象学会季刊》, 149:1769–88, 2023. B. D. O. Anderson 和 J. B. Moore. 最优滤波. Prentice Hall, 恩格尔伍德克利夫斯, 新泽西州, 1979. J. Anderson, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, 和 A. Avellano. 数据同化研究试验台：一项社区设施。美国气象学会通报, 90:1283–96, 2009. J. L. Anderson. 用于数据同化的集合调整卡尔曼滤波器。《月刊天气评论》, 129:2884–2903, 2001. J. L. Anderson. 使用分层集合滤波器探究集合数据同化中局地化的必要性。《物理学*D*》, 230:99–111, 2007a. J. L. Anderson. 集合滤波器的自适应协方差膨胀误差校正算法。《泰勒斯》, 59A:210–24, 2007b. J. L. Anderson 和 S. L. Anderson. 非线性滤波问题的蒙特卡罗实现，用于生成集合资料同化结果和预报。《月刊天气评论》, 127(12): 2741–2758, 1999年. D.P. Bertsekas. 非线性规划. Athena Scientific, 贝尔蒙特, 1995. A. Beskos, D. Crisan, 和 A. Jasra. 关于高维序贯蒙特卡罗方法的稳定性。《应用概率年刊》, 24(4):1396–1445, 2014年. C. H. Bishop 和 D. Hodyss. 对虚假集合相关性的流适应性调节及其在基于集合的数据同化中的应用。《英国皇家气象学会季刊》, 133:2029–44, 2007. C.H. Bishop, B.J. Etherton, 和 S.J. Majumdar. 使用集合变换卡尔曼滤波器进行自适应采样。第一部分：理论方面。《月刊天气评论》, 129:420–36, 2001. D. M. Blei, A. Kucukelbir, 和 J. D. McAuliffe. 变分推断：统计学家综述。美国统计学会杂志, 112:859 – 77, 2017. M. Bocquet. 局地化与迭代集合卡尔曼平滑器。《英国皇家气象学会季刊》, 142:1075–89, 2016. G. Burgers, P. Jan van Leeuwen, 和 G. Evensen. 集合卡尔曼滤波器中的分析方案。《月刊天气评论》, 126:1719–24, 1998年. M. A. Chappell, A. R. Groves, B. Whitcher, 和 M. W. Woolrich. 非线性前向模型的变分贝叶斯推断。*IEEE*信号处理汇刊, 57(1):223–36, 2009年1月. D. Crisan 和 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(3):736–46, 2002.

M. De La Chevrotière 和 J. Harlim. 一种改进串行集合卡尔曼滤波器相关性估计的数据驱动方法. 《月刊天气评论》, 145:985–1001, 2017年. A. Doucet, S.J. Godsill, and C. Andrieu. 关于贝叶斯滤波的序贯蒙特卡洛采样方法. 《统计与计算》, 10:197–208, 2000年. C. A. Edwards, A. M. Moore, I. Hoteit, and B. D. Cornuelle. 区域海洋数据同化.《海洋科学年度评论》, 7:21–42, 2015年. G. Evensen. 使用蒙特卡洛方法和非线性准地转模型进行序贯数据同化，以预报误差统计. 《地球物理研究杂志》, 99(C5):143–62, 1994年. G. Evensen. EnKF 的采样策略与平方根分析方案. 《*Ocean Dy- namics*》, 54:539–60, 2004年. G. Evensen. 《数据同化：集合卡尔曼滤波器》. 纽约：施普林格, 2006, 版次 1. G. Evensen. 《数据同化：集合卡尔曼滤波器》. 纽约：施普林格, 第 2 版, 2009年. E. J. Fertig, B. R. Hunt, E. Ott, and I. Szunyogh. 使用局地集合卡尔曼滤波器同化非局地观测. 《*Tellus A*》, 59:719–30, 2007年. A. Fraysse and T. Rodet. 一种用于大维问题的测度论变分贝叶斯算法. 《*SIAM*影像科学杂志》, 7(4):2591–2622, 2014年. R. Furrer and T. Bengtsson. 卡尔曼滤波器变体中高维先验与后验协方差矩阵的估计. 《多元分析杂志》, 98(2):227–55, 2007年. G. Gaspari and S. E. Cohn. 构造二维和三维相关函数.《英国皇家气象学会季刊》, 125(554):723–757, 1999年. ISSN 1477- 870X. N. J. Gordon, D. J. Salmond, and A. F. M. Smith. 非线性/非高斯贝叶斯状态估计的新方法. 《*IEE*学报-*F*辑》, 140:107–113, 1993年. T. M. Hamill, J. S. Whitaker, and C. Snyder. 在集合卡尔曼滤波器中对背景误差协方差估计进行距离依赖滤波. 《月刊天气评论》, 129:2776–90, 2001年. I. Hoteit, D.T. Pham, and J. Blum. 一种简化的降阶卡尔曼滤波及其在热带太平洋高度计数据同化中的应用. 《海洋系统杂志》, 36(1–2):101– 127, 2002年. I. Hoteit, D-T. Pham, G. Triantafyllou, and G. Korres. 一种用于气象学和海洋学数据同化的最优非线性滤波器的新近似解. 《月刊天气评论》, 136(1):317–334, 2008年. I. Hoteit, D.-T. Pham, M. E. Gharamti, and X. Luo. 减轻随机集合卡尔曼滤波中的观测扰动抽样误差. 《月刊天气评论》, 143:2918–36, 2015年. I. Hoteit, X. Luo, M. Bocquet, A. Kohl, and B. Ait-El-Fquih. 《业务海洋学新前沿》, 第17章：海洋学中的数据同化：现状与新方向, 页码 465–512. GODAE Ocean View, 2018年. P. Houtekamer and F. Zhang. 集合卡尔曼滤波在大气数据同化中的综述. 《月刊天气评论》, 144:4489–4532, 2016.

P. L. Houtekamer and H. L. Mitchell. 使用集合卡尔曼滤波技术进行数据同化. 《月刊天气评论》, 126:796–811, 1998年. P. L. Houtekamer and H. L. Mitchell. 一种用于大气数据同化的序贯集合滤波器. 《月刊天气评论》, 129:123–37, 2001年. P. L. Houtekamer and H. L. Mitchell. 集合卡尔曼滤波. 《英国皇家气象学会季刊》, 131:3269–89, 2005年. C.-C. Hu and P. Jan van Leeuwen. 一种用于高维系统应用的粒子流滤波器.《英国皇家气象学会季刊》, 147:2352–74, 2021年. C.-C. Hu, P. Jan van Leeuwen, and J. L. Anderson. 粒子流滤波器在大气模型中的实现. 《月刊天气评论》, 152:2247–64, 2024年. B.R. Hunt, E.J. Kostelich, and I. Szunyogh. 时空混沌的高效数据同化：局部集合变换卡尔曼滤波器. 《物理学*D*》, 230:112–26, 2007年. T.S. Jaakkola. 变分近似方法教程. 收录于 M. Opper and D. Saad 编, 《高级平均场方法：理论与实践》, 页码 129–59. 麻省理工学院出版社, 2001年. T.S. Jaakkola and M.I. Jordan. 通过变分方法进行贝叶斯参数估计. 《统计与计算》, 10:25–37, 2000年. M.I. Jordan, Z. Ghahramani, T.S. Jaakkola, and L.K. Saul. 图模型变分方法导论. 《机器学习》, 37:183–233, 1999年. R. E. Kalman. 线性滤波与预测问题的新方法. 《美国机械工程师学会会刊：基础工程杂志，*D*辑》, 82(1):35–45, 1960. H. Künsch. 状态空间与隐马尔可夫模型. 收录于 O. E. Barndorff-Nielsen, D. R. Cox, and C. Klüppelberg 编, 《复杂随机系统》, 第3章, 页码 109–173. CRC出版社, 2001年. B. Liu, B. Ait-El-Fquih, and I. Hoteit. 高效的基于核的集合高斯混合滤波. 《月刊天气评论》, 144:781–800, 2015年. Edward N. Lorenz and Kerry A. E manuel. 补充天气观测的最优站点：使用小型模型进行模拟. 《大气科学杂志》, 55(3):399–414, 1998年. X. Luo, T. Bhakta, and G. Nædal. 基于相关性的自适应局地化及其在集合方法四维地震历史拟合中的应用. 石油工程师学会, 23:396–427, 2018年. J. Martens. 关于自然梯度法的新见解与新视角. 《机器学习研究杂志》, 1:5753–5828, 2020. C. Di Mauro, R. Hostache, P. Matgen, R. Pelich, M. Chini, P.Jan van Leeuwen, N. Nichols, and G. Blöschl. 一种温度化粒子滤波器，用于增强将 SAR 派生的洪水范围图同化到洪水预报模型中. 《水资源研究》, 58(8):e2022WR031940, 2022. H. L. Mitchell and P. L. Houtekamer. 一种自适应集合卡尔曼滤波器. 《月刊天气评论》, 128:416–33, 2000年. T. Miyoshi. 用于自适应协方差膨胀的高斯方法及其与局地集合变换卡尔曼滤波器的结合实现. 《月刊天气评论》, 139:1519–35, 2011.

D.T. Pham. 强非线性系统序贯数据同化的随机方法。《月刊天气评论》，129:1194–1207，2001年。J. Poterjoy. 面向高维非线性系统的局部化粒子滤波器。《月刊天气评论》，144(1):59–76，2016。R. Potthast、A. Walter 和 A. Rhodin. 业务数值天气预报框架中的局部化自适应粒子滤波器。《月刊天气评论》，147(1):345–62，2019年。M. Pulido 和 P. Jan van Leeuwen. 基于核嵌入映射的序贯蒙特卡罗：映射粒子滤波器。《计算物理学杂志》，396:400–15，2019年。N. F. Raboudi、B. Ait-El-Fquih、C. Dawson 和 I. Hoteit. 将混合方法与一步超前平滑相结合，利用集合卡尔曼滤波器高效预报短期风暴潮。《月刊天气评论》，147:3283–3300，2019年。R. Ranganath、D. Tran 和 D. M. Blei. 层次变分模型。载于国际机器学习会议论文集，第324–33页，纽约，美国，2016。R. H. Reichle、D. B. McLaughlin 和 D. Entekhabi. 使用集合卡尔曼滤波器进行水文数据同化。《月刊天气评论》，130:103–14，2002年。M. Sato. 基于变分贝叶斯的在线模型选择。《神经计算》，13(7):1649–81，2001年7月。F. Septier 和 G. W. Peters. 高维空间中贝叶斯滤波蒙特卡罗方法的最新进展综述，第2章："时空建模的理论方面"。SpringerBriefs丛书—JSS统计学研究系列，G. W. Peters 和 T. Matsui 编，2015年。V. Smidl 和 A. Quinn. 信号处理中的变分贝叶斯方法。施普林格，2006。V. Smidl 和 A. Quinn. 变分贝叶斯滤波。*IEEE*信号处理汇刊，56(10):5020–5030，2008年。C. Snyder、T. Bengtsson、P. Bickel 和 J. Anderson. 高维粒子滤波的障碍。《月刊天气评论》，136(12):4629–40，2008年。J. S. Whitaker、T. M. Hamill、J. L. Anderson 和 C. Snyder. 面向强非线性系统的西格玛点卡尔曼滤波器数据同化方法。《大气科学杂志》，66:3498–3500，2009年。Michael Tippett、Jeffrey Anderson、Craig Bishop、Thomas Hamill 和 Jeffrey Whitaker. 集合平方根滤波器。《月刊天气评论》，131(7):1485–1490，2003年。D. Urozayev、B. Ait-El-Fquih、I. Hoteit 和 D. Peter. 一种用于高效地下成像的降阶变分贝叶斯方法。《国际地球物理学杂志》，229:838–52，2022。P. Jan van Leeuwen. 集合卡尔曼滤波器随机版本的一致解释。《英国皇家气象学会季刊》，146:2815–25，2020。H.L. van Trees. 检测、估计与调制理论：第一部分。约翰·威利父子公司，纽约，1968年。X. Wang、C. H. Bishop 和 S. J. Julier. 正负对集合与中心球面单纯形集合，哪种更好？《月刊天气评论》，132(7):1590–1605，2004年。

J. S. Whitaker 和 T. M. Hamill. 评估在集合数据同化中处理系统误差的方法。《月刊天气评论》，140:3078–89，2012。J. S. Whitaker 和 T. S. Hamill. 不使用扰动观测的集合数据同化。《月刊天气评论》，130:1913–24，2002年。J. S. Whitaker、H. T. Hamill、X. Wei、Y. Song 和 Z. Toth. 使用NCEP全球预报系统进行集合数据同化。《月刊天气评论》，136:463–81，2008年。Y. Zhang 和 D. S. Oliver. 通过自助抽样改进卡尔曼增益的集合估计。《数学地球科学》，42:327–45，2010。