arXiv:2602.01770v1 [stat.CO] 2 Feb 2026

一种带预筛选的多保真度近似贝叶斯计算

Xuefei Cao[1], Shijia Wang[2*] 和 Yongdao Zhou[1*][1]南开大学统计与数据科学学院，中国[2]上海科技大学数学科学研究所，中国

2026年2月3日

摘要

近似贝叶斯计算（ABC）方法通常需要大量模拟，导致计算成本高昂。本文聚焦于多保真模拟模型，并提出一种预筛选分层重要性采样算法。在温和的假设条件下，我们从理论上证明了所提算法满足后验集中性质，并刻画了误差上界以及算法效率与预筛选准则之间的关系。此外，我们提供了一种实用策略，用于评估多保真模型是否适用于所提方法。最后，我们提出了一种采用自适应预筛选策略的多保真 ABC序贯蒙特卡罗方法。通过数值实验展示了所提方法的有效性。我们开发了一个 R软件包，可通过以下链接获取：https://github.com/caofff/MAPS。

关键词：近似贝叶斯计算、多保真度、预筛选、序贯蒙特卡洛

1 引言

贝叶斯推断通过似然函数将先验知识与观测数据相结合，为参数估计和不确定性量化提供了一个有原则的框架。

*通信地址：Shijia Wang 博士（wangshj1@shanghaitech.edu.cn）和 Yongdao Zhou 博士（ydzhou@nankai.edu.cn）。

然而，对于系统生物学、流行病学和宇宙学等领域中遇到的许多复杂模型，直接计算似然在计算上可能代价过高，甚至不可行。这一挑战促使无似然推断技术的发展（Papamakarios 和 Murray，2016；Papamakarios 等人，2019；Hermans 等人，2020；Boelts 等人，2024），其中一种重要方法是近似贝叶斯计算（ABC）（Pritchard 等人，1999；Beaumont 等人，2002；Sisson 等人，2018 年）。ABC 通过从模型中模拟数据，并将模拟数据集与观测数据进行比较，避免了显式的似然计算，从而为复杂情境下的参数推断提供了一种有效方法。

ABC 后验估计的准确性在很大程度上取决于贯穿整个参数空间的大量抽样和模拟。为克服这一低效问题，人们开发了各种算法，通过设计定制的提议分布，更有效地探索参数空间，从而减少在后验概率较低区域进行的模拟。例如，ABC马尔可夫链蒙特卡罗（ABC-MCMC）（Marjoram 等人，2003年；Wegmann 等人，2009）通过一个平稳分布为目标后验分布的马尔可夫链生成样本，并通过在当前粒子附近提出新参数来提高效率。重要性采样方法校正提议分布与目标分布之间的不匹配，其算法性能取决于两个分布的接近程度。ABC序贯蒙特卡罗（ABC-SMC）（Sisson 等人，2007年；Del Moral 等人，2012年）是重要性采样的扩展，它通过由一系列递减的容差阈值定义的中间分布序列传播样本，并利用重要性加权和重采样来近似后验分布。

为应对ABC中高昂模拟成本带来的挑战，一条研究路线是采用多保真算法：这些算法利用成本较低的低保真模拟，在进行高成本高保真（HF）运行之前，尝试提前接受或拒绝参数提议。例如，Picchini (2014年提出了一种专用于随机微分方程参数估计的早期拒绝ABC-MCMC算法，将Metropolis-Hastings接受率分解为两个阶段，以节省计算资源。 Everitt 和 Rowi´nska (2021年) 将延迟接受（DA）MCMC（Christen 和 Fox，2005）整合到ABC框架中，首先利用成本低廉的近似方法筛选样本。此外，Cao 等 (2025年) 开发了一种高斯过程模型来估计模拟

差异，并在完整模拟之前拒绝不太可能的参数。

在许多实际应用中，由于复杂性和计算成本，模拟模型通常具有多个保真度级别。将多保真模型纳入近似贝叶斯推断，对于平衡计算速度与准确性至关重要。低保真模型可快速进行近似评估，而高保真模型则能提供精确但成本高昂的结果。忽视这一层级结构可能导致计算量过大或推断效果不佳。因此，多保真方法对于复杂应用中可扩展且可靠的贝叶斯推断至关重要。由Prescott 和 Baker (2020年) 提出的多保真近似贝叶斯计算（MF-ABC）框架，通过结合低保真和高保真模型模拟来估计似然，从而实现提前接受和提前拒绝。在此框架基础上，Prescott 和 Baker(2021年) 将 MF-ABC 与序贯蒙特卡罗（SMC）技术相结合。此外，Krouglova等人 (2025年) 提出了一种多保真迁移学习框架，该框架使用计算高效的低保真模拟对神经后验估计器进行预训练，随后利用数量有限的高保真数据对其进行微调。

其他方法则采用方差缩减技术，利用低保真度估计来降低估计量的方差。Warne等人 (2018 年) 通过利用多保真度模拟，并将多层蒙特卡罗方差缩减技术直接应用于拒绝抽样，加速近似贝叶斯计算推断。多层序贯蒙特卡罗（Jasra等人，2019年）采用基于序贯蒙特卡罗的重要性采样方案，在不同容差水平间重复利用粒子，并利用多层恒等式实现精准的近似贝叶斯计算推断，同时降低计算成本，从而无需显式耦合各层。

本文考虑输出具有不同保真度级别的模拟模型，其中高保真模拟准确性更高，但计算成本也更高；低保真模拟准确性较低，但计算成本更低。例如，在随机微分方程中，较小的时间步长能得到更精确的结果，但成本更高；较大的时间步长则会降低准确性，但能加快模拟速度并降低计算开销。我们提出一种预筛选分层重要性采样方法，先利用低保真模拟筛除不太可能属于后验分布的参数，再进行高保真模拟，从而减少生成高成本高保真数据的需求。我们从理论上证明，在温和的假设条件下，

通过预筛选分层重要性采样得到的后验分布满足后验集中性质。此外，我们推导出算法准确性、效率与预筛选标准之间的显式关系，揭示了筛选条件与估计精度之间的权衡。我们还提供实用策略，用于评估多保真模型是否适用于所提出的算法。在Del Moral 等人(2012年提出的自适应序贯蒙特卡罗（ASMC）框架基础上，我们进一步开发了一种自适应预筛选 SMC 算法，可在迭代过程中动态选择筛选阈值，在准确性与计算效率之间取得平衡。数值实验表明，与标准自适应序贯蒙特卡罗方法相比，我们的方法将高保真模拟的数量减少了约 40%。

本文结构如下：第2 节回顾近似贝叶斯计算和 SMC 算法的基础知识。第3节介绍所提出的多保真近似贝叶斯计算模型及其性质，并提出一种评估多保真模型适用性的策略。第4 节介绍自适应预筛选序贯蒙特卡罗算法。第5节展示数值实验，第6 节总结全文。附录 S.4 汇总了本文使用的记号。

2 预备知识

本节回顾近似贝叶斯计算和自适应序贯蒙特卡罗（ABC-SMC）算法，这些内容构成了应用于多保真模拟器的预筛选重要性采样框架的基础。

2.1 近似贝叶斯计算

我们考虑一个输入参数为 $\boldsymbol{\theta} \in \Theta$ 、输出随机变量为 $\boldsymbol{x} \in \mathcal{X}$ 的模拟模型 $P_{\boldsymbol{\theta}}$ 。似然函数 $p(\boldsymbol{x} \mid \boldsymbol{\theta})$ 难以处理，但我们可以从该模型生成模拟数据，例如 $\boldsymbol{x} \sim p(\cdot \mid \boldsymbol{\theta})$ 或 $\boldsymbol{x} \sim P_{\boldsymbol{\theta}}$。

给定观测数据 a $\boldsymbol{y}$ 和先验分布 $\pi(\boldsymbol{\theta})$，近似贝叶斯计算后验分布定义为

as

$$\pi_\varepsilon \equiv \pi_\varepsilon(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\theta}) p_\varepsilon(\boldsymbol{\theta} \mid \boldsymbol{y}), \tag{1}$$

with

$$p_\varepsilon(\boldsymbol{\theta} \mid \boldsymbol{y}) = \int \mathbb{I}\big(\Delta(\boldsymbol{x}, \boldsymbol{y}) \leq \varepsilon\big)\, p(\boldsymbol{x} \mid \boldsymbol{\theta})\, d\boldsymbol{x}, \tag{2}$$

其中 $\varepsilon > 0$ 是容差阈值，$\mathbb{I}(\cdot)$ 是一个指示函数：条件成立时取 1，否则取 0；而 $\Delta(\boldsymbol{x}, \boldsymbol{y})$ 表示 $\boldsymbol{x}$ 与 $\boldsymbol{y}$之间的差异。为简便起见，且不致引起混淆，下文将 $p_\varepsilon(\boldsymbol{\theta} \mid \boldsymbol{y})$ 简记为 $p_\varepsilon(\boldsymbol{\theta})$。这意味着，只有生成的模拟数据与观测数据足够接近（即差异在容差 ε范围内）的参数值才会被接受，从而得到对真实后验的近似。当 $\varepsilon \to 0$时，近似贝叶斯计算后验分布收敛于精确的贝叶斯后验。式（2）中的积分可以解释为似然 $p(\boldsymbol{y}_{obs} \mid \boldsymbol{\theta})$ 的近似。

在实践中，$\Delta(\cdot,\cdot) = d(S(\cdot), S(\cdot))$，其中 $d(\cdot,\cdot)$ 是距离函数。当观测数据 $\boldsymbol{y}$ 和模拟数据 $\boldsymbol{x}$ 具有高维或结构复杂的特征时，通常会采用汇总统计量 $S(\cdot)$。使用汇总统计量可以降低数据维度，从而促进在 ABC框架内对模拟数据与观测数据进行更高效且更易处理的比较。理想情况下，这些汇总统计量应尽可能保留有关模型参数的信息。信息性汇总统计量的选择与构建仍是一个活跃的研究领域（Sisson 和 Fan，2010；Marin 等，2012；Blum 等，2013）。本文不讨论汇总统计量的构建，并假设它们是给定的。式（1）的积分通过蒙特卡洛（MC）抽样进行近似，即

$$\hat{p}_\varepsilon(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\big(\Delta(\boldsymbol{x}_i, \boldsymbol{y}_{obs}) \leq \varepsilon\big), \tag{3}$$

其中，对于 $i = 1, \ldots, n$，$\boldsymbol{x}_i \sim p(\cdot \mid \boldsymbol{\theta})$。模拟数据 n 的数量越多，对似然函数 $p(\boldsymbol{y} \mid \boldsymbol{\theta})$ 的估计就越准确。

2.2 自适应近似贝叶斯计算序贯蒙特卡罗算法

最简单的ABC算法采用基本的拒绝抽样。然而，当先验分布与后验分布之间存在显著差异时，该方法的效率会变得极低。尽管存在更有效的重要性采样方案，但通常

很难设计出高质量的提议分布。常用方法（如ABC-MCMC）往往会陷入局部区域，且收敛速度较慢，尤其是在容差 ε 设得较小时。近似贝叶斯计算序贯蒙特卡洛（ABC-SMC）（Sis- son 等人，2007年）算法通过引入一系列阈值递减的中间目标分布，缓解了这些问题。ABC-SMC结合了重要性采样与ABC-MCMC，既减轻了构建合适提议分布的难度，也缓解了ABC-MCMC在阈值较小时收敛缓慢的问题。此外，自适应近似贝叶斯计算序贯蒙特卡洛算法（ABC-ASMC）（Del Moral 等人，2012年）提出了一种自适应策略，用于确定递减阈值序列。下文将详细介绍ABC-ASMC算法。

近似贝叶斯计算-SMC 方法采用一系列逐渐减小的容差阈值 $\{\varepsilon_0, \ldots, \varepsilon_T\}$，其中 $\varepsilon_0 = \infty$ 和 $\varepsilon_T = \varepsilon$。相应的中间分布记为 $\pi_{\varepsilon_t}(\boldsymbol{\theta} \mid \boldsymbol{y})$。在每次迭代 t 中，利用前一次迭代的目标分布 $\pi_{\varepsilon_{t-1}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ 构建重要性提议分布，并据此抽取新样本。这一迭代过程确保加权样本逐步近似目标后验分布，缩小提议分布与目标分布之间的差异，从而提高采样效率。

令 $\{(\boldsymbol{\theta}_{t-1}^{(i)}, W_{t-1}^{(i)})\}_{i=1}^{N}$ 表示迭代 $t-1$ 时的加权粒子集。在自适应序贯蒙特卡洛框架中，选择容差 ε_t，使上一迭代中固定比例 α 的“活跃”粒子得以保留。具体而言，ε_t 设定为满足

$$PA(\{W_t^{(i)}\}; \varepsilon_t) = \alpha \, PA(\{W_{t-1}^{(i)}\}),$$

其中 $PA(\{W_{(i)t}\}; \varepsilon_t) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}\{W_t^{(i)} > 0\}$ 表示在容差 ε_t 下具有正权重的粒子所占比例，而 $PA(\{W_{(i)t-1}\}) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}\{W_{t-1}^{(i)} > 0\}$ 其中 $\left\{W_{(i}_{t-1}\right\}_{i=1}^{)}{}_{:N}$ 已给定。这里，$\mathbb{I}\{\cdot\}$ 表示指示函数。对 ε_t 进行这种自适应选择，既能维持粒子多样性，又能逐步收紧后验近似。随后按如下方式更新粒子权重：

$$W_t^{(i)} \propto W_{t-1}^{(i)} \frac{\sum_k \mathbb{I}\big(\Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \varepsilon_t\big)}{\sum_k \mathbb{I}\big(\Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \varepsilon_{t-1}\big)}. \tag{4}$$

随后，我们执行重采样步骤，以稳定算法并缓解粒子退化。重采样后，所有粒子的权重都相等。

为更充分地探索参数空间，我们应用一个以 π_{ε_t} 为目标的马尔可夫核，通常通过梅特ropolis–黑斯廷斯（MH）步骤实现。给定当前状态（$\boldsymbol{\theta}$，$\boldsymbol{x}_{1:n}$），从 $q_t(\boldsymbol{\theta}^* \mid \boldsymbol{\theta}) \prod_{k=1}^{n} p(\boldsymbol{x}_k^* \mid \boldsymbol{\theta}^*)$ 中提出一个候选状态（$\boldsymbol{\theta}^*$，$\boldsymbol{x}_{1:n}^*$）。提议分布以一定概率被接受：

$$\alpha((\boldsymbol{\theta}, \boldsymbol{x}_{1:n}), (\boldsymbol{\theta}^*, \boldsymbol{x}_{1:n}^*)) = \min\left\{1, \frac{\pi(\boldsymbol{\theta}^*) \sum_{k=1}^{n} \mathbb{I}(\Delta(\boldsymbol{x}_k^*, \boldsymbol{y}) < \varepsilon_t)}{\pi(\boldsymbol{\theta}) \sum_{k=1}^{n} \mathbb{I}(\Delta(\boldsymbol{x}_k, \boldsymbol{y}) < \varepsilon_t)} \frac{q_t(\boldsymbol{\theta} \mid \boldsymbol{\theta}^*)}{q_t(\boldsymbol{\theta}^* \mid \boldsymbol{\theta})}\right\}. \tag{5}$$

ABC-ASMC 的详细内容见补充算法 S.1。

3 面向多保真度模拟模型的预筛选分层重要性采样

本节提出一种专为多保真度模拟模型设计的预筛选算法。首先介绍多保真度仿真框架，并概述其关键原理。接着，我们提出一种预筛选分层重要性采样方法，该方法利用分层保真度级别提高采样效率。随后，我们介绍一种自适应预筛选 SMC 算法，该算法自适应地利用低保真（LF）模拟进行预筛选，以减少所需的高保真（HF）模拟次数。最后，讨论所提出算法的理论性质。

3.1 多保真度模拟器

在多保真度框架中，会采用具有不同保真度级别的模拟模型。高保真模型生成准确性较高的模拟数据，但计算成本也很高。相比之下，低保真模型生成的数据准确性较低，但所需的计算量要少得多。这类低保真模型可能自然产生，例如通过微分方程求解器中较粗的时间离散化得到；也可能通过构建代理模型获得。低保真模型的开发不在本文讨论范围之内。接下来几节将介绍低保真模拟需要满足的条件。

在多保真度 ABC 框架中，主要模型 $P_{\boldsymbol{\theta}}$ 由与高保真模拟器对应的似-˜然函数 $p(\cdot \mid \boldsymbol{\theta})$ 描述。低保真模拟器记为 $P_{\boldsymbol{\theta}}$,

其对应的似然为 $\tilde{p}(\cdot \mid \boldsymbol{\theta})$。对于任意参数 $\boldsymbol{\theta} \in \Theta$，可以独立生成模拟输出 $\boldsymbol{x} \sim p(\cdot \mid \boldsymbol{\theta})$ 和 $\tilde{\boldsymbol{x}} \sim \tilde{p}(\cdot \mid \boldsymbol{\theta})$。低保真近似贝叶斯计算后验分布为

$$\tilde{\pi}_{\tilde{\varepsilon}}(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\theta}) \tilde{p}_{\tilde{\varepsilon}}(\boldsymbol{\theta} \mid \boldsymbol{y}), \tag{6}$$

其中，近似的低保真似然定义为

$$\tilde{p}_{\tilde{\varepsilon}}(\boldsymbol{\theta} \mid \boldsymbol{y}) = \int \mathbb{I}(\Delta(\tilde{\boldsymbol{x}}, \boldsymbol{y}) < \tilde{\varepsilon}) \tilde{p}(\tilde{\boldsymbol{x}} \mid \boldsymbol{\theta}) d\tilde{\boldsymbol{x}}.$$

为简洁起见，后文将 $\tilde{p}_{\tilde{\varepsilon}}(\boldsymbol{\theta} \mid \boldsymbol{y})$ 简记为 $\tilde{p}_{\tilde{\varepsilon}}(\boldsymbol{\theta})$。

我们假设观测数据 $\boldsymbol{y}$ 可以通过映射$M(\cdot)$映射到低保真数据空间，使得 $\tilde{\boldsymbol{y}} = M(\boldsymbol{y})$。借助这一映射，可以通过差异函数 $\Delta(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})$计算低保真模拟与观测值之间的差异。例如，如果 $\boldsymbol{y}$ 表示时间序列，那么 $\tilde{\boldsymbol{y}}$ 可以是该时间序列的稀疏版本，而 M 表示下采样操作。在本研究中，我们进一步简化假设，认为数据空间相同，即 $\boldsymbol{y} = \tilde{\boldsymbol{y}}$，且差异函数满足 $\Delta(\tilde{\boldsymbol{x}}, \boldsymbol{y}) = \Delta(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})$。

3.2 预筛选分层重要性采样

在近似贝叶斯计算重要性采样算法中，使用先验分布作为重要性提议分布简单直观。然而，当先验分布与后验分布之间存在显著差异时，这种方法往往效率低下，通常会导致接受率较低，并浪费大量模拟。为了提高效率，理想的做法是设计一个能够紧密近似 ABC 后验分布的重要性提议分布。精心选择的提议分布能在固定的计算预算下生成更多有效样本，在处理计算开销大的模拟器时尤其有益。尽管如此，构建这样的重要性提议分布仍是一项艰巨任务，其复杂程度与直接对 ABC 后验分布进行推断相当。

本文利用低保真模拟与高保真模拟之间的一致性，借助计算成本较低的低保真近似贝叶斯计算后验分布来改进重要性提议分布。详细算法流程总结于算法1。最初，参数样本$\boldsymbol{\theta}_{1:N}$ 抽取自重要性提议分布 $q(\cdot)$，进行低保真模拟，并且

赋予相应的低保真权重:

$$W^{(i)} = \frac{\pi(\boldsymbol{\theta}^{(i)})}{q(\boldsymbol{\theta}^{(i)})} \sum_{k=1}^{n_L} \mathbb{I}\left(\Delta(\tilde{\boldsymbol{x}}_k^{(i)}, \boldsymbol{y}) < \tilde{\varepsilon}\right), \tag{7}$$

其中 $\tilde{\boldsymbol{x}}_k^{(i)}$ 表示第 i个参数样本的第 k个低保真模拟输出， $\tilde{\varepsilon}$ 为相应的低保真容差阈值。对于满足 $W^{(i)} > 0$的样本，进一步执行高保真模拟以更新权重:

$$W^{(i)} = W^{(i)} \cdot \frac{\sum_{k=1}^{n_H} \mathbb{I}\left(\Delta(\boldsymbol{x}_k^{(i)}, \boldsymbol{y}) < \varepsilon\right)}{\sum_{k=1}^{n_L} \mathbb{I}\left(\Delta(\tilde{\boldsymbol{x}}_k^{(i)}, \boldsymbol{y}) < \tilde{\varepsilon}\right)} = \frac{\pi(\boldsymbol{\theta}^{(i)})}{q(\boldsymbol{\theta}^{(i)})} \sum_{k=1}^{n_H} \mathbb{I}\left(\Delta(\boldsymbol{x}_k^{(i)}, \boldsymbol{y}) < \varepsilon\right), \tag{8}$$

其中 $\boldsymbol{x}_k^{(i)}$ 表示第 i个参数的第 k个高保真模拟输出， ε 为高保真容差阈值。

算法1 一种预筛选分层重要性抽样**输入：** 先验 π(·)，重要性提议分布 q(·)，低保真模拟器 $\tilde{p}$(·$\mid \boldsymbol{\theta}$)，高保真模拟器p(·$\mid \boldsymbol{\theta}$)，观测数据 $\boldsymbol{y}$，容差阈值 $\varepsilon, \tilde{\varepsilon}$，样本数 N，低保真模拟次数n_L，高保真模拟次数 n_H。

输出： 加权样本 $\{(W^{(i)}, \boldsymbol{\theta}^{(i)})\}_{i=1}^{N}$。

1：对 $i = 1$，.....， N执行 2：抽取 $\boldsymbol{\theta}^{(i)} \sim q$(·)。3：生成 n_L 次低保真模拟 $\tilde{\boldsymbol{x}}_1^{(i)}$ ： $_{n_L} \sim \tilde{p}$(·$\mid \boldsymbol{\theta}^{(i)}$)，并用式(7)计算 $W^{(i)}$ 。4：若 $W^{(i)} > 0$ 则 5：生成 n_H 次高保真模拟 $\boldsymbol{x}_1^{(i)}$ ： $_{n_H} \sim p$(·$\mid \boldsymbol{\theta}^{(i)}$)，并根据式(8)更新 $W^{(i)}$ 。6：结束如果 7：结束循环 8：归一化权重 $\{W^{(i)}\}_{i=1}^{N}$。

算法 1 中的预筛选策略仅依靠低保真模拟，在进行计算成本高昂的高保真模拟之前，筛除那些不太可能做出显著贡献的样本。值得注意的是，该方法不使用低保真近似贝叶斯计算后验分布 $\tilde{\pi}_{\tilde{\varepsilon}}$ 来构建完整的重要性提议分布。这是因为 $\tilde{\pi}_{\tilde{\varepsilon}}$ 通常通过蒙特卡洛方法估计，从而导致重要性权重 $\pi_{\varepsilon}/\tilde{\pi}_{\tilde{\varepsilon}}$的方差较大，因为 $\tilde{\pi}_{\tilde{\varepsilon}}$ 出现在

分母中。这种方差会对估计准确性和整体算法性能产生不利影响。算法 1 有效利用低保真模拟来改进提议分布，同时避免这一问题。

在算法1中，用于预筛选分层重要性抽样的参数样本权重为 $\{\boldsymbol{\theta}_i\}_{i=1}^N$

$$W^{(i)} \propto \frac{\pi(\boldsymbol{\theta}^{(i)})}{q(\boldsymbol{\theta}^{(i)})} \mathbb{I}\left(\min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)}, \boldsymbol{y}) < \tilde{\varepsilon}\right) \cdot \sum_{k=1}^{n_H} \mathbb{I}\left(\Delta(\boldsymbol{x}_k^{(i)}, \boldsymbol{y}) < \varepsilon\right).$$

该表述意味着，引入低保真预筛选步骤会产生估计偏差。以下小节将对此进行理论论证。

3.3 理论性质

本节将展示所提出的多保真模拟器的一些性质。命题1 给出了算法1所产生的近似后验分布，命题2分析了低保真预筛选对估计准确性和计算效率的影响。定理1证明了所提出算法的后验集中性。

命题1 运行算法*1* 所获得的加权样本近似于多保真近似贝叶斯计算后验分布，该分布定义为

$$\pi_{\varepsilon,\tilde{\varepsilon}} \equiv \pi_{\varepsilon,\tilde{\varepsilon}}(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\theta})\, p_\varepsilon(\boldsymbol{\theta})\, Q_{\tilde{\varepsilon}}(\boldsymbol{\theta}), \tag{9}$$

其中

$$Q_{\tilde{\varepsilon}}(\boldsymbol{\theta}) = 1 - (1 - \tilde{p}_{\tilde{\varepsilon}}(\boldsymbol{\theta}))^{n_L},$$

且参数 n_L 表示低保真模拟的数量。

假设 1 对于任意 $\varepsilon > 0$，存在一个 $\tilde{\varepsilon} > 0$ 和常数 $a_0 \in (0,1)$ 使得 $a_L < 1$，其中

$$a_L = \int \pi_\varepsilon(\boldsymbol{\theta})\, (1 - \tilde{p}_{\tilde{\varepsilon}}(\boldsymbol{\theta}))^{n_L}\, d\boldsymbol{\theta}. \tag{10}$$

假设 1 规定了预筛选过程中被排除的后验样本比例，从而为算法的输出集非空提供了必要保证。

定理 1 如果满足 *Frazier* 等人*(2018* 年*)* 的假设 *1-3*，以及本文所述的假设 *1* 和 *S.4*，那么针对 (9)的预筛选分层重要性采样仍然具有后验集中性质。

定理1 表明，在对低保真模拟器作出合理假设的条件下，所提出的算法仍能保持后验集中。这一性质至关重要，因为对于任意子集 $A \subset \Theta$，ABC 后验概率 $\Pi_\varepsilon\{A \mid \boldsymbol{y}\}$ 通常会偏离精确后验，其中 $\Pi_\varepsilon(\cdot \mid \boldsymbol{y})$ 表示密度为 $\pi_\varepsilon(\cdot \mid \boldsymbol{y})$ 的 ABC 后验测度。在缺乏精确后验保证的情况下，$\Pi_\varepsilon\{\cdot \mid \boldsymbol{y}\}$ 围绕真实参数 $\boldsymbol{\theta}_0$ 的集中，为量化 $\boldsymbol{\theta}$中的不确定性提供了稳健的依据。详细的技术讨论和证明见附录 S.2。

命题 2 *(i)* 对于任意满足 $a_L < 1$ 的 ε 和 $\tilde{\varepsilon}$，在预筛选分层重要性采样框架中纳入低保真筛选步骤所引入的近似误差有界于

$$\left\|\pi_{\varepsilon,\tilde{\varepsilon}} - \pi_\varepsilon\right\|_1 < \frac{1}{1-a_L} - (1-a_L). \tag{11}$$

(ii) 分层重要性采样算法的期望预筛选率为

$$1 - \alpha_L = \int q(\boldsymbol{\theta})(1 - Q_\varepsilon(\boldsymbol{\theta}))d\boldsymbol{\theta}., \tag{12}$$

其中 $q(\boldsymbol{\theta})$ 为重要性提议分布，满足 $q(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta})$。

命题 2 (i) 给出了由两阶段低保真与高保真 ABC 方法得到的联合后验分布 L_1 与原始 ABC 后验分布 $\pi_{\varepsilon\text{之间}\tilde{\varepsilon}}$ 距离的显式上界 π_ε。命题 2 (ii) 给出了低保真层级的预筛选率，反映了可提前舍弃的参数样本比例，从而避免高成本高保真模拟所带来的计算成本。

这一结果为将低保真模型用作低计算成本筛选步骤提供了理论保证：尽管低保真模型通过不完善的筛选机制改变样本分布，从而引入一定偏差，但由此产生的误差是可控且有界的。这类

误差界确保多保真 ABC 方法能够降低计算成本，同时保持合理的准确性。

假拒绝率 a_L 应尽可能小，而预筛选率$1-\alpha_L$ 应尽可能大。对于固定的低保真模型， $Q_{\tilde{\varepsilon}}$ 会随着样本量参数 n_L 和阈值参数 $\tilde{\varepsilon}$的增大而单调增加。随着 n_L 和 $\tilde{\varepsilon}$增大，算法精度会提高，而效率则相应降低。因此， n_L和 $\tilde{\varepsilon}$体现了准确性与效率之间的权衡。对于固定的低保真模型，调整阈值参数 $\tilde{\varepsilon}$ 可以有效控制 $Q_{\tilde{\varepsilon}}=0$所在区域的大小。因此，与调整样本量 n_L相比，调整 $\tilde{\varepsilon}$ 对算法性能的影响更为关键。在第4节中，我们将把所提出的方法嵌入序贯蒙特卡洛框架，从而能够自适应地选择阈值 $\tilde{\varepsilon}$，进一步提升算法性能和鲁棒性。

3.4 模拟模型适用性评估

在采用多保真度方法之前，有必要验证所选或构建的低保真模型能否作为有效的预筛选模拟器。本节提出一种度量，用于衡量低保真模拟模型与高保真模拟模型之间的一致性。该度量的基本思路是在真正率固定的情况下，评估预筛选过程的 II 类错误率（假阴性比率）。这里，真阳性是指被高保真模拟器接受且通过预筛选的粒子；假阴性是指被高保真模拟器拒绝但通过预筛选的粒子。这需要选择一组合适的阈值（ε_0, $\tilde{\varepsilon}_0$），并评估假阴性比率。

首先，我们从提议分布$q(\cdot)$中抽取 N_0 （例如， $N_0=5000$）个参数样本 $\{\boldsymbol{\theta}^{(i)}\}_{i=1}^{N_0}$。对于每个参数样本，我们生成一组低保真模拟输出 $\tilde{\boldsymbol{x}}_1^{(i)}:_{n_L}\sim\tilde{p}(\cdot\mid\boldsymbol{\theta}^{(i)})$和高保真模拟输出 $\boldsymbol{x}_1^{(i)}:_{n_H}\sim p(\cdot\mid\boldsymbol{\theta}^{(i)})$。

评估 ε_0**：** 我们通过求解 $PA(\{W_i\}_{i=1}^{N_0},\varepsilon_0)=\kappa$（例如 $\kappa=0.1$）来确定高保真距离阈值 ε_0；如果 $\varepsilon_0<\varepsilon$，则设为 $\varepsilon_0=\varepsilon$，其中每个样本的权重为

$$W_i\propto\frac{\pi(\boldsymbol{\theta}_i)}{q(\boldsymbol{\theta}_i)}\sum_{k=1}^{n_H}\mathbb{I}\left(\Delta(\boldsymbol{x}_k^{(i)},\boldsymbol{y})<\varepsilon_0\right),$$

其中 $\pi(\cdot)$ 表示先验分布， $\mathbb{I}(\cdot)$ 表示指示函数。在实际应用中，当 ε 非常小时，经过 N_0 次试验后接受样本的数量可能不足，从而导致评估不可靠。选择 ε_0 是为了避免使用过小的 ε ，导致仅保留极少数样本。

评估 $\tilde{\varepsilon}_0$： 低保真阈值 $\tilde{\varepsilon}_0$ 设定为所有正权重抽样 $W_i > 0$中，其到观测值的最小低保真距离的最大值：

$$\tilde{\varepsilon}_0 = \max_{i:W_i>0} \min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)}, \boldsymbol{y}).$$

选择 $\tilde{\varepsilon}_0$ 的直觉是确保所有被高保真模型接受的粒子都能通过预筛选。因此，在这种设定下，真正率为 1。

计算假阴性： 然后，对于每个参数样本，我们统计距离小于 $\tilde{\varepsilon}_0$的低保真模拟次数，记为

$$\tilde{w}_i = \sum_{k=1}^{n_L} \mathbb{I}\left(\Delta(\tilde{\boldsymbol{x}}_k^{(i)}, \boldsymbol{y}) \leq \tilde{\varepsilon}_0\right).$$

请注意， $\tilde{w}_i = 0$ 仅适用于被两个阈值同时拒绝的粒子。最后，我们计算度量

$$E = \frac{1}{(1-\kappa)N_0} \sum_{i=1}^{N_0} \mathbb{I}(\tilde{w}_i > 0),$$

它表示假阴性比率。

该度量用于衡量低保真模拟模型与高保真模拟模型之间的一致性。 E 的值越低，表示一致性越好。在实际应用中，阈值 ε_0 通常设为 ε ，以获得最佳性能。

4 一种自适应预过滤多保真近似贝叶斯计算-SMC 算法

本节将预筛选分层重要性采样嵌入自适应序贯蒙特卡罗（SMC）框架，并提出带自适应预过滤序贯蒙特卡罗的多保真近似贝叶斯计算（MAPS）算法。在标准序贯蒙特卡罗方法的递减阈值序列 $\{\varepsilon_t\}$框架基础上，我们

引入一个递减的辅助阈值序列 $\{\tilde{\varepsilon}_t\}$。这些阈值交替递减，以定义一系列中间目标分布

$$\pi_{\varepsilon_0}, \pi_{\varepsilon_0,\tilde{\varepsilon}_1}, \pi_{\varepsilon_1,\tilde{\varepsilon}_1}, \pi_{\varepsilon_1,\tilde{\varepsilon}_2}, \pi_{\varepsilon_2,\tilde{\varepsilon}_2}, \ldots,$$

并逐步收敛到 $\pi_{\varepsilon_T,\ \tilde{\varepsilon}_T}$，其中 $\varepsilon_T = \varepsilon$ 是最终目标阈值。

算法3 概述了所提出的 MAPS算法。在第 t次 SMC 迭代中，该算法在确定临界值、辅助加权、选择辅助阈值、重采样、提议分布、加权 和选择阈值 之间迭代，以实现从 $\pi_{\varepsilon_{t-1},\ \tilde{\varepsilon}_{t-1}}$对 $\pi_{\varepsilon_t,\ \tilde{\varepsilon}_t}$ 的近似。下面详细介绍这些步骤。

确定临界值：辅助阈值序列 $\tilde{\varepsilon}_{1:T}$ 会影响滤波性能和后验分布的最终估计。为避免选择过于保守的阈值（例如 由离群值导致的阈值），我们提出为辅助阈值确定一个下界。通过控制 a_L 来选择临界值 $\tilde{\varepsilon}$ ，其中 a_L表示距离高于临界值的低保真模拟所占的比例。该 a_L 通常设为较小的正值，例如 0.001。 a_L 的数值估计为

$$a_L \approx \sum_{i=1}^{N} \tilde{w}_i \, \mathbb{I}\left(\min_k \Delta(\tilde{\boldsymbol{x}}_{t-1,k}^{(i)}, \boldsymbol{y}) \geq \underline{\tilde{\varepsilon}}\right),$$

其中 $\{(\boldsymbol{\theta}_t^{(i)}, \tilde{w}_i)\}_{i=1}^N$ 表示用于近似 ABC后验分布π_{ε_T}的加权粒子。对 π_{ε_T} 的这一近似通过以中间目标为提议分布的重要性采样过程实现，并采用权重更新函数

$$\tilde{w}_i \propto W_{t-1}^{(i)} \frac{\sum_k \mathbb{I}\big(\Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \varepsilon_T\big)}{\sum_k \mathbb{I}\big(\Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \varepsilon_t\big)},$$

其中 $W_{t-1}^{(i)}$ 表示第 i 个粒子在第 t 次迭代时的原始权重。在实践中，临界值 $\tilde{\varepsilon}$ 被选为加权最小距离的 $(1 - a_L)$ 分位数，即

$$\underline{\tilde{\varepsilon}} = Q_{1-a_L}\left(\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_{t-1,k}^{(i)}, \boldsymbol{y}), \tilde{w}_i\right\}\right).$$

随着 a_L 趋近于零，算法的估计误差会减小。然而，在实际应用中，通常更倾向于将 a_L 设为严格大于零，因为将阈值设为分位数而非最大值，既能有效控制估计误差，又能减轻离群值的影响。这种方法可避免阈值过于保守，从而

提高预筛选步骤的效率。在每次迭代中，阈值 $\tilde{\varepsilon}$ 都会更新，因为随着算法的推进，样本集 $\{(\boldsymbol{\theta}_t^{(i)}, \tilde{w}_i)\}_{i=1}^N$ 的质量不断提高，使 $\tilde{\varepsilon}$ 逐渐收敛至真实值。

辅助权重与阈值的更新：随后我们降低辅助阈值 $\tilde{\varepsilon}$。新的辅助阈值 $\tilde{\varepsilon}_t$ 通过从先前的加权粒子集 $\left\{(\boldsymbol{\theta}_{t-1}^{(i)}, W_{t-1}^{(i)})\right\}_{i=1}^N$ 中保留固定比例 α 的活跃粒子来确定。具体而言，选择 $\tilde{\varepsilon}_t$ 使得

$$\mathrm{PA}\left(\{\tilde{W}_t^{(i)}\}; \tilde{\varepsilon}_t\right) = \alpha_L \,\mathrm{PA}\left(\{W_{t-1}^{(i)}\}\right), \tag{13}$$

其中更新后的权重定义为

$$\tilde{W}_t^{(i)} \propto W_{t-1}^{(i)} \cdot \mathbb{I}\left\{\min_k \Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \tilde{\varepsilon}_t\right\}. \tag{14}$$

由此得到加权粒子集 $\left\{(\boldsymbol{\theta}_{t-1}^{(i)}, \tilde{W}_t^{(i)})\right\}_{i=1}^N$，其近似服从分布$\pi_{\varepsilon_{t-1},\tilde{\varepsilon}_t}$。辅助阈值的降低大约移除了 $1-\alpha_L$ 比例的活跃粒子。最后，我们令 $\tilde{\varepsilon}_t = \max\{\tilde{\varepsilon}_t, \tilde{\varepsilon}\}$。该机制在确保高效筛选的同时，防止因活跃粒子数量不足而导致算法崩溃。

自适应重采样：为缓解粒子退化，当有效样本量（ESS）低于阈值 N_T时执行动态重采样:

$$\mathrm{ESS} = 1/\sum_{i=1}^{N}(W_t^{(i)})^2 < N_T.$$

重采样步骤完成后，所有粒子都被赋予相同的权重。

提议分布： 我们构造一个平稳分布为$\pi_{\varepsilon_{t-1},\tilde{\varepsilon}_t}$ 的马尔可夫核，以传播这些加权粒子并增强样本多样性。每次迭代中，给定当前状态$\boldsymbol{z}_{t-1} = (\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)})$，根据提议分布 $q(\cdot \mid \boldsymbol{\theta}_{t-1}^{(i)})$ 提出一个新的参数 $\boldsymbol{\theta}^*$。从当前状态 $\boldsymbol{z}_{t-1}$ 转移到提议状态 $\boldsymbol{z}^{(i)*} = (\boldsymbol{\theta}^{(i)*}, \boldsymbol{x}_{1:n_H}^{(i)*}, \tilde{\boldsymbol{x}}_{1:n_L}^{(i)*})$ 的 MH 接受概率为

$$\tilde{\alpha}\left\{\boldsymbol{z}_{t-1}^{(i)}, \boldsymbol{z}^{(i)*}\right\} = \frac{\pi(\boldsymbol{\theta}^{(i)*}) \sum_{k=1}^{n_H} \mathbb{I}\left\{\Delta(\boldsymbol{x}_k^{(i)*}, \boldsymbol{y}) < \varepsilon_{t-1}\right\} q(\boldsymbol{\theta}_{t-1}^{(i)} \mid \boldsymbol{\theta}^{(i)*}) \mathbb{I}\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)*}, \boldsymbol{y}) < \tilde{\varepsilon}_t\right\}}{\pi(\boldsymbol{\theta}_{t-1}^{(i)}) \sum_{k=1}^{n_H} \mathbb{I}\left\{\Delta(\boldsymbol{x}_k^{(i)*}, \boldsymbol{y}) < \varepsilon_{t-1}\right\} q(\boldsymbol{\theta}^{(i)*} \mid \boldsymbol{\theta}_{t-1}^{(i)}) \mathbb{I}\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)*}, \boldsymbol{y}) < \tilde{\varepsilon}_t\right\}},$$

其中，高保真模拟为 $\boldsymbol{x}_1^{(i)*}{}_{:n_H} \sim (\cdot \mid \boldsymbol{\theta}^{(i)*})$，低保真模拟为 $\tilde{\boldsymbol{x}}_1^{(i)*}{}_{:n_L} \sim \tilde{p}(\cdot \mid \boldsymbol{\theta}^{(i)*})$。由于当前状态已从 $\pi_{\varepsilon_{t-1},\tilde{\varepsilon}_t}$ 中抽取，因此分母中的因子 $\mathbb{I}\{\min_k \Delta(\tilde{\boldsymbol{x}}_k, \boldsymbol{y}) < \tilde{\varepsilon}\}$ 始终为一。因此，MH 接受概率可简化为

$$\tilde{\alpha}\left\{\boldsymbol{z}_{t-1}^{(i)}, \boldsymbol{z}^{(i)*}\right\} = \frac{\pi(\boldsymbol{\theta}^{(i)*}) \sum_{k=1}^{n_H} \mathbb{I}\left\{\Delta(\boldsymbol{x}_k^{(i)*}, \boldsymbol{y}) < \varepsilon_{t-1}\right\} q(\boldsymbol{\theta}_{t-1}^{(i)} \mid \boldsymbol{\theta}^{(i)*}) \mathbb{I}\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)*}, \boldsymbol{y}) < \tilde{\varepsilon}_t\right\}}{\pi(\boldsymbol{\theta}_{t-1}^{(i)}) \sum_{k=1}^{n_H} \mathbb{I}\left\{\Delta(\boldsymbol{x}_k^{(i)*}, \boldsymbol{y}) < \varepsilon_{t-1}\right\} q(\boldsymbol{\theta}^{(i)*} \mid \boldsymbol{\theta}_{t-1}^{(i)})}. \quad (15)$$

由式（15）可知，当 $\min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)*}, \boldsymbol{y}) \geq \tilde{\varepsilon}_t$时，接受概率为零。因此，可以先对所提出的参数 $\boldsymbol{\theta}^{(i)*}$进行低保真模拟。仅当低保真差异满足 $\min_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)*}, \boldsymbol{y}) < \tilde{\varepsilon}_t$ 时，才应进行高保真模拟。在这种情况下，候选参数以式（15）给出的 Metropolis–Hastings 接受概率被接受。否则，可以立即拒绝所提出的参数，而无需进行高保真模拟。详细步骤总结于算法 2。

算法 2 预筛选 ABC-MCMC 的单阶段

过程 PF-ABC-MCMC($(\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1}^{(i)}{}_{,1:n_H}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}$ ˜), π, q, P, P, $\tilde{\varepsilon}$, ε)

˜**输入：** 先验：$\pi(\cdot)$，低保真模拟器 $P_{\boldsymbol{\theta}}$，高保真模拟器 $P_{\boldsymbol{\theta}}$，观测数据 $\boldsymbol{y}$，阈值 ε_{t-1} 和$\tilde{\varepsilon}_t$，当前状态 $(\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1}^{(i)}{}_{,1:n_H}, \tilde{\boldsymbol{x}}_{t-1}^{(i)}{}_{,1:n_L})$。

输出： $(\boldsymbol{\theta}_t^{(i)}, \boldsymbol{x}_{t\,1}^{(i)}{}_{,:n_H}, \tilde{\boldsymbol{x}}_{t\,1}^{(i)}{}_{,:n_L})$

1: 抽样 $\boldsymbol{\theta}^{(i)*} \sim q(\cdot \mid \boldsymbol{\theta}_{t-1}^{(i)})$ 和 $\tilde{\boldsymbol{x}}_k^{(i)*} \sim \tilde{P}_{\boldsymbol{\theta}^{(i)*}}$，其中 $k = 1, \ldots, n_L$。

2: **如果** 最小值$_k \Delta(\tilde{\boldsymbol{x}}_k^{(i)*}, \boldsymbol{y}) > \tilde{\varepsilon}_t$ **则**

3: Set $\left\{\boldsymbol{\theta}_t^{(i)}, \boldsymbol{x}_{t,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t,1:n_L}^{(i)}\right\} = \left\{\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}\right\}$.

4: **否则**

5: Sample $\boldsymbol{x}_{t,1:n_H}^{(i)*} \sim P_{\boldsymbol{\theta}^{(i)*}}$, set $\left\{\boldsymbol{\theta}_t^{(i)}, \boldsymbol{x}_{t,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t,1:n_L}^{(i)}\right\} = \left\{\boldsymbol{\theta}^{(i)*}, \boldsymbol{x}_{1:n_H}^{(i)*}, \tilde{\boldsymbol{x}}_{1:n_L}^{(i)*}\right\}$ with probability Eq.(15); otherwise set $\left\{\boldsymbol{\theta}_t^{(i)}, \boldsymbol{x}_{t,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t,1:n_L}^{(i)}\right\} = \left\{\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}\right\}$.

6: **结束如果**

权重和阈值的更新： 我们通过令 PA ; PA $\left(\left\{\tilde{W}_{()\,t}\right\}\right)$来确定阈值

ε_t $\left(\left\{W_t^{(i)}\right\}\ \varepsilon_t\right) =$，其中 α i

$$W_t^{(i)} \propto \tilde{W}_t^{(i)} \frac{\sum_k \left\{\mathbb{I}(\Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \varepsilon_t\right\}}{\sum_k \left\{\mathbb{I}(\Delta(\boldsymbol{x}_{t-1,k}^{(i)}, \boldsymbol{y}) < \varepsilon_{t-1}\right\}}. \quad (16)$$

在确定新阈值 ε_t 之前执行马尔可夫核转移，因为辅助阈值 $\tilde{\varepsilon}_t$ 是通过相较于前一阈值ε_{t-1}收紧条件得到的。鉴于 $\tilde{\varepsilon}_t$ 更为严格，针对分布$\pi_{\varepsilon_{t-1},\ \tilde{\varepsilon}_t}$ 的马尔可夫核转移能够更充分地受益于低保真模拟的预筛选效果。这使粒子细化更加高效，并加快序贯蒙特卡洛框架中的收敛。

实际上，算法3 第 4 步中的高保真模拟可以省略。这是因为，在 $t = 1$且 $\varepsilon_0 = \infty$的情况下，算法第8 步中的Metropolis–Hastings接受概率为

$$\tilde{\alpha}\{\mathbf{z}, \mathbf{z}^*\} = \frac{\pi(\boldsymbol{\theta}^*)q(\boldsymbol{\theta} \mid \boldsymbol{\theta}^*)\mathbb{I}\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_k^*, \boldsymbol{y}) < \tilde{\varepsilon}_t\right\}}{\pi(\boldsymbol{\theta})q(\boldsymbol{\theta}^* \mid \boldsymbol{\theta})}.$$

此阶段不需要高保真模拟数据。只需在第一次马尔可夫链蒙特卡洛移动后，对接受样本进行高保真模拟即可，以便在后续步骤中使用这些高保真数据。图1 展示了MAPS算法的流程图（对应算法3中的步骤6–10），有助于更直观地理解算法流程。

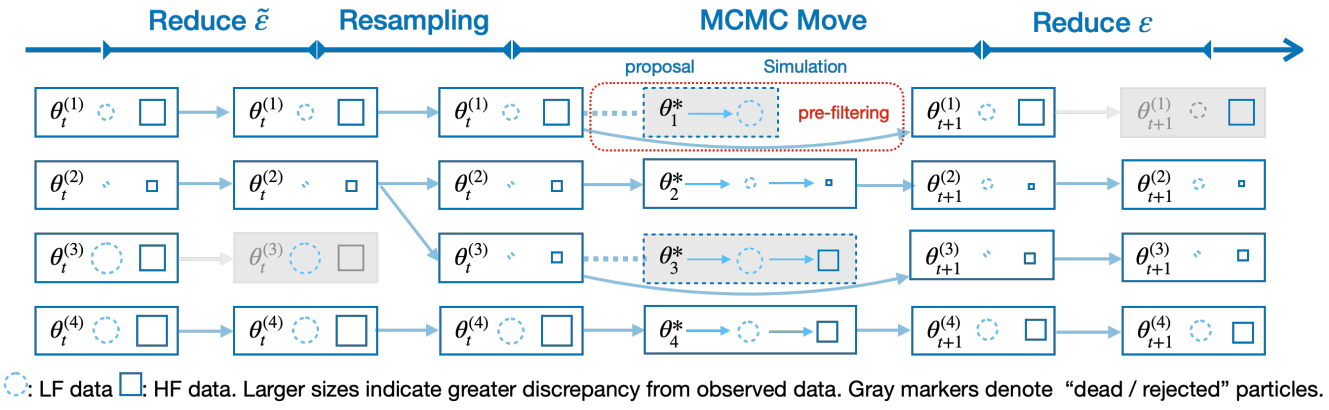


图1：MAPS算法概览（算法 3 的步骤6–10）。给定一组加权粒子 $\boldsymbol{\theta}_t^{1:N}$，MAPS算法通过执行以下步骤来近似 $\boldsymbol{\theta}_{t+1}^{1:N}$：(1) 降低辅助阈值，以筛除低活性粒子；(2) 根据有效样本量进行自适应重采样；(3) 采用低保真预筛选和高保真细化进行多保真马尔可夫链蒙特卡洛移动；(4) 最终降低阈值。

Algorithm 3 MAPS : Mu低保真elity ABC with pre-filtering SMC

输入： 先验： $\pi(\cdot)$，低保真模拟器 $\tilde{p}(\cdot \mid \cdot)$，高保真模拟器 $p(\cdot \mid \cdot)$，观测数据 $\boldsymbol{y}$，目标阈值 ε_T，每次迭代的粒子数 N，有效样本量阈值 N_T，保留的活跃粒子比例 α（基于高保真数据）和 α_L（基于低保真数据）。

输出： $\left\{\boldsymbol{\theta}_t^{(i)}, W_t^{(i)}\right\}_{i=1}^{N}$。1: 初始化 $\varepsilon_0 = \infty$, $\tilde{\varepsilon}_0 = \infty$, $t = 0$。2: 对于 $i = 1, \ldots, N$，执行 3: 抽样 $\boldsymbol{\theta}_0^{(i)} \sim \pi(\cdot)$ 并设置 $W_0^{(i)} = 1/N$。%{$(\tilde{W}_0^{(i)}, \boldsymbol{\theta}_0^{(i)})\}_{i=1:N} \sim \pi_{\varepsilon_0}$。4: 生成 $\tilde{\boldsymbol{x}}_{0,1:n_L}^{(i)} \sim \tilde{p}(\cdot \mid \boldsymbol{\theta}_0^{(i)})$ 和 $\boldsymbol{x}_{0,1:n_H}^{(i)} \sim p(\cdot \mid \boldsymbol{\theta}_0^{(i)})$。

5: **结束循环**

6: **当** $\varepsilon_t > \varepsilon_T$ **时**

7: 设定 $\tilde{w}_i \propto W_t^{(i)} \sum_k \mathbb{I}(\Delta(\boldsymbol{x}_{t,k}^{(i)}, \boldsymbol{y}) < \varepsilon_T) / \sum_k \mathbb{I}(\Delta(\boldsymbol{x}_{t,k}^{(i)}, \boldsymbol{y}) < \varepsilon_t)$ $\tilde{\varepsilon} = Q_{1-\alpha_L}\left(\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_{i,k}^{(i)}, \boldsymbol{y}) \ \tilde{w}_i\right\}\right)$。8: 设定 $t = t + 1$，通过令 $\mathrm{PA}\left(\left\{\tilde{W}_t^{(i)}\right\}, \tilde{\varepsilon}_t\right) = \alpha_L \mathrm{PA}\left(\left\{W_{t-1}^{(i)}\right\}\right)$，（若 $\tilde{\varepsilon}_t < \tilde{\varepsilon}$）来确定 $\tilde{\varepsilon}_t$，并设定 $\tilde{\varepsilon}_t = \tilde{\varepsilon}_T$，其中 $\tilde{W}_t^{(i)} \propto W_{t-1}^{(i)} \mathbb{I}\left\{\min_k \Delta(\tilde{\boldsymbol{x}}_{t-1,k}^{(i)}, \boldsymbol{y}) < \tilde{\varepsilon}_t\right\}$。%{$(\tilde{W}_t^{(i)}, \boldsymbol{\theta}_{t-1}^{(i)})\}_{i=1:N} \sim \pi_{\varepsilon_{t-1}, \tilde{\varepsilon}_t}$。9: 若有效样本量 $\left\{\tilde{W}_t^{(i)}\right\} < N_T$），则 10: 从 $\left\{\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}\right\}$ 中按权重 $\left\{\tilde{W}_t^{(i)}\right\}$ 重采样 N 个粒子，也不严格地记作 $\left\{\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}\right\}$，并设定 $\tilde{W}_t^{(i)} = \frac{1}{N}$。11: 否则 12: 设定 $\left\{\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}\right\} = \left\{\boldsymbol{\theta}_{t-1}^{(i)}, \boldsymbol{x}_{t-1,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t-1,1:n_L}^{(i)}\right\}$，且 $\left\{\tilde{W}_t^{(i)}\right\} = \left\{\tilde{W}_t^{(i)}\right\}$ 13: 结束条件判断 14: 对 $i = 1, \ldots, N$ 执行 15: 通过运行 PF-ABC-MCMC（算法 2）获得 $\left\{\boldsymbol{\theta}_t^{(i)}, \boldsymbol{x}_{t,1:n_H}^{(i)}, \tilde{\boldsymbol{x}}_{t,1:n_L}^{(i)}\right\}$。16: 结束循环 17: 通过令 $\mathrm{PA}\left(\left\{W_t^{(i)}\right\}, \varepsilon_t\right) = \alpha \mathrm{PA}\left(\left\{\tilde{W}_t^{(i)}\right\}\right)$ 来确定 ε_t，若 $\varepsilon_t < \varepsilon_T$，则设定 $\varepsilon_t = \varepsilon_T$，其中 %{$(\tilde{W}_t^{(i)}, \boldsymbol{\theta}_t^{(i)})\}_{i=1:N} \sim \pi_{\varepsilon_t, \tilde{\varepsilon}_t}$

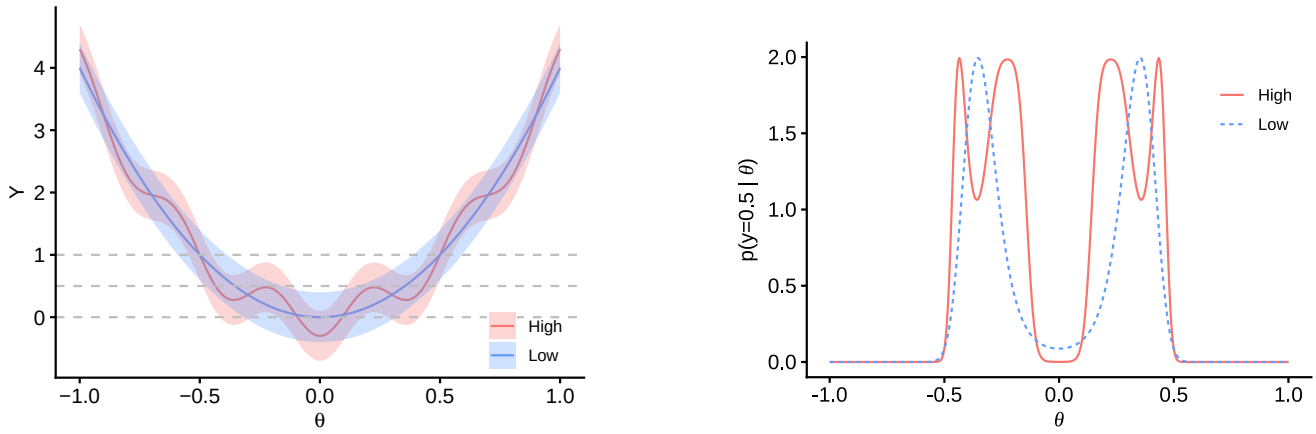
$$W_t^{(i)} \propto \tilde{W}_t^{(i)} \frac{\sum_k \mathbb{I}(\Delta(\boldsymbol{x}_{t,k}^{(i)}, \boldsymbol{y}) < \varepsilon_t)}{\sum_k \mathbb{I}(\Delta(\boldsymbol{x}_{t,k}^{(i)}, \boldsymbol{y}) < \varepsilon_{t-1})}.$$

18: **结束循环**

5 数值实验

5.1 玩具示例

在本节中，我们利用一个玩具示例，展示所提出的 MAPS 算法相较于标准 ABC-ASMC 算法的优势。高保真模拟模型定义为 $x \mid \theta \sim N(4\theta^2 + 0.3\cos(5\pi\theta)\, 0.2_2)$，低保真模拟模型定义为 $x \mid \theta \sim N(4\theta^2, 0.2_2)$。我们将先验分布设为 $\mathcal{U}[-2, 2]$，目标阈值为 $\varepsilon_T = 0.1$。



(a) (b)

图 2：(a) HF 和 LF 模拟数据随参数 θ的变化。实线表示均值；阴影区域表示 95% 置信区间。灰色虚线标示观测值 $y_{\text{obs}} = 0$、0.5 和 1。(b) 两种保真度模型在$y_{\text{obs}} = 0.5$处的似然函数。

图2（a）展示了低保真和高保真模拟数据随参数 θ的变化。总体而言，低保真模拟可作为高保真模拟的近似，但在某些区域存在明显差异。我们考虑了三组观测数据：$y_{\text{obs}} = 0$（高保真似然为多峰，而低保真似然为单峰）、$y_{\text{obs}} = 0.5$（高保真和低保真似然的最大值点不同），以及 $y_{\text{obs}} = 1$（该数据揭示了高保真似然显著大于零、而低保真似然趋近于零的区域）。正文中仅展示 $y_{\text{obs}} = 0.5$的结果，相应的似然函数见图2（b），其余结果见附录 S.2.1 节。

差异函数定义为 $\Delta(x,\ y)\ =(x-y)^2,,$ 每次迭代中的粒子数设为 $N=5120$。在标准 ABC-ASMC 算法中，每个参数的模拟次数为 $n=10$，阈值缩减参数为 $\alpha=0.7$。在 MAPS 算法中，每个参数的模拟次数为 $n_H\ =\ 10,\ \ n_L\ =\ 20$ 阈值缩减参数为 $\alpha=0.7$。我们研究了辅助阈值缩减参数 α_L和辅助临界值参数 a_L 对算法效率和结果准确性的影响。结果见附录 S.2.1 节的图 S.1 和 S.2。根据数值结果，我们建议将 $\alpha_L=\alpha=0.7$ 和 $a_L=0.001$设为相应值。

图 3 展示了 ABC-ASMC 和 MAPS 在各次迭代中获得的后验密度估计，以及真实密度估计。如图所示，通过引入低保真模拟进行预筛选，与 ABC-ASMC 相比，MAPS方法能够使后验密度更快地收敛于真实的 ABC 后验分布。

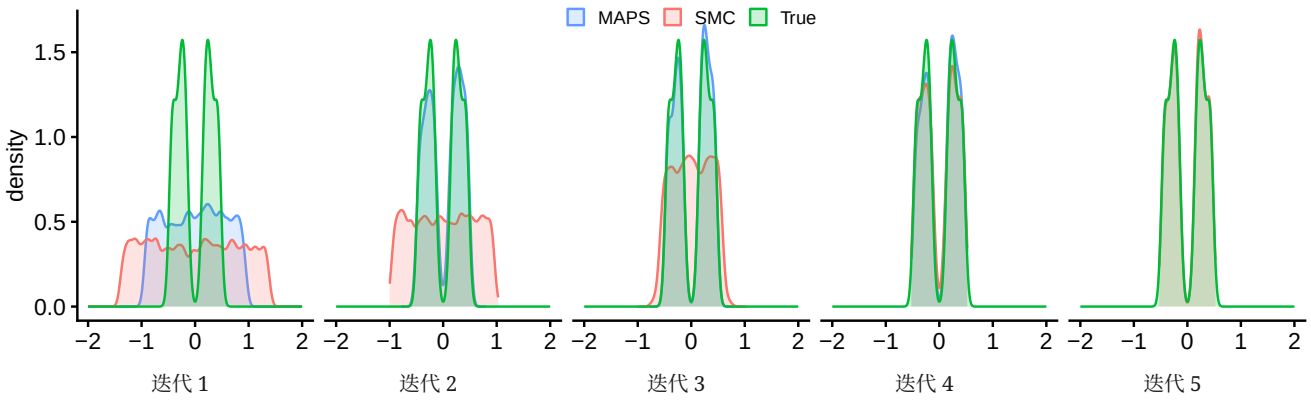


图3：当 $y_{obs}=0.5$ 时，ABC-ASMC 和 MAPS 方法在各次迭代中的后验密度估计与真实后验密度的比较。

图 4 展示了对MAPS和ABC-ASMC重复运行50次所得结果摘要的比较，其中真实后验通过解析计算获得。结果表明，在有效样本量和KL散度方面，MAPS方法优于ABC-ASMC。此外，MAPS所需的高保真模拟次数显著少于ABC-ASMC。中间面板展示了MAPS和ABC-ASMC的后验均值估计值与模拟次数之间的关系，进一步表明，MAPS在大幅减少

高保真模拟次数的同时，能够达到相当甚至更高的推断精度。图4 的最后两个面板展示了阈值ε和高保真模拟次数随迭代次数的变化。由于引入了预筛选，MAPS的阈值 ε 更快收敛到目标值；具体而言，ABC-ASMC需要五次迭代，而MAPS在四次迭代内即可收敛。

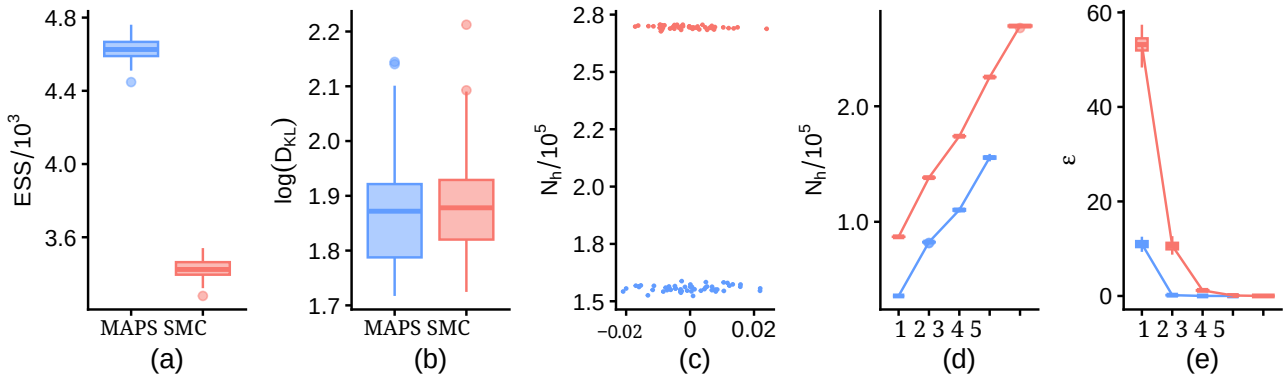


图4：ABC-ASMC 和 MAPS 方法进行 50 次重复实验的结果比较，含$y_{\text{obs}} = 0.5$。(a) 有效样本量，(b) KL散度，(c) 后验均值和高保真模拟总次数，(d) 阈值演化，以及 (e) 高保真模拟次数随迭代次数的变化。

表1 总结了 MAPS 与 ABC-ASMC 方法在三个不同观测数据点上的比较 $y_{\text{obs}} = 1$, $y_{\text{obs}} = 0.5$和 $y_{\text{obs}} = 0$。评估指标包括 Kullback–Leibler（KL）散度、有效样本量和高保真模拟总次数。所有报告值均为 50 次独立重复试验的平均值。百分比变化表示 MAPS 相对于 ABC-ASMC 的提升幅度。结果表明，在所有测试情形中，MAPS 均能显著减少高保真模拟次数。具体而言，与 ABC-ASMC 相比，MAPS 将高保真模拟次数减少了约 34% 至 42%，体现出计算效率的显著提升。在推断准确性方面，以 KL 散度衡量时，MAPS 的表现相当或有所改善，其中 $y_{\text{obs}} = 0.5$的改善幅度最高达 21%。MAPS 的有效样本量通常也更高，表明其后验估计更可靠，其中两种情形的增幅超过 30%。值得注意的是，对于 $y_{\text{obs}} = 0$这一情形，尽管 KL 散度和有效样本量相近或略差，该方法仍显著降低了计算成本。总之，这些结果证实

MAPS 提供了一个效率更高的推断框架，在不损害后验估计质量的前提下显著减少了对高成本高保真模拟的依赖，某些情况下甚至还能提升后验估计质量。

表 1：在不同观测值 y_{obs} 下，SMC 与 MAPS 方法在 KL散度、有效样本量和高保真模拟总次数方面的比较。数值为 50 次重复实验的平均值。百分比变化表示 MAPS 相较于 SMC 的相对改进。

	KL散度			有效样本量			高保真模拟次数		
	SMC	MAPS		SMC	MAPS		SMC	MAPS	
$y_{obs} = 1$	0.04	0.039	**-2.4%**	1157	1614	**39.4%**	327652	196979	**-39.9%**
$y_{obs} = 0.5$	0.071	0.056	**-21.1**%	3426	4628	**35.1**%	269479	155677	**-42.2**%
$y_{obs} = 0$	0.152	0.153	0.1%	3816	3621	-5.1%	319667	210058	**-34.3**%

5.2 Ornstein–Uhlenbeck过程

I在本小节中，我们考虑由下式定义的 Ornstein–Uhlenbeck过程

$$dX_t = \gamma(\mu - X_t)\,dt + \sigma\,dW_t,$$

其中，初始条件的分布为 $X(0) \sim \mathcal{N}(\mu + \mu_{\text{offset}}, 0.1_2)$。这里，$\mu$ 表示长期均值，σ 为扩散系数，γ 表示均值回归速率，μ_{offset}量化初始状态相对于 μ 的偏差。高保真模拟采用 Euler–Maruyama格式近似：

$$X(t + \delta_t) = X(t) + \gamma(\mu - X_t)\,\delta_t + \sigma\sqrt{\delta_t}\,z,$$

wh其中 $z \sim \mathcal{N}(0, 1)$。低保真模型假设其平稳分布为

$$X_t \sim \mathcal{N}\left(\mu, \left(\frac{\sigma}{2.5\gamma}\right)^2\right).$$

参数的先验分布指定如下：

$$\mu \sim \mathcal{U}(0.1, 3), \quad \sigma \sim \mathcal{U}(0.1, 1), \quad \gamma \sim \mathcal{U}(0.1, 2), \quad \mu_{\text{offset}} \sim \mathcal{U}(2, 6).$$

高保真数据记为 $\boldsymbol{x}$，它们是在给定一组参数的情况下，使用时间步长 $\delta_t = 0.01$ ，在区间 $[0, 30]$ 内通过高保真模型生成的。通过按长度为 $t = 0.1$ 的间隔对模型输出进行采样来收集观测值，得到 $\boldsymbol{x} = (x_1, x_2, \ldots\ldots, x_{301})$，其中每个$x_j$ 对应于相应时间点的模型状态。通过使用参数值 $\mu = 2$、$\sigma = 0.5$、$\gamma = 1$ 和 $\mu_{\text{offset}} = 3$ 进行高保真模拟，生成了观测数据 $\boldsymbol{y}$ 。我们将轨迹 $\boldsymbol{x}$ 的汇总统计量 $\boldsymbol{S}(\boldsymbol{x})$ 定义为：

$$S_1(\boldsymbol{x}) = \frac{1}{150}\sum_{j=151}^{301} x_j, \quad S_2(\boldsymbol{x}) = 10 \times \text{sd}(x_{151:301}), \quad S_3(\boldsymbol{x}) = x_1 - S_1(\boldsymbol{x}), \quad S_4(\boldsymbol{x}) = x_1 - x_{21},$$

其中，sd(·) 表示样本标准差。这四个汇总统计量分别作为四个感兴趣参数的粗略估计量。第二个统计量包含一个乘法因子 10，以确保其尺度与其他汇总统计量相当。高保真模拟数据 $\boldsymbol{x}$ 与观测数据 $\boldsymbol{y}$ 之间的差异度量定义为

$$\Delta(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4}\left\|\boldsymbol{S}(\boldsymbol{x}) - \boldsymbol{S}(\boldsymbol{y})\right\|_2^2.$$

对于给定的参数集，低保真数据 $\tilde{\boldsymbol{x}}$ 由低保真模型生成，其中包含 200˜ 个模拟点。相应的汇总统计量 $\boldsymbol{S}(\tilde{\boldsymbol{x}})$ 定义为

$$S_1(\tilde{\boldsymbol{x}}) = \frac{1}{200}\sum_{i=1}^{200} \tilde{x}_i, \quad S_2(\tilde{\boldsymbol{x}}) = 10 \times \text{sd}(\tilde{\boldsymbol{x}}).$$

差异　　低保真模拟数据与观测数据之间的差异通过以下方式度量：

$$\Delta(\tilde{\boldsymbol{x}}, \boldsymbol{y}) = \frac{1}{2}\left\|\tilde{\boldsymbol{S}}(\tilde{\boldsymbol{x}}) - \boldsymbol{S}_{1,2}(\boldsymbol{y})\right\|_2^2,$$

其中$\boldsymbol{S}_{1,\ 2}(\boldsymbol{y})$ 表示 $\boldsymbol{S}(\boldsymbol{y})$ 的前两个分量。图5 展示了使用真实参数值通过高保真模型和低保真模型生成的模拟数据。

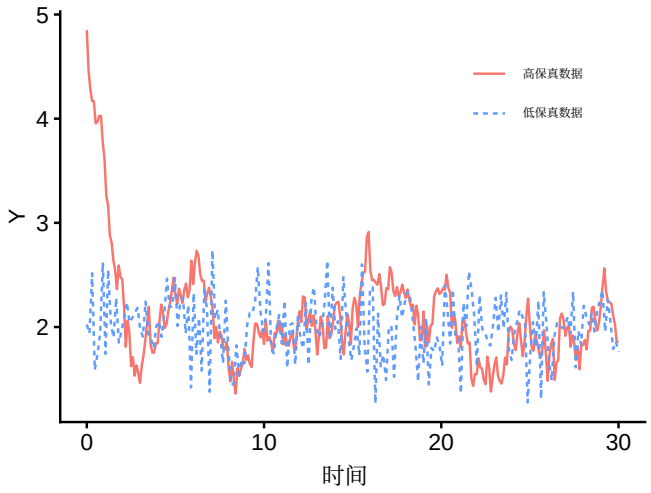


图5：使用参数 $\mu = 2$、 $\sigma = 0.5$、 $\gamma = 1$ 和 $\mu_{\text{offset}} = 3$ 生成的高保真和低保真模拟数据。

图6比较了 MAPS 和 ABC-ASMC 算法的性能。为确保统计稳健性，两种算法均使用 5,120 个粒子，并重复运行 50 次。对于 MAPS 算法，高保真模型和低保真模型的每个参数模拟次数分别设为 $n_H = 10$ 和 $n_L = 20$，阈值缩减参数设为 $\alpha_H = \alpha_L = 0.7$，基准阈值为 $a_L = 0.001$。标准 ABC-ASMC 算法的每个参数模拟次数为 $n = n_H = 10$，阈值缩减参数为 $\alpha = 0.7$。通过 ABC-MCMC 近似真实 ABC 后验分布：使用 10 条链，每条链迭代 500,000 次，并将前 100,000 次迭代作为预热期丢弃。

从图 6所示结果可以看出，MAPS算法以显著减少的高保真模拟次数取得了更优的推断结果，比 ABC-ASMC 少约 44%。这一改进体现为相较于真实ABC后验分布，KL散度始终更低，表明后验近似精度有所提升。此外，图 7显示，MAPS后验分布与真实后验及 ABC-ASMC估计均高度吻合，证实了 MAPS 方法的有效性。

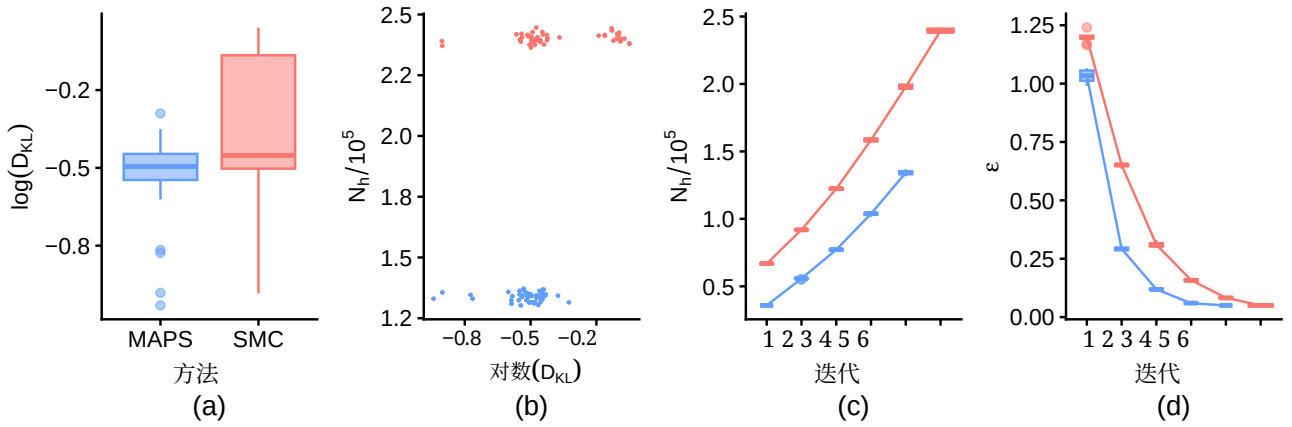
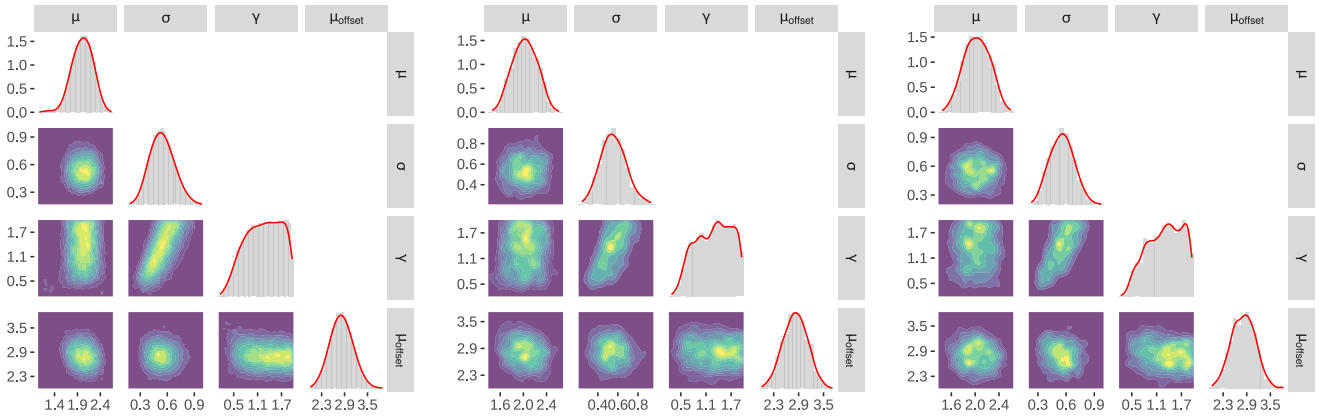


图 6：ABC-ASMC 与 MAPS 方法 50 次重复实验结果的比较。(a) KL散度；(b) KL散度和高保真模拟总次数；(c) 阈值演化；(d) 各次迭代中的高保真模拟次数。



(a) 真实后验　　(b) ABC-ASMC　　(c) MAPS

图 7：比较　　二维和一维边缘后验分布　　.

5.3 Kuramoto 振荡器网络

本节考虑 Prescott 和 BakerPrescott 和 Baker(2021年年）研究的 Kuramoto 振荡器模型。该模型由一个全连接网络构成，包含 M 个振荡器，每个振荡器的相位为 $\varphi_i(t)$，其演化由下式控制

$$\dot{\varphi}_i = \omega_i + \frac{K}{M}\sum_{j=1}^{M}\sin(\varphi_j - \varphi_i), \quad i = 1, \ldots, M, \tag{17}$$

其中， ω_i 是从柯西分布(ω_0, γ)中独立抽取的样本， K 表示耦合强度。初始相位设为零，模拟在 $t \in [0, T]$上进行。

多保真度方法通过监测Daido序参量，利用耦合振子系统的低维表示，这些序参量定义如下：

$$Z_n(t) = \frac{1}{M} \sum_{j=1}^{M} \exp(in\varphi_j(t)),$$

其中， n 是正整数， i 表示虚数单位。具体而言，一阶序参量 $Z_1(t)$通常表示为其幅值 $R(t)$ $= |Z_1(t)|$以及相位 $\Phi(t)$ =arg$(Z_1(t))$。

在Ott–Antonsen拟设下，关系 $Z_n(t) = [Z_1(t)]^n$ 成立，因此系统可简化为一个支配 Z_1的常微分方程。由此，方程（ M）中描述的原17维常微分方程系统可由以下二维常微分方程系统有效近似：

$$\dot{\tilde{R}} = \left(\frac{K}{2} - \gamma\right)\tilde{R} - \frac{K}{2}\tilde{R}^3, \quad \dot{\tilde{\Phi}} = \omega_0, \tag{18}$$

˜ ˜初始条件为 $R(0) = 1$ 和 $\Phi(0) = 0$。这种约化在保留振子系综基本动力学的同时，实现了高效计算。

我们的目标是根据通过模拟 $M = 32$ 个振荡器、使其服从 $\omega_i \sim$ 柯西分布(ω_0, γ) 并在 $t \in (0, 20]$ 上生成的合成数据，推断参数 (K, ω_0, γ)。为对轨迹进行汇总以便比较，我们使用了三种统计量：

$$S_1 = \left(\frac{1}{20}\int_0^{20} R(t)dt\right)^2, \quad S_2 = \frac{\Phi(20) - \Phi(0)}{20}, \quad S_3 = R(T_{1/2}),$$

其中， $T_{1/2}$ 是 R_{obs} 首次达到其初始值与长期平均值中点的时间。

高保真仿真数据（R， Φ）由式（17）通过经典四阶˜ ˜龙格-库塔方法生成，时间步长为0。1。低保真模拟数据（R， Φ）由式（18）采用欧拉法获得，时间步长同样为0。1。观测轨迹R_{obs}（t）和 Φ_{obs}（t）通过在参数值$K = 2$， $\omega_0 = \pi/3$和 $\gamma = 0.1$下模拟高保真模型（式（17））生成。模拟数据与观测数据之间的差异由以下度量量化：

度量为

$$\Delta\big((R,\Phi),(R_{\text{obs}},\Phi_{\text{obs}})\big) = \big\|S\big((R,\Phi)\big) - S\big((R_{\text{obs}},\Phi_{\text{obs}})\big)\big\|_2^2,$$

其中， $S(\cdot)$ 表示汇总统计量向量。

真实参数值下，由低保真模型和高保真模型生成的观测数据与模拟结果如附录第 S.2.2 节的图 S.6 所示。请注意，低保真模拟是确定性的，而高保真模拟包含随机性。进行 5，000 次高保真模拟需要 55.75 秒，而 5，000 次低保真模拟仅需 2.017 秒。

图 8 比较了 MAPS 和 ABC-ASMC 算法的性能。两种算法均使用 2048 个粒子，并重复运行 50 次，以确保统计稳健性。对于 MAPS 算法，每个参数在 $n_H = 5$ 高保真模型中模拟 $n_L = 1$ 次，在低保真模型中模拟 $\alpha_H = \alpha_L = 0.9$ 次；阈值缩减参数为 $a_L = 0.001$，基准阈值为 $n = n_H = 10$。标准 ABC-ASMC 算法每个参数使用 $\alpha = 0.9$次模拟，阈值缩减参数为 $\alpha = 0.9$。真实 ABC 后验分布通过 ABC-MCMC 近似得到，该方法使用 10 条链，每条链运行 500,000 次迭代，并将前 200,000 次迭代作为预热期舍弃。附录图 S.7 和 S.8 展示了双变量边际后验密度随迭代的变化，表明 MAPS 收敛更快。

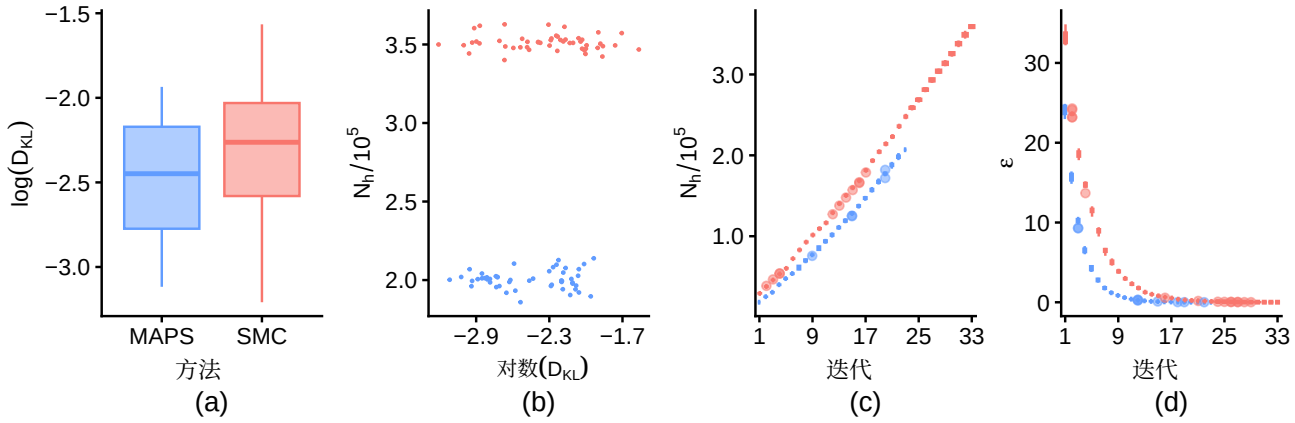
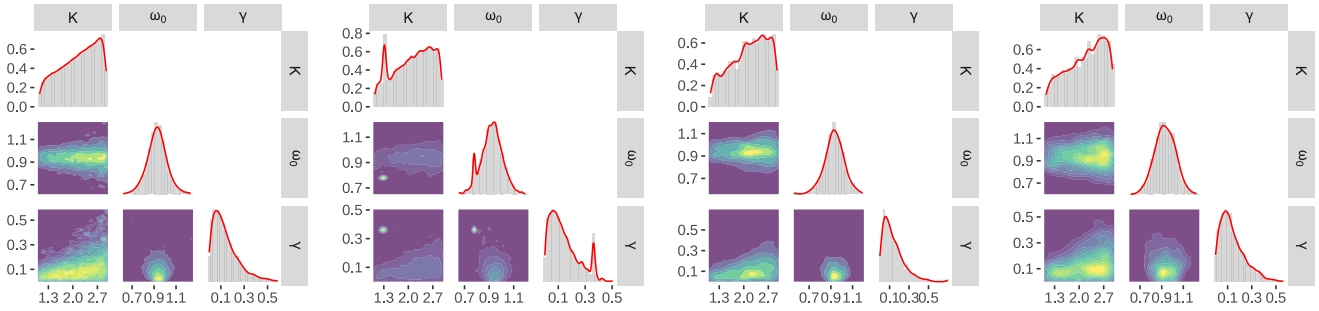


图 8：比　　较 ABC-ASMC 和 MAPS 方法 50 次重复运行的结果　　S.

在图 9中，我们展示了 ABC-ASMC、MAPS 方法、ABC-MCMC 算法以及真实后验分布所得到的后验分布。ABC-MCMC

结果基于 200，000 次迭代，每个参数进行 10 次模拟，因此高保真模拟的总次数与 MAPS 算法相当。前 100，000 次迭代被作为预热期舍弃。与基于 SMC 的方法相比，ABC-MCMC 算法往往会陷入局部众数，从而对其收敛产生负面影响。



(a) 真实后验　(b) ABC-MCMC　(c) ABC-ASMC　(d) MAPS

图 9：不同方法得到的后验估计。

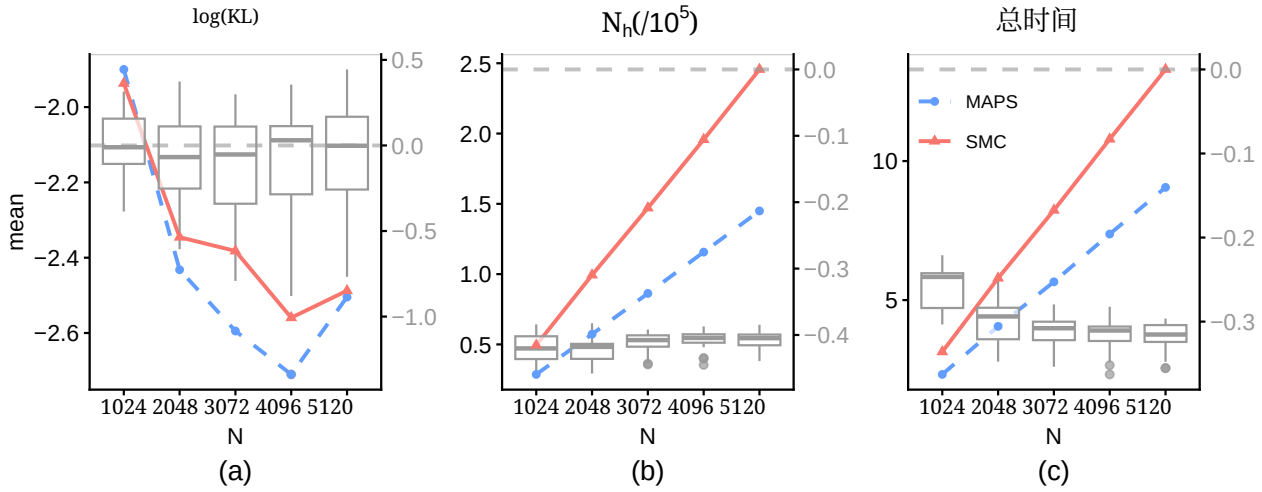


图10：MAPS 与 SMC 随 N with $\alpha = 0.8$ 的变化比较（20 次运行）。曲线表示均值，箱线图表示 MAPS 相对于 SMC 的性能。

此外，我们研究了每次迭代粒子数（N）和阈值缩减参数（α）对SMC推断的影响。当将阈值缩减参数固定为 $\alpha = 0.8$时，图10 表明，随着 N 增大，两者的推断精度

ABC-ASMC与MAPS算法均有所提高，而高保真模拟的成本与总计算时间则呈线性增长。当固定 $N = 3072$时，图11 表明，随着 α 增大，ABC-ASMC与MAPS的高保真仿真成本和计算时间均增加；MAPS的推断精度有所提高，而ABC-ASMC则未呈现明显趋势。图10 和11 的箱线图展示了MAPS相对于SMC的相对性能。结果表明，MAPS达到了与ABC-ASMC相当或更优的推断精度，同时将高保真仿真成本降低约40%，并将总计算时间减少约30%。

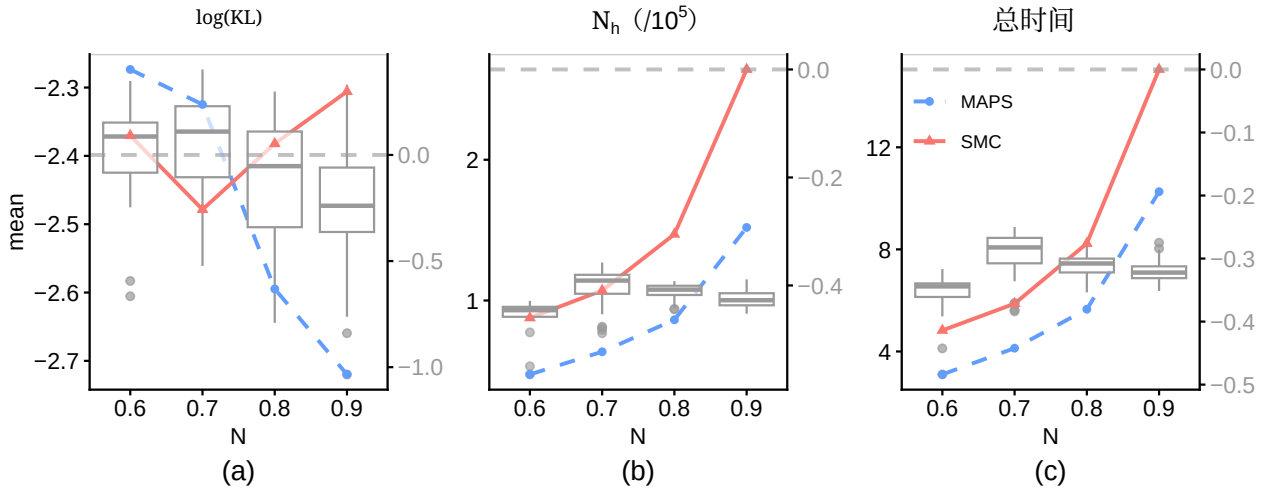


图11：MAPS 与 SMC 随 α with $N = 3072$ 的比较（20 次运行）。线条表示均值，箱线图表示 MAPS 相对于 SMC 的相对性能。

6 讨论

我们研究模拟成本高昂模型的近似贝叶斯推断。我们将多保真度模拟模型整合到近似贝叶斯计算算法中，并提出一种预筛选分层重要性抽样算法，利用低成本低保真模拟减少成本高昂的高保真模拟次数。我们还提出一种策略，用于评估给定多保真度模型是否适用于我们的算法。基于该多保真度模型，我们开发了 MAPS 算法，以自适应地选择预筛选标准。在温和的正则性条件下，我们证明了所提出算法的后验集中性质，并证明超参数 a_L

决定推断误差上界。数值实验验证了这些理论结果，并表明我们的方法在保持与标准近似序贯蒙特卡罗相当的准确性的同时，可将高保真模型评估次数减少约 40%。此外，所提出的方法既适用于低保真模拟具有随机性的模型（第 5.1节和 5.2节），也适用于低保真模拟具有确定性的模型（第 5.3节）。

MAPS算法引入了与预筛选阶段相关的三个参数：每个参数的低保真模拟次数 n_L、辅助阈值的衰减因子 $\tilde{\varepsilon}$（记为 α_L），以及控制辅助阈值边界的因子 a_L。超参数的选择在实践中可能会影响算法性能。我们建议选择衰减因子 α_L，使其与主阈值 ε的尺度保持一致，并将 a_L 设为 0.001。关于 n_L，其选择可依据相对模拟成本来指导：对于每个参数，生成n_L 次低保真模拟的成本不应超过进行 n_H 次高保真模拟成本的0.4倍。

我们提出的方法与Prescott 和 Baker（2020年，2021年）的多保真ABC算法在若干关键方面有所不同。首先，Prescott 和 Baker（2020年）通过结合低保真数据与高保真数据构造似然的无偏估计量，并据此设计了提前接受和提前拒绝算法，以减少高保真模拟的次数。我们的方法则基于利用低保真模型筛除权重极有可能为零的参数。在此基础上，Prescott 和 Baker（2021年）提出了一种多保真近似贝叶斯计算序贯蒙特卡罗算法，该算法依赖于固定的阈值序列，并采用核密度估计来近似当前样本分布，将其作为下一个中间目标分布的重要性提议分布。相比之下，我们的方法通过控制活跃粒子的数量自适应地降低阈值。随后相应地更新粒子权重，并采用重采样与马尔可夫链蒙特卡罗传播相结合的方式，从后续中间目标分布中生成加权样本。Prescott 和 Baker（2021年）的多保真 ABC 方法在理论上是严谨的，并已证明其有效性，但在实践中可能会产生带有负权重的样本。

未来工作中还可以进行多项改进和扩展。本文提出的方法不涉及低保真模型的开发，并假设该模型已先验已知。然而，在一些实际应用中，推导低保真模型可能并不容易。未来，值得探索利用机器学习方法构建低保真模型，并将其结合

近似贝叶斯计算用于推断。文献中的多保真度推断方法通常依赖单个低保真模型来辅助推断（Prescott 和 Baker，2020年；Cao 等，2025年）。然而，准确性与计算速度之间存在权衡。更简单的模型通常速度更快，但准确性较低。另一个未来研究方向是将多个保真度级别的模型纳入近似贝叶斯推断框架。

致谢

本研究得到国家自然科学基金委员会（12131001 和 12101333）、上海科技大学的启动经费和机构发展基金及高性能计算平台、中央高校基本科研业务费、LPMC 和 KLMDASR 的资助。作者按姓氏字母顺序排列。

利益披露声明

作者声明不存在需要披露的竞争性利益。

补充材料

附录 S.1： 正文中所呈现的所有理论结果的证明。

附录 S.2： 正文中未展示的一些数值结果。

Appendix S.3： 正文中未展示的一些算法说明 .

附录 S.4： 论文中使用的记号汇总。

参考文献

Beaumont, M. A., W. Zhang, 和 D. J. Balding（2002年12月）。群体遗传学中的近似贝叶斯计算。*Genetics 162*(4)，2025–2035。

Blum, M. G., M. A. Nunes, D. Prangle, 和 S. A. Sisson（2013）。近似贝叶斯计算中降维方法的比较综述。《统计科学》*28*(2)，189–208。´Boelts, J., M. Deistler, M. Gloeckler, A. Tejero-Cantero, J.-M. Lueckmann, G. Moss, P. Steinbach, T. Moreau, F. Muratore, J. Linhart 等（2024）。sbi reloaded：用于基于模拟的推断工作流的工具包。*arXiv*预印本 *arXiv:2411.17337*。Cao, X., S. Wang, 和 Y. Zhou（2025年）。利用提前拒绝的马尔可夫链蒙特卡罗和高斯过程加速ABC方法。计算与图形统计学杂志 *34*(2)，395–408。

Christen, J. A. 与 C. Fox（2005年）。使用近似的马尔可夫链蒙特卡洛。《计算与图形统计学杂志》*14*(4)，795–810。

Del Moral, P.、A. Doucet 与 A. Jasra（2012年）。用于近似贝叶斯计算的自适应序贯蒙特卡洛方法。《统计与计算》*22*(5)，1009–1020。

Everitt, R. G. 与 P. A. Rowi´nska（2021年）。延迟接受ABC-SMC。《计算与图形统计学杂志》*30*(1)，55–66。

Frazier, D. T.、G. M. Martin、C. P. Robert 与 J. Rousseau（2018 年）。近似贝叶斯计算的渐近性质。《生物计量学》*105*(3)，593–607。

Hermans, J.、V. Begy 与 G. Louppe（2020年）。基于摊销近似比率估计器的无似然 MCMC。载于国际机器学习会议，第4239–4248页。PMLR。

Jasra, A.、S. Jo、D. Nott、C. Shoemaker 和 R. Tempone（2019年）。近似贝叶斯计算中的多层蒙特卡洛。《随机分析与应用》*37*(3), 346–360。

Krouglova, A. N.、H. R. Johnson、B. Confavreux、M. Deistler 和 P. J. Gonc̸alves（2025 年）。面向计算开销大的模拟器的多保真基于模拟的推断。*arXiv* 预印本 *arXiv:2502.08416* 。

Marin, J.-M., P. Pudlo, C. P. Robert 和 R. J. Ryder（2012年）。近似贝叶斯计算方法。《统计与计算》*22*(6), 1167–1180。

Marjoram, P., J. Molitor, V. Plagnol 和 S. Tavar´e（2003）。无需似然的马尔可夫链蒙特卡洛。美国国家科学院院刊 *100*(26), 15324–15328。Papamakarios, G. 和 I. Murray（2016）。利用贝叶斯条件密度估计对模拟模型进行快速的无 ε 推断。神经信息处理系统进展 *29*。Papamakarios, G., D. Sterratt 和 I. Murray（2019）。序贯神经似然：利用自回归流进行快速的无似然推断。载于第 *22* 届人工智能与统计国际会议，第 837–848 页。PMLR。

Picchini, U.（2014）。通过近似贝叶斯计算对随机微分方程模型进行推断。《计算与图形统计学杂志》*23*(4), 1080–1100。

Prescott, T. P. 和 R. E. Baker（2020）。多保真度近似贝叶斯计算。*SIAM/ASA* 不确定性量化杂志 *8*(1), 114–138。

Prescott, T. P. 和 R. E. Baker（2021）。采用序贯蒙特卡洛参数抽样的多保真度近似贝叶斯计算。*SIAM/ASA* 不确定性量化杂志 *9*(2), 788–817。

Pritchard, J. K., M. T. Seielstad, A. Perez-Lezaun 和 M. W. Feldman（1999, 12）。人类 Y 染色体的种群增长：一项关于 Y 染色体微卫星的研究。《分子生物学与进化》*16*(12), 1791–1798。

Sisson, S. A. 和 Y. Fan（2010）。无似然马尔可夫链蒙特卡洛。载于 S. Brooks、G. L. Jones 和 X.-L. Meng（编），《马尔可夫链蒙特卡洛手册》。Chapman and Hall/CRC Press。即将出版。

Sisson, S. A.、Y. Fan 和 M. Beaumont（2018 年）。《近似贝叶斯计算手册》。CRC出版社。

Sisson, S. A., Y. Fan 和 M. M. Tanaka（2007年）。无似然的序贯蒙特卡洛。美国国家科学院院刊 *104*(6)，1760–1765。Warne, D. J., R. E. Baker 和 M. J. Simpson（2018 年）。用于近似贝叶斯计算的多层拒绝抽样。计算统计学与数据分析 *124*，71–86。

Wegmann, D., C. Leuenberger 和 L. Excoffier（2009年，08月）。结合马尔可夫链蒙特卡洛且无需似然的高效近似贝叶斯计算。*Genetics 182*(4)，1207–1218。