Statistica Sinica **34** (2024), 1-22
doi:https://doi.org/10.5705/ss.202022.0256

DIFFERENTIABLE PARTICLE FILTERS WITH SMOOTHLY JITTERED RESAMPLING

Yichao Li[*#1,2], Wenshuo Wang[#2], Ke Deng[1] and Jun S. Liu[2]

[1] *Tsinghua University and* [2] *Harvard University*

Abstract: Particle filters, also known as sequential Monte Carlo, are a powerful computational tool for making inference with dynamical systems. In particular, it is widely used in state space models to estimate the likelihood function. However, estimating the gradient of the likelihood function is hard with sequential Monte Carlo, partially because the commonly used reparametrization trick is not applicable due to the discrete nature of the resampling step. To address this problem, we propose utilizing the smoothly jittered particle filter, which smooths the discrete resampling by adding noise to the resampled particles. We show that when the noise level is chosen correctly, no additional asymptotic error is introduced to the resampling step. We support our method with simulations.

Key words and phrases: Reparametrization trick, resampling, sequential Monte Carlo, state space models.

1. Introduction

Computing or estimating the likelihood function and its gradient for a statistical model with hidden variables is a challenge with applications in many areas of statistics and machine learning (Andrieu, Doucet and Tadic (2005); Kingma and Welling (2013); Le et al. (2017); Mohamed et al. (2020)). In a general form, such a likelihood function can be written as

$$L(\theta \mid y) = \int p(y, z; \theta) dz, \tag{1.1}$$

where $p(y, z; \theta)$ is the joint likelihood of the observed variables y and the unobserved latent variables z with θ as parameters. In practice, the integral is usually estimated by a Monte Carlo method, typically a form of importance sampling:

$$\hat{L}(\theta \mid y) = \frac{1}{n} \sum_{i=1}^{n} \frac{p(y, z_i; \theta)}{q(z_i; \theta)} \quad \text{with} \quad z_i \sim q, \tag{1.2}$$

where q is the sampling distribution of z_i's which could potentially depend on θ. The gradient of the likelihood function can be estimated by differentiating $\hat{L}(\theta \mid y)$ with respect to θ, provided that the sampling distribution q is differentiable (in

*Corresponding author.
#Contributed equally to this work.