Resampling and genealogies in sequential Monte Carlo algorithms

Susanna Elizabeth Brown

A thesis submitted for the degree of
Doctor of Philosophy in Statistics

University of Warwick, Department of Statistics

Janurary 2021

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank...

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

Parts of this thesis have been published by the author:

Abstract

2 Background

Anyone who considers arithmetical methods of producing
random digits is, of course, in a state of sin.

John von Neumann

2.1 State space models

[Introduction of state space models: target tracking example (1D train); modelling assumptions for this example; why Bayesian? Scenarios where model is tractable: Kalman filter, extended KF, etc. Brief discussion of scope of SMC beyond SSMs.]

2.2 Sequential Monte Carlo

[Motivation in context of SSMs (IS, SIS, SIR). Generic SMC algorithm (possibly introduce bootstrap PF first). Discussion of each step. Theoretical justifications for SMC (e.g. convergence results).]

2.3 Coalescent theory

[Review of literature from population genetics, introducing the relevant population models (Wright–Fisher, Moran, Cannings) and Kingman's coalescent / n-coalescent. Domain of attraction of Kingman's coalescent, as far as previous works have shown. Properties of such models (neutrality, Markov property) that may be violated by SMC systems.]

2.4 SMC genealogies & ancestral degeneracy

[Description of how genealogies are induced by SMC algorithms and how this is related to the performance of the algorithms (ancestral degeneracy, variance estimation, storage cost). Existing results characterising these genealogies.

Ways to mitigate ancestral degeneracy (low-variance resampling, adaptive resampling, backward sampling).]

2.5 Resampling

[Definition of a 'valid' resampling scheme and justification for these restrictions. Tour of key resampling schemes (multinomial, residual, stratified, systematic, ...), with discussion of their properties, implementation and usage in practice. Idea of 'optimality' in resampling, description of so-called optimal schemes. Existing results and conjectures comparing the performance of different schemes. Introduction of stochastic rounding as a class of resampling schemes. Adaptive resampling.
Examples and discussion of resampling schemes that violate the three properies; optimal transport resampling, along with the others I mentioned in previous writings.]

2.6 Conditional SMC & particle MCMC

Motivation and definition of particle Gibbs algorithm, and how CSMC crops up in it. Ancestor sampling.