Doctopic: Optimization and Control ARTICLE IN PRESS YJMAA:20986

J. Math. Anal. Appl. • • • (• • • •) • • •–• • •

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Convergence rates for residual branching particle filters ☆

Michael A. Kouritzin

Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada

A R T I C L E I N F O

Article history:
Received 13 March 2016
Available online xxxx
Submitted by V. Pozdnyakov

Keywords:
Branching process
Central limit theorem
Coupling
McKean–Vlasov
Sequential Monte Carlo
Marcinkiewicz strong laws of large numbers

A B S T R A C T

A large class of proven discrete-time branching particle filters with Bayesian model selection capabilities and effective resampling is analyzed mathematically. The particles interact weakly in the branching procedure through the total mass process in such a way that the expected number of particles can remain constant. The weighted particle filter, which has no resampling, and the fully-resampled branching particle filter are included in the class as extreme points. Otherwise, selective residual branching is used allowing any number of offspring. Each particle filter in the class is coupled to a McKean–Vlasov particle system, corresponding to a reduced, unimplementable branching particle filter, for which Marcinkiewicz strong laws of large numbers (Mllns) and the central limit theorem (clt) can be written down. Coupling arguments are used to show the reduced system can be used to predict performance of and to transfer the Mllns to the real weakly-interacting residual branching particle filter. This clt is also shown transferable when (a few) extra particles are used.

1. Introduction

Sequential Monte Carlo (SMC) algorithms are used in diverse problems like tracking, prediction, parameter estimation, model calibration, classification, Bayesian model selection and imaging (see e.g. [18, 17,13,9] for sample applications). Branching SMC algorithms have the advantage that offspring generation only depends upon the parent not the whole population and the disadvantage of having randomly-varying populations (i.e. particle numbers). Recently, Kouritzin [10] introduced four new classes of branching sequential Monte Carlo algorithms that were designed to limit wide particle variations. The tracking and model selection performance of all four algorithms was shown experimentally to be superior to a collection of popular resampled particle algorithms and these four branching algorithms have even greater advantages when it comes to distributed implementations (see Kouritzin [12]). However, there is little theory to back up these experimental findings. Theoretical rate-of-convergence results are desired to understand why these

☆ Supported in part by an NSERC Discovery Grant.
E-mail address: michaelk@ualberta.ca.

http://dx.doi.org/10.1016/j.jmaa.2016.12.046
0022-247X/

Please cite this article in press as: M.A. Kouritzin, Convergence rates for residual branching particle filters, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2016.12.046

algorithms perform so well and what their weaknesses might be. Unfortunately, the branching algorithms lack the independence and fixed particle numbers of many resampled algorithms so their analysis is necessarily difficult and the desired convergence results hard to come by. Herein, we start the theoretical study by establishing Marcinkiewicz strong laws of large numbers (Mllns) and a central limit theorem (clt) for the residual branching algorithm, which is the simplest of the four branching algorithms introduced in [10]. We get around the lack of independence by using exchangeability techniques and by introducing an unrealizable approximate McKean–Vlasov particle system (originally motivated by the work of McKean [16]) which has independence.

The weighted particle filter, largely credited to Handschin [6] as well as Handschin and Mayne [7], approximates the unnormalized filter, denoted σ_n below. This weighted particle filter is the most basic particle filter and is embarrassingly computer parallelizable. However, it is well known to suffer particle spread issues that have to be corrected by branching or resampling. Branching particle filters, like those introduced by Crisan and Lyons [3], can have effective resampling yet still be highly parallelizable. Nonetheless, these early branching particle filters generally have very unstable particle numbers, which affect performance adversely. Recently, Kouritzin [10] introduced four successively more refined branching particle filters with the aim of reducing particle number fluctuations and thereby improving performance and reliability. Even the simplest of these four, the Residual Branching Particle Filter, was shown in [10,12] to avoid wild particle swings and to outperform many popular sequential Monte Carlo methods by a large amount. Herein, we analyze this Residual Branching Particle Filter by way of Marcinkiewicz strong laws of large numbers (Mllns) and the central limit theorem (clt). As a consequence, we also layout a framework for further analysis of the Residual Branching filter as well as the three more-complicated improvements of this filter given in [10].

The *bootstrap particle filter algorithm* was introduced in 1993 by Gordon, Salmond and Smith [5]. It has been improved by using residuals and stratified random variables. This collection *resampled* particle filters is one of the big breakthroughs in big data sequential estimation and their convergence properties have been thoroughly studied by many authors (see e.g. Douc et al. [4]). In particular, Chopin [2] obtained a clt for the residual improvement of the bootstrap algorithm. However, these particle filters approximate the actual filter π_n not the unnormalized σ_n, do not have the (same degree of) ancestral dependence as the Residual Branching filter and base their resampling decisions upon the (locations of the) whole population. Hence, their analysis is quite different from what is required for the Residual Branching particle filter.

In terms of convergence results for branching filters to the unnormalized filter, Kouritzin and Sun [11] obtain L_2-rates of convergence for a partially-resampled branching algorithm. However, no other results were attained and their results are in a specific setting. From a mathematical perspective our work might be closest to Kurtz and Xiong [14,15]. Their work applies to a more general setting than nonlinear filtering but in the non-linear filtering setting it only considers the weighted particle filter. Consequently, substantially new methods are required herein. We make use of classical exchangeability works like Weber [19] and McKean–Vlasov equations as in [16]. However, several new (at least to particle filtering) ideas including branching particle filter coupling, use of infinite branching particle systems, use of tracking systems and Hoeffding-inequality-based particle system bounding are also utilized.

For motivational purposes, we consider tracking a non-observable, random, dynamic signal X given the history of a distorted, corrupted partial observation process Y living on the same probability space $(\Omega, \mathcal{F}, P)$ as X. For many practical problems the signal is a time-homogeneous discrete-time Markov process $\{X_n,\ n = 0, 1, 2, ...\}$, living on some complete, separable metric space (E, ρ), with initial distribution π_0 and transition probability kernel K. The observation process takes the form ($Y_0 = 0$ and) $Y_n = h(X_{n-1}) + V_n$ for $n \in \mathbb{N}$, where $\{V_n\}_{n=1}^{\infty}$ are independent random vectors with common *strictly positive, bounded* density g that are independent of X, and the sensor function h is a measurable mapping from E to $\mathbb{R}^d$. (Such g still allows popular observation noise like Gaussian or Cauchy distributed ones.) Then, the objective of filtering is to compute the conditional expectations $\pi_n(f) = E^P\left(f(X_n)\middle|\mathcal{F}_n^Y\right)$ for all bounded, measurable functions $f : E \to \mathbb{R}$, where $\mathcal{F}_n^Y \doteq \sigma\{Y_l, l = 1, ..., n\}$ is the information obtained from the back observations.

Please cite this article in press as: M.A. Kouritzin, Convergence rates for residual branching particle filters, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2016.12.046