Statistica Sinica **34** (2024), 1-21
doi:https://doi.org/10.5705/ss.202022.0243

A DIVIDE-AND-CONQUER SEQUENTIAL MONTE CARLO APPROACH TO HIGH-DIMENSIONAL FILTERING

Francesca R. Crucinio* and Adam M. Johansen

University of Warwick

Abstract: We propose a divide-and-conquer approach to filtering. The proposed approach decomposes the state variable into low-dimensional components, to which standard particle filtering tools can be successfully applied, and recursively merges them to recover the full filtering distribution. This approach is less dependent on factorizing transition densities and observation likelihoods than are competing approaches, and can be applied to a broader class of models. We compare the performance of the proposed approach with that of state-of-the-art methods on a benchmark problem, and show that the proposed method is broadly comparable in settings in which the other methods are applicable, and that it can be applied in settings in which they cannot.

Key words and phrases: Data assimilation, marginal particle filter, particle filtering, spatio-temporal models, state-space model.

1. Introduction

Particle filters (PFs), an instance of sequential Monte Carlo (SMC) methods, are a popular class of algorithms for performing state estimation for state-space models (SSM), also known as general-state-space hidden Markov models. We consider the class of SSMs with a latent $\mathbb{R}^d$-valued process $(X_t)_{t\geq 1}$ and conditionally independent $\mathbb{R}^p$-valued observations $(Y_t)_{t\geq 1}$. Such an SSM $(X_t, Y_t)_{t\geq 1}$ is defined by the transition density $f_t(x_{t-1}, x_t)$ of the latent process, with the convention that $f_1(x_0, x_1) \equiv f_1(x_1)$, and by the observation likelihood $g_t(y_t|x_t)$. In this work, we are interested in approximating the sequence of filtering distributions, $(p(x_t|y_{1:t}))_{t\geq 1}$, that is, at each time t, the distribution of the latent state at that time given the observations obtained by that time.

Basic PF algorithms are known to suffer from the curse of dimensionality, requiring an exponential increase in computational requirements as the dimension d grows, limiting their applicability to large systems (Rebeschini and Van Handel (2015); Bengtsson, Bickel and Li (2008)). Although the ensemble Kalman filter (Evensen (2009)) can handle high-dimensional problems, it involves approximations that do not disappear, even asymptotically and does not perform well if the model is far from being linear and Gaussian (Lei, Bickel and Snyder (2010)).

To extend the use of PFs to high-dimensional problems, it is natural to

*Corresponding author