Journal of Computational Physics 368 (2018) 154–178

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Multilevel Sequential^2 Monte Carlo for Bayesian inverse problems

Jonas Latz [a], Iason Papaioannou [b], Elisabeth Ullmann [a,*]

[a] Chair of Numerical Analysis, TU München, Boltzmannstr. 3, 85748 Garching b.M., Germany
[b] Engineering Risk Analysis Group, TU München, Theresienstr. 90, 80333 München, Germany

A R T I C L E   I N F O

Article history:
Received 27 September 2017
Accepted 6 April 2018
Available online 22 April 2018

Keywords:
Uncertainty quantification
Partial differential equation
Finite element method
Particle filter
Sequential importance sampling
Tempering

A B S T R A C T

The identification of parameters in mathematical models using noisy observations is a common task in uncertainty quantification. We employ the framework of Bayesian inversion: we combine monitoring and observational data with prior information to estimate the posterior distribution of a parameter. Specifically, we are interested in the distribution of a diffusion coefficient of an elliptic PDE. In this setting, the sample space is high-dimensional, and each sample of the PDE solution is expensive. To address these issues we propose and analyse a novel Sequential Monte Carlo (SMC) sampler for the approximation of the posterior distribution. Classical, single-level SMC constructs a sequence of measures, starting with the prior distribution, and finishing with the posterior distribution. The intermediate measures arise from a tempering of the likelihood, or, equivalently, a rescaling of the noise. The resolution of the PDE discretisation is fixed. In contrast, our estimator employs a hierarchy of PDE discretisations to decrease the computational cost. We construct a sequence of intermediate measures by decreasing the temperature or by increasing the discretisation level at the same time. This idea builds on and generalises the multi-resolution sampler proposed in P.S. Koutsourelakis (2009) [33] where a bridging scheme is used to transfer samples from coarse to fine discretisation levels. Importantly, our choice between tempering and bridging is fully adaptive. We present numerical experiments in 2D space, comparing our estimator to single-level SMC and the multi-resolution sampler.

1. Introduction

In science and engineering we use mathematical models to simulate and understand physical processes. These models require input parameters. Once the parameters are specified we can solve the so-called forward problem to obtain output quantities of interest. In this work we focus on models that involve partial differential equations (PDEs). To date approximate forward solvers are available for many PDE-based models, and output quantities of interest can be approximated efficiently. In contrast, the identification of input parameters (the inverse problem) is more challenging. Often the physical process is only given implicitly by observations (data, measurements). These measurements are typically noisy and/or sparse, and do not contain sufficient information on the underlying parameter or are disturbed in such a way that the true parameter cannot be recovered at all. The inverse problem is ill-posed.

* Corresponding author.
E-mail addresses: jonas.latz@ma.tum.de (J. Latz), iason.papaioannou@tum.de (I. Papaioannou), elisabeth.ullmann@ma.tum.de (E. Ullmann).

https://doi.org/10.1016/j.jcp.2018.04.014

A classical example is the simulation of steady-state groundwater flow to assess the safety of proposed long-term radioactive waste repositories. The quantity of interest is the travel time of radioactive particles to the boundary of a safety zone. The simulation requires the hydraulic conductivity of the ground; it can be observed implicitly by pumping tests, and by pressure measurements. The objective of the groundwater flow inverse problem is the identification of the conductivity. In this example, the mathematical model involves an elliptic PDE. The groundwater flow inverse problem is well known, see e.g. [13,14,36,45,47].

In contrast to deterministic regularisation techniques, the Bayesian approach to inverse problems uses the probabilistic framework of Bayesian inference. Bayesian inference is built on *Bayes' Formula* in the formulation given by Laplace [34, II.1]. We remark that other formulations are possible, see e.g. the work by Matthies et al. [38]. We make use of the mathematical framework for Bayesian Inverse Problems (BIPs) given by Stuart [48]. Under weak assumptions – which we will give below – one can show that the BIP is well-posed. The solution of the BIP is the conditional probability measure of the unknown parameter given the observations.

The Bayesian framework is very general and can handle different types of forward models. However, in this work we consider PDE-based forward models, and in particular an elliptic PDE. The exact solution of the associated BIP is often inaccessible for two reasons: (i) there is no closed form expression for the posterior measure, and (ii) the underlying PDE cannot be solved analytically. We focus on (i), and study efficient approximations to the *full* posterior measure. Alternatively, one could also only approximate the expectation of output quantities of interest with respect to the posterior measure, or estimate the model evidence, the normalization constant of the posterior measure.

Typically, BIPs are approached with sampling based methods, such as Markov Chain Monte Carlo (MCMC) or Importance Sampling. Classical MCMC samplers are the algorithms suggested by Metropolis et al. [39] and the generalisation by Hastings [26]. Advanced MCMC methods for BIP settings are Hamiltonian Monte Carlo [7] and preconditioned Crank–Nicholson MCMC [6,11]. A disadvantage of MCMC samplers is the fact that it is often difficult to assess their convergence after an initial burn-in phase. Importance Sampling [1] on the other hand does not require burn-in. However, Importance Sampling is inefficient if the sampling density differs significantly from the target density. For these reasons we employ Sequential Monte Carlo (SMC) [10,15,41] to approximate the posterior measure. SMC was initially developed to approximate sequences of measures which arise from time-dependent estimation problems in data assimilation. In our setting, since the elliptic PDE models a steady-state process the SMC sequences are constructed artificially such that, starting from the prior measure, they gradually approach the posterior measure. Artificial sequences of measures arise also in simulated annealing [15], the estimation of rare events [43], model selection [50], and bridging [20].

In some situations it is convenient to determine the artificial sequences "on the fly" during the execution of the algorithm. The associated method is termed *adaptive SMC*; see [19,29] for a discussion, and [2] for a careful analysis. A well-known drawback is the fact that adaptive SMC returns a biased model evidence estimate, however, the model evidence is not the major focus of our work. The estimation of the model evidence with non-adaptive SMC is discussed in [20,42].

The major advantage of SMC is its dimension-independent convergence which is often observed in practise and which can be proved e.g. for uniformly bounded update densities [5]. Thus SMC can be used in high- and infinite dimensional settings. See [44] for a discussion of this point. Similar results are also known for the Ensemble Kalman Filter (EnKF) applied to linear inverse problems with a finite number of particles [47]. The EnKF is a linearised version of SMC and has been applied to linear and nonlinear inverse problems (see [28]).

SMC has already been used to solve BIPs where the forward model is an elliptic [5] or Navier–Stokes equation [31]. The computational challenge is that PDE-based forward solves are in general very expensive. Thus every sample, required by standard solvers such as MCMC or SMC, is expensive. The total computational budget might allow only a few samples and thus the sample error can be considerably large. We handle this problem by constructing a *multilevel* SMC sampler. To do this we assume that the PDE can be discretised with multiple levels of accuracy. In our work these levels are associated with different mesh sizes in a spatial domain. However, it is also possible to consider e.g. different time step sizes, or target accuracies of Newton's method.

Multilevel samplers enjoy considerable attention at the moment, and are available for various tasks in uncertainty quantification. Multilevel Monte Carlo is widely used in forward uncertainty quantification; see [24] for an overview. In the pioneering work by Giles [23] the multilevel idea is combined with standard Monte Carlo in a forward setting. However, it can be used with other samplers such as MCMC, SMC, and the EnKF, for the estimation of rare events, for filtering problems in data assimilation, and to solve Bayesian Inverse Problems. For example, multilevel Ensemble Kalman Filters have been proposed in [9,27]. The authors in [27] consider continuous-time data assimilation with multiple time-step discretisations. In contrast, the work in [9] considers data assimilation for spatially extended models e.g. time dependent stochastic partial differential equation. The multilevel estimation of rare events with an SMC type method has been proposed in [49].

For Bayesian Inverse Problems a multilevel MCMC method has been introduced in [18]. Multilevel Sequential Monte Carlo is introduced in [4] and further discussed in [3,17,16]. Both the multilevel MCMC and multilevel SMC use coarse PDE discretisations for variance reduction with the help of a telescoping sum expansion. Moreover, these multilevel samplers are built to integrate output quantities of interest with respect to the posterior. In contrast, in our work we do not rely on a telescoping sum, and we construct approximations to the full posterior measure.

We build on and generalise the work by Koutsourelakis in [33]. As in [33] we combine SMC with tempering on a fixed PDE discretisation level, and bridging between two consecutive discretisation levels. The major novel contribution of our work is a fully adaptive algorithm to decide when to increase the discretisation accuracy (bridging) and when to proceed