LATENT SPACE PARTICLE FILTER
PhD proposal

Scientific advisors: Elise Arnaud (elise.arnaud@univ-grenoble-alpes.fr)
Arthur Vidard (arthur.vidard@inria.fr)
Websites: https://team.inria.fr/airsea/en/

Functional area: Laboratoire Jean Kuntzmann
Inria Project Team AIRSEA
IMAG Building, 700, avenue centrale
38401 Saint Martin d’Hères

Keywords: data assimilation, particle filter, Latent spaces, Neural Networks

Context. One of the many achievements of computer aided simulations has been the sharp improvement in the human prediction capabilities of weather and climate, with undeniable benefits for society: agricultural planning, extreme weather preparedness, mitigation of climate change effects, etc. The quality of the simulations rely in particular on the ability to define a sensible starting point (initial and/or boundary conditions) for said simulation. This is usually achieved by gathering observational data about the present and the past state and and gradually assimilating these data in the numerical model to make reliable predictions. This is referred as data assimilation.

Goals. Data assimilation can be seen as the art of compromise. It refers to a large variety of methods that allow to combine all sources of information available about a given system : mathematical equation (physical laws), observations (measures of reality) and error statistics.

Among all data assimilation approaches, particle filters are theoretically-validated Monte-Carlo based techniques that do not require Gaussian assumptions on the model. In that framework, the data assimilation problem is written as a large-scale Bayesian inverse problem. It requires the estimation of the probability distribution of the state variables, at a given time, conditioned on all the previously observed measurement data. Particle filter algorithms receive a growing attention from the data assimilation community [1]. They consists in sampling the state space in places where the data could appear. The (huge) downside is that they perform poorly for large state space dimension [1,2].

Large scale modelling systems like ocean and atmosphere represent the discretisation of physical laws and generally represent the state of this system on a discrete grid containing millions of unknown. However, the size of the space of possible states is much smaller than that (not all combinations are physically relevant). This space of possible states is often called latent space.

The goal of this phd is to combine IA techniques for latent space estimation and surrogate modelling with particle filter. Auto-encoders, for instance, seem very relevant for this type of application [3]. They can be compared to more traditional approaches used in data assimilation, such as PCA or singular vectors. Evidences show that the properly defined error representation will allow for a better exploration of the state space, and thus is a possible way to address the dimension issue. After a literature study and some theoretical developments, we will test the feasibility of the idea on a test case that implement a simplified 2D ocean model at coarse resolution in a double gyre setting. This configuration is meant to represent an idealized North Atlantic circulation (Gulf Stream). In a second phase an application to realistic ocean models is foreseen. This work is part of an ongoing collaboration with ATOS.

Prerequisites. Applied math skills (optimisation, numerical analysis, probability / statistics) and programming skills (Matlab, python, C or Fortran)

Bibliography