INTERACTING PARTICLES: THEORETICAL ANALYSIS AND APPLICATION TO AN EM INVERSE PROBLEM

F. Giraud

(joint with P. Minvielle and P. Del Moral)

CEA-CESTA, Le Barp, France
INRIA Bordeaux Sud-Ouest
francois.giraud@cea.fr

ABSTRACT

Sequential Monte Carlo sampling methods are powerful tools, especially to solve inverse problems by sampling from posterior distributions. Their convergence is clearly related to the stability of Feynman-Kac’s dynamic distribution system. By analysing this stability under particular assumptions, we deduced theorems about the algorithm’s tuning, to ensure the convergence, with special attention devoted to the particular case of Gibbs distributions’ simulation. We finally focus on a application to an electromagnetic (EM) inverse problem, from the measurement of reflection coefficients of a 1-D multi-layered radar absorber.

1. INTRODUCTION

Sequential Monte Carlo (SMC) is a genetic algorithm to sample from complex high-dimensional probability distributions. It has found a wide range of applications in applied probability, Bayesian statistics and information engineering, especially through problems of filtering and stochastic optimization. The principle (e.g., [1]) is to approximate the target probability distributions by a large cloud of random samples termed particles \((\zeta^k_n)_{1 \leq k \leq N} \in \mathbb{E}^N, E\) being called the state space, defining at each generation \(n\) the occupation distribution: \(\eta^N_n = \frac{1}{N} \sum_{k=1}^N \delta_{\zeta^k_n}\). We run from generation \((\zeta^k_{n-1})\) to generation \((\zeta^k_n)\) through a selection step using a positive selection function \(g_n\) on \(E\), and a mutation step, using a Markov kernel \(M_n\).

\[ \zeta_n \in \mathbb{E}^N \xrightarrow{\text{selection}} \hat{\zeta}_n \in \mathbb{E}^N \xrightarrow{\text{mutation}} \zeta_{n+1} \in \mathbb{E}^N \]

More precisely, the genetic type evolution of the system is described by the following synthetic picture:

\[ \begin{array}{c}
\zeta^1_n \\
\vdots \\
\zeta^N_n
\end{array} \xrightarrow{S_{n,\eta^N_n}} \begin{array}{c}
\zeta^1_n \\
\vdots \\
\zeta^N_n
\end{array} \xrightarrow{M_{n+1}} \begin{array}{c}
\zeta^1_{n+1} \\
\vdots \\
\zeta^N_{n+1}
\end{array} \]

with the selection Markov transition:

\[ S_{n,\eta^N_n}(\zeta^i_n, dx) := g_n(\zeta^i_n) \delta_{\zeta^i_n}(dx) + (1 - g_n(\zeta^i_n)) \sum_{1 \leq j \leq N} \frac{g_n(\zeta^j_n)}{\sum_{1 \leq k \leq N} g_n(\zeta^k_n)} \delta_{\zeta^i_n}(dx) \]

Evolving this way, the occupation distribution \(\eta^N_n\) approximates for each \(n\) the theoretical distribution \(\eta_n\) defined recursively by the Feynman-Kac formulae (e.g., [3]), associated with the potentials \(g_n\) and kernels \(M_n\).

2. THEORETICAL RESULTS

Our work is to quantify the distance between \(\eta^N_n\) and \(\eta_n\) according to the following quantities: the logarithmic oscillation \(g_n\) of potentials \(g_n\) defined by \(g_n := e^{\text{osc} \log g_n}\) and the Dobrushin coefficients \(b_n := \beta(M_n)\) (e.g., [3]). As an example, we prove that under the assumption \(g_n \leq M\), choosing \(M_n\) such that \(b_n \leq \frac{a}{M+a}\) (where \(a < 1\)) ensures the \(L^p\) mean error:

\[ \|\eta^N_n(f) - \eta_n(f)\|_p \leq \frac{B_p}{\sqrt{N}} \frac{1}{1-a} \]

and the following concentration inequality (e.g., [2]) (for all uniformly bounded by 1 function \(f\) on \(E\)):

\[ \mathbb{P}\left( |\eta^N_n(f) - \eta_n(f)| \geq r_n^N + r_n^N y \right) \leq e^{-y} \]
In the special case of the optimization of a potential function \(V\), \(\eta_n\) can be chosen as a Gibbs distribution associated with an increasing sequence of "temperatures" (\(\beta_n\)). One can then turn the previous results into conditions on the temperature scheme and Markov kernels iteration numbers \(m_n\). For example, if one use Metropolis Markov kernels \(K_{\beta_n}\) with Dobrushin coefficients satisfying 

\[ \beta(R^k_{\beta}) \leq 1 - \delta e^{-\beta \text{osc}(V)} \text{ (e.g., [3])}, \]

the choice of a sequence \(\beta_n\) such that \((\beta_n - \beta_{n-1})\) is bounded by \(\Delta\), and iteration numbers \(m_n\) satisfying \(m_n \geq C \exp (\beta_n \text{osc}(V))\) with \(C = \frac{1}{\Delta} \log(\frac{\Delta \text{osc}(V)}{\alpha + \beta})\), ensures that \(\forall \varepsilon > 0\), the proportion \(p_n^N(\varepsilon)\) of particles \(\zeta_n^k\) such that \(V(\zeta_n^k) > V_{\min} + \varepsilon\) satisfies \(\forall y > 0\):

\[ P \left( p_n^N(\varepsilon) \geq \frac{e^{-\varepsilon \beta_n}}{m_{\varepsilon'}} + \frac{r_{\varepsilon'}^y}{N^2} \right) \leq e^{-y}. \]

\(N\) being the number of particles, \(\forall \varepsilon' < \varepsilon\), \(m_{\varepsilon'}\), \(r_{\varepsilon'}^x\) and \(r_{\varepsilon'}^y\) being constants.

3. APPLICATION

Let us consider a 1-D multi-layered radar absorber composed with \(N_e\) layers of uniform materials (e.g., [4]). For several incidence angles \(\theta\) and frequencies \(f\), the multilayer object is illuminated with a monochromatic plane wave. For each couple \((\theta, f)\) and each polarization TM or TE, we measure (with uncertainty) the global reflection coefficient (see figure [1]). Each material \(i\) can be characterized by its permittivity \(\varepsilon_i\) and permeability \(\mu_i\). The inverse problem is to determine the materials properties \(\varepsilon_i\) and \(\mu_i\) (with associated uncertainties) from the measured reflection coefficients.

Figure 1: multilayer material

Quantities \(\varepsilon_i\) and \(\mu_i\) being dependent on the frequency \(f\), we use a frequency biased parameterizing according to which each possible solution is represented by \(4N_e\) positive numbers. The relation between the sequence of measured reflection coefficients \(y\) and the unknown multilayer material parameters \(x\) can be modeled by:

\[ y = f_{MD}(x) + n \]

where \(f_{MD}\) is an analytic function in the microwave multilayer field, and \(n\) a Gaussian random vector (noise). In a Bayesian statistics viewpoint, the inverse problem solution is then defined by the posterior distribution \(p(x|y)\), that can be derived from the Bayes rule \(p(x|y) \propto p(x)p(y|x)\). Here, the likelihood term \(p(y|x)\) is fixed by the direct model equation and the noise model, and the prior distribution \(p(x)\) is chosen non-informative. Fixing an increasing sequence from 0 to 1 \((\alpha_n)_{1 \leq n \leq n_f}\), and denoting \(\eta_n := p(x) [p(y|x)]^{\alpha_n}\), we turn the target distribution into the last distribution of a sequence \(\eta_n\) satisfying the Feynman-Kac formula, so that the SMC algorithm can be performed to approximate it. Figure 2 shows the final cloud of particles, represented as histograms respectively of the 2nd of the 8th component, in the case of \(N_e = 2\).

Figure 2: Posterior distribution approximation

4. CONCLUSIONS

Our theoretical analysis provides bounds far from being perfectly relevant for the practical use of SMC methods. Yet, they highlight interesting interpretations of the algorithm and help to understand, predict and tune its behavior. It has been evaluated on several multilayer materials with \(N_e = 2\) or \(3\), and gives good results with reasonable calculus time (\(\simeq 1\) hour) using a single processor.

5. REFERENCES