SIMINOLE Kick-Off Meeting

27 septembre 2010

The Team

Our Contribution

The Team

Olivier Cappé (mngr., 30%) DR CNRS *Computational statistics, Machine learning, Monte Carlo methods*

Gersende Fort (20%) CR CNRS *Monte Carlo methods, Markov chains, Bayesian modelling and inference*

Eric Moulines (10%) Prof. Télécom ParisTech *Mathematical Statistics, Monte Carlo methods, Time series analysis*

With, on average, one PhD student and one postdoc working on topics related to the SIMINOLE project (currently : one second year student, a good potential PhD candidate for sep. 2011)

Recent and Ongoing Related Collaborative Projects

ECOSSTAT *Measuring cosmological parameters from large heterogeneous surveys* 2005–2009 with Cérémade, Institut d'Astrophysique de Paris, Laboratoire d'Astrophysique de Marseille — mngr. O. Cappé

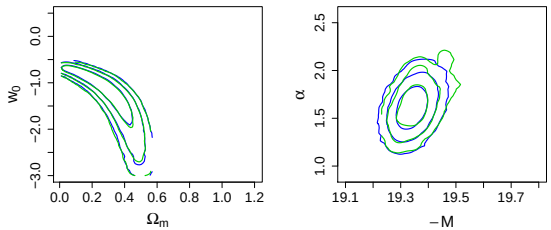
ADAP'MC *Adaptive Monte Carlo methods* 2006–2008 with Cérémade, Cermics (Ecole des Ponts), CMAP (Ecole Polytechnique) — project mngr. G. Fort

---

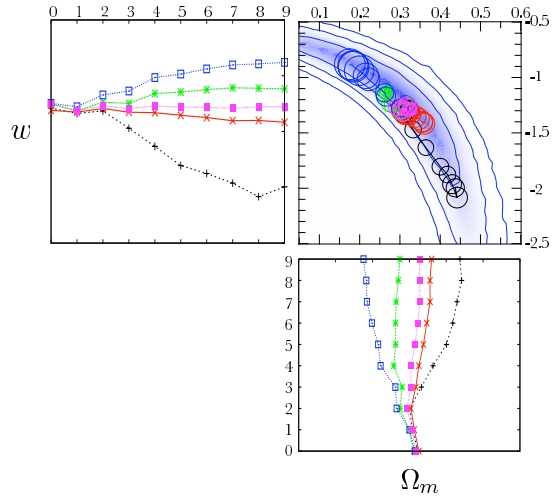
BigMC *Issues in large scale Monte Carlo* 2009–2012 with Cérémade, Cermics (Ecole des Ponts) — project mngr. G. Fort

A Few More Words on the ECOSSTAT Project

Goal Estimation of parameters and comparison of **cosmological models** (*CDM) from heterogeneous data surveys (WMAP CMB, SNLS supernovae Ia, CFHTLS cosmic shear)



Challenge Monte Carlo exploration in moderate size problem (5 to 15 dimensions) where likelihood evaluations are very costly (several seconds)



Main Contribution Adaptive importance sampling approach that makes use of parallel computing (implemented on IAP cluster)

The Team

Our Contribution

Our Contribution

- We will be focussing mostly on Simulation-based probabilistic inference (task 2) and its application to the Auger experiment (task 5)
- Interested in other tasks as well, in particular simulation-based stochastic optimization (task 3)
- Already some familiarity with the probabilistic model developed by Balázs and Rémi for the Auger experiments

Scientific Challenges

Adaptive algorithms with an unknown number of "parameters" The model for the Auger experiment involves a varying number of hidden variables: adaptive (i.e. self-tuning) Markov Chain Monte Carlo (MCMC) algorithms applicable to this context are almost inexistent

Parallelizable MCMC implementations Cannot re-use the solution developed for the ECOSSTAT project due to the scale of the problem. Need to develop MCMC-based algorithms that can make use of parallelization (challenging, but some first ideas regarding the independent sampler)

Sequential breakdown of the inference problem Capitalizing on the breakthrough paper of [Andrieu, Doucet & Holenstein, 2010] which suggests promising solutions for the model used for the Auger experiment