Estimation of cosmological parameters using adaptive importance sampling

Gersende FORT

LTCI, CNRS / TELECOM ParisTech

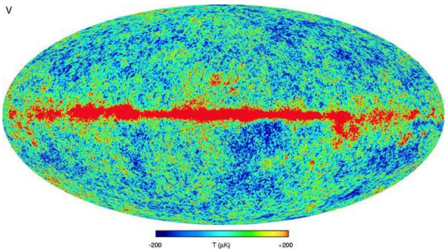
This work was supported by the *french National Research Agency* under the program ECOSSTAT (Jan. 06 - Dec. 08)

Exploration du modèle cosmologique
par fusion statistique de grands relevés hétérogènes

Members:

- IAP Institut d'Astro-Physique de Paris (F. Bouchet et Y. Mellier), Paris.
- LAM Laboratoire d'Astro-Physique de Marseille (O. Le Fèvre), Marseille.
- LTCI Laboratoire Traitement et Communication de l'Information (O. Cappé), Paris.
- CEREMADE Centre de Recherche en Mathématique de la Décision (C.P. Robert), Paris.

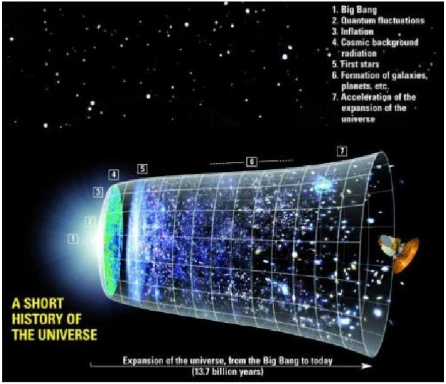
Objectives of the project:
Combine three deep surveys of the universe to set new constraints on the evolution scenario of galaxies and large scale structures, and the fundamental cosmological parameters.



Example of survey: WMAP (or Planck) for the Cosmic Microwave Background (CMB) radiations = temperature variations are related to fluctuations in the density of matter in the early universe and thus carry out information about the initial conditions for the formation of cosmic structures such as galaxies, clusters and voids for example.

Some questions in cosmology

- will the universe expand for ever, or will it collapse?
- what is the shape of the universe?
- Is the expansion of the universe accelerating rather than decelerating?
- Is the universe dominated by dark matter and what is its concentration?



Today, talk about

Estimation of cosmological parameters using adaptive importance sampling

A work in collaboration with

- Darren WRAITH and Martin KILBINGER (CEREMADE/IAP)
- Karim BENABED, François BOUCHET, Simon PRUNET (IAP)
- Olivier CAPPE, Jean-François CARDOSO, Gersende FORT (LTCI)
- Christian ROBERT (CEREMADE)

A work published in Phys.Rev. D. 80(2), 2009.

Model (I)

- Observational data from
  - the CMB Cosmic Microwave Background ⟶ five-year WMAP data.
  - the observation of weak gravitational shear ⟶ CFHTLS-Wide third release.
- explained by some cosmologic parameters

TABLE II: Parameters for the cosmology likelihood. C=CMB, S=SNIa, L=lensing.

Symbol	Description	Minimum	Maximum	Experiment
Ω_b	Baryon density	0.01	0.1	C L
Ω_m	Total matter density	0.01	1.2	C S L
w	Dark-energy eq. of state	-3.0	0.5	C S L
n_s	Primordial spectral index	0.7	1.4	C L
Δ_R^2	Normalization (large scales)			C
σ_8	Normalization (small scales)[a]			C L
h	Hubble constant			C L
τ	Optical depth			C
M	Absolute SNIa magnitude			S
α	Colour response			S
β	Stretch response			S
a				L
b	galaxy z-distribution fit			L
c				L

Model (II)

This yields:

- a likelihood of the data given the parameters: some of them computed from publicly available codes ex. WMAP5 code for CMB data
- combined with a priori knowledge: uniform prior on hypercubes.

Therefore, statistical inference consists in the exploration of the a posteriori density of the parameters, **a challenging task** due to

- potentially high dimensional parameter space (not really considered here: sampling in $\mathbb{R}^d$, $d \sim 10$ to 15)
- immensely slow computation of likelihoods,
- non-linear dependence and degeneracies between parameters introduced by physical constraints or theoretical assumptions.

Monte Carlo algorithms for the exploration of the a posteriori density π

- **(naive) Monte Carlo** methods: i.i.d. samples under π. Here, NO: π is only known through a "numerical box"

- **Importance Sampling** methods: i.i.d. samples $\{X_k, k \geq 0\}$ under a proposal distribution q and

$$\sum_{k=1}^{n} \frac{\omega_k}{\sum_{j=1}^{n} \omega_j} \mathbb{I}_\Delta(X_k) \approx \mathbb{P}_\pi(X \in \Delta) \qquad \text{with} \qquad \omega_k = \frac{\pi(X_k)}{q(X_k)}$$

- **Markov chain Monte Carlo** methods: a Markov chain with stationary distribution π

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_\Delta(X_k) \approx \mathbb{P}_\pi(X \in \Delta)$$

- $\cdots$

Importance sampling or MCMC?

All of these sampling techniques, require **time consuming** evaluations of the a posteriori distribution π for each new draw

- Importance sampling: allow for parallel computation.
- MCMC: can not be parallelized. well, say, most of them

The efficiency of these sampling techniques depend on design parameters

- Importance sampling: the proposal distribution.
- Hastings-Metropolis type MCMC: the proposal distribution.

$\hookrightarrow$ towards adaptive algorithms that learn on the fly how to modify the value of the design parameters.

Monitoring convergence

- Importance sampling: criteria such as Effective Sample Size (ESS) or the Normalized Perplexity.
- MCMC: no such explicit criterion.

Therefore, we decided to

- run an adaptive Importance Sampling algorithm: **Population Monte Carlo** [Robert et al. 2005]
- compare it to an adaptive MCMC algorithm: **Adaptive Metropolis algorithm** [Haario et al. 1999]

Population Monte Carlo (PMC) algorithm

- Idea: choose the **best** proposal distribution among a set of (parametric) distributions.
  Criterion based on the Kullback-Leibler divergence

$$q_\star = \mathrm{argmax}_{q \in \mathcal{Q}} \int \log q(x)\ \pi(x)\ dx$$

- In order to have a / to approximate the solution of this optimization problem
  - choose $\mathcal{Q}$ as the set of mixtures of Gaussian distributions (or t-distributions).
  - solve the optimization by applying the same updates as when iterating an Expectation-Maximimzation algorithm for fitting mixture models on i.i.d. samples $\{Y_k, k \geq 0\}$

$$\mathrm{argmax}_{q \in \mathcal{Q}} \frac{1}{n} \sum_{k=1}^{n} \log q(Y_k)$$

except that it requires integration w.r.t. π !!

Population Monte Carlo (PMC) algorithm (II)

Iterative algorithm:

- initialization: choose an initial proposal distribution $q^{(0)}$ and draw weighted points $\{(w_k, X_k)\}_k$ that approximate π

- Based on these samples, update the proposal distribution

$$q^{(1)} = \mathrm{argmax}_{q \in \mathcal{Q}} \sum_{k=1}^{n} \frac{\omega_k}{\sum_{j=1}^{n} \omega_j} \log q(X_k)$$

  and draw weighted points $\{(w_k, X_k)\}_k$ that approximate π.

- Repeat until $\cdots$ further adaptations do not result in significant improvements of the KL divergence. e.g. compute the Normalized Effective Sample Size at each iteration

$$ESS = \frac{1}{n} \left(\sum_{k=1}^{n} \left(\frac{\omega_k}{\sum_{j=1}^{n} \omega_j} \right)^2 \right)^{-1}$$

Adaptive Metropolis

- Symmetric Random Walk Metropolis algorithm

- with Gaussian proposal distribution, with "mysterious" (but famous) scaling matrix
$$\mathcal{N}\left(0, \frac{2.38^2}{d}\Sigma_\pi\right)$$
where Σ_π is the **unknown** covariance matrix of π. [Roberts et al. 1997]

- "unknown"?! estimate it on the fly, from the samples of the algorithm $\longrightarrow$ adaptive Metropolis algorithm

Simulations

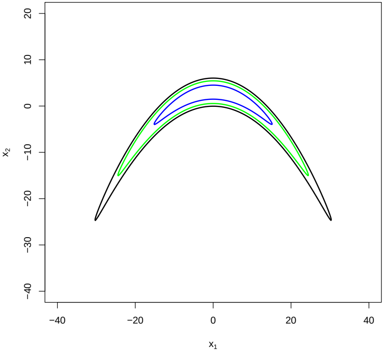
on

1. simulated data, from a "banana" density

2. real data.

Simulated data

The target distribution in $\mathbb{R}^{10}$. Below marginal distribution of (x_1,x_2)



and $(x_3,\cdots,x_{10})$ are independent $\mathcal{N}(0,1)$.

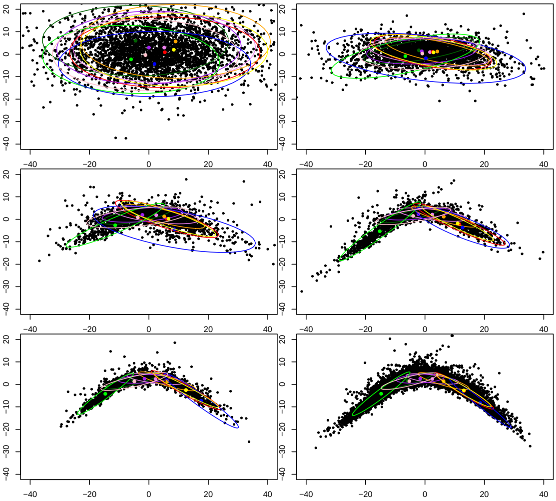


FIG.: Iterations 1,3,5,7,9,11. 10k points per plot, except 100k in the lase one. Mixture of 9 t-distributions, with 9 degrees of freedom

Monitoring convergence: the *Normalized perplexity (top panel)* and the *Normalized Effective Sample size* (bottom panel)

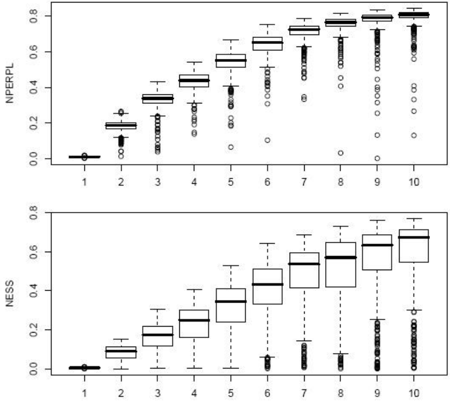


FIG.: for the first 10 iterations, over 500 simulation runs.

Comparison of adaptive MCMC and PMC:

$$f_a(x) = x_1 \qquad f_b(x) = x_2$$

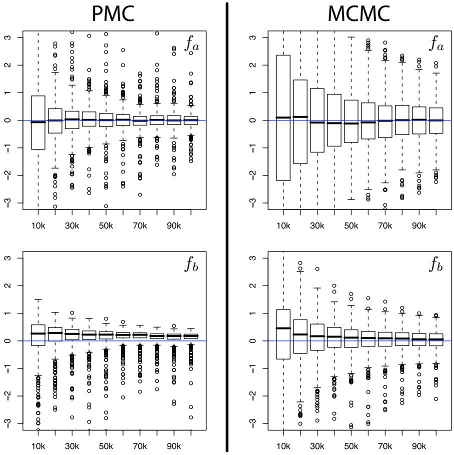
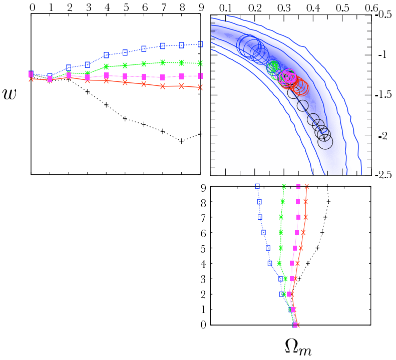


FIG.: for the first 10 iterations, over 500 simulation runs.

Application to cosmology

Evolution of the PMC algorithm: *the likelihood is from the SNIa data*



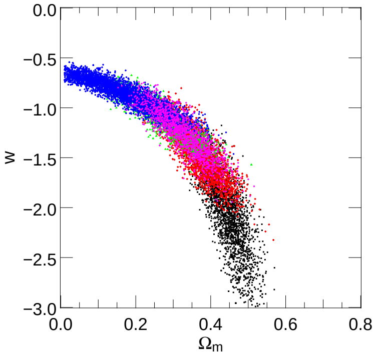


FIG.: [left] evolution of the Gaussian mixtures with 5 components. [right] samples at the last PMC iteration, from the 5 components

Evolution of the weights: *the likelihood is WMAP5 for a flat ΛCDM model with six parameters*

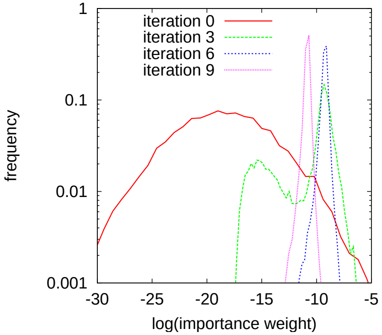


FIG.: Histogram of the normalized weihts for four iterations

- Monitoring convergence: *the likelihood is WMAP5 for a flat ΛCDM model with six parameters*

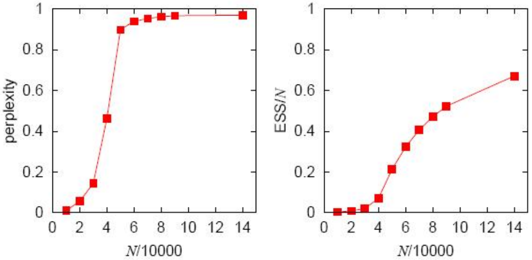


FIG.: perplexity (left) and ESS (right) as a function of the cumulative sample size

- After $150k$ evaluations of π: ESS is about 0.7; mean acceptance rate in MCMC about 0.25.

Comparison of MCMC and PMC: *the likelihood is from the SNIa data*

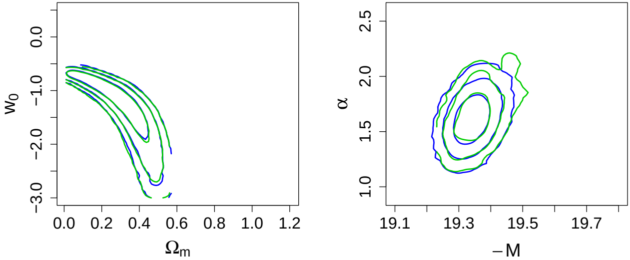


FIG.: Marginalized likelihoods (68%,95%,99.7% contours are shown) for PMC (solid blue) and MCMC (dashed green)

Estimates of cosmological parameters: *from the WMAP5 data (left) and from the lensing+SNIa+CMB data sets (right)*

Parameter	PMC	MCMC
Ω_b	$0.04424^{+0.00321}_{-0.00290}$	$0.04418^{+0.00321}_{-0.00294}$
Ω_m	$0.2633^{+0.0340}_{-0.0282}$	$0.2626^{+0.0359}_{-0.0280}$
τ	$0.0878^{+0.0181}_{-0.0160}$	$0.0885^{+0.0181}_{-0.0160}$
n_s	$0.9622^{+0.0145}_{-0.0143}$	$0.9628^{+0.0139}_{-0.0145}$
$10^9\Delta^2_R$	$2.431^{+0.118}_{-0.113}$	$2.429^{+0.123}_{-0.108}$
h	$0.7116^{+0.0271}_{-0.0261}$	$0.7125^{+0.0274}_{-0.0268}$

Parameter	PMC	MCMC
Ω_b	$0.0432^{+0.0027}_{-0.0024}$	$0.0432^{+0.0026}_{-0.0023}$
Ω_m	$0.254^{+0.018}_{-0.017}$	$0.253^{+0.018}_{-0.016}$
τ	$0.088^{+0.018}_{-0.016}$	$0.088^{+0.019}_{-0.015}$
w	-1.011 ± 0.060	$-1.010^{+0.059}_{-0.060}$
n_s	$0.963^{+0.015}_{-0.014}$	$0.963^{+0.015}_{-0.014}$
$10^9\Delta^2_R$	$2.413^{+0.098}_{-0.093}$	$2.414^{+0.098}_{-0.092}$
h	$0.720^{+0.022}_{-0.021}$	$0.720^{+0.023}_{-0.021}$
a	$0.648^{+0.040}_{-0.041}$	$0.649^{+0.043}_{-0.042}$
b	$9.3^{+1.4}_{-0.9}$	$9.3^{+1.7}_{-0.9}$
c	$0.639^{+0.084}_{-0.070}$	$0.639^{+0.082}_{-0.070}$
$-M$	19.331 ± 0.030	$19.332^{+0.029}_{-0.031}$
α	$1.61^{+0.15}_{-0.14}$	$1.62^{+0.16}_{-0.14}$
$-\beta$	$-1.82^{+0.17}_{-0.16}$	-1.82 ± 0.16
σ_8	$0.795^{+0.028}_{-0.030}$	$0.795^{+0.030}_{-0.027}$

FIG.: Means and 68% confidence intervals

Conclusion

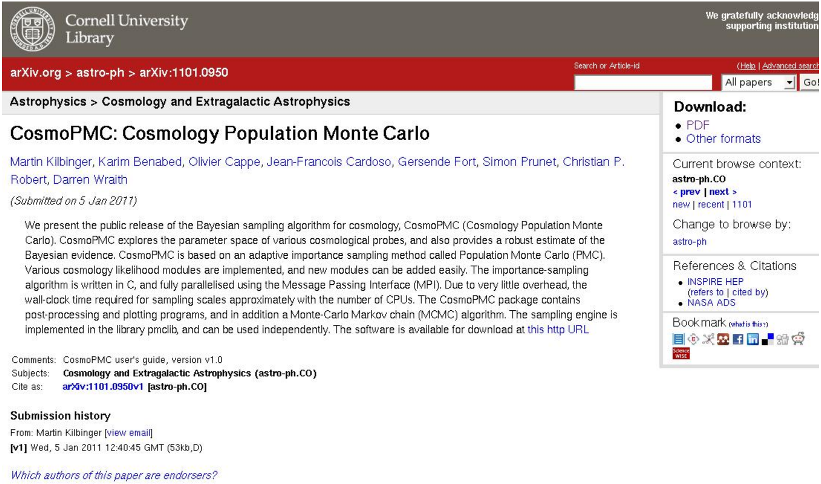
Cosmology provides challenging problems for Bayesian inference:

- large dimension of the parameter space
- time consuming likelihood

Open questions:

- parallelization of Monte Carlo methods
- methods robust to the dimension

Public release of the Bayesian sampling algorithm for cosmology, CosmoPMC (Martin KILBINGER and Karim BENABED)



arXiv.org > astro-ph > arXiv:1101.0950

Search or Article-id (Help | Advanced search) All papers Go!

Astrophysics > Cosmology and Extragalactic Astrophysics

CosmoPMC: Cosmology Population Monte Carlo

Martin Kilbinger, Karim Benabed, Olivier Cappe, Jean-Francois Cardoso, Gersende Fort, Simon Prunet, Christian P. Robert, Darren Wraith

(Submitted on 5 Jan 2011)

We present the public release of the Bayesian sampling algorithm for cosmology, CosmoPMC (Cosmology Population Monte Carlo). CosmoPMC explores the parameter space of various cosmological probes, and also provides a robust estimate of the Bayesian evidence. CosmoPMC is based on an adaptive importance sampling method called Population Monte Carlo (PMC). Various cosmology likelihood modules are implemented, and new modules can be added easily. The importance-sampling algorithm is written in C, and fully parallelised using the Message Passing Interface (MPI). Due to very little overhead, the wall-clock time required for sampling scales approximately with the number of CPUs. The CosmoPMC package contains post-processing and plotting programs, and in addition a Monte-Carlo Markov chain (MCMC) algorithm. The sampling engine is implemented in the library pmclib, and can be used independently. The software is available for download at this http URL

Comments: CosmoPMC user's guide, version v1.0
Subjects: **Cosmology and Extragalactic Astrophysics (astro-ph.CO)**
Cite as: **arXiv:1101.0950v1 [astro-ph.CO]**

Submission history

From: Martin Kilbinger [view email]
[v1] Wed, 5 Jan 2011 12:40:45 GMT (53kb,D)

Which authors of this paper are endorsers?

Download:
- PDF
- Other formats

Current browse context:
astro-ph.CO
< prev | next >
new | recent | 1101

Change to browse by:
astro-ph

References & Citations
- INSPIRE HEP (refers to | cited by)
- NASA ADS

Bookmark (what is this?)