BAT

BAYESIAN ANALYSIS TOOLKIT

mpp.mpg.de/bat

Introduction to BAT

Frederik Beaujean, C2PAP
Frederik.Beaujean@lmu.de

25.07.2013

Bayesian probability theory

Logical reasoning with uncertainties = probability theory

What is the probability that

- it rains tomorrow?
- the SM is right?
- the Higgs has a mass of 125 GeV?

Analogy to Newton's equations

Newton (1643-1727)

- Describe state of particle(s): position, velocity...
- Must specify initial conditions of particles
- Include all (relevant) forces that act upon the particles
- Time evolution

$\Rightarrow$ State of particles after time has elapsed

Bayes (1702–1761)

- Describe state of knowledge: Probability (Degree-of-belief)
- Must specify prior knowledge before data is collected
- Write down a statistical model of all possible observations
- New data comes in

$\Rightarrow$ State of knowledge after learning from the data

Bayes' Theorem (Inverse Probability)

$$\boxed{P_1(\vec{\lambda}\,|\,\vec{D},M) = \frac{P(\vec{D}\,|\,\vec{\lambda},M)\,P_0(\vec{\lambda}\,|\,M)}{\int P(\vec{D}\,|\,\vec{\lambda},M)\,P_0(\vec{\lambda}\,|\,M)\,d\vec{\lambda}}}$$

$P_0(\vec{\lambda}\,|\,M)$	Prior belief in parameters given the model
$P_1(\vec{\lambda}\,|\,\vec{D},M)$	Posterior belief in parameters given model&data
$P(\vec{D}\,|\,\vec{\lambda},M)$	Probability (frequency) of the data assuming model&parameters
$L(\vec{\lambda}\,|\,\vec{D},M)\equiv P(\vec{D}\,|\,\vec{\lambda},M)$	Likelihood (not a probability!)

$$\boxed{\text{posterior} \sim \text{likelihood x prior}}$$

Bayesian Analysis Toolkit

http://mpp.mpg.de/bat

Motivation:

- Bayes' theorem simple on paper, but numerics can be hard
- Implementing standard algorithms by yourself a time waster, error prone
- create a toolkit (C++ library) to take care of that so user can focus on the problem, uses ROOT, (optionally) used by roostats
- Doc, tutorials, examples... on web page

Algorithms

- Optimization: Minuit, Simulated annealing
- Integration: sample mean, importance sampling, cuba(Vegas, cubature...)
- Sampling: Markov chain Monte Carlo

Components

Separate the common parts from the rest

- case specific: the model and the data
- common tools: all the rest

USER DEFINED

- create model
- read-in data

COMMON TOOLS
(common tools)

- `Normalize()`
- `FindMode()`
- `CalculatePValue...()`
- `MarginalizeAll()`
- compare models: `BayesFactor()`
- provide nice output: `PrintAllMarginalized()` `PrintKnowledgeUpdatePlots()`

Define `MyModel : BCModel`

- `AddParameter("lambda", 0, 1)` $\vec{\lambda}$
- `double LogLikelihood(pars)` $p(\vec{D}\,|\,\vec{\lambda})$
- `double LogAPrioriProbability(pars)` $p_0(\vec{\lambda})$

Read DATA

- from text file, ROOT tree, user defined (anything)

Bayes formula

$$p_1(\vec{\lambda}\,|\,\vec{D}) = \frac{p(\vec{D}\,|\,\vec{\lambda})\,p_0(\vec{\lambda})}{\int p(\vec{D}\,|\,\vec{\lambda})\,p_0(\vec{\lambda})\,d\vec{\lambda}}$$

Key idea:
draw random numbers from posterior

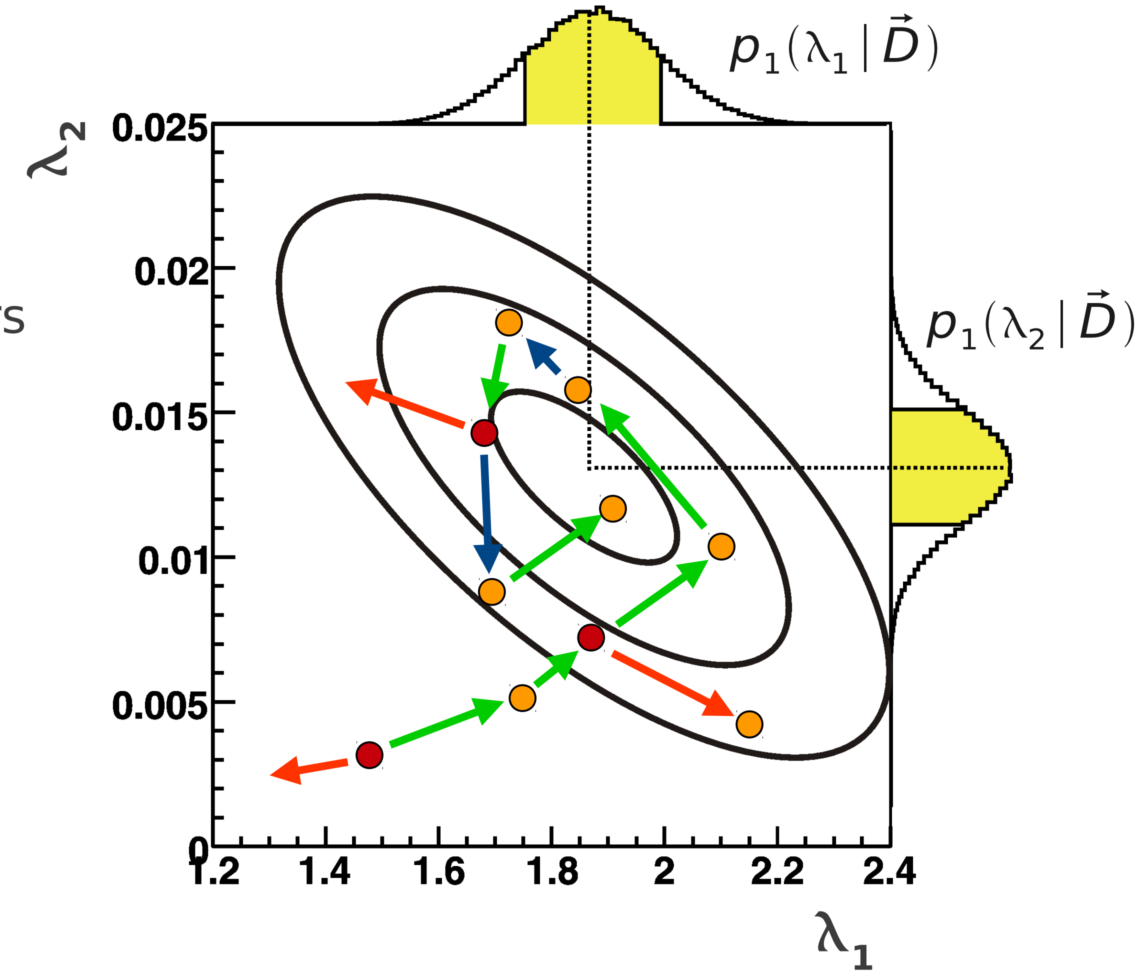
- In BAT use
  Markov Chain Monte Carlo (MCMC)

$$\vec{\lambda} \sim p_1(\vec{\lambda} \mid \vec{D}) = \frac{p(\vec{D} \mid \vec{\lambda})\, p_0(\vec{\lambda})}{\int p(\vec{D} \mid \vec{\lambda})\, p_0(\vec{\lambda})\, d\vec{\lambda}}$$

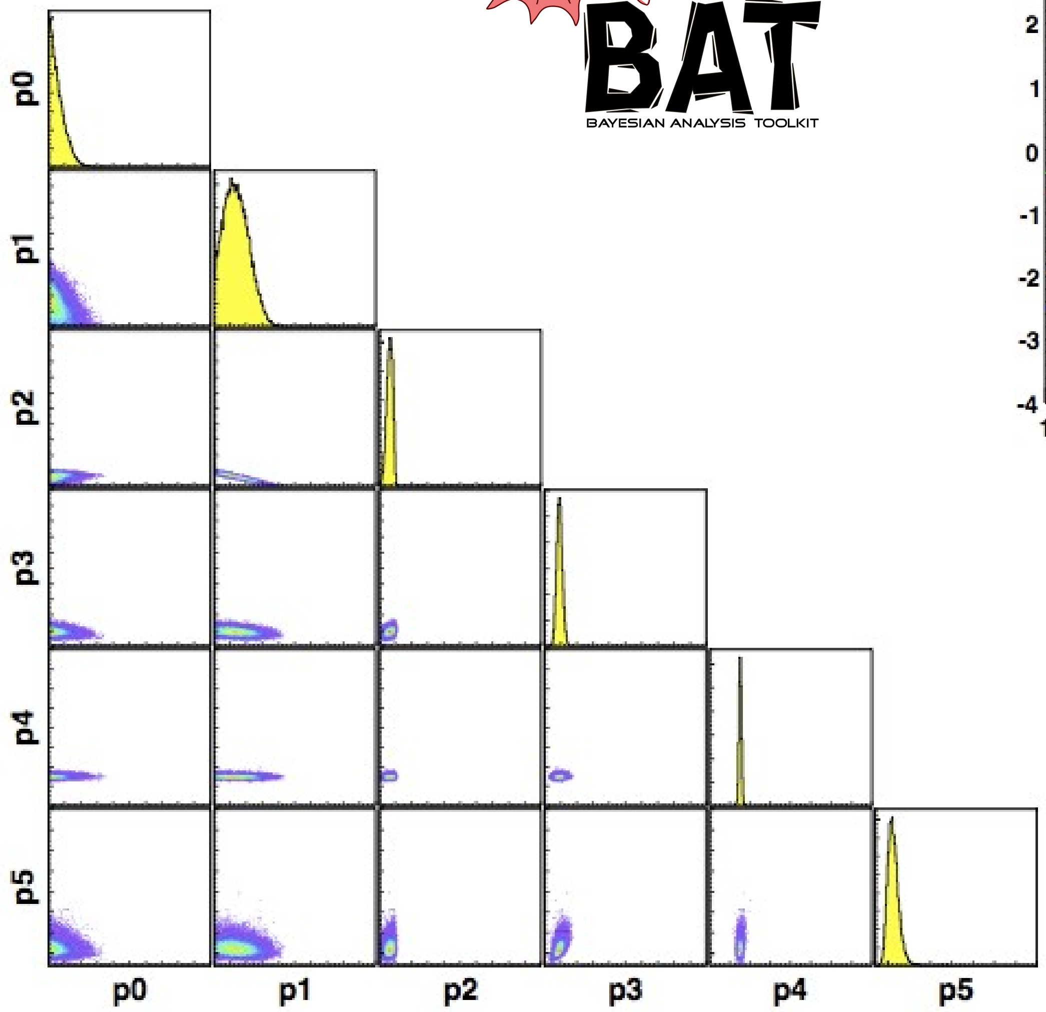
- Marginalize wrt. individual parameters while walking
  → obtain

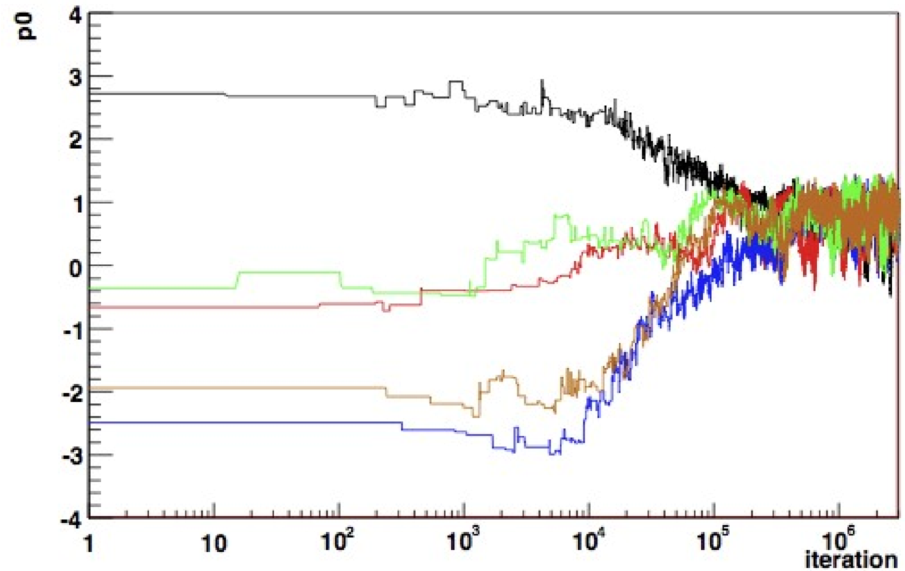
$$p_1(\lambda_i \mid \vec{D}) = \int p(\vec{D} \mid \vec{\lambda})\, p_0(\vec{\lambda})\, d\vec{\lambda}_{j \neq i}$$

- Find maximum: "best" parameters
- Uncertainty propagation: distribution of $f(\vec{\lambda})$



Markov chain output



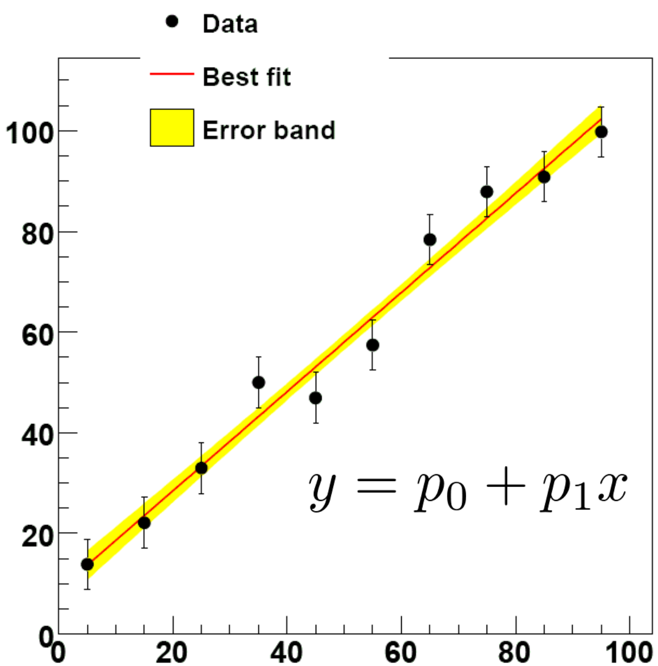


Multiple chains mixing (multithreaded)

Posterior probability distributions from analysis of data – how to use these to make statements on your model and parameters ?

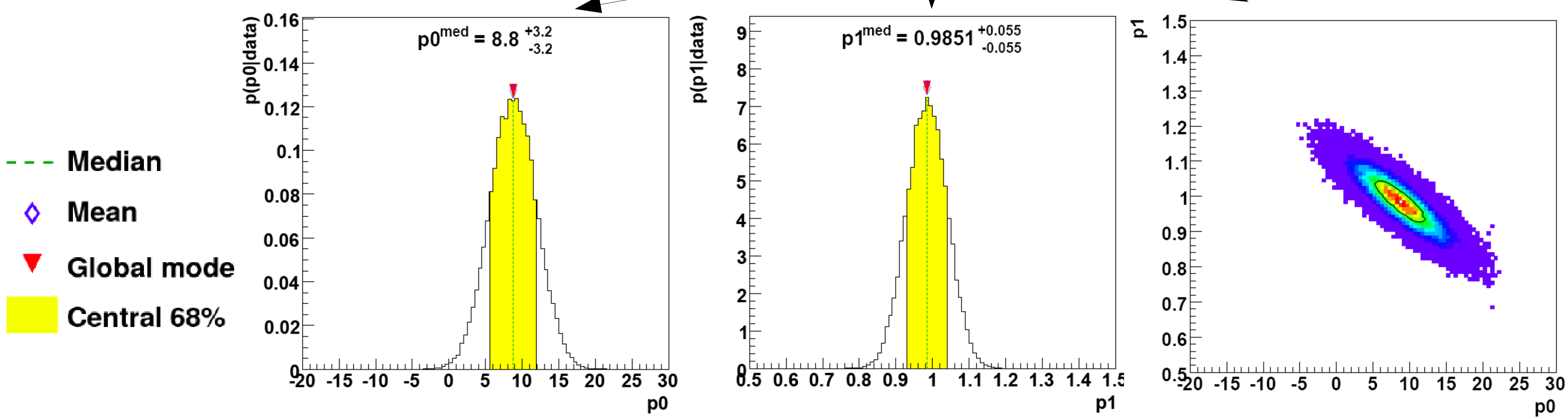
Graphical output

straight line fit



Fit and the error band representing the central 68% probability interval

Marginalized posterior probability densities



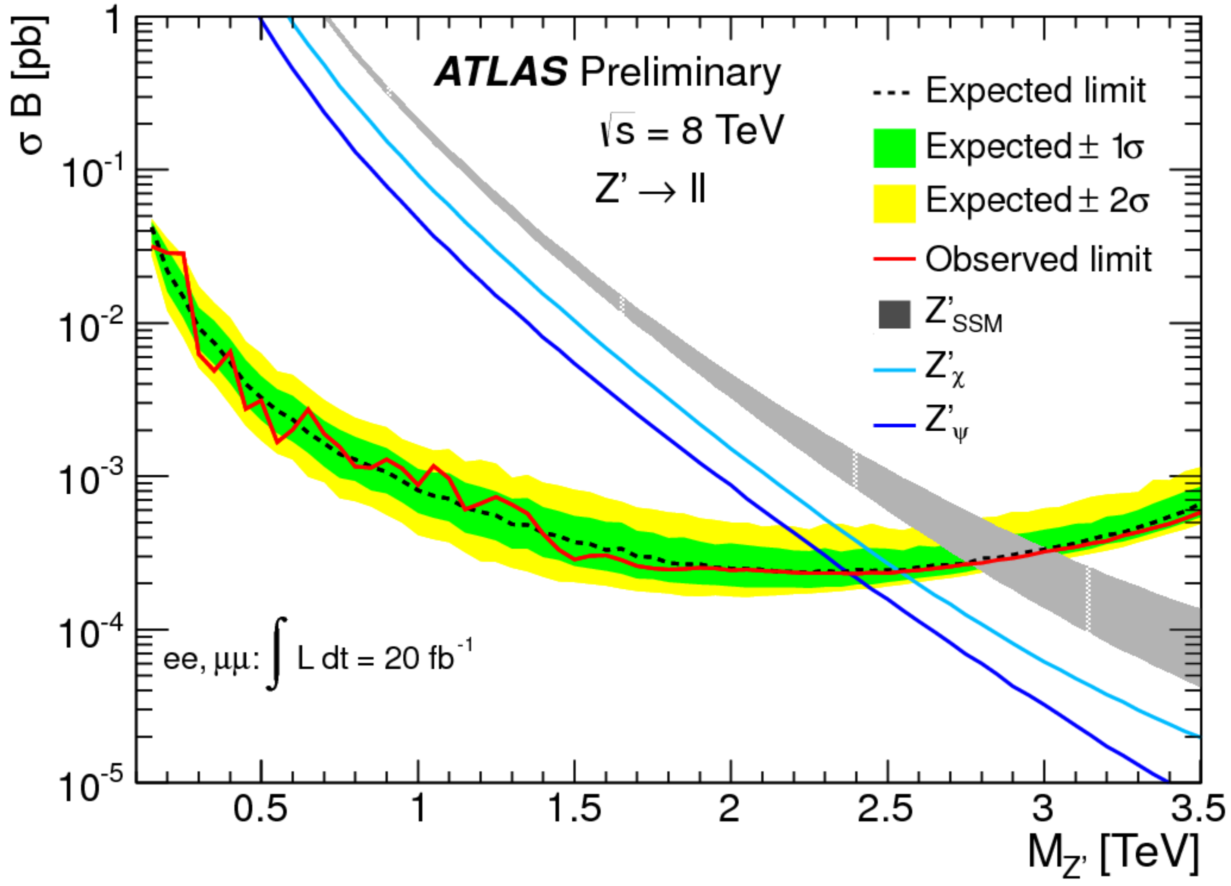
Search for high-mass dilepton resonances in 20 fb−1 of pp collisions at √s=8 TeV with the ATLAS experiment

ATLAS-CONF-2013-017

- Search for $Z' \to \ell^+\ell^-$ in various models
- Poisson in single bin $P(n|\lambda) = \dfrac{\lambda^n e^{-\lambda}}{n!}$
- Signal and bkg decomposition $\lambda = \lambda_s + \lambda_b$
- Signal $\lambda_s = \sigma\mathcal{B}\mathcal{L}\varepsilon$ depends on nuisance $\vec{\nu}_1$
- Background $\lambda_b(\vec{\nu}_1, \vec{\nu}_2)$
- Input: only positive signal $\mathrm{P}_0(\lambda_s) = \begin{cases} \text{const} & \lambda_s \geq 0 \\ 0 & \lambda_s < 0 \end{cases}$
- Posterior $P(\lambda_s|D) \propto \int \mathrm{d}\vec{\nu} P(D|\lambda_s, \lambda_b(\vec{\nu}))P_0(\lambda_s, \vec{\nu})$
- Uncertainty propagation $P(\lambda_s|D) \Rightarrow P(\sigma\mathcal{B}|D)$

Search for high-mass dilepton resonances in 20 fb−1 of pp collisions at √s=8 TeV with the ATLAS experiment

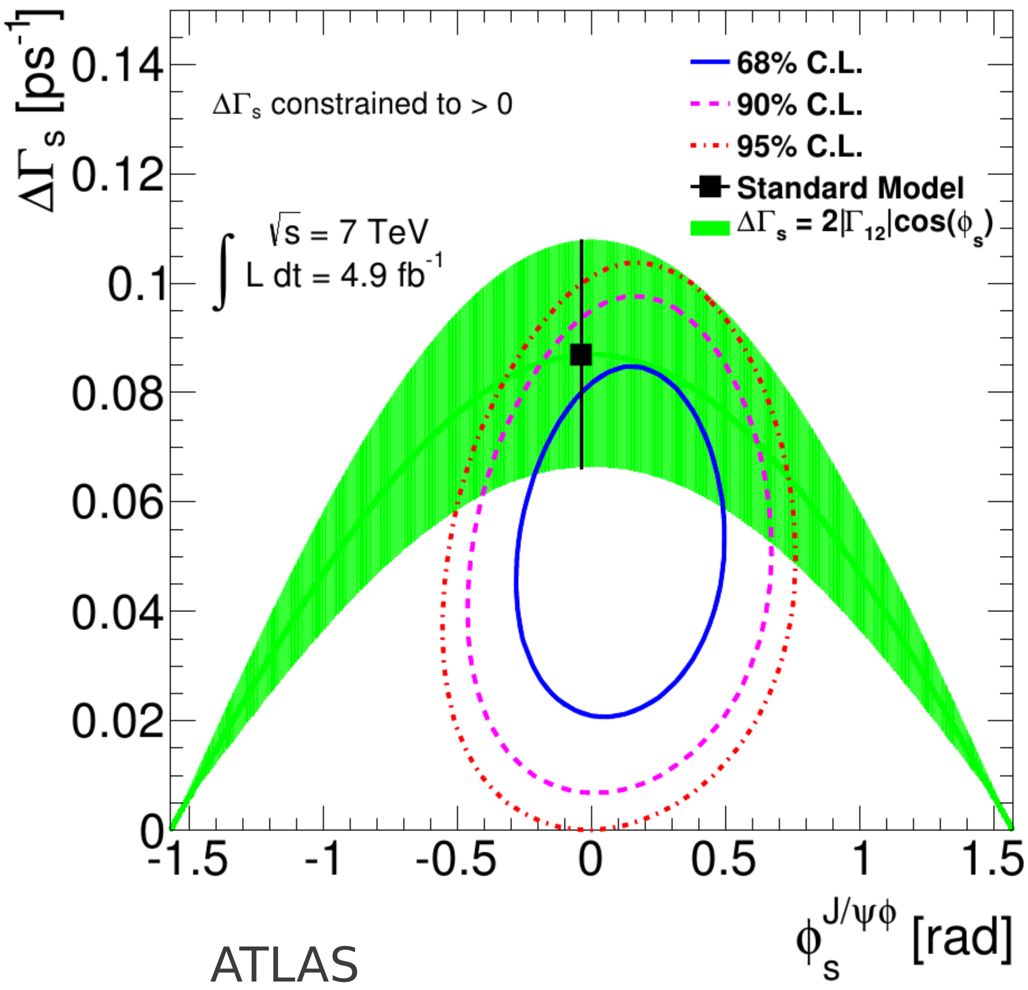
ATLAS-CONF-2013-017



Time-dependent angular analysis of $B_s \to J/\Psi\phi$

ATLAS-CONF-2013-039

Fit with B_s flavour-charge tagging



- Extract $\Delta\Gamma_s, \Phi_s$
- Challenge: bkg pdf systematics

- Single fit: 2h
- 50x50 grid in each plane

Improving the speed

$$\mathcal{L} \sim \prod_{1}^{N} w_i \cdot [f_s \cdot \boxed{\mathcal{F}_s(m_i, t_i, \Omega_i)} + f_s \cdot f_{B^0} \cdot \mathcal{F}_{B^0}(m_i, t_i, \Omega_i) + (1 - f_s - f_s \cdot f_{B^0}) \cdot \mathcal{F}_{bck}(m_i, t_i, \Omega_i)]$$

$$\mathcal{L} = \prod_{1}^{N} w_i \cdot [f_s \cdot \boxed{P_s(m_i|\sigma_{m_i}) \cdot P_s(\sigma_{m_i}|p_{t_i}) \cdot \boxed{P_s(\Omega_i, t_i|\sigma_{t_i})} \cdot P_s(\sigma_{t_i}|p_{Ti}) \cdot \boxed{A(\Omega_i, p_{Ti})} \cdot P_s(p_{Ti}) \cdot P_s(p_{B|Q_i}|\mathcal{M}_i) \cdot P_s(\mathcal{M}_i)} + \ldots$$

$$\boxed{P_s(\Omega, t|\sigma)} = \int \mathrm{d}\tau \, \frac{\mathrm{d}\Gamma}{\mathrm{d}t\,\mathrm{d}\Omega} \mathcal{N}(\tau|t, \sigma) \qquad \frac{\mathrm{d}\Gamma}{\mathrm{d}t\,\mathrm{d}\Omega} = \sum_k \mathcal{O}^k(t) g^k(\theta, \Psi, \varphi)$$

k	$\mathcal{O}^{(k)}(t)$	$g^{(k)}(\theta_T, \psi_T, \varphi)$
1	$\frac{1}{2}\|A_0(0)\|^2 \left[(1+\cos\phi_s)\, e^{-\Gamma_L^{(s)} t} + (1-\cos\phi_s)\, e^{-\Gamma_H^{(s)} t} \pm 2e^{-\Gamma_s t} \sin(\Delta m_s t) \sin\phi_s\right]$	$2\cos^2\psi_T (1 - \sin^2\theta_T \cos^2\varphi_T)$

Normalization: Convolute with acceptance in p_T bins

$$\int \mathrm{d}\Omega\,\mathrm{d}t \, \boxed{P_s(\Omega, t|\sigma)} \boxed{A(\Omega, p_T)}$$

Previous: 3D histogram with 50^3 bins per p_T

New: only 1000 – 30000 calls with cubature (cuba library). Trouble with Minuit convergence