A Bayesian approach to gravitational lens model selection

Irène Balmès, Pier-Stefano Corasaniti
LUTh, Observatoire de Paris–Meudon, Meudon, France

Objective

Devising a Bayesian statistical discrimination of lens models, for the purpose of using time delays measurements as a cosmological probe.

Introduction

In a strong gravitational lens, each image is the result of a different light-path (see figure 1). As a result, if the source behind the lens has a variable luminosity, it will manifest with a **time delay** between the two images.

This time delay Δt depends on the gravitational potential of the lens, and the underlying cosmological model.

The time delay between two images A and B is:

$$\Delta t_{A,B} = (1+z_l)\frac{d_l d_s}{d_{ls}}\left(\frac{1}{2}((\theta_A-\beta)^2-(\theta_B-\beta)^2) + \psi(\theta_A)-\psi(\theta_B)\right)$$

where $\Delta t_{A,B}$, z_l, θ_A and θ_B are observables, β, $\psi(\theta_A)$ and $\psi(\theta_B)$ depend on the lens model and d_l, d_s and d_{ls} depend on cosmology.

Using the above relation, we can derive constraints on cosmological parameters, provided we assume a lens model. Time delays are particularly sensitive to the value of the Hubble constant H_0.

Unfortunately, a change in the lens model can shift the inferred value of H_0 by a factor of two. Hence, the modeling of the lens, as well as a robust discrimination between lens models, is critical to the study of time delays.

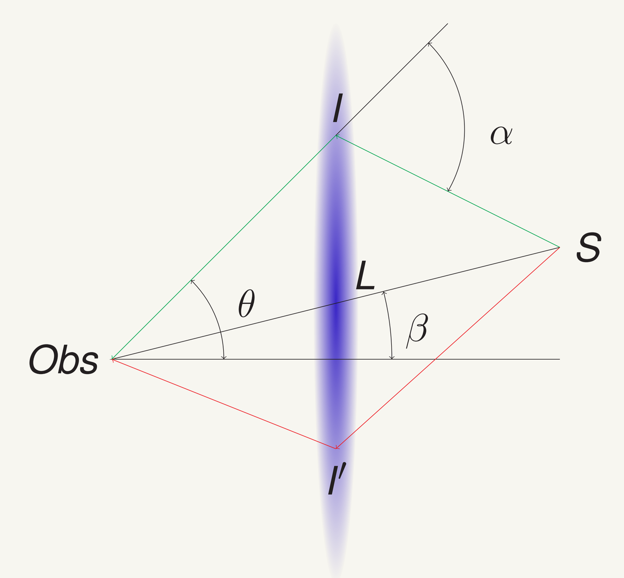

Figure 1: The lens *L* bends the light from the source *S*, so that the observer in *Obs* sees two (or more) images, *I* and *I′*, of the same source.

Various lens models

A large number of lens models have been considered in a vast litterature. Therefore, to start with, we restrict ourselves to simple examples, characterized by few parameters, given the fact that observables are limited.

Constraints

These are:

positions of the images;

time delays;

flux ratios between images.

Our lens models

We consider realistic models for lenses with two images, so called double lenses (cf Kochanek et al., 2004).

(*a*) **Isothermal lens**: this is the simplest model for a lens, which has the advantage of simplifying the time delay expression.

(*b*) **Power-law model:** a generalization of the previous case, is given by assuming a density profile ρ as:

$$\rho \propto r^{-n}$$

It has one parameter n, which for $n=2$, coincides with the isothermal lens. We adopt two different priors, a large prior, $0<n<3$, and a more restrictive one $1<n<3$.

(*c*) **Power-law model with external shear:** in addition to the previous model, we include a shear that accounts for environmental effect on the lens. This adds two parameters: the strength of the shear γ, and its direction. Expected values for the shear vary up to $\gamma \simeq 0.1$, therefore we adopt three different priors on γ: $\gamma<0.1$, $\gamma<0.2$ and $\gamma<0.5$ respectively. This allow us to test the shear strength up to nearly unrealistic values.

This three models are characterized by different numbers of parameters and thus different prior parameters volumes.

Bayes factor

This Bayes factor provides a robust statistical approach to quantify the hability of models, with respect to others, to account for the observations.
The Bayes factor between models i and j is:

$$B_{i,j} = \frac{P(M_i|D)}{P(M_j|D)} = \frac{P(D|M_i)}{P(D|M_j)}$$

where D is the data. The last equality is obtained through Bayes theorem, with the assumption that the prior probabilities on the models are taken equal. We can here recognize the evidence, that we can calculate through:

$$P(D|M_i) = \int d\theta P(D|\theta, M_i)P(\theta|M_i)$$

where θ are the parameters to the model. In the integrand, we can identify $P(D|\theta, M_i)$, the likelihood and $P(\theta|M_i)$ the prior probability on the parameters.

We obtain the likelihood by the usual formula:

$$P(D|\theta, M_i) = \frac{1}{\sqrt{2\pi}\sigma}\exp -\frac{(D_{th}(\theta, M_i)-D)^2}{2\sigma^2}$$

Results

Up to now, there are 11 double lenses with measured time delays. We present results from the comparison between models (*b*) and (*c*) for each of these lenses in figure 2. For this analysis, we adopt a concordance ΛCDM cosmology, with $H_0 = 70\,\mathrm{km/s\,Mpc^{-1}}$ and $\Omega_m = 0.27$.

We adopt the criterium for conclusiveness of the Bayes factor presented in Trotta (2008): if $\ln B_{i,j} > 5$, the evidence is strong in favor of model i compared to j, while in the case $\ln B_{i,j} < -5$, model j is favored compared to model i. The case $-5 < \ln B_{i,j} < 5$ is considered inconclusive. For visual purposes, in figure 2, we therefore plot $y = \mathrm{sign}(x)(\log_{10}|x| + 0.5)$ with $x = \ln B_{b,c}$.
We also computed the Bayes factor for different combinations of the observable constraints.

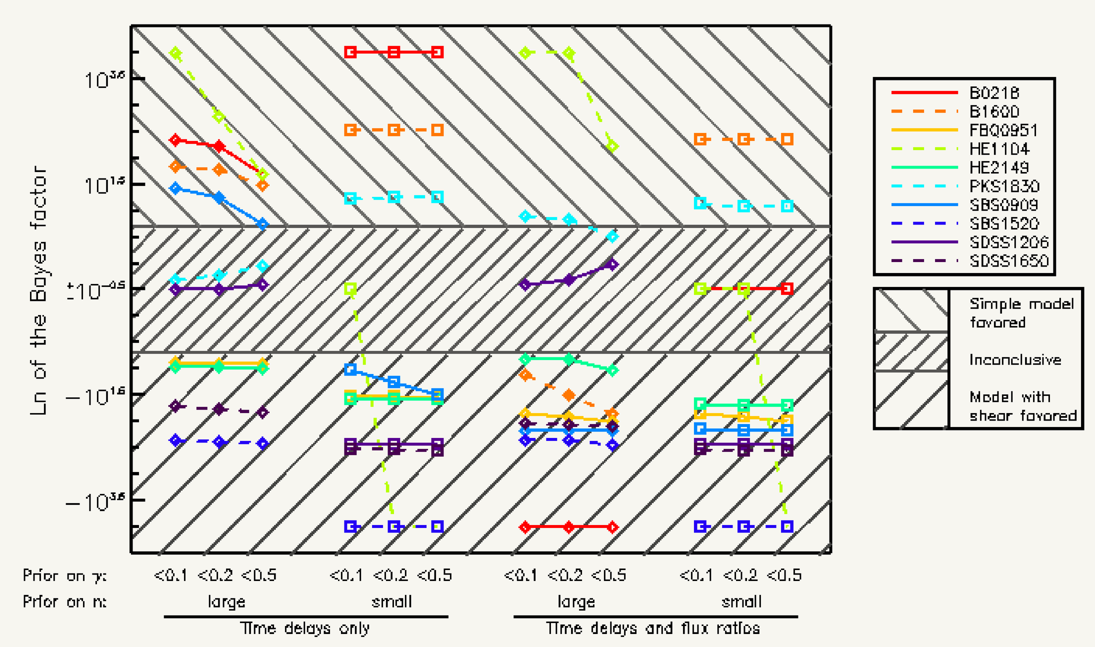

Figure 2: Bayes factor between the power-law model and the power-law plus shear model, for all double lenses and with different priors. The lens Q0957 has been omitted in this graph since all draws for this lens give $\ln B_{b,c} = -\infty$.

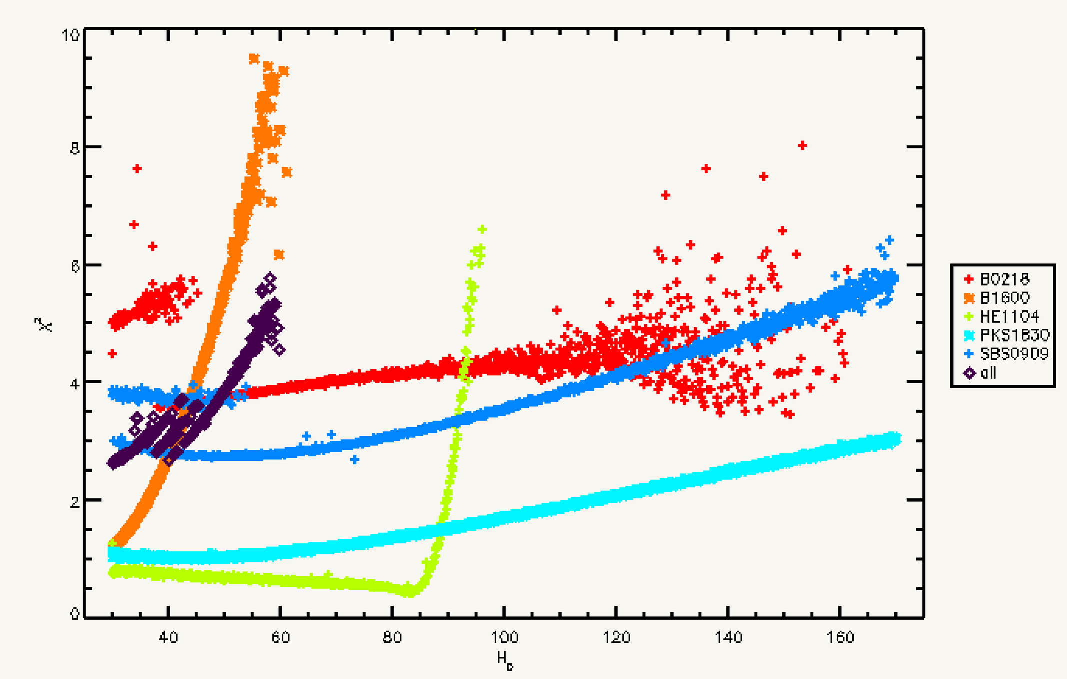

Figure 3: χ^2 of the Hubble parameter with the subset of double lenses selected through our Bayesian analysis.

Interpretation

Effect of the prior on n
The lens data set is mainly composed of galaxies, which we expect to be nearly isothermal. Nevertheless, our analysis show that few of them are in fact well described by a power-law model with a reduced prior on n, while the majority favors the more complex model, which also include the shear. In the case of the larger prior, $0<n<3$, the number of lenses well described by the power-law increases.

Effect of the prior on γ
In more than half of the cases, allowing higher (unrealistic) shear strength does not change the Bayes factor. This is a consequence of Occam's razor: as the parameter space grows, the fit gets better and better, but this effect is compensated for by Bayes factor. In a quarter of the cases, widening the prior on γ favors the more complex model, as the fit gets sufficiently better to over-compensate for the Occam's razor term.

Effect of the flux ratios
Time delays depend on the gravitational potential of the lens, whereas flux ratios depend on its second derivative.
Furthermore, they are subject to a number of local phenomena (microlensing, absorption...) that do not affect time delays. Therefore, flux ratios can be hardly described with a simple model, eventually requiring a more complex modeling than needed by time delays. This is consistent with our findings: in fact, adding flux ratios as a constraint leads to having less lenses accurately described by model (*b*), since $\ln B_{b,c}$ decreases.

Conclusions and preliminary results on H_0

In figure 3, we show preliminary results on the MCMC likelihood analysis of H_0 in a concordance ΛCDM model, inferred from lenses which are compatible with a power law lens model. However, we have reason to believe that the structure of the likelihood is not well sampled by standard Monte Carlo approaches, thus requiring a grid analysis of the prior parameter space.

References

1. C.S. Kochanek, P. Schneider, and J. Wambsganss, Part 2. In G. Meylan, P. Jetzer and P. North, editors, *Gravitational Lensing: Strong, Weak & Micro, Proceedings of the 33rd Saas-Fee Advanced Course, Berlin Springer-Verlag*, 2004
2. R. Trotta, Bayes in the sky: Bayesian inference and model selection in cosmology, *Contemporary Physics*, 49:71–204, Mar 2008

Mail: irene.balmes@obspm.fr