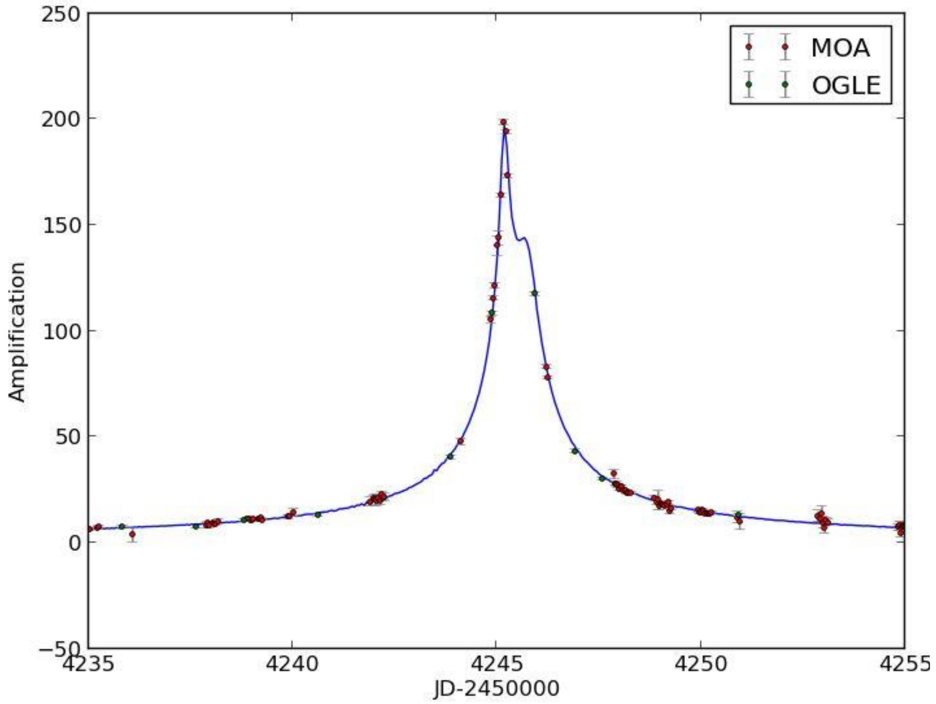
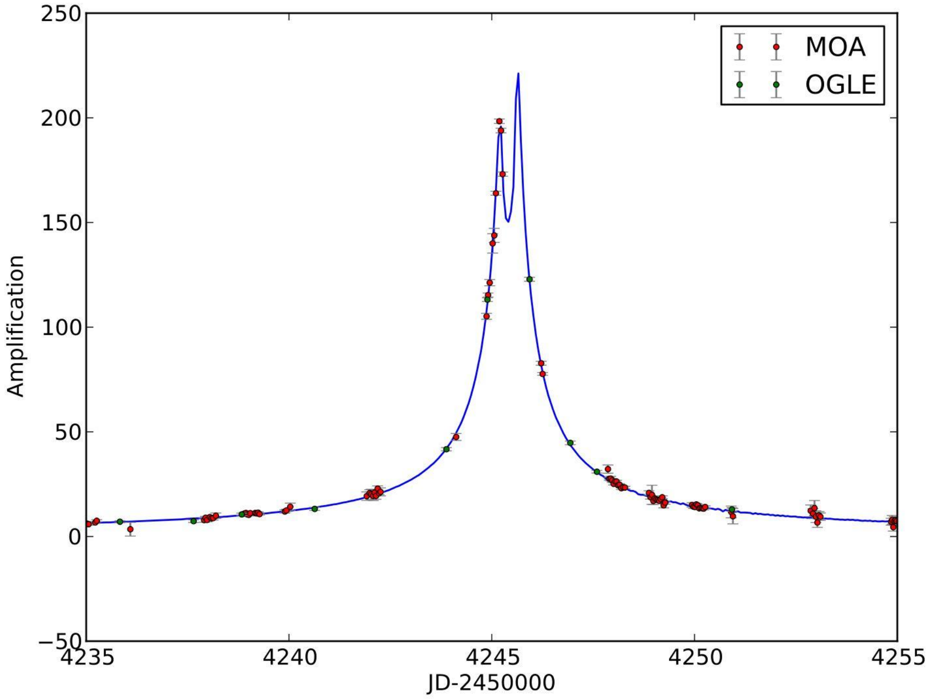
Bayesian Analysis of Microlensing Events using MultiNest

PhD: Ashna Sharan
Supervisor: Dr. Nicholas Rattenbury
Department of Physics, UOA

Model Selection - Which Curve?





Data Modelling

There are two distinct requirements for a complete analysis.

- **Parameter Estimation**

Find the parameter values that achieve closest fit to the data.

- **Model Selection**

Choose the best model between competing/alternative models.

Model Selection

- **Traditional Method:**

Chi-squared (χ^2) Goodness of Fit

+

Qualitative analysis to avoid physically implausible parameter values, overfitting. (Apply Occam's Razor, experimental & theoretical arguments from prior knowledge to inform the choice...)

- **Alternative**

Bayesian Inference

Bayesian Data Modelling

Bayes Theorem:

Posterior × Evidence = Likelihood × Prior

- ★ Concise statement about our state of knowledge before and after data is considered.

Prior

Posterior × Evidence = Likelihood × Prior

What we know about
the parameters before
considering the data.

Likelihood

Posterior × Evidence = Likelihood × Prior

Quantifies the degree to which the model prediction and data agree.

Example Function:

Log-likelihood = Constant - $\chi^2 / 2$

Posterior

Posterior × Evidence = Likelihood × Prior

What we know about the parameters after considering the data.

➔ Parameter Estimation

Evidence

Posterior × Evidence = Likelihood × Prior

Probability that a particular model gave rise to the data (irrespective of the parameter values.)

→ Model Selection

Evidence

- An integration over the entire parameter space of the model $\longrightarrow$ prior-weighted average of the likelihood.

- More complicated models with larger parameter spaces get penalised $\longrightarrow$ Occam's Razor quantitatively implemented.

- Straightforward model selection but expensive to compute!

Nested Sampling

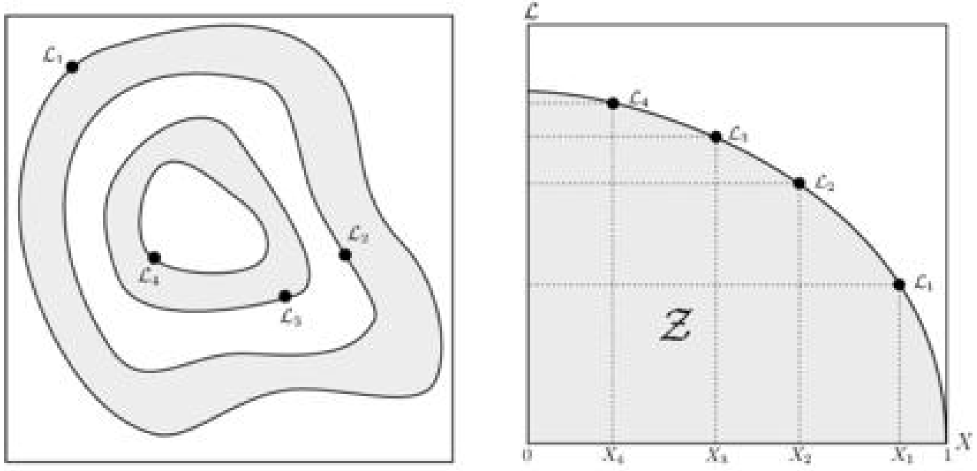
- John Skilling (2004).
- Computes the evidence integral, numerically, as a summation, thus affordable. Posteriors computed as a by-product.
- Essentially, N live points are sampled from a prior space, sorted according to their log-likelihood values and at each iteration the lowest log-likelihood value point is replaced by one with a higher log-likelihood.

Multi-modal Nested Sampling

MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics

F. Feroz*, M.P. Hobson and M. Bridges
Astrophysics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK

More efficient by clustering the live points into ellipsoids and sampling new points only from these ellipsoids.



(a) (b)

Image Credit: Feroz, et al.

Advantages

- Quantitative implementation of Occam's Razor.
- Easy to avoid implausible physical parameter estimations by constraining Bayesian priors using prior knowledge of typical parameter values.
- Simultaneous model selection and parameter estimation as a by-product.

MNRAS **000**, 000–000 (2016) Preprint 17 October 2016 Compiled using MNRAS LaTeX style file v3.0

Bayesian Analysis of Microlensing Events using MultiNest

A. Sharan,[1*] N. J. Rattenbury,[1]and C. Ling,[2]

Affiliations appear at the end of the paper

Accepted .... Received ....; in original form ......

ABSTRACT

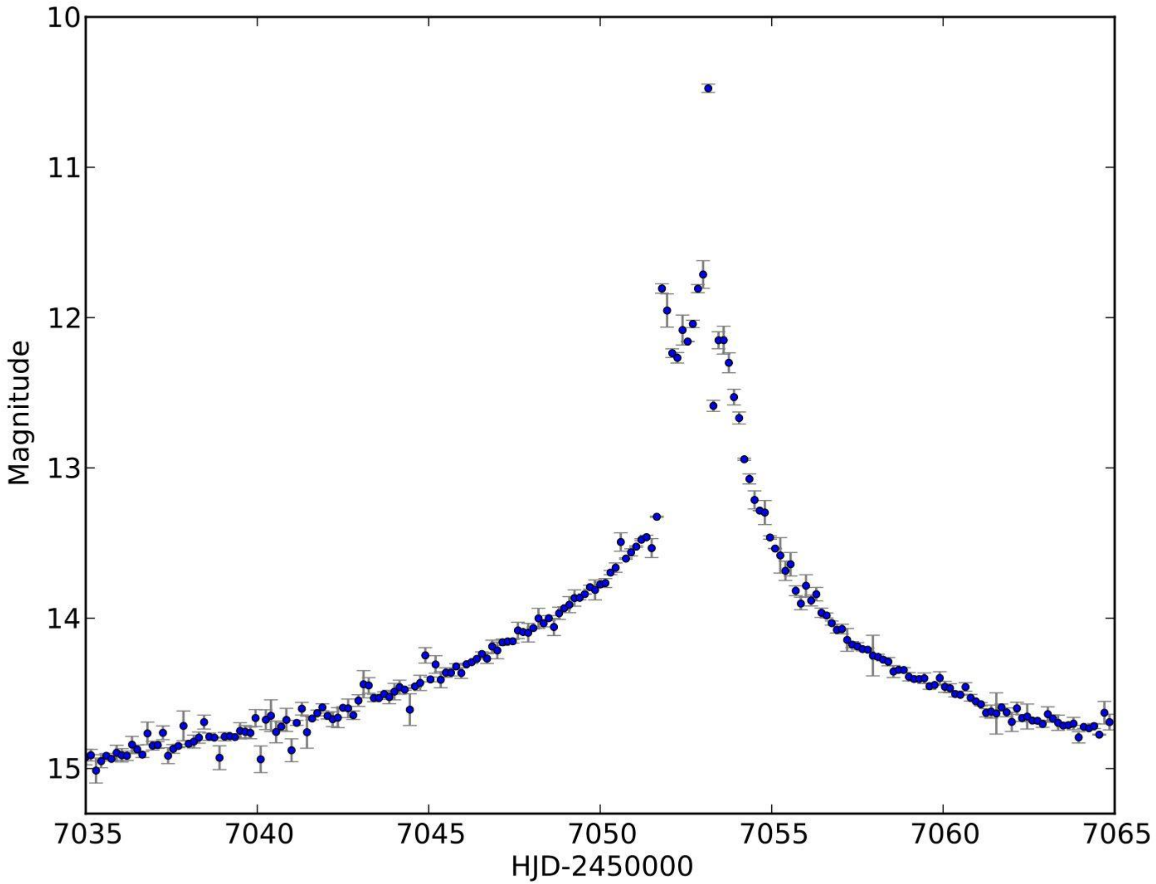
We present the analysis of microlensing events using our newly developed methodology employing the MultiNest algorithm. MultiNest is based on the principles of Bayesian inference, which allows us to solve the model selection and parameter estimation problems simultaneously. The focus is placed on the model selection problem since a Bayesian based algorithm such as MultiNest allows us to shift the approach to model selection from qualitative arguments to a quantitative quality factor.

We demonstrate our methodology by testing a finite-source point-lens model versus a finite-source binary-lens model, and for presence of parallax effects. We do this for a simulated synthetic event and for a real event, OGLE-2011-BLG-0251.

Nested Sampling and its variant algorithms such as MultiNest have been tried and tested in many fields of study. By demonstrating MultiNest on a real microlensing event for the first time, we aim to provide an impetus for said algorithms to find their place in the microlensing community as well.

Key words: gravitational lensing: micro – methods: data analysis – methods: statistical – techniques: miscellaneous

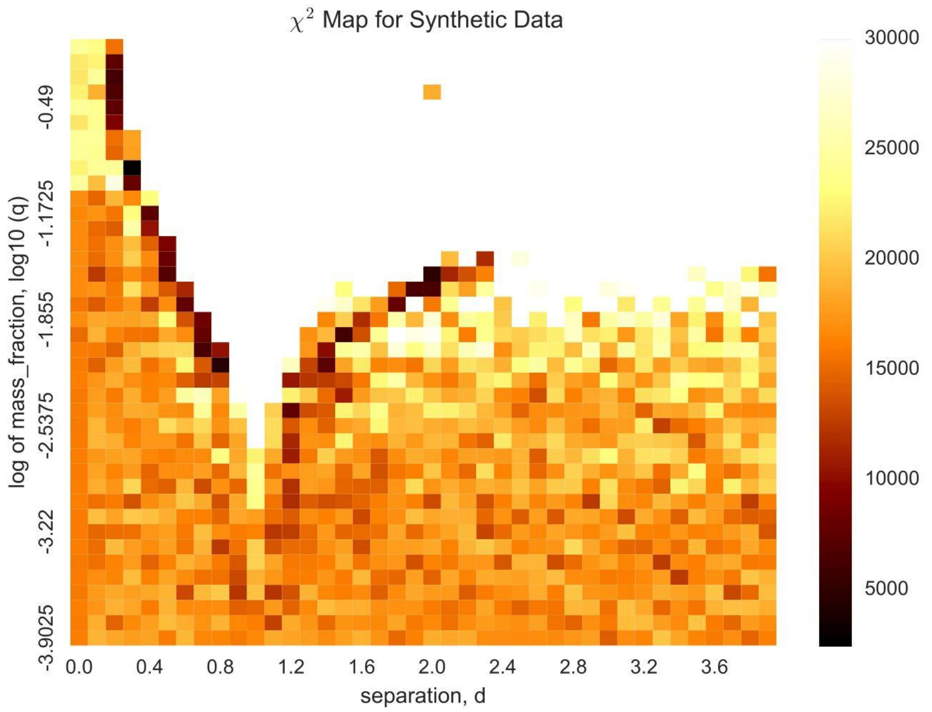
Simulated Event



Parameter	Actual Values
q	**0.03**
d	**2.0**
$t0$	**7050.0**
tE	**65.0**
$\mu 0$	**0.01**
ρ	**0.0005**
α	**3.0**

Simulated Event

Blind search in a prior space of full range of typical parameter values → not feasible, so:



→

Variable	Prior Intervals
q	[1e-2, 0.2]
d	[0.01, 1.0], [1.0, 2.0]
$t0$	[7050.0,7055.0]
tE	[60.0,80.0]
$\mu 0$	[0.005,0.007]
ρ	[7e-5,2e-4]
α	[2.5,3.0]
π_{EN}	[-0.5, 0.5]
π_{EE}	[-0.5, 0.5]

Simulated Event

Parameter	Actual Values	MLE
q	**0.03**	**0.02**
d	**2.0**	**1.9**
*t*0	**7050.0**	**7050.61**
*t*E	**65.0**	**78.32**
μ0	**0.01**	**0.01**
ρ	**0.0005**	**0.0002**
α	**3.0**	**3.0**

- **MultiNest run for 4 days.**
- **Estimated parameter values resemble the actual values used to generate the data (more or less).**

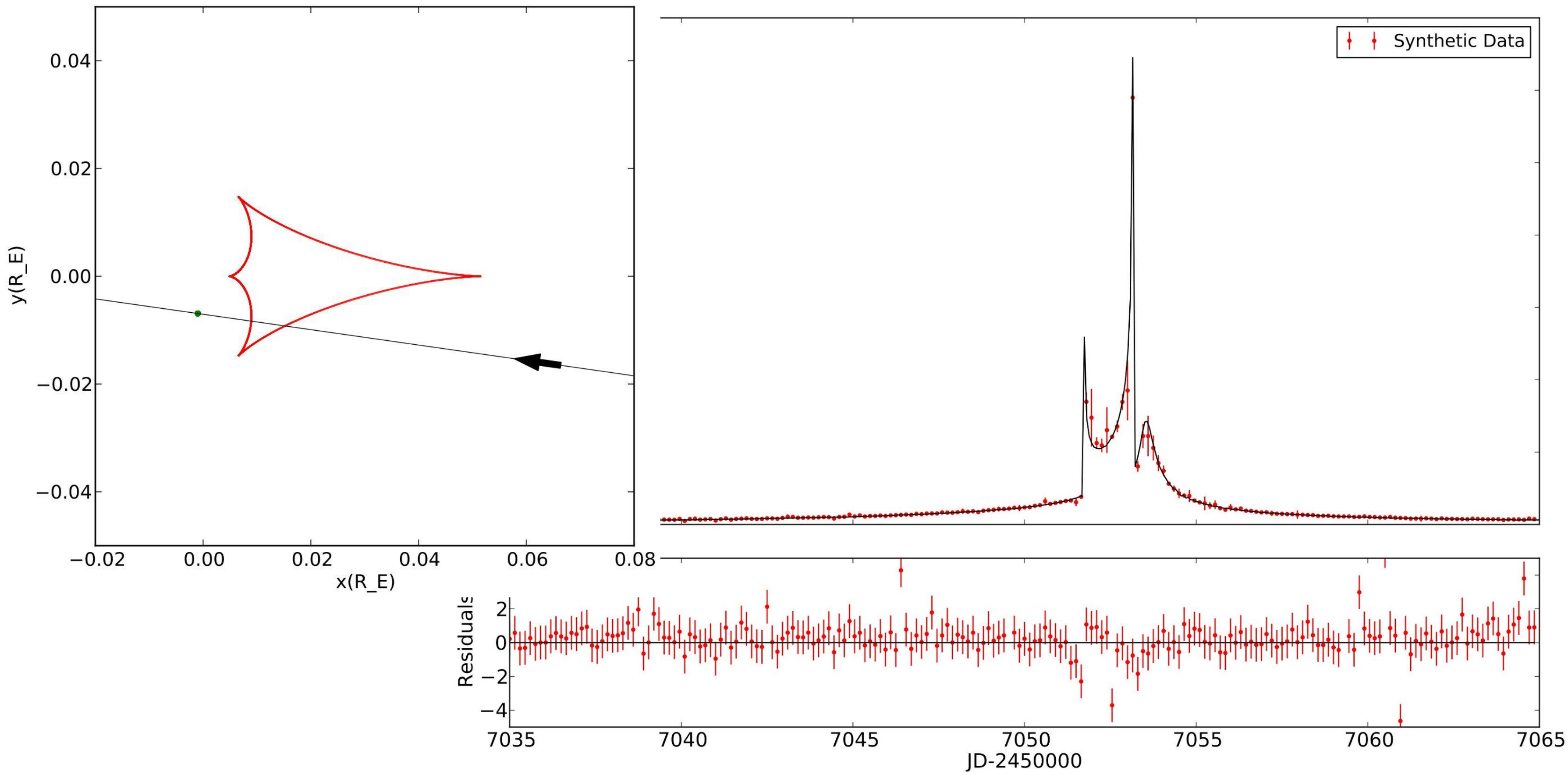
Simulated Event

Finite Source Binary Lens Model Solutions	**ℓn Z**	**χ^2 /dof**
With Parallax (Wide Orbit)	9936.95 ± 0.55	199.89 / 191
Without Parallax (Wide Orbit)	**9936.62 ± 0.56**	**310.96 / 193**
With Parallax (Close Orbit)	9906.42 ± 0.58	249.14 / 191
Without Parallax (Close Orbit)	9904.95 ± 0.58	393.52 / 193

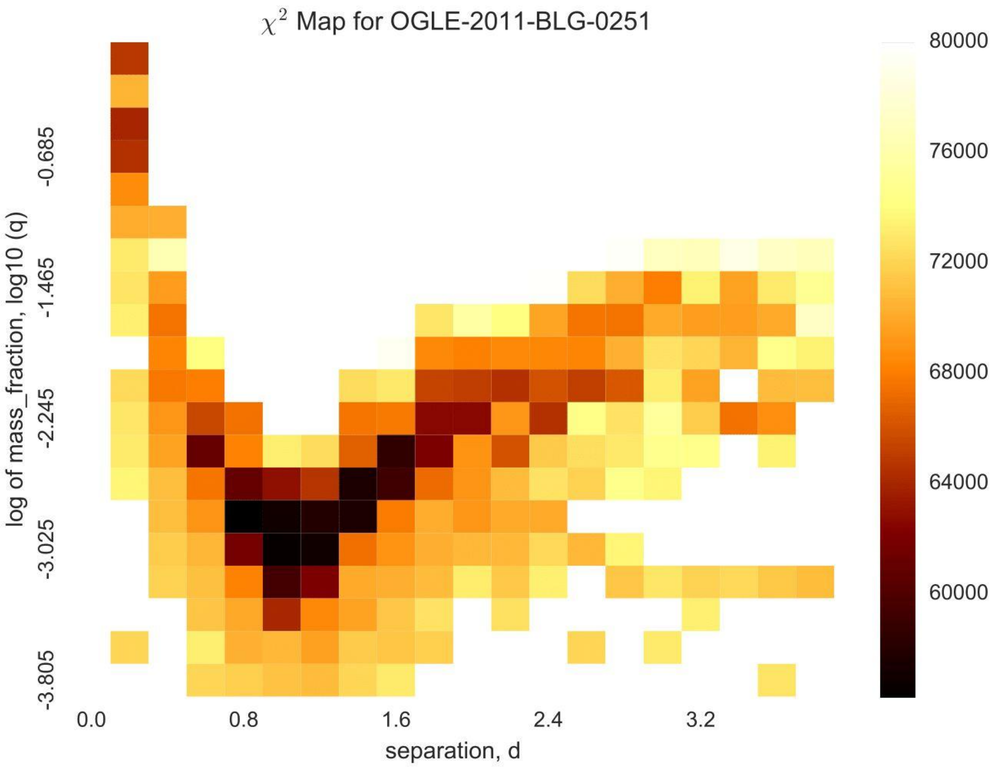
χ^2 method - choose the model with the parallax included because of the large improvement in χ^2.

Bayesian Evidence method - indistinguishable by evidence value comparison alone so we choose the simpler model with 7 parameters - without parallax.

Simulated Event



Real Event - OB110251



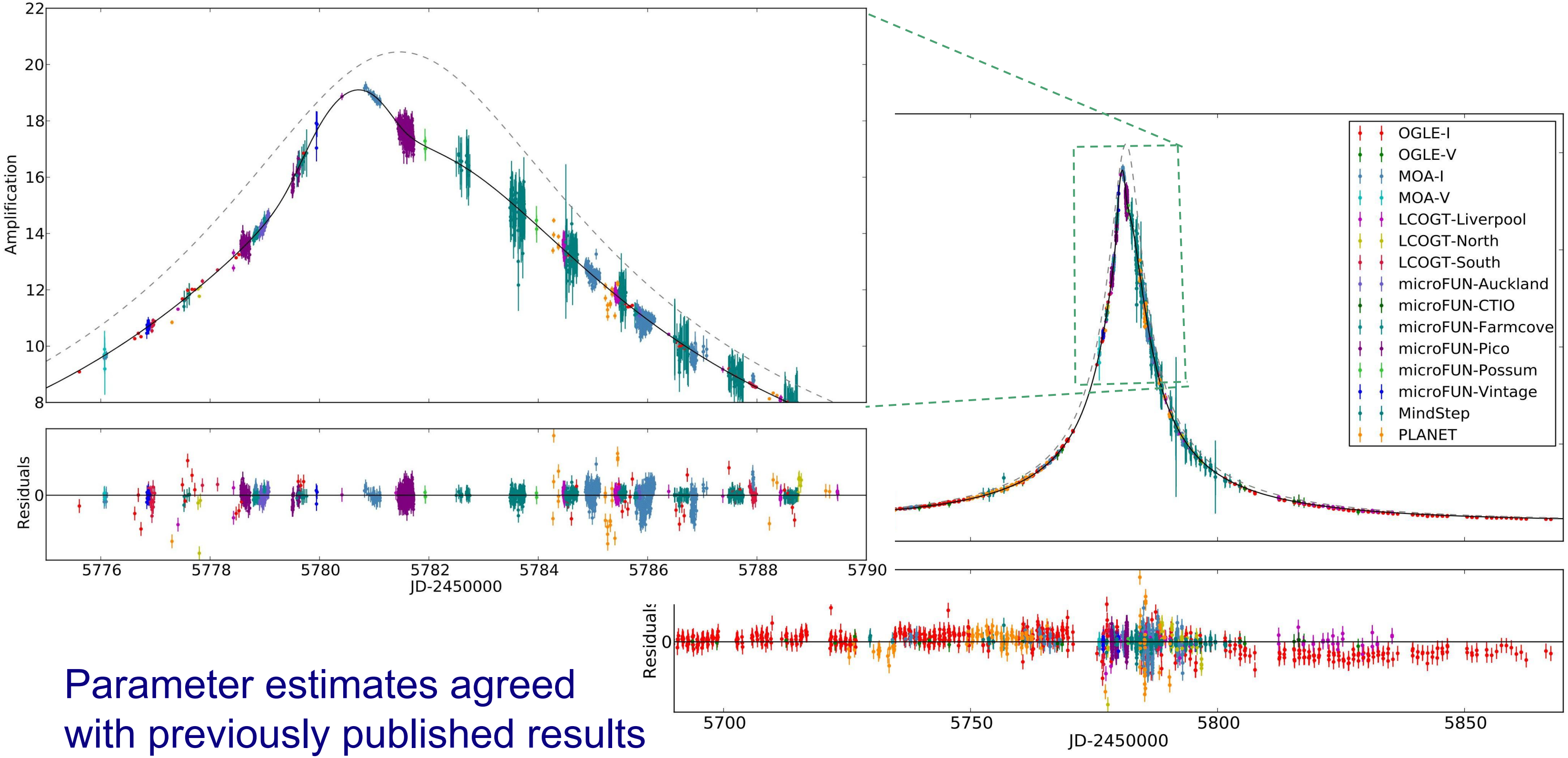
Variable	Prior Intervals
q	[5e-4, 5e-3]
d	[0.4, 1.0] & [1.0, 2.5]
$t0$	[5781.3, 5781.5]
tE	[60.0, 70.0]
$\mu 0$	[0.04, 0.06]
ρ	[0.001, 0.02]
α	[4.0, 5.0]
π_{EN}	[-0.5, 0.5]
π_{EE}	[-0.5, 0.5]

Real Event - OB110251

Finite Source Binary Lens Model Solutions	$\ell n\ Z$	χ^2 /dof
With Parallax (Wide Separation)	**8107.01 ± 0.52**	**3726.53 / 3729**
Without Parallax (Wide Orbit)	8103.50 ± 0.47	3740.67 / 3731
With Parallax (Close Orbit)	8094.53 ± 0.46	3763.37 / 3729
Without Parallax (Close Orbit)	8093.87 ± 0.46	3760.51 / 3731

- Most favourable model with highest log-evidence value also has the lowest χ^2 value. The 2 methods agree.
- Wide separation model with the parallax is favored over the model without the parallax by a factor of $e^{3.51}$

Real Event - OB110251



Parameter estimates agreed with previously published results to within **2σ.** Some discrepancies.

Challenges ...

- Gain: Straightforward and Quantitative Model Selection

Challenges:

- Higher computation time
  - With increase in prior range.
  - For complex real events.
- Doesn’t work (so far) for some events with sharp peaks due to caustic crossings.
- Is using Gridsearch to eliminate lower likelihood regions justified in the Bayesian framework? (reviewer’s comments)

Acknowledgments

- Dr. Nicholas Rattenbury
- C. Ling (Joe)
- Dr. Brendon Brewer
- Feroz Khan
- Johannes Buchner
- N. Kains, R. A. Street

A giant planet beyond the snow line in microlensing event OGLE-2011-BLG-0251

N. Kains[1,⋆], R. A. Street[2], J.-Y. Choi[3], C. Han[3,⋆], A. Udalski[4], L. A. Almeida[5], F. Jablonski[5], P. J. Tristram[6], U. G. Jørgensen[7,8], and

Thank You!