Probabilistic Catalogs *and beyond...*

Brendon J. Brewer

Department of Statistics
The University of Auckland

https://www.stat.auckland.ac.nz/~brewer/

Probabilistic catalogs

So, what is a *probabilistic catalog*? And what's *beyond*?

Finite Mixture Models

Finite Mixture Models

- These are what statisticians call **finite mixture models**. They have a lot of potential in astronomy.

Finite Mixture Models

- These are what statisticians call **finite mixture models**. They have a lot of potential in astronomy.

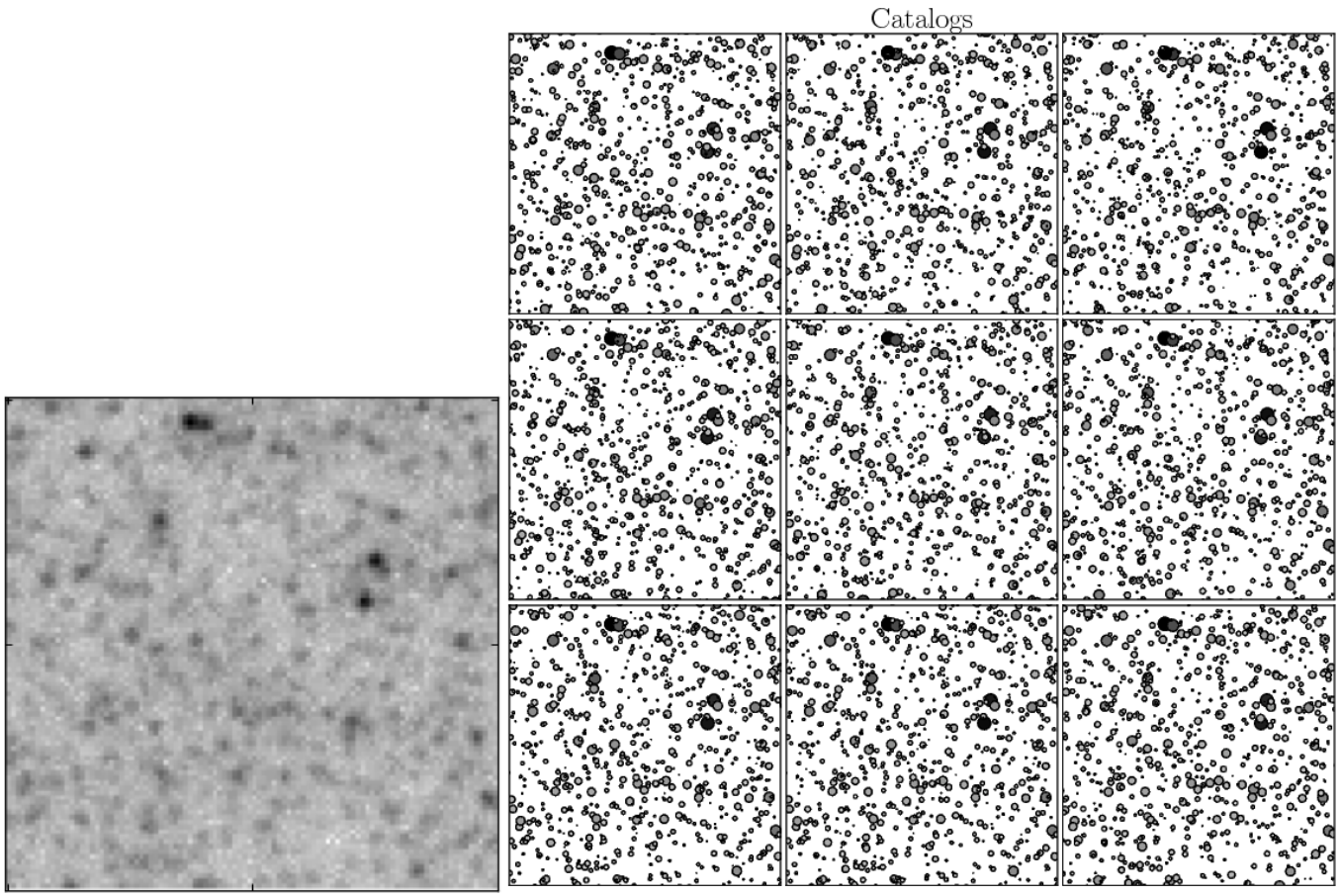
- The data have something to do with a **sum** of an unknown number of “components”, whose properties we want to know.

Finite Mixture Models

- These are what statisticians call **finite mixture models**. They have a lot of potential in astronomy.
- The data have something to do with a **sum** of an unknown number of "components", whose properties we want to know.
- I calculate the posterior distribution for the number of components N, and their parameters.

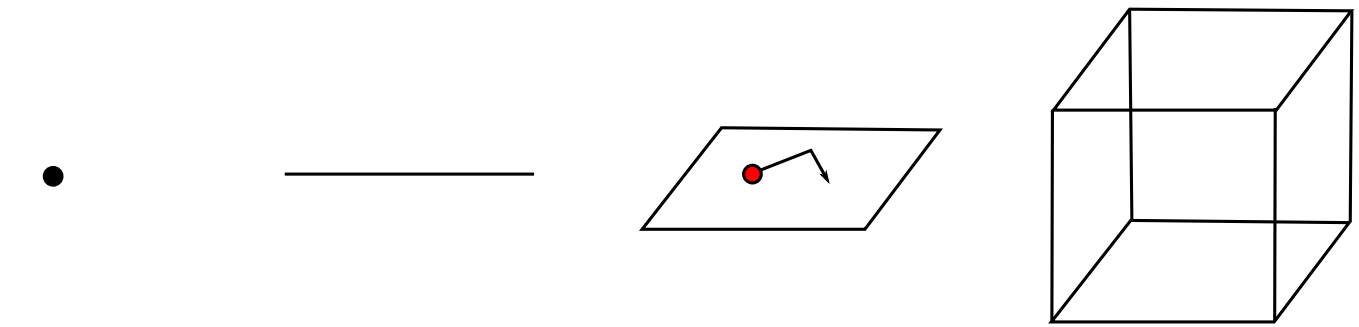
Probabilistic catalogs

Brewer, Foreman-Mackey, and Hogg, 2013, AJ, 146, 7.
arXiv: 1211.5805



(a) Data (b) Nine “plausible” catalogs

The hypothesis space

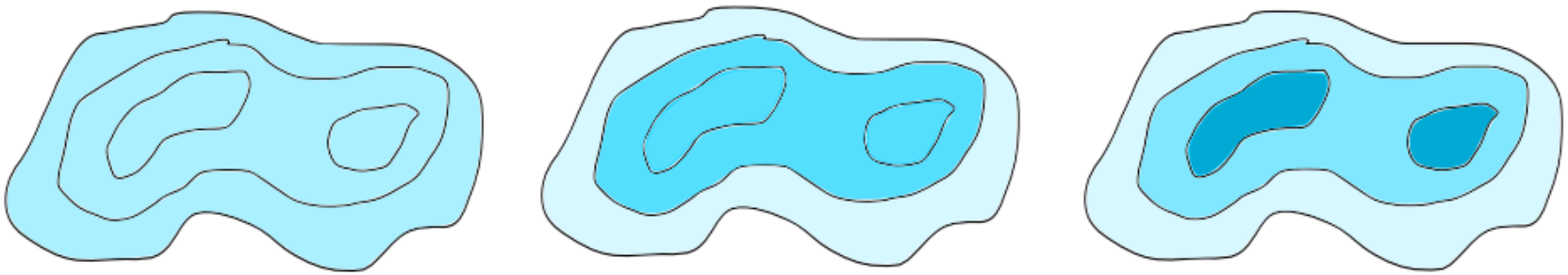


N=0 *N=1* *N=2* *N=3*

Computation

I use **DNest4** (Diffusive Nested Sampling)
Brewer, Pártay, and Csányi, 2011, Statistics and Computing, 21, 4, 649-656.

arXiv: 0912.2380



The target distribution is the *mixture of constrained priors*:

$$p(\theta) = \pi(\theta) \sum_{j=0}^{j_{\max}} \frac{w_j \mathbb{1}\left[L(\theta) > \ell_j\right]}{X(\ell_j)}. \quad (1)$$

CENTER FOR COSMOLOGY
UNIVERSITY OF CALIFORNIA, IRVINE

Statistical Methods

Bayesian Analysis Methods

Integration
Monte Carlo Integration
Rejection Method
Barker-Hastings Algorithm
Metropolis-Hastings Algorithm
Slice Sampling
Nested Sampling
MultiModal Nested Sampling
Metropolis Nested Sampling
Hamiltonian Sampling
Diffusive Nested Sampling

Method Complexity

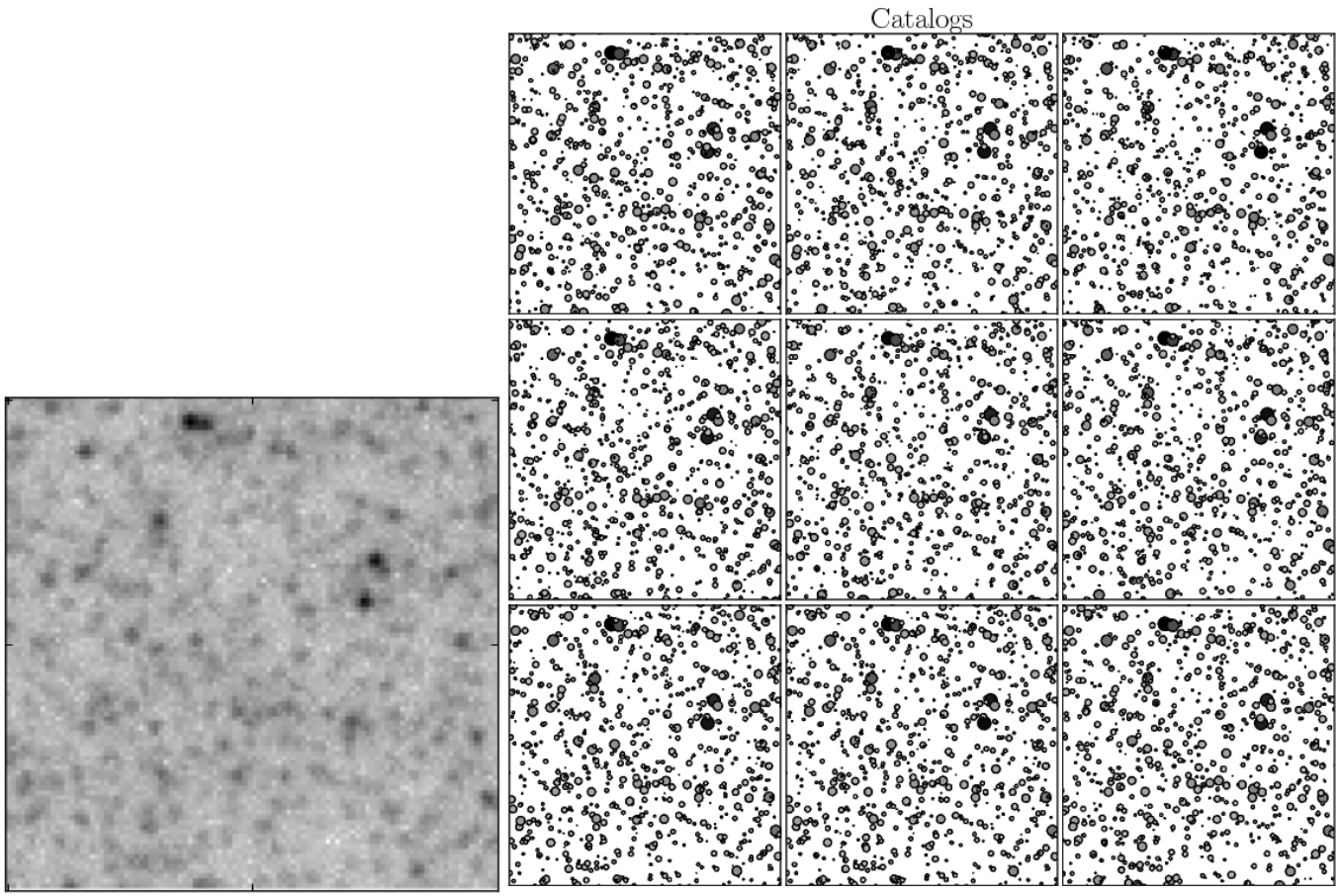
"Sure, I'll do it"

"You want me to do what?"

"I have no clue what's going on?"

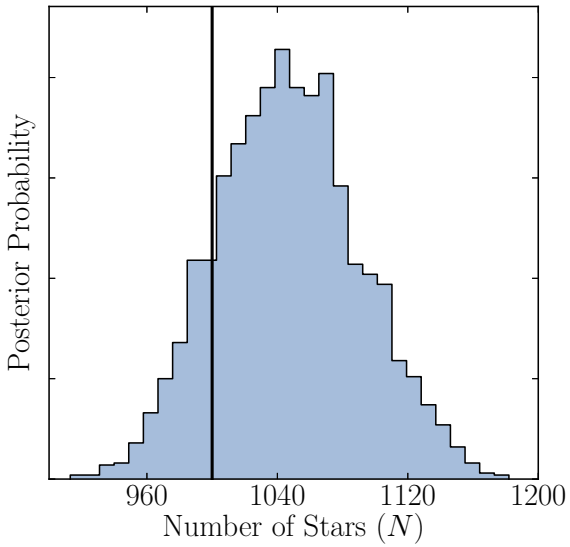
Probabilistic catalogs

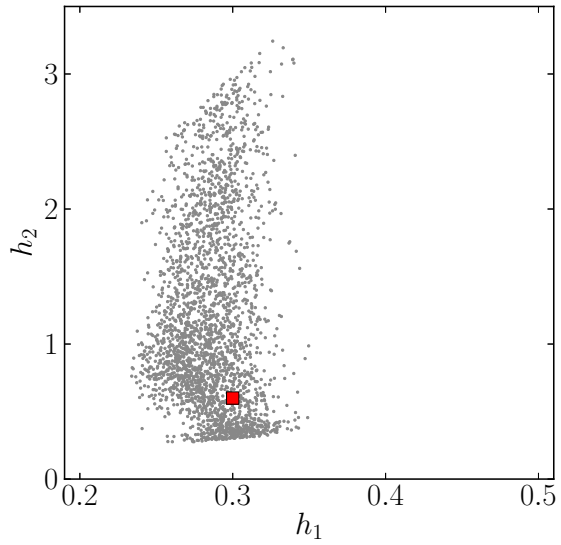
Brewer, Foreman-Mackey, and Hogg, 2013, AJ, 146, 7.
arXiv: 1211.5805

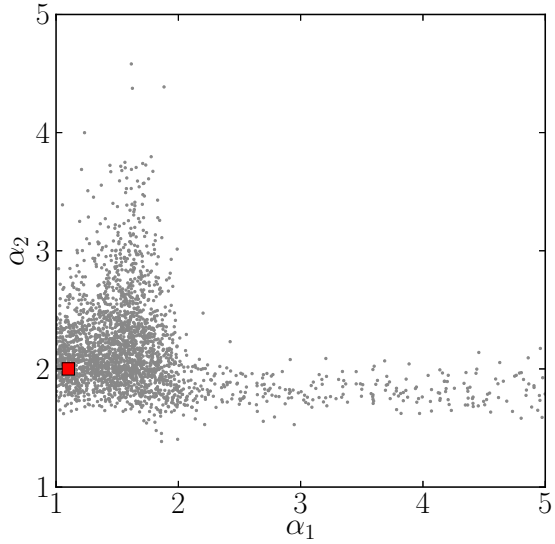


(c) Data (d) Nine “plausible” catalogs

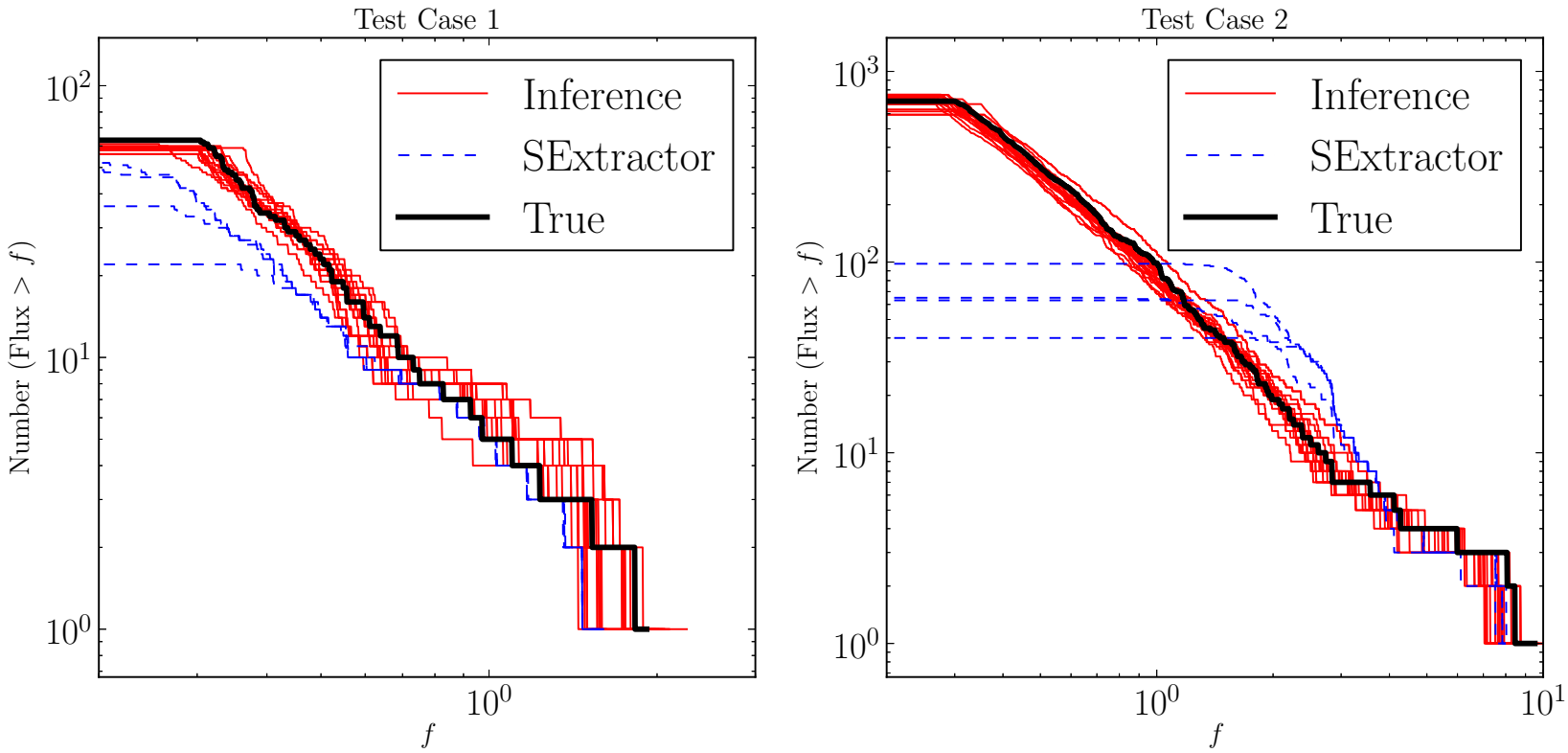
Results: Hyperparameters







Results: Luminosity Function



First real application!

- A candidate dwarf around a nearby galaxy (all other information **suppressed** by me not being a proper astronomer)..

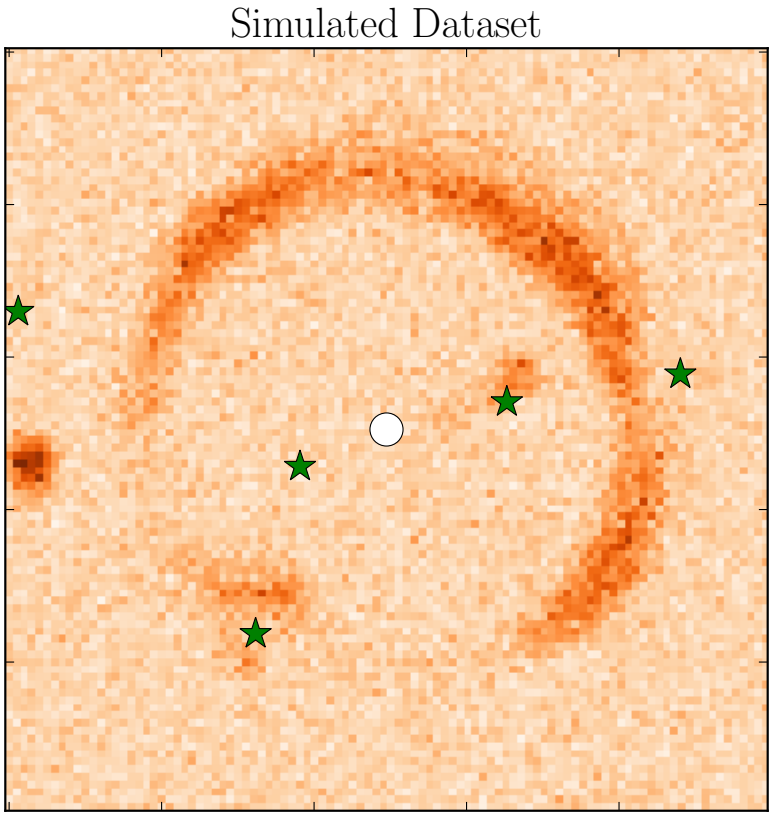
First real application!

- A candidate dwarf around a nearby galaxy (all other information **suppressed** by me not being a proper astronomer)..
- Four bands, unknown (5-parameter) PSF in each band, fluxes different in each band, hierarchical prior for fluxes.

First real application!

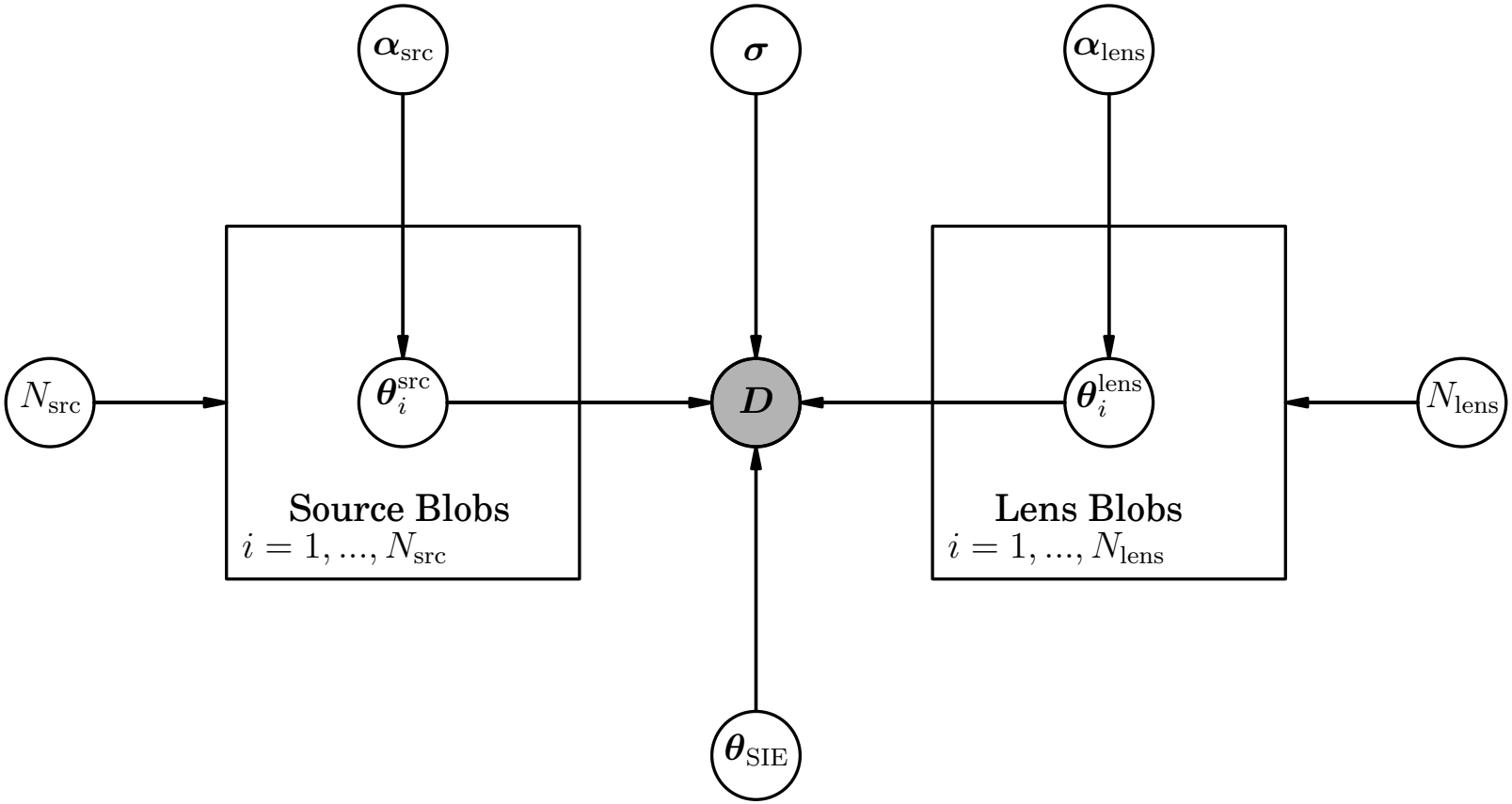
- A candidate dwarf around a nearby galaxy (all other information **suppressed** by me not being a proper astronomer)..
- Four bands, unknown (5-parameter) PSF in each band, fluxes different in each band, hierarchical prior for fluxes.
- Play movie.mkv

Gravitational lensing and dark substructures

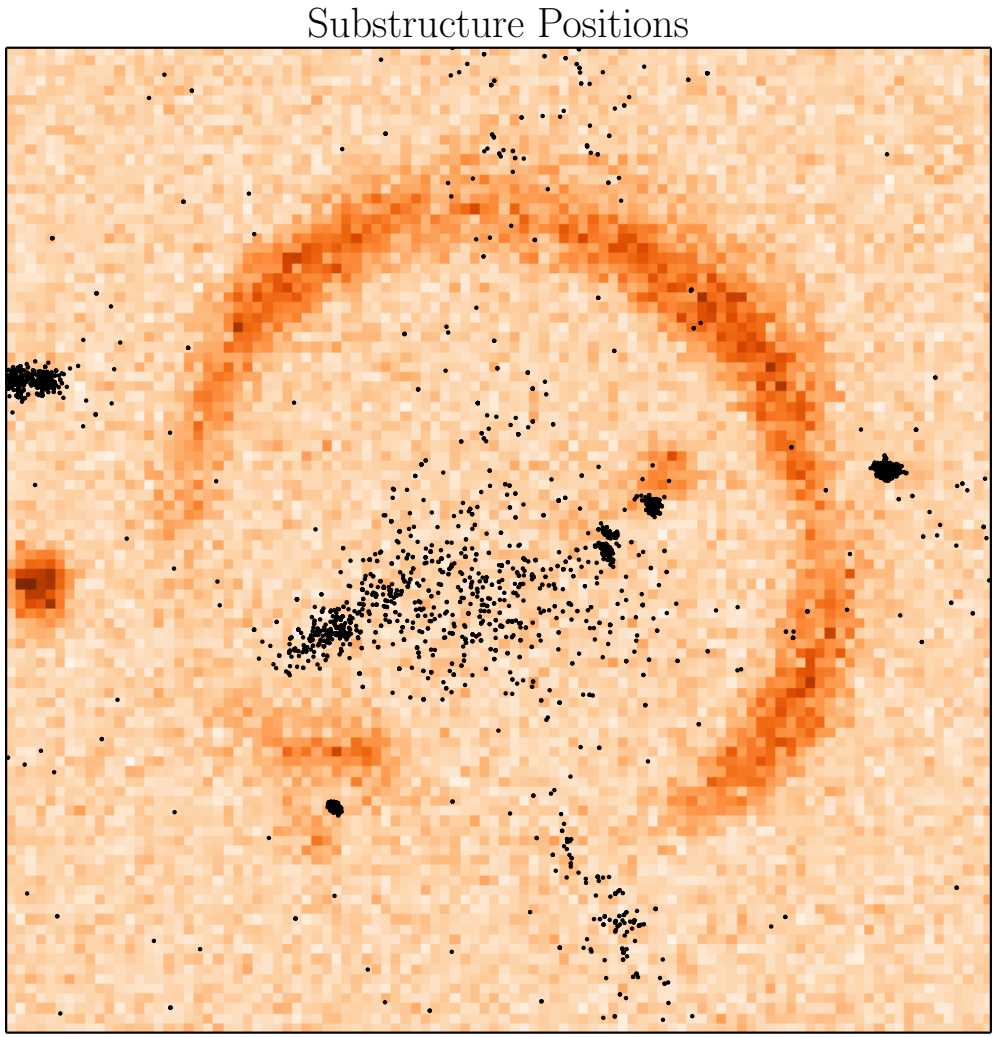


Brewer, Huijser, and Lewis, MNRAS, 455, 1819-1829. arXiv: 1508.00662

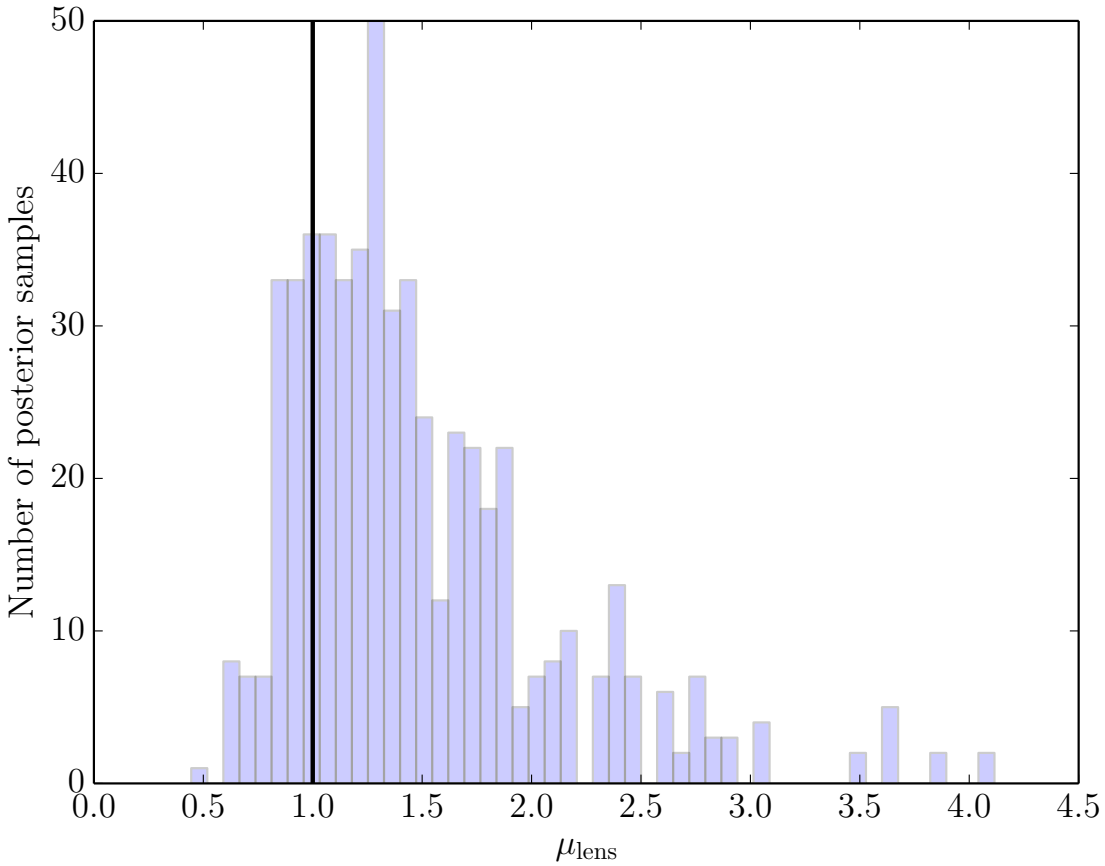
The prior information



Yes we can!



Sí, ¡se puede!


Number of posterior samples
0
10
20
30
40
50
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
$\mu_{\rm lens}$

The “Jackpot”

Was discovered in the Sloan Digital Sky Survey [spectrum had two redshifts]. HST follow-up as part of SLACS (SLoan Lens ACS Survey). Has *two* Einstein rings!

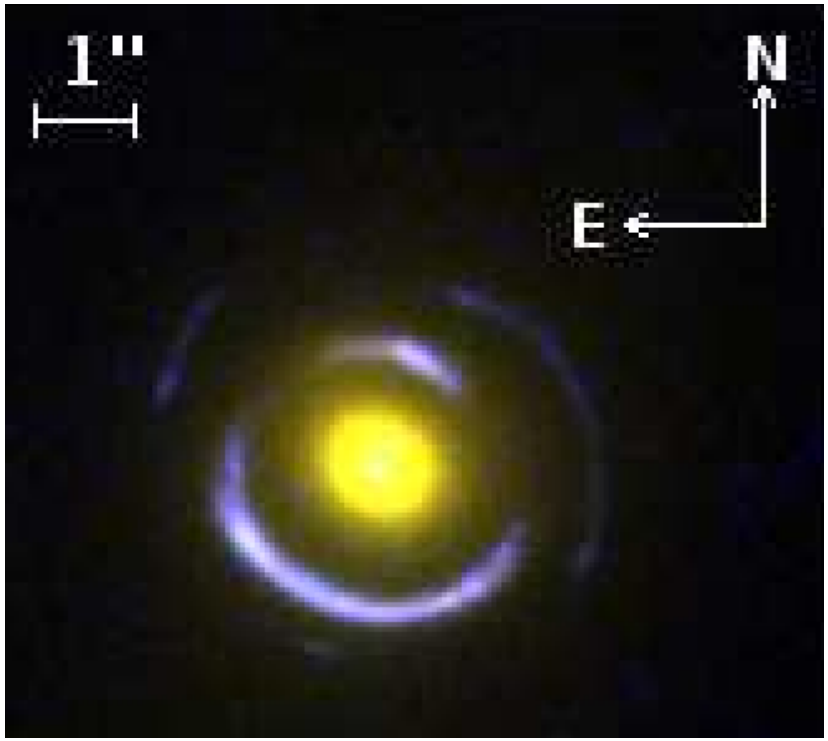
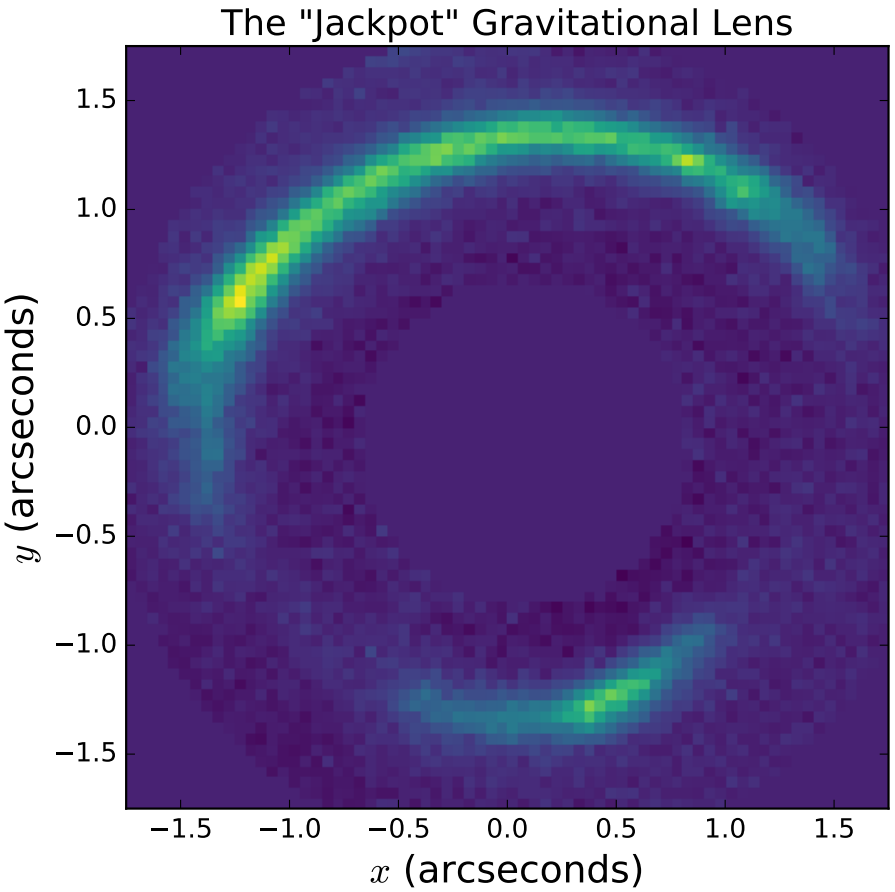


Figure: Sonnenfeld et al, 2012, ApJ 752, 163. arXiv: 1111.4215

The “Jackpot”

Sadly, the redshift of the second source couldn’t be measured. But we can do plenty with the first.



The “Jackpot”

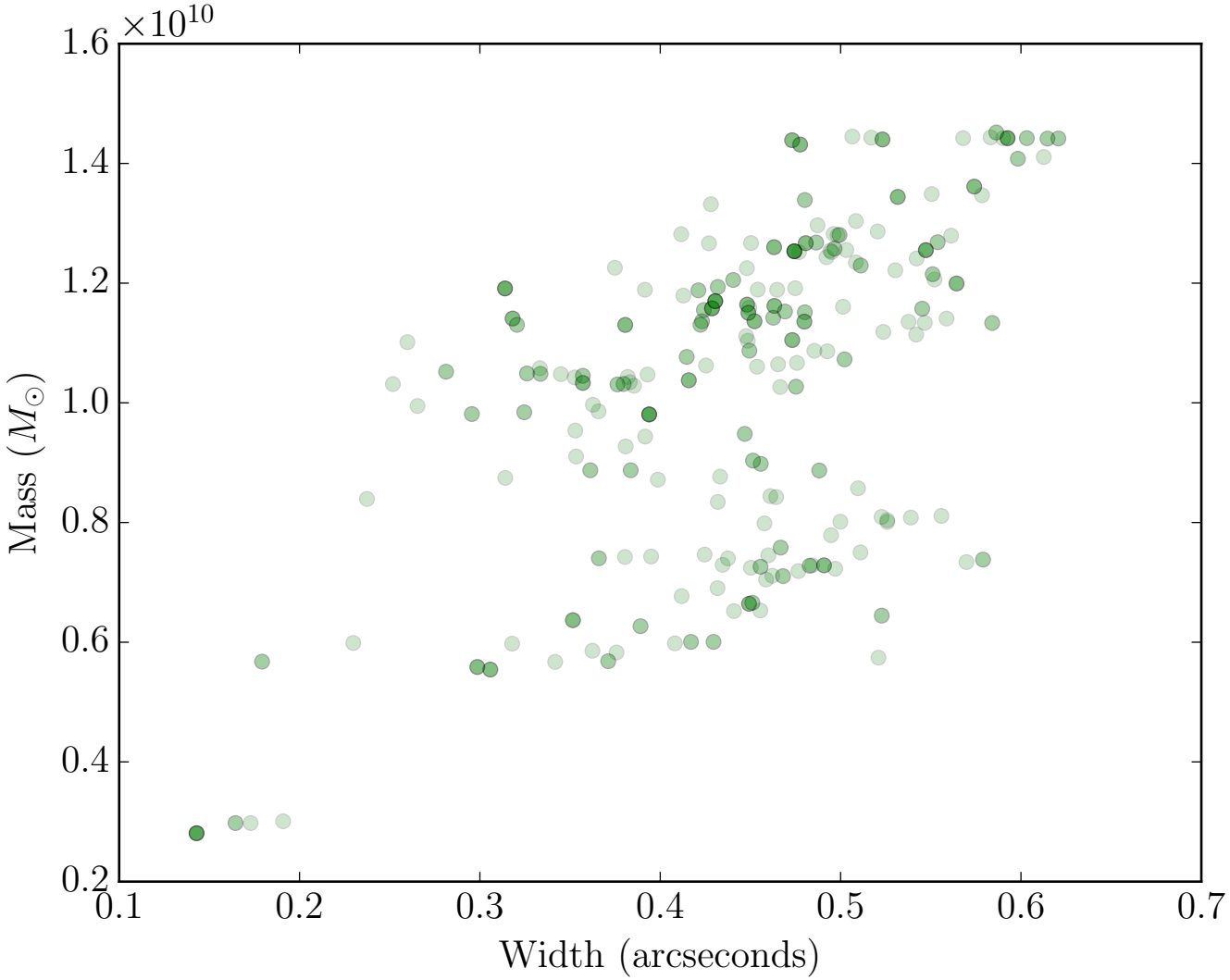
Previous authors found a “dark” substructure with mass $M_{\rm sub} = (3.51 \pm 0.15) \times 10^9 M_\odot$.

Vegetti et al, 2010, MNRAS, 408, 1969-1981. arXiv: 0910.0760

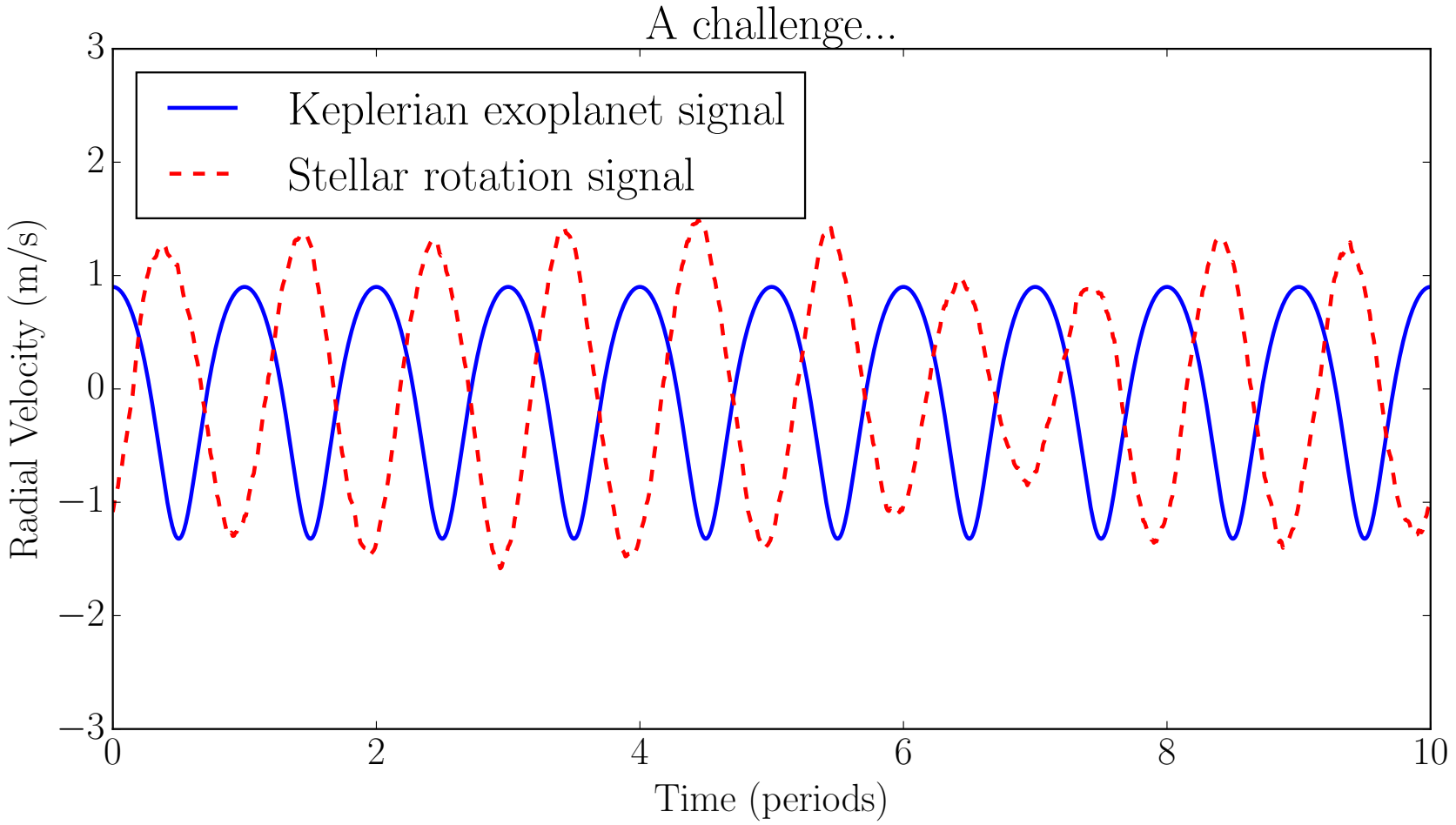
Samples of lenses

Show movie2.mkv

Measuring the mass?



A challenge for the radial velocity technique



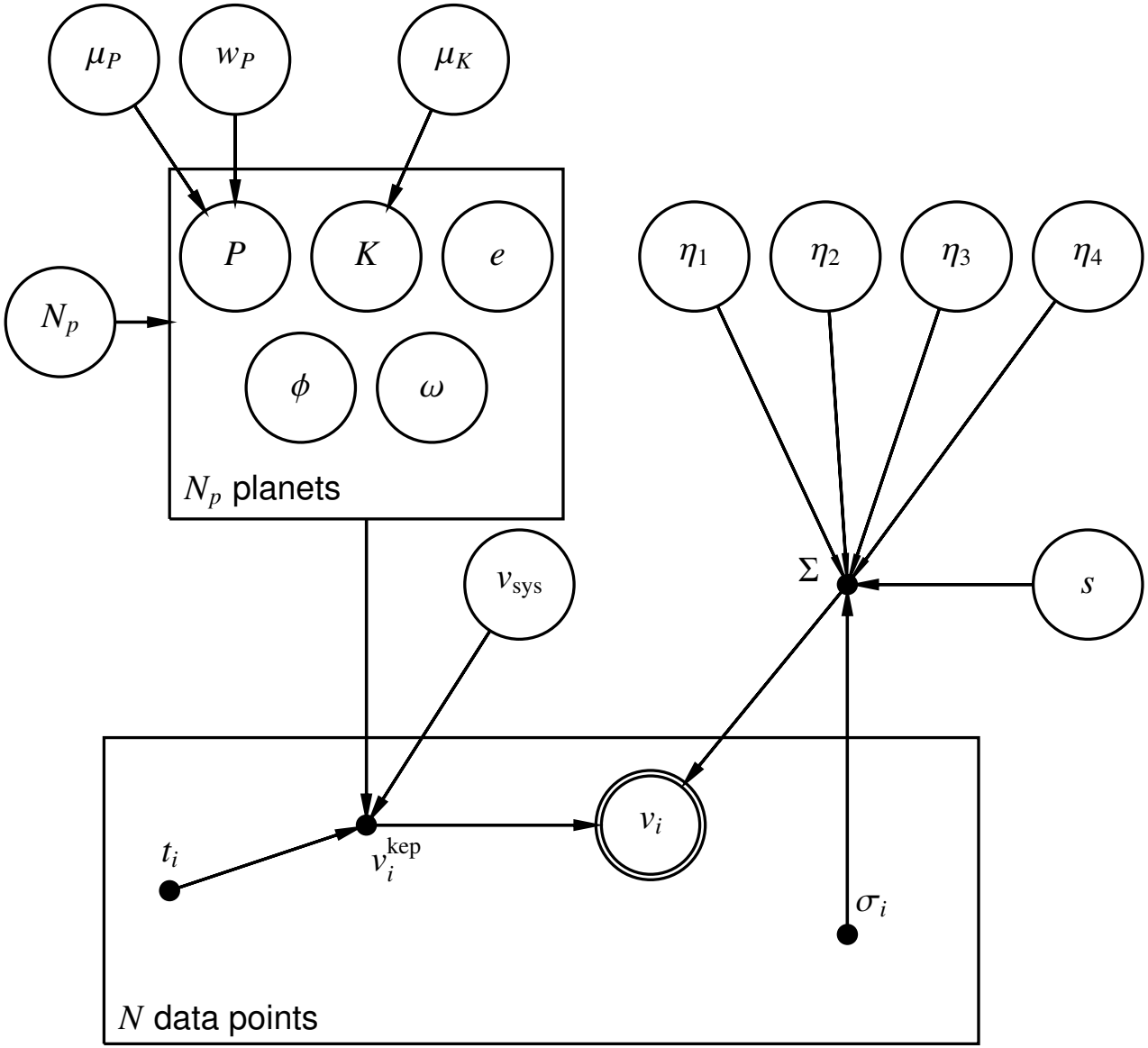
A challenge for the radial velocity technique

Can we distinguish these?

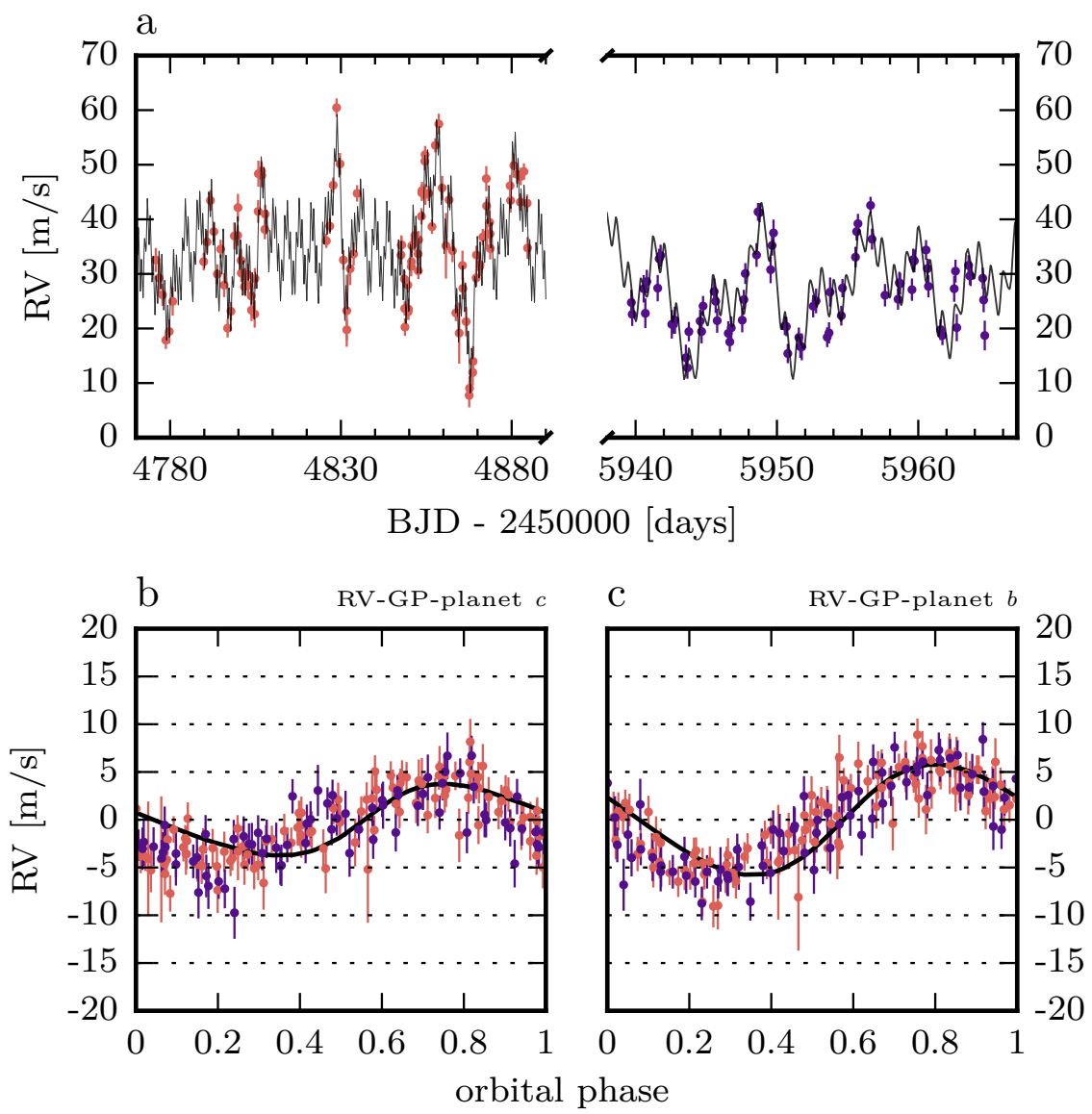
Depends. In the limit (eccentricity $\rightarrow$ 0, stellar rotation signal $\rightarrow$ sinusoid), there is a genuine ambiguity. Away from this limit, it depends on how informative the data is about the signal shape.

Inference methods should tell us about the ambiguity, not "explain" the stellar rotation signal with multiple planets.

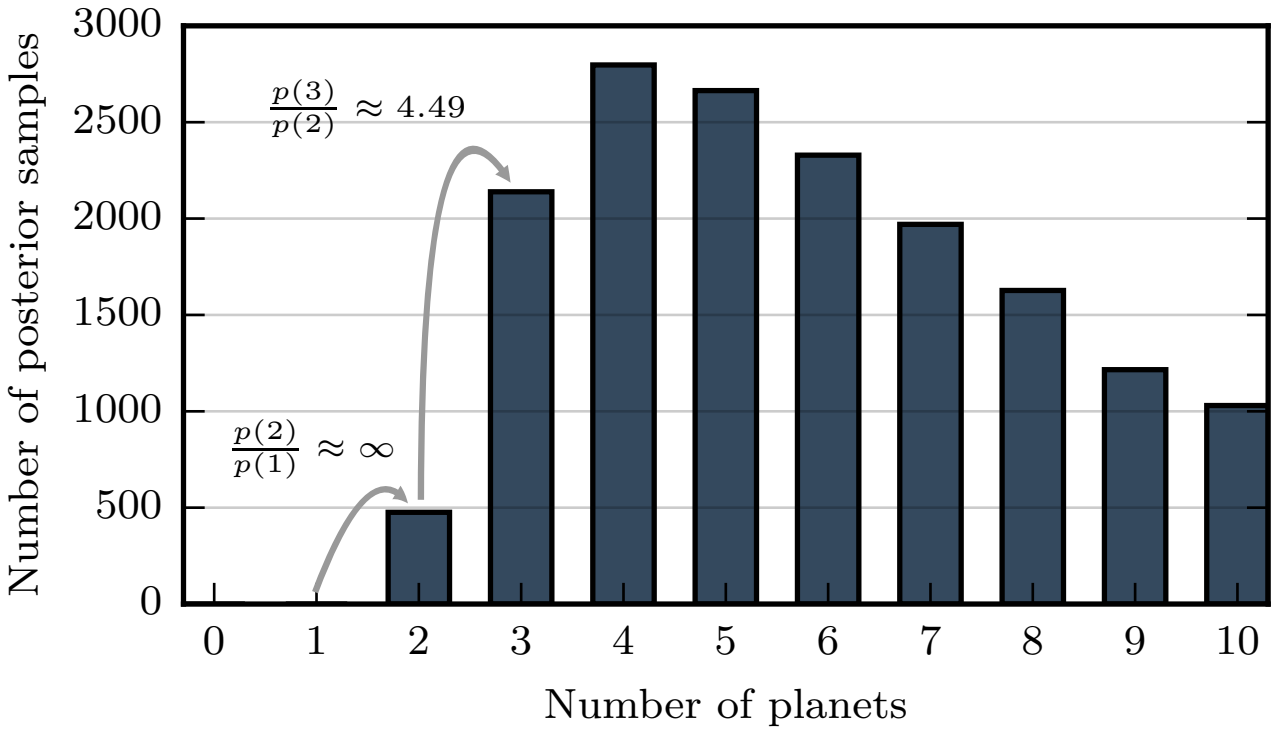
The Prior Information



HARPS Data


a
RV [m/s]
BJD - 2450000 [days]
b
RV-GP-planet c
c
RV-GP-planet b
RV [m/s]
orbital phase

How many planets?



Faria et al, 2016, A&A 588, A31 arXiv: 1601.07495

Have a mixture modelling problem?

Or any Bayesian computation for that matter (including ABC).
https://github.com/eggplantbren/DNest4

Software paper on the way (draft in repo — currently **30 pages**!).