Bayesian inference

Paul Lasky

MONASH University

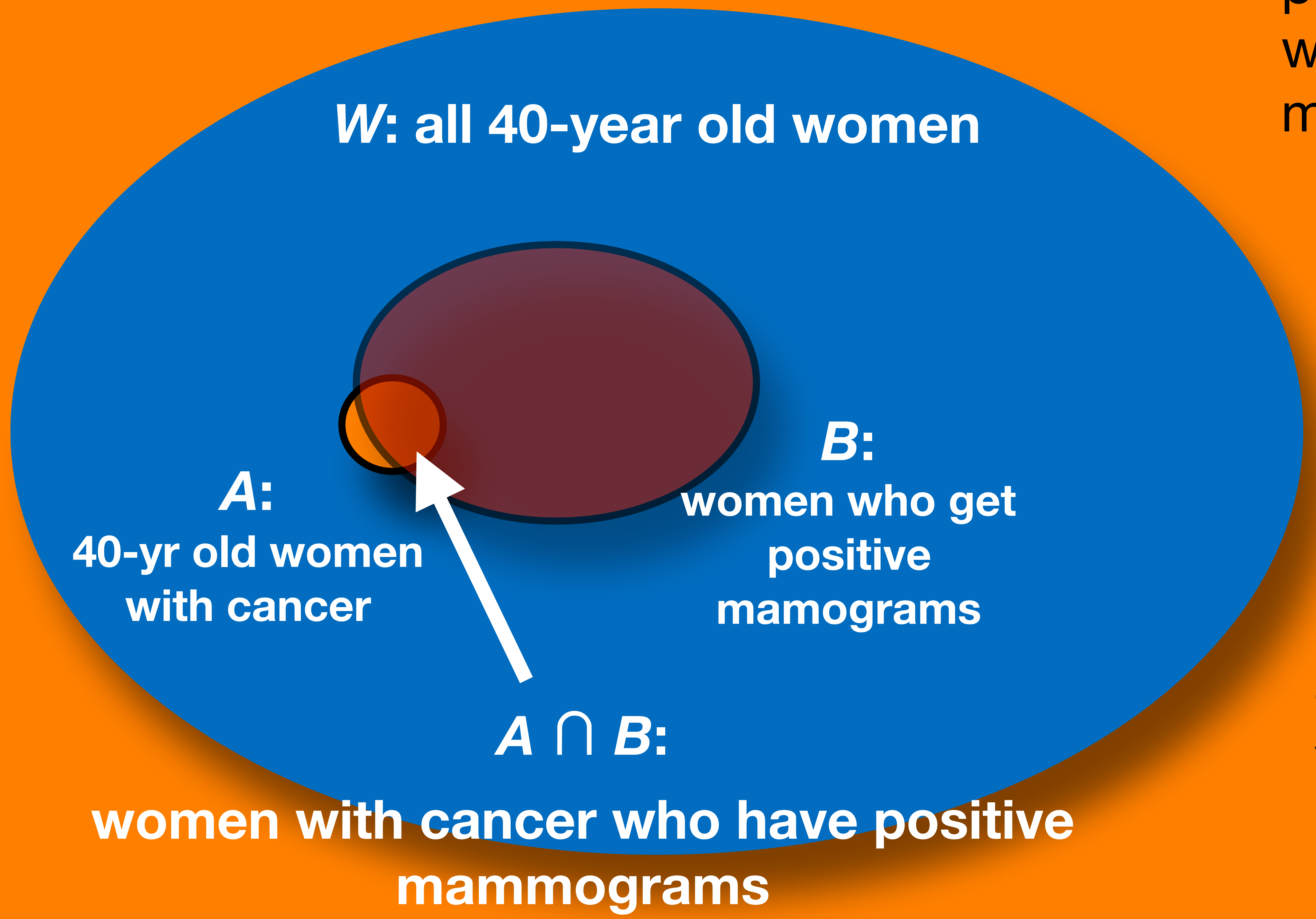
OzGrav

"Understanding the underlying processes that create data streams is a ubiquitous problem across the sciences"

Matilda B. Bilby*

*not a real quote
(also not a real Bilby)

- 1% of women at age forty have breast cancer.
- 80% of women with breast cancer have positive mammograms.
- 9.6% of women without breast cancer also get positive mammograms (false positive).
- A woman in this age group had a positive mammogram in a routine screening. What is the probability that she has breast cancer?



probability of picking a random 40-yr old woman that has cancer:

$$p(A) = \frac{|A|}{|W|}$$

probability of picking a random 40-yr old woman that will have a positive mammogram (with or without cancer):

$$p(B) = \frac{|B|}{|W|}$$

$$p(A \cap B) = \frac{|A \cap B|}{|W|}$$

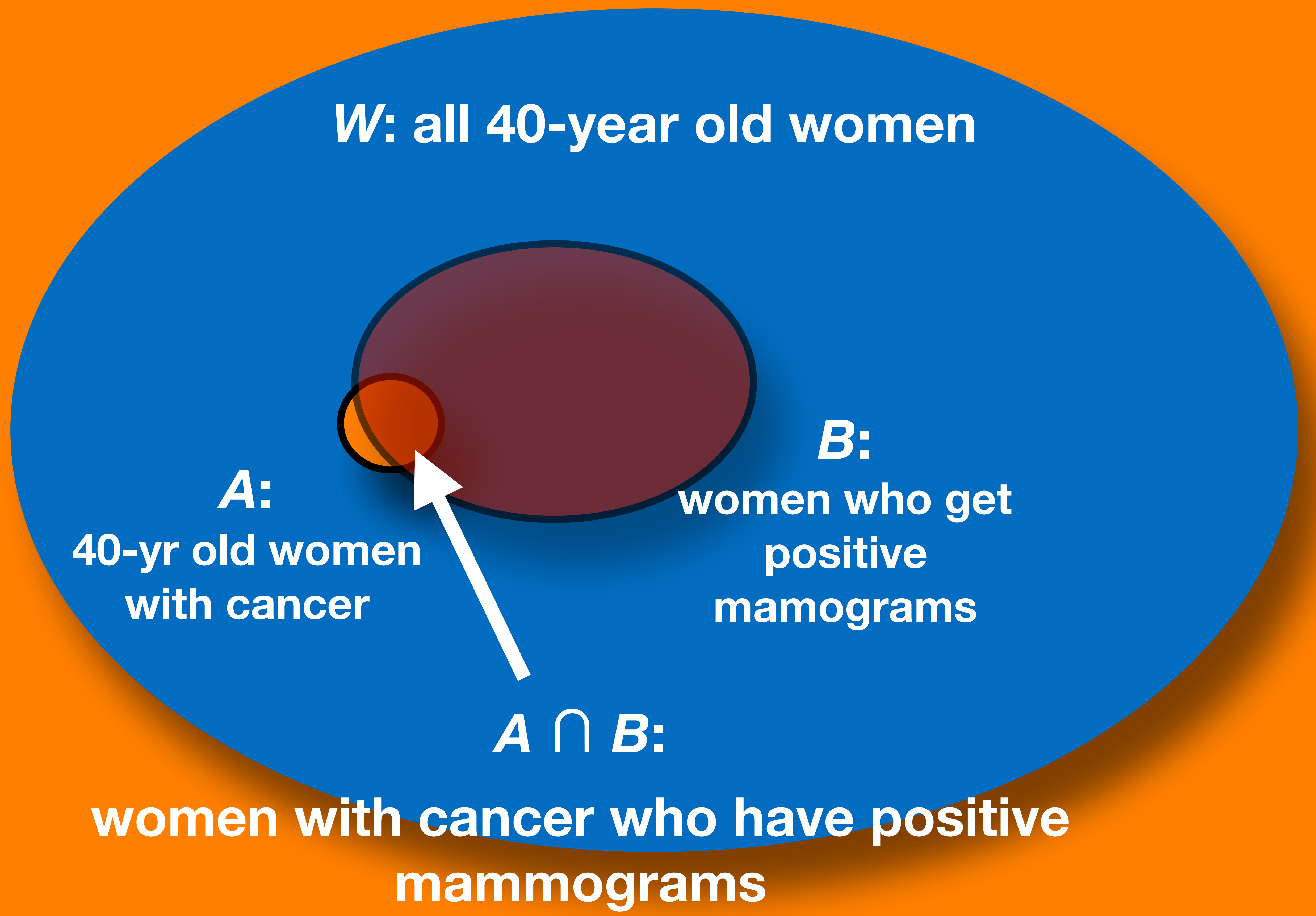
given you've got cancer, what's the likelihood you'll also have a positive mammogram:

$$p(B|A) = \frac{|A \cap B|}{|A|} \times \frac{|W|}{|W|} = \frac{p(A \cap B)}{p(A)}$$

given you've had a positive mammogram, what's the probability you've got cancer:

$$p(A|B) = \frac{|A \cap B|}{|B|} \times \frac{|W|}{|W|} = \frac{p(A \cap B)}{p(B)}$$

- 1% of women at age forty have breast cancer.
- 80% of women with breast cancer have positive mammograms.
- 9.6% of women without breast cancer also get positive mammograms (false positive).
- A woman in this age group had a positive mammogram in a routine screening. What is the probability that she has breast cancer?



$$p(B|A) = \frac{p(A \cap B)}{p(A)} \qquad p(A|B) = \frac{p(A \cap B)}{p(B)}$$

$$\boxed{p(A|B) = \frac{p(B|A)p(A)}{P(B)}}$$

$$p(B|A) = 0.8$$

$$p(A) = 0.01$$

$$p(B) = 0.8p(A) + 0.096(1 - p(A)) \approx 0.1$$

$$\Rightarrow \; p(A|B) \approx 0.078$$

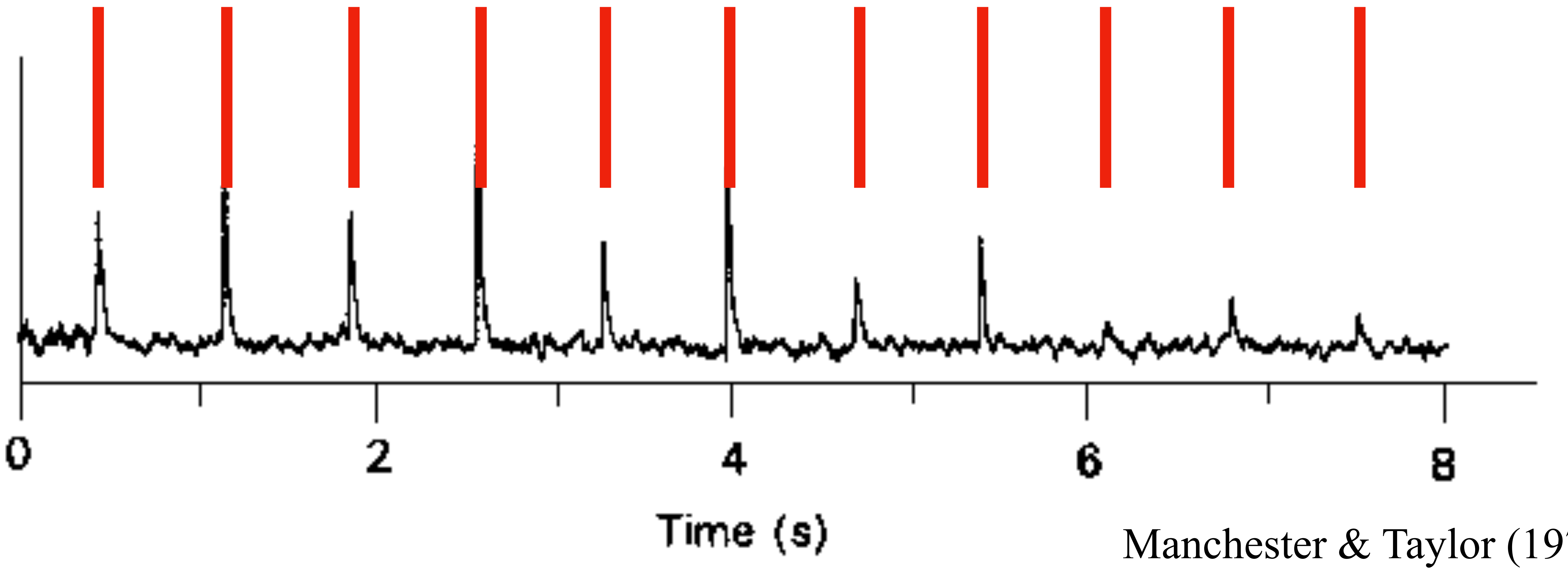
If you're a 40-year old woman, and have a positive mammogram, there's only a ~7.8% chance you have cancer!

$$\boxed{p(A|B) = \frac{p(B|A)p(A)}{P(B)}}$$

What about pulsar astronomy?

data: d

(pulse times of arrival)



Manchester & Taylor (1977)

Model: M

with parameters: θ

e.g., pulsar spin frequency, f

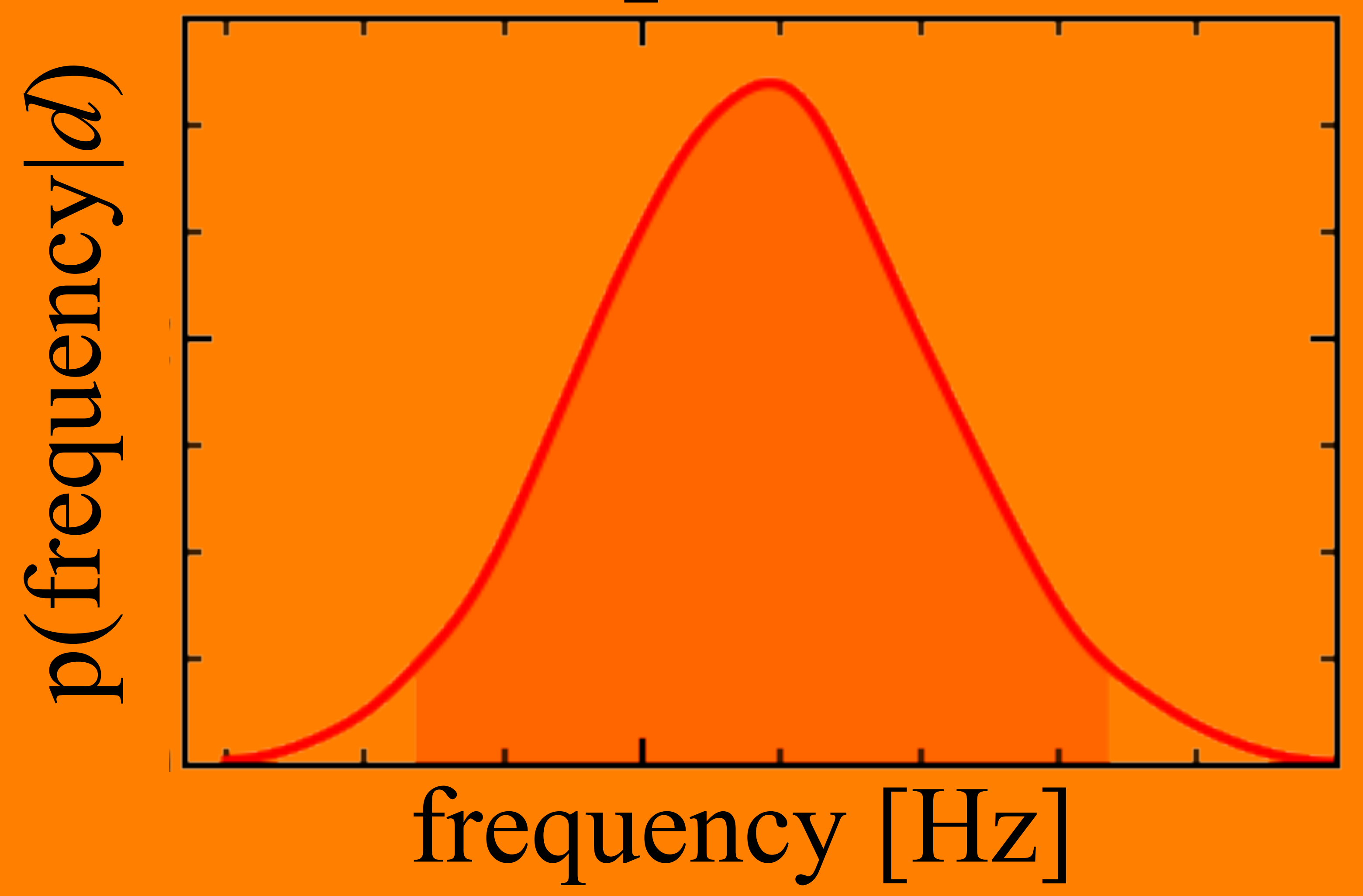
We want to calculate the probability of certain model parameters given the data: $p(\vec{\theta}|d)$

Likelihood: probability of the data, given the model.

Prior: our prior knowledge of the parameters of the model

$$p(\theta|d) = \frac{L(d|\theta)\pi(\theta)}{Z(d)}$$

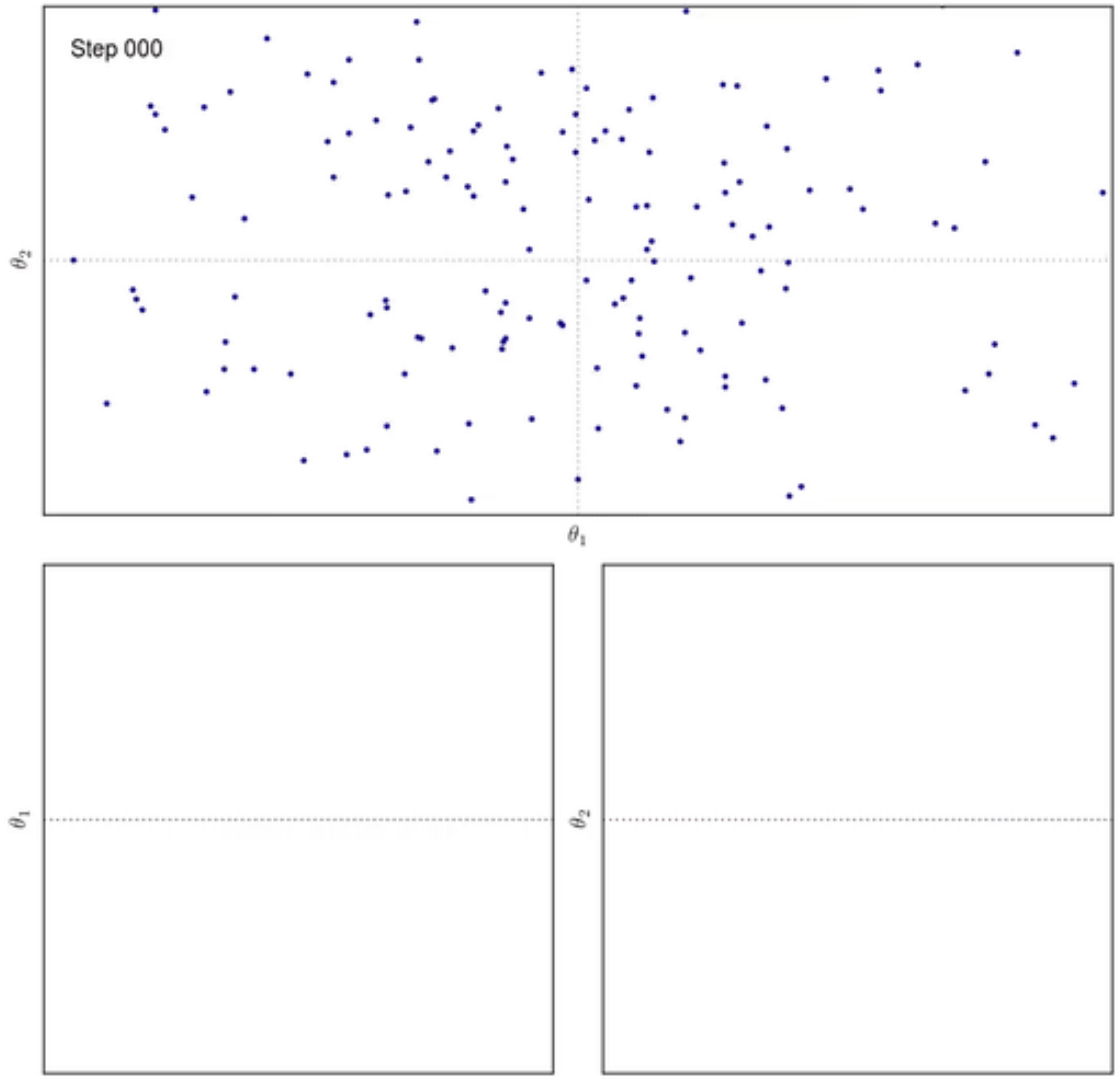
The game: Evaluate this equation for all values of the model parameters



Evidence: (only depends on the data, for parameter estimation is just a normalisation constant; let's get back to this)

in reality, for large dimensional spaces, direct integration is computationally infeasible

- Markov chain Monte Carlo
- Nested sampling
- etc



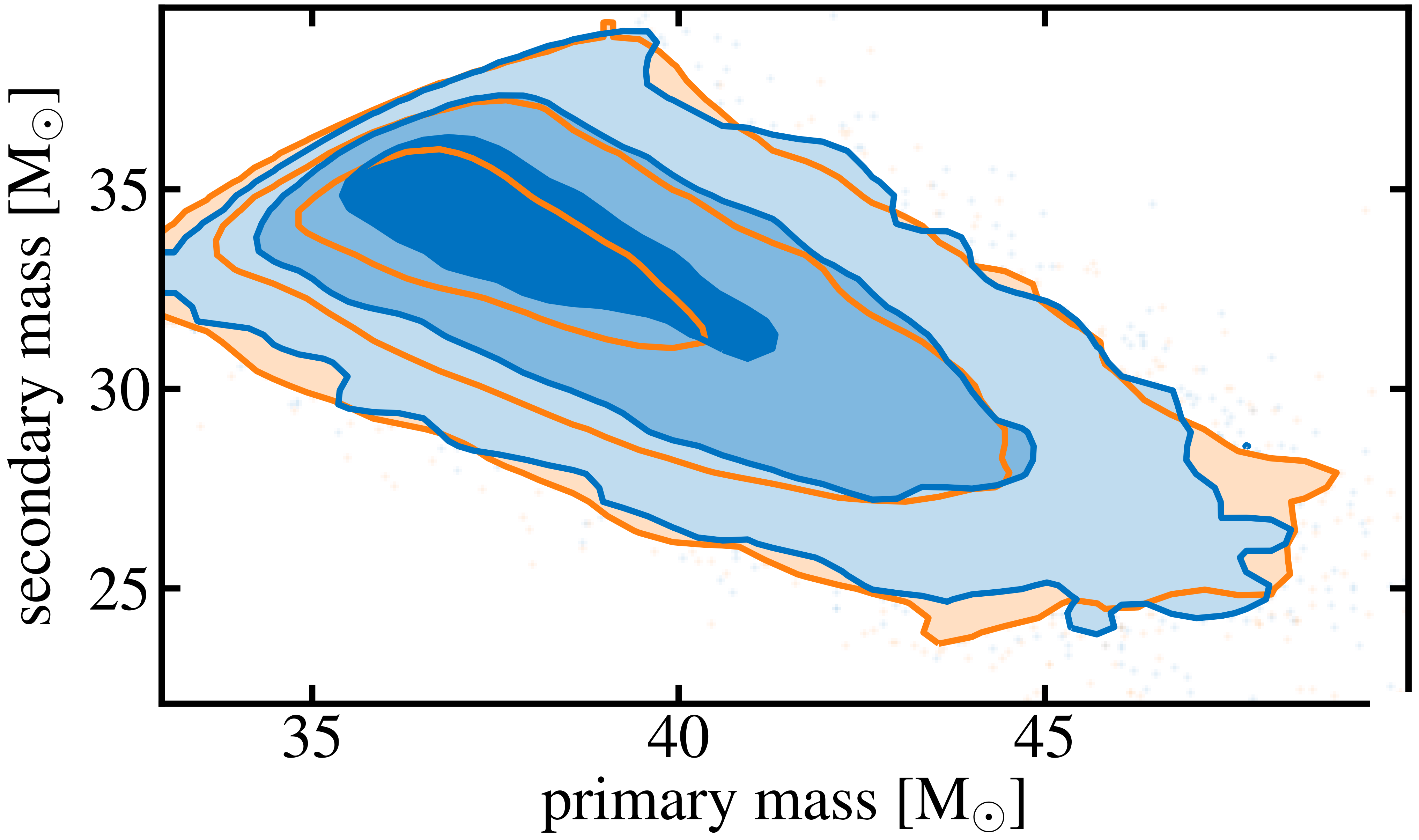
r/dataisbeautiful

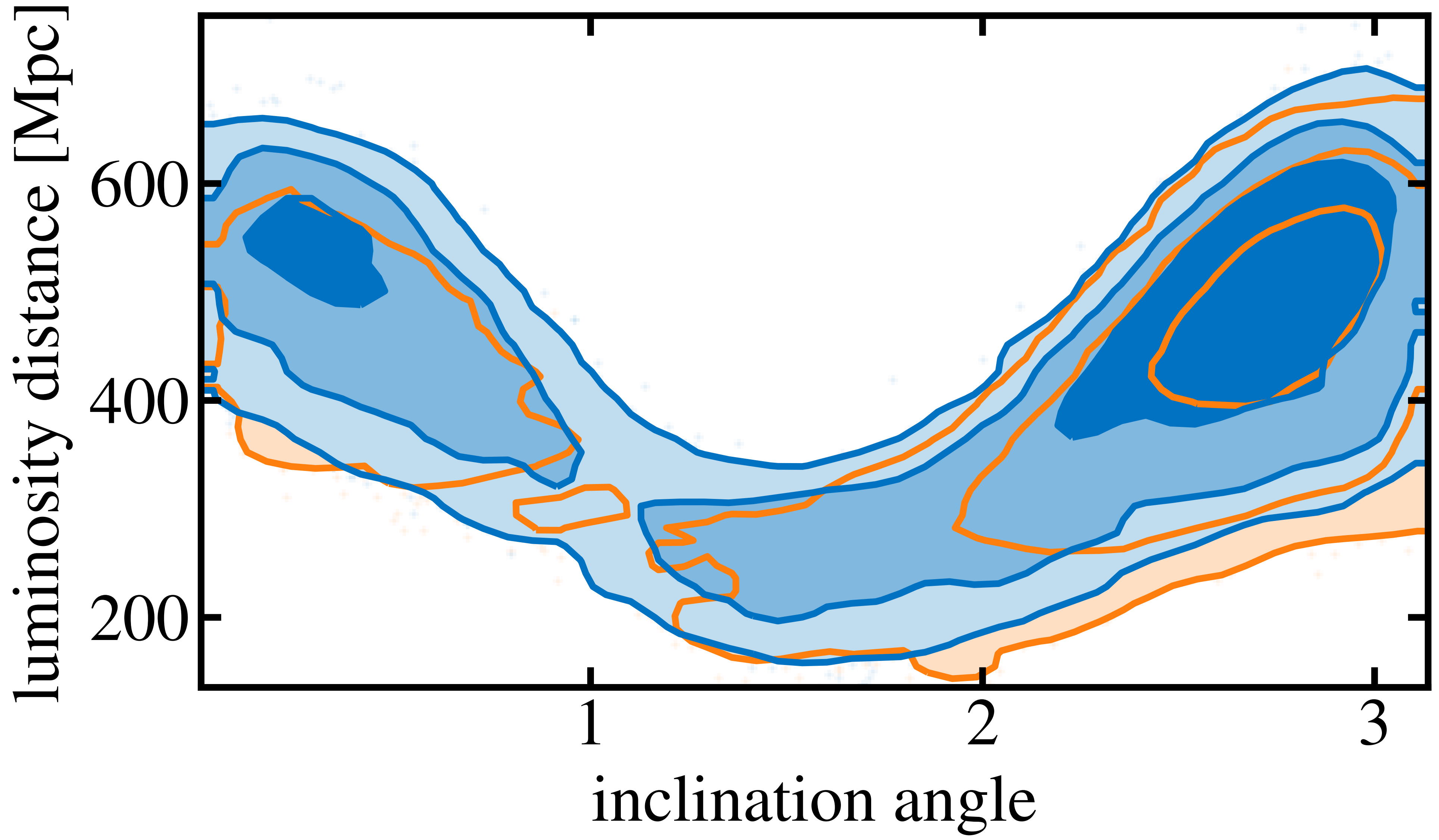
The user-friendly Bayesian inference library

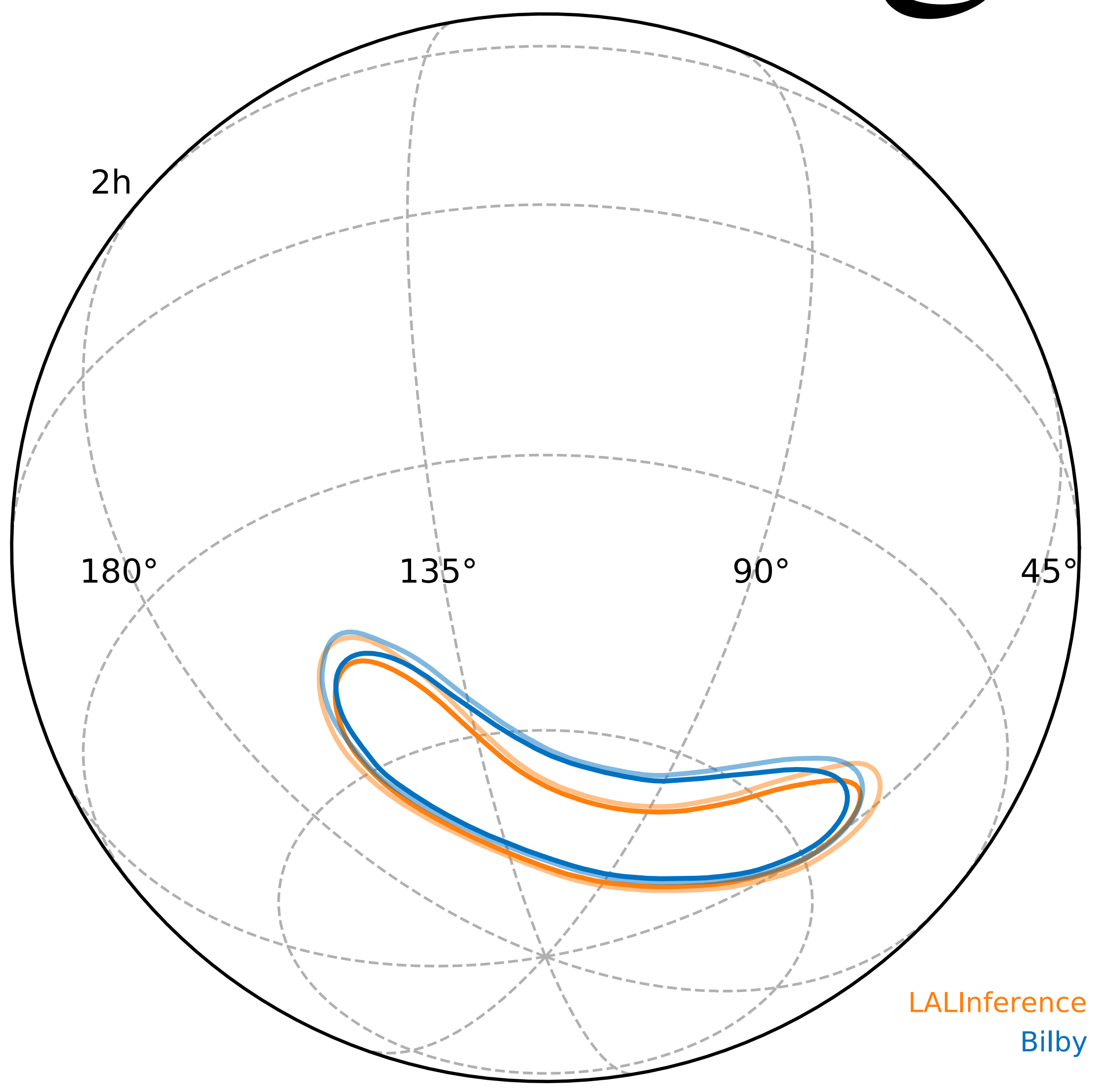
A versatile parameter-estimation code being adopted for production science in next LIGO observing run

git.ligo.org/lscsoft/bilby/

Ashton, Hübner, PL, Talbot + (2019)



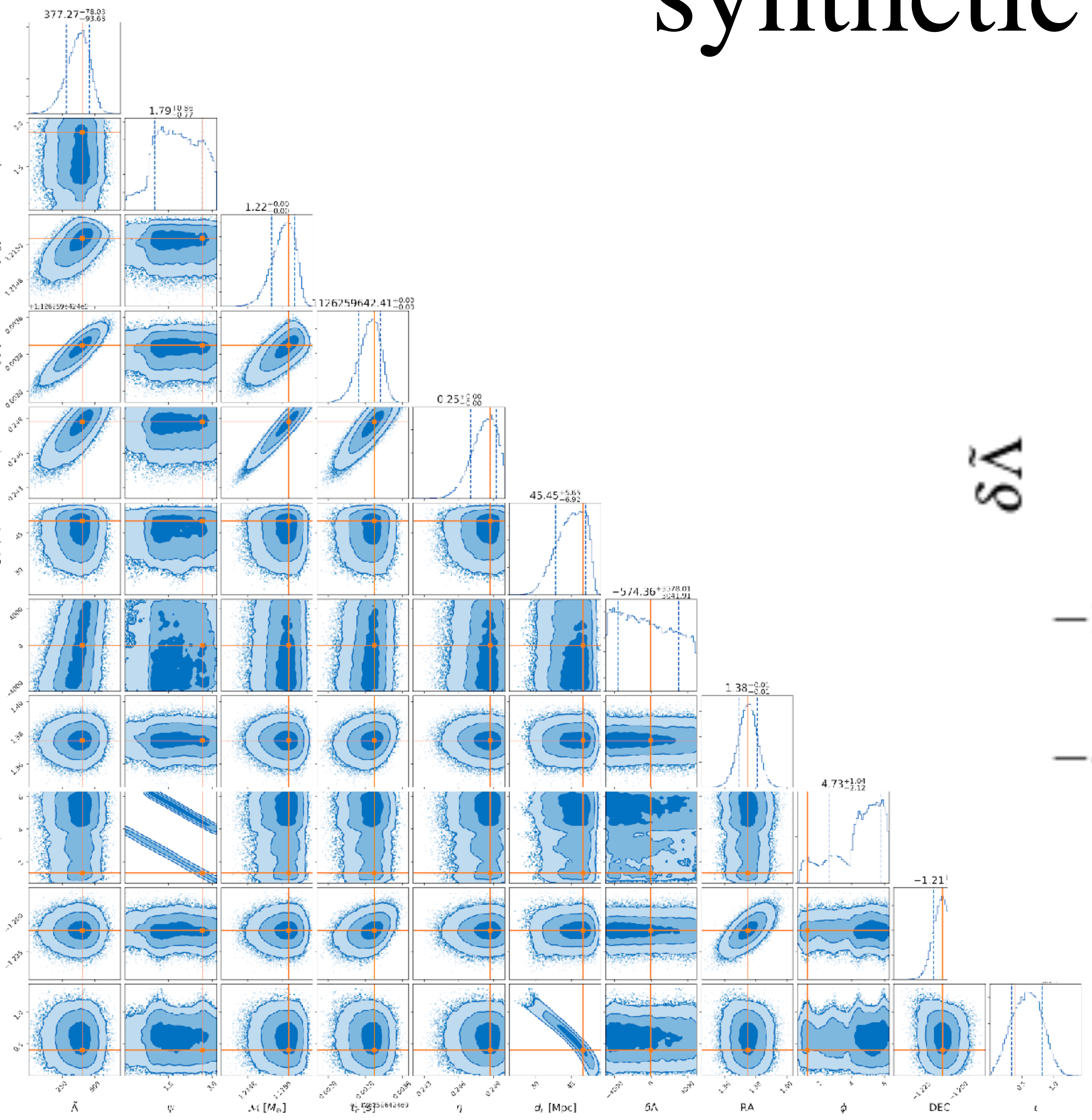


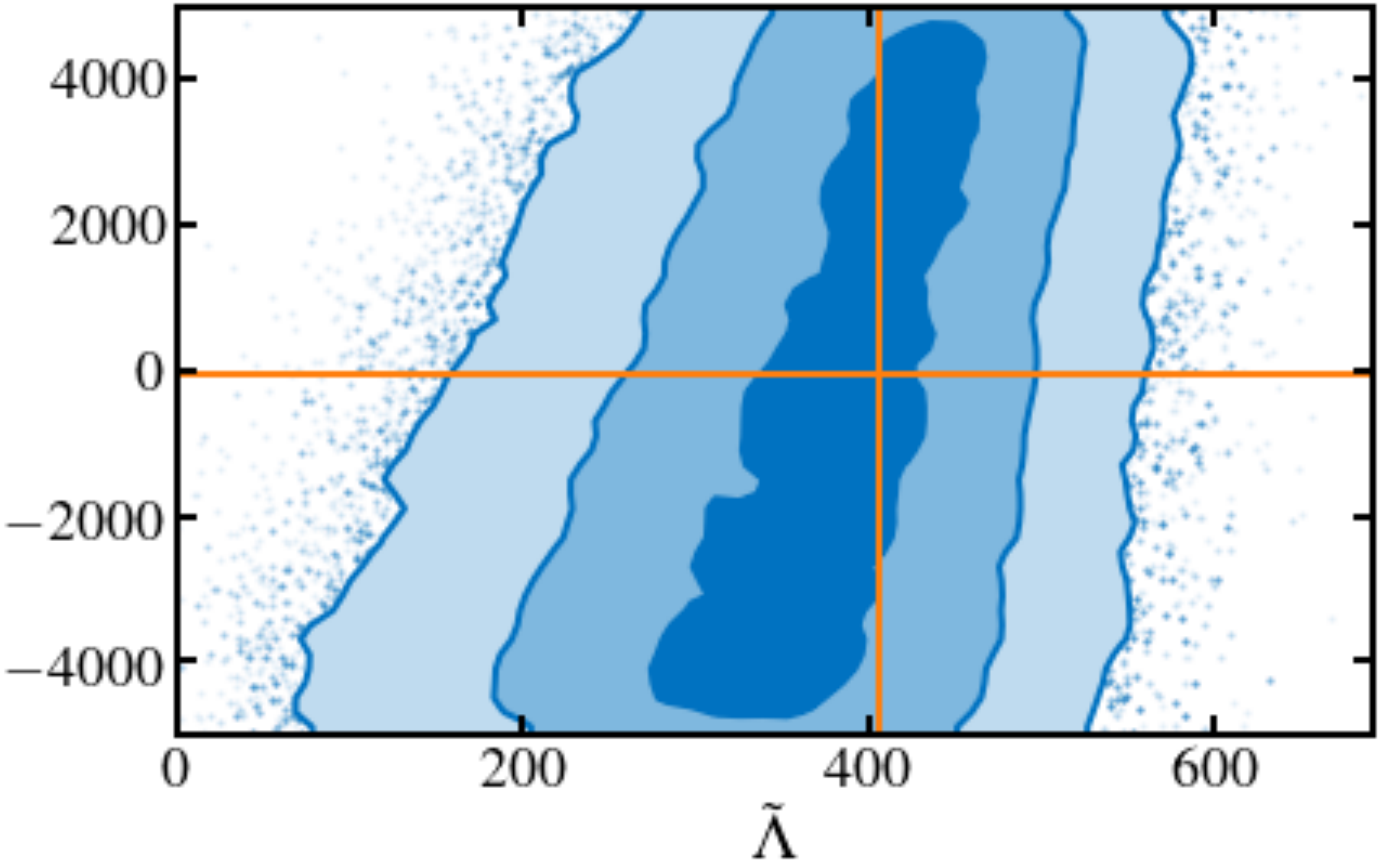


ground-based gravitational-wave experiments

Ashton, Hübner, PL, Talbot + (2019)

synthetic neutron stars

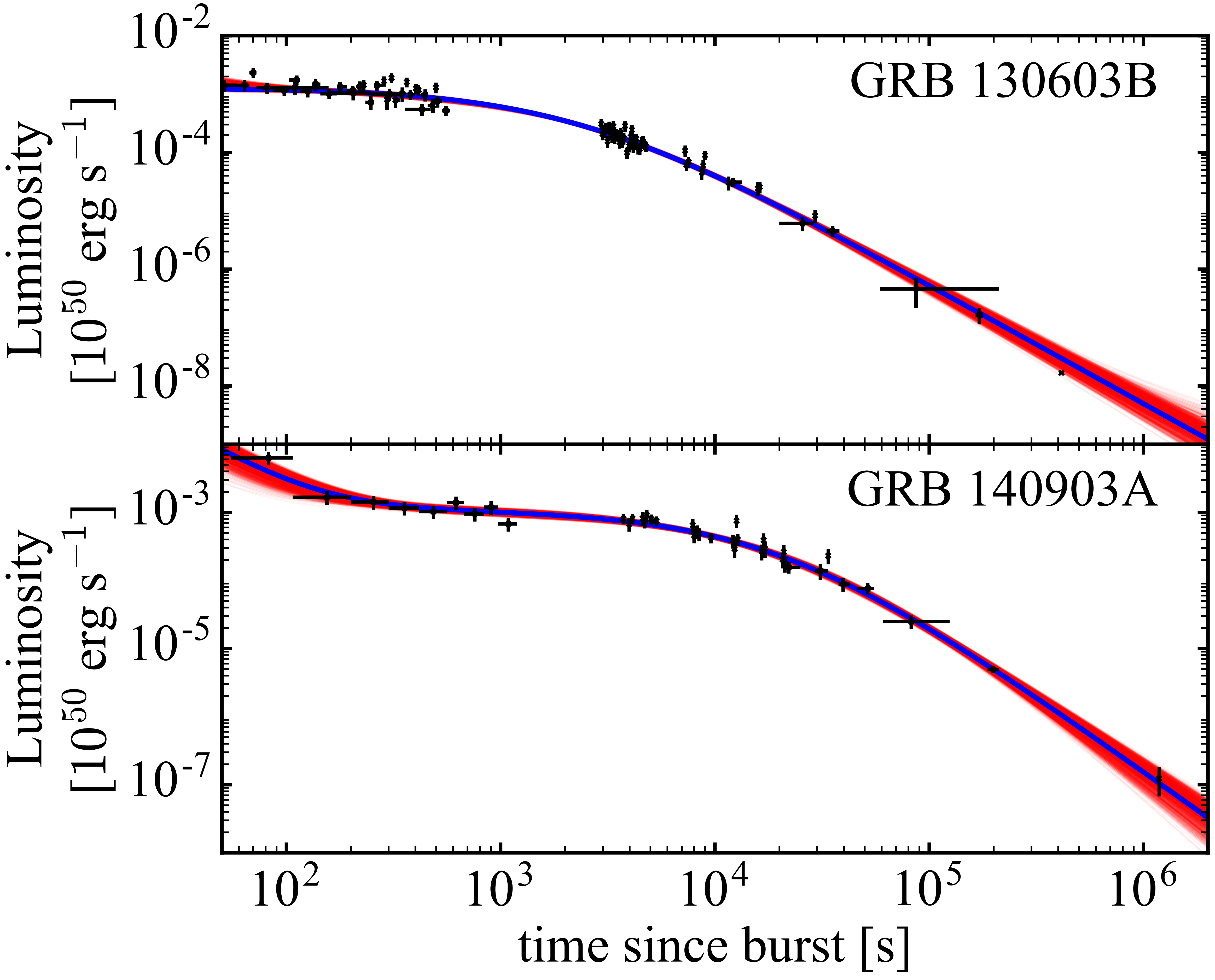


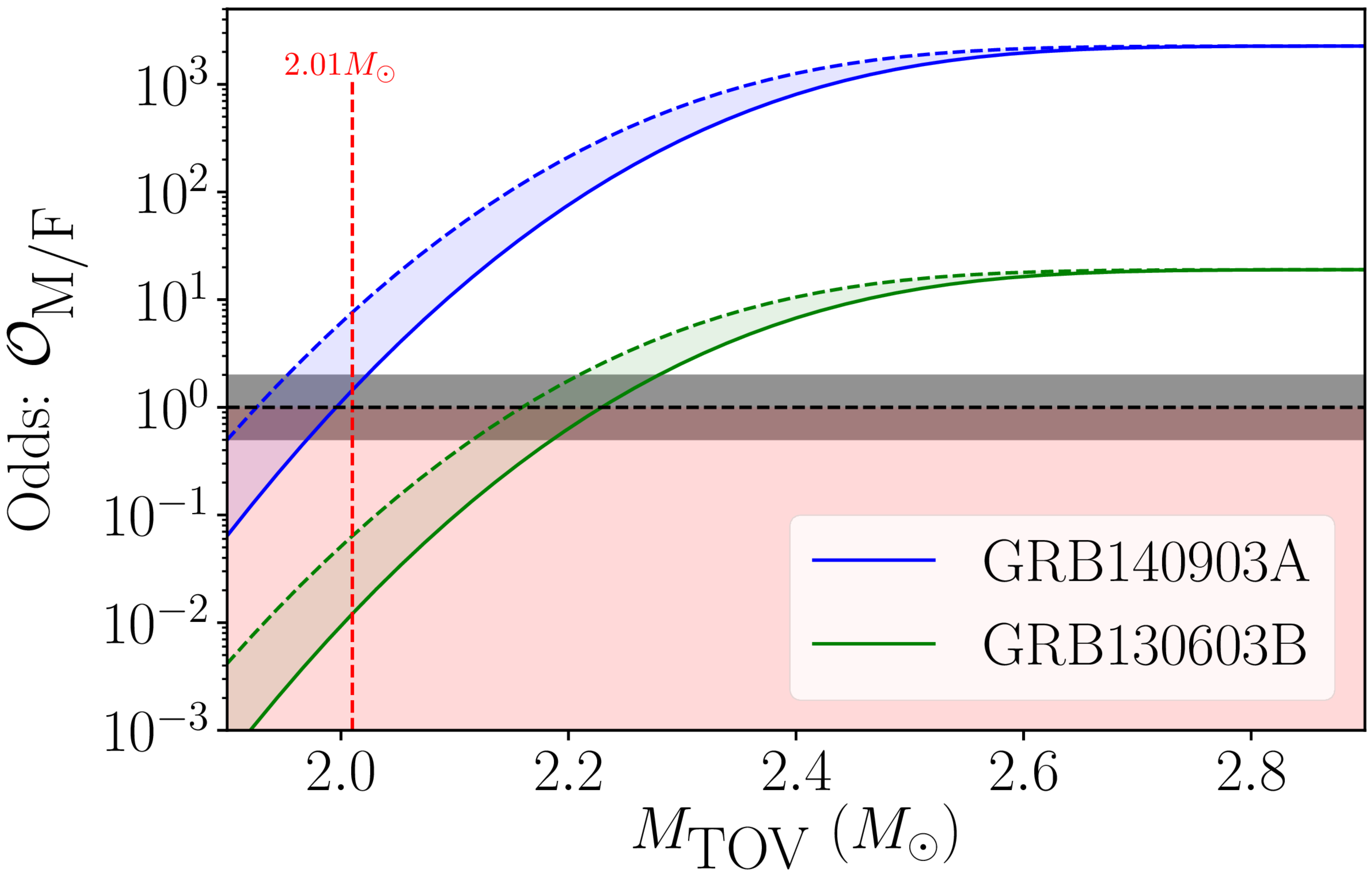


Ashton, Hübner, PL, Talbot + (2019)

x-ray light curves of gamma-ray bursts

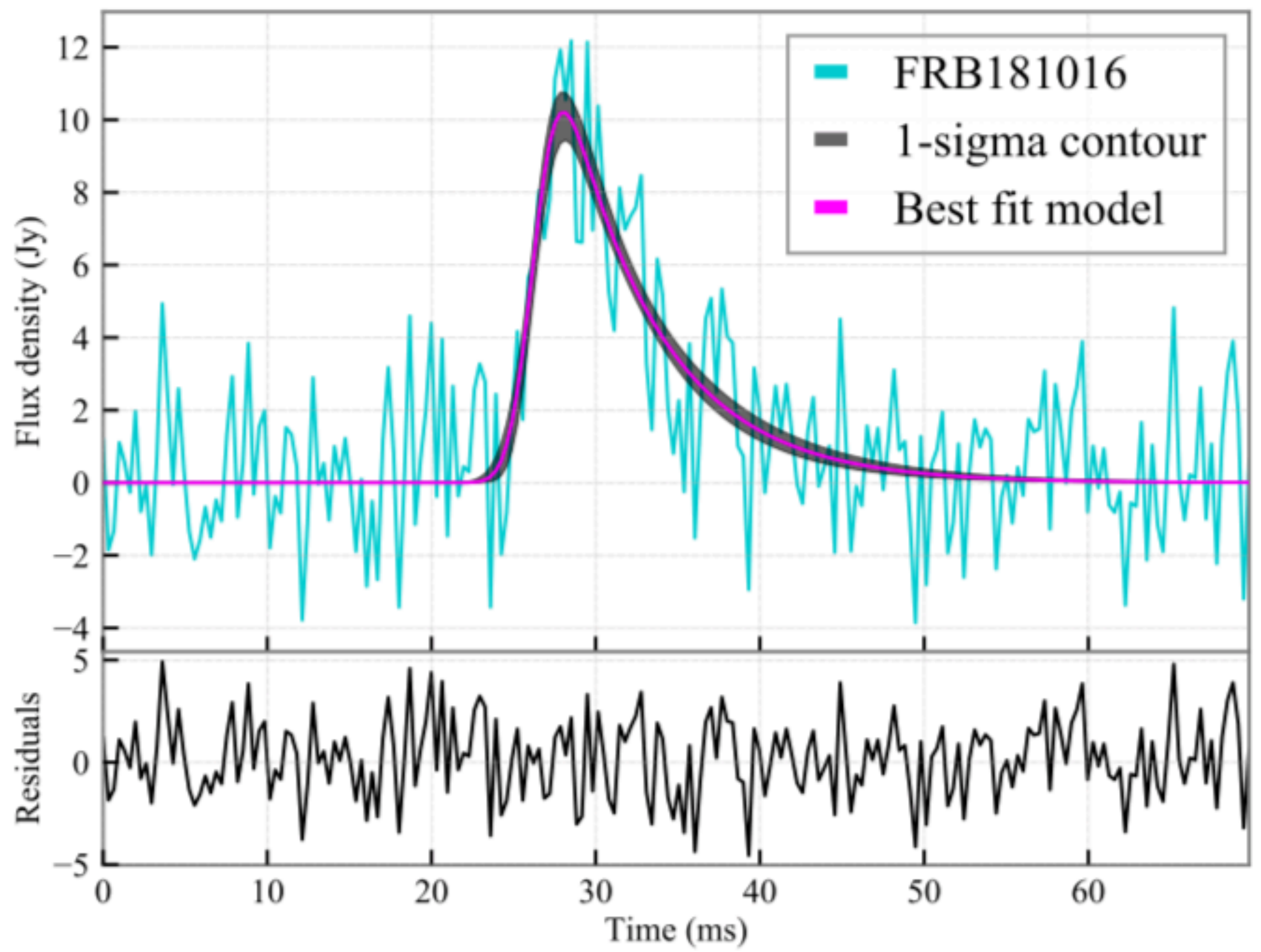
do millisecond magnetars exist?

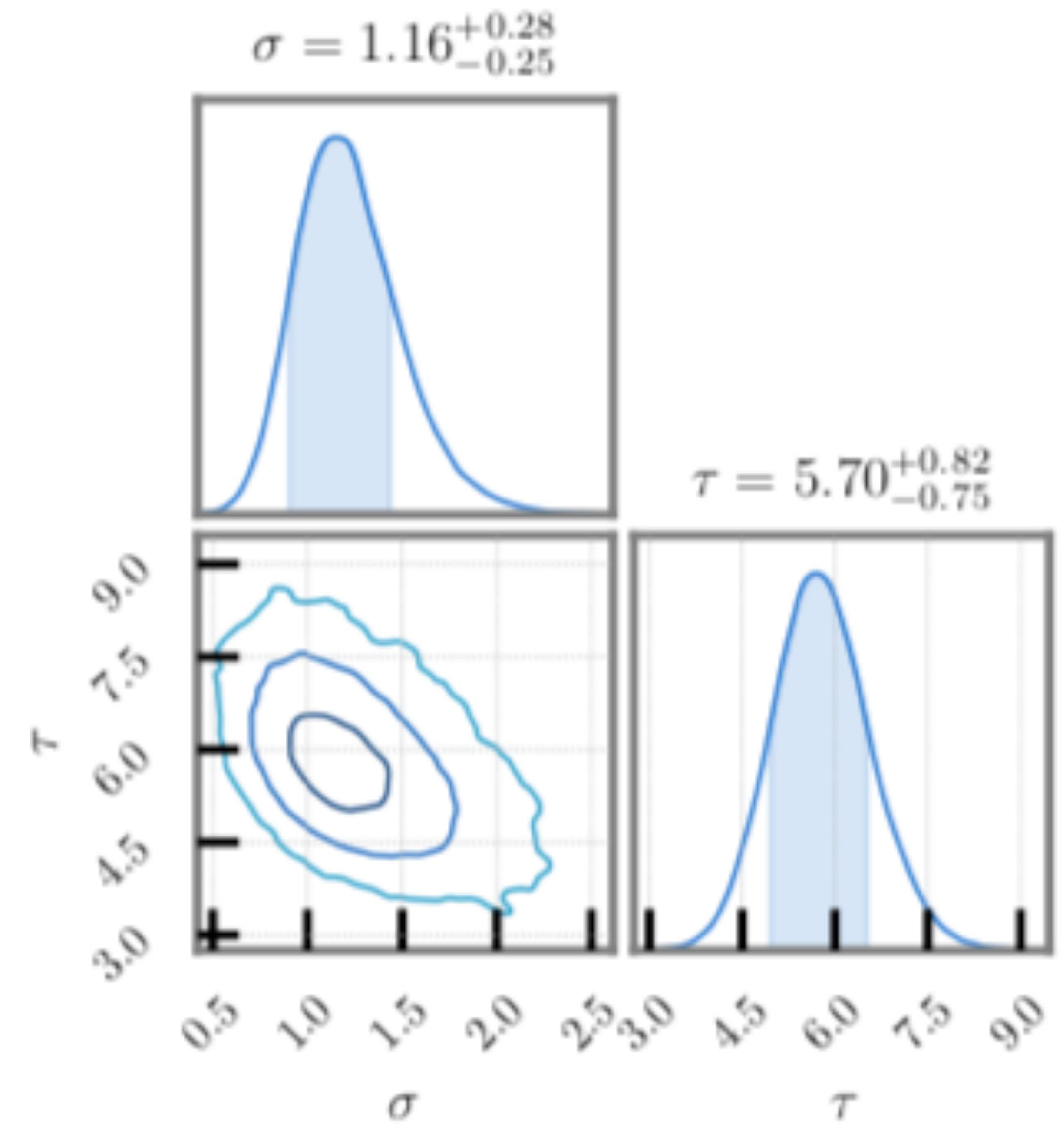


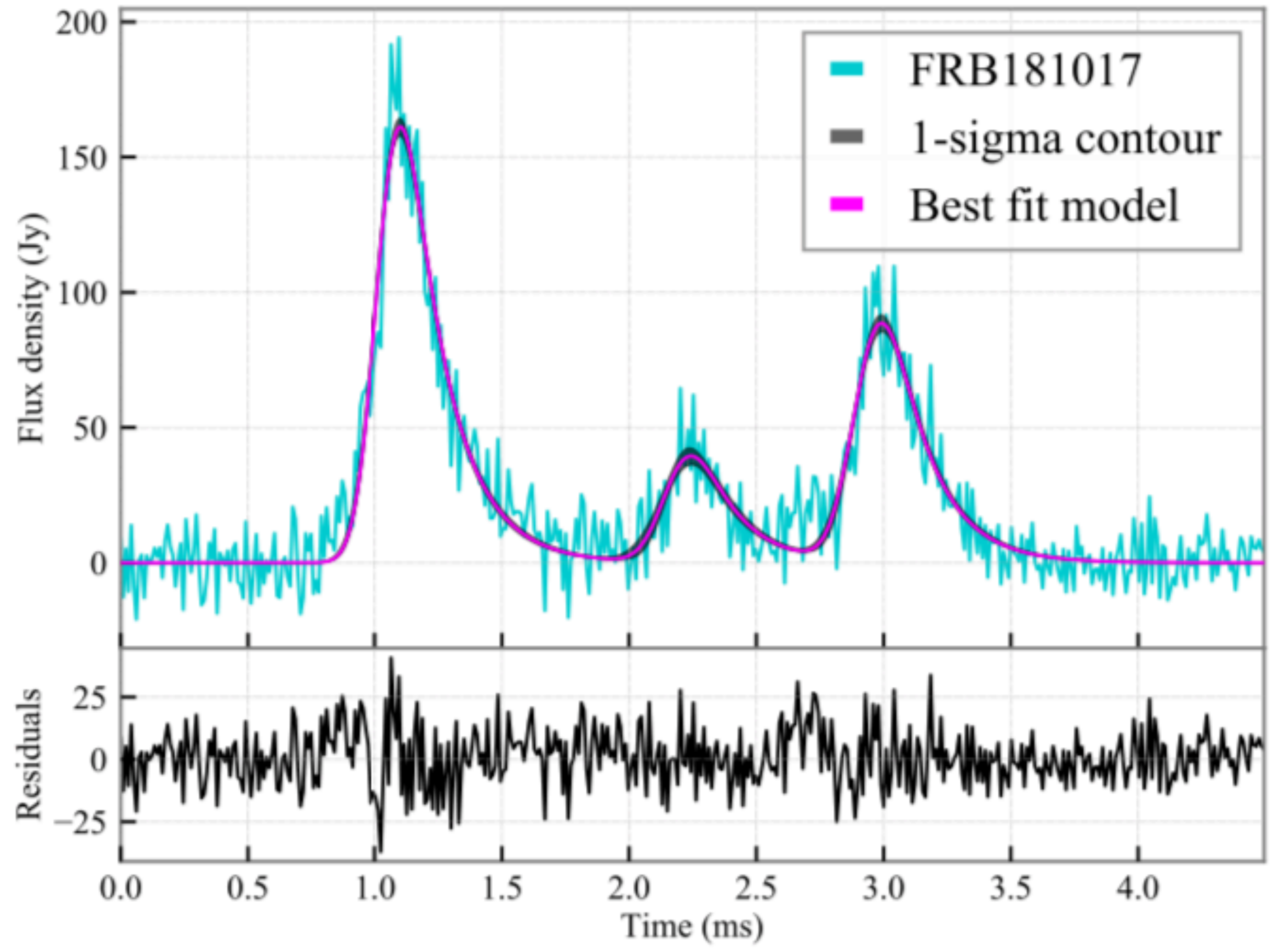


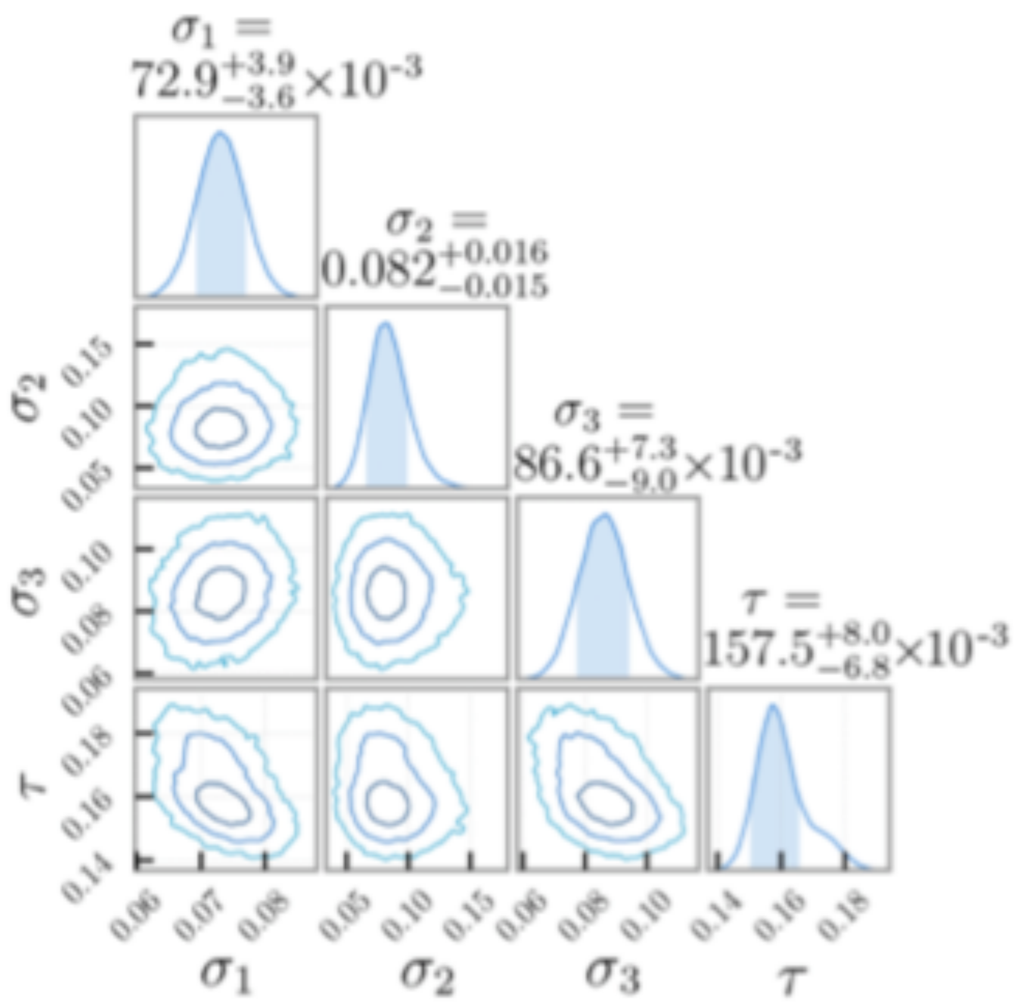
PL, Leris, Rowlinson & Glampedakis (2017)

Sarin, PL, Ashton (2019)







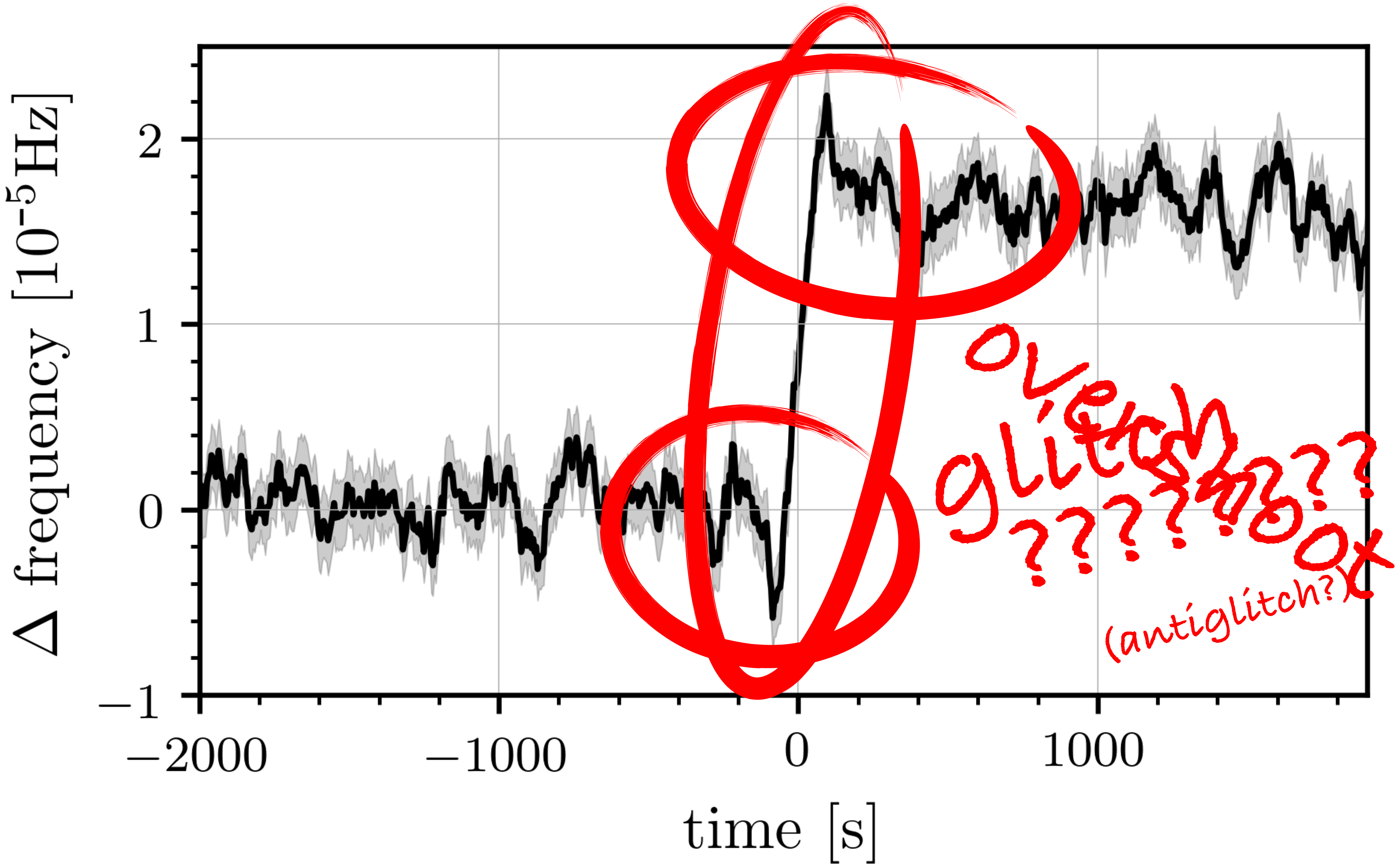


Fast Radio Bursts

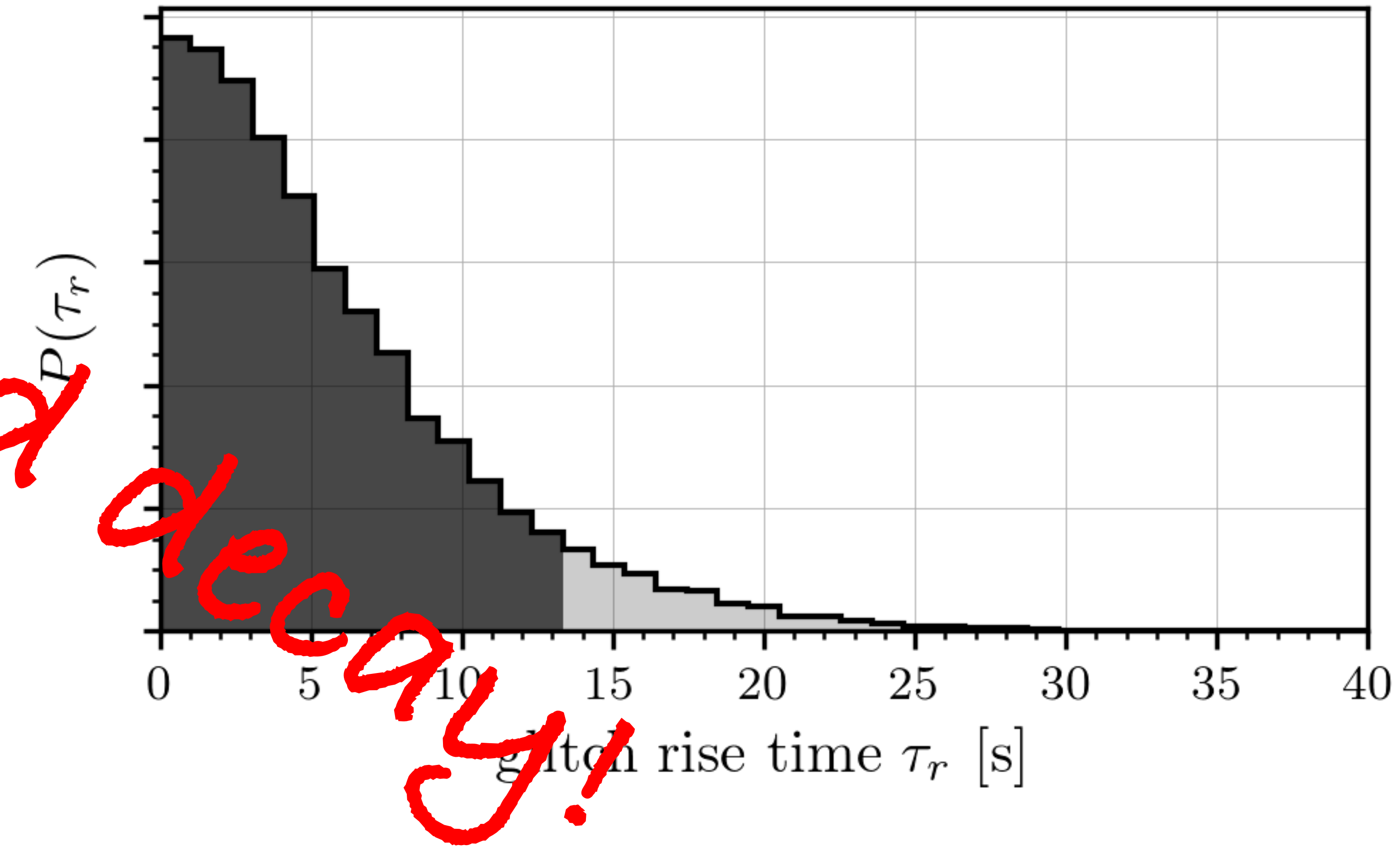
Farah et al. 2019

finally back to pulsars!

The 2016-vela glitch



- 200-s sliding windows
- each window, Bayesian inference using a constant-frequency model
- fit timing model with basic glitch, derive posterior for the rise time



Ashton, PL, Graber, Palfreyman (submitted)

$$p(\theta|d) = \frac{L(d|\theta)\pi(\theta)}{Z(d)}$$

Evidence: used for doing model selection!

$$\int d\theta p(\theta|d) = 1 \quad \longrightarrow \quad Z(d) = \int d\theta \, L(d|\theta)\pi(\theta)$$

Given two models: $M_1(\theta_1)$ and $M_2(\theta_2)$, the Bayes factor is defined as the ratio of the two evidences:

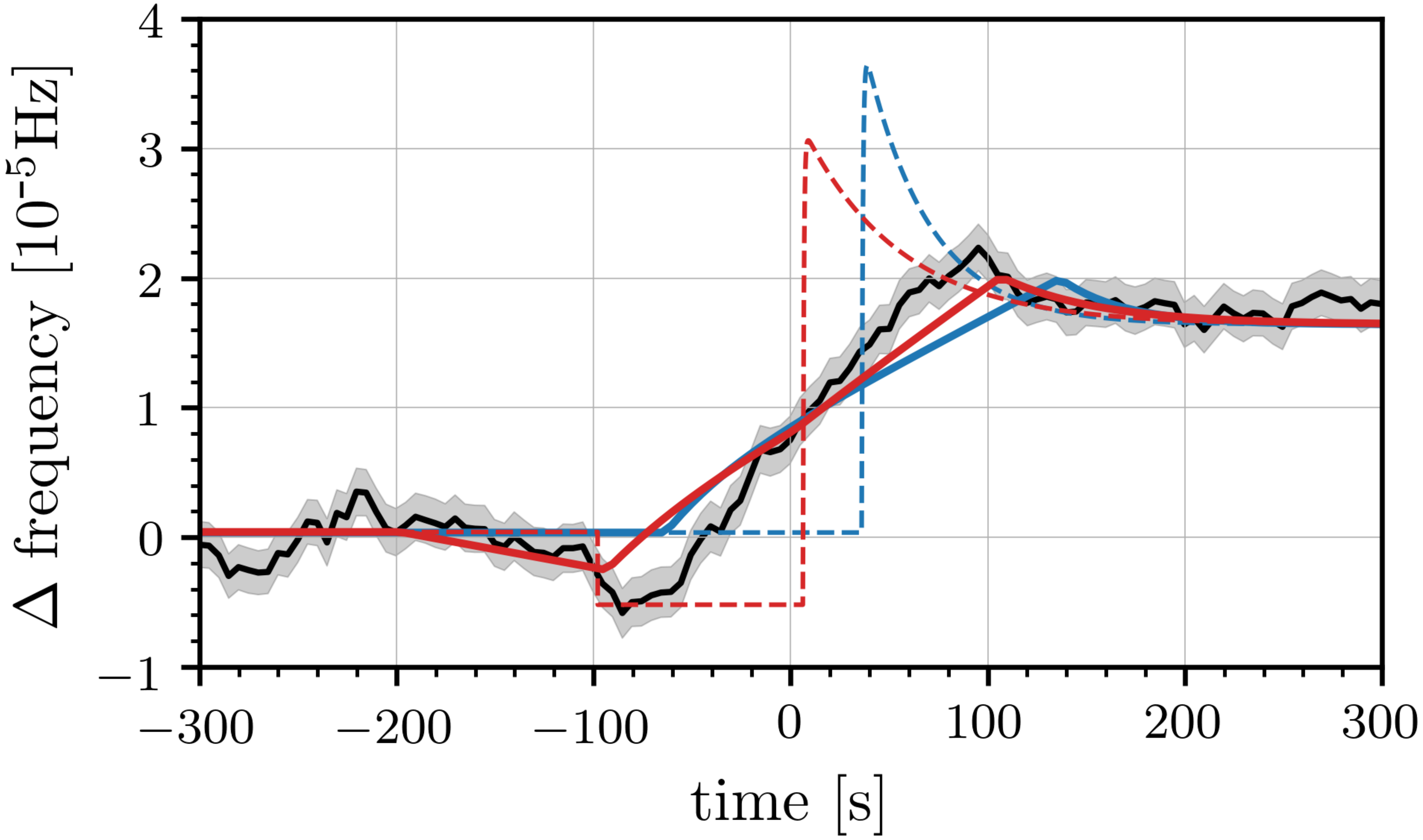
$$BF = \frac{Z_1}{Z_2} = \frac{\int d\theta_1 \, L(d|\theta_1)\pi(\theta_1)}{\int d\theta_2 \, L(d|\theta_2)\pi(\theta_2)}$$

The Bayes factor (almost*) gives the odds between any two models. e.g., model 1 is preferred over model 2 with XX Bayes factor…

* technically, it's the "Odds", which multiplies the Bayes factor by the prior odds. But for simple situations, the prior odds equal one!

Bayesian model selection

The 2016-vela glitch



- does the overshoot-decay exist?

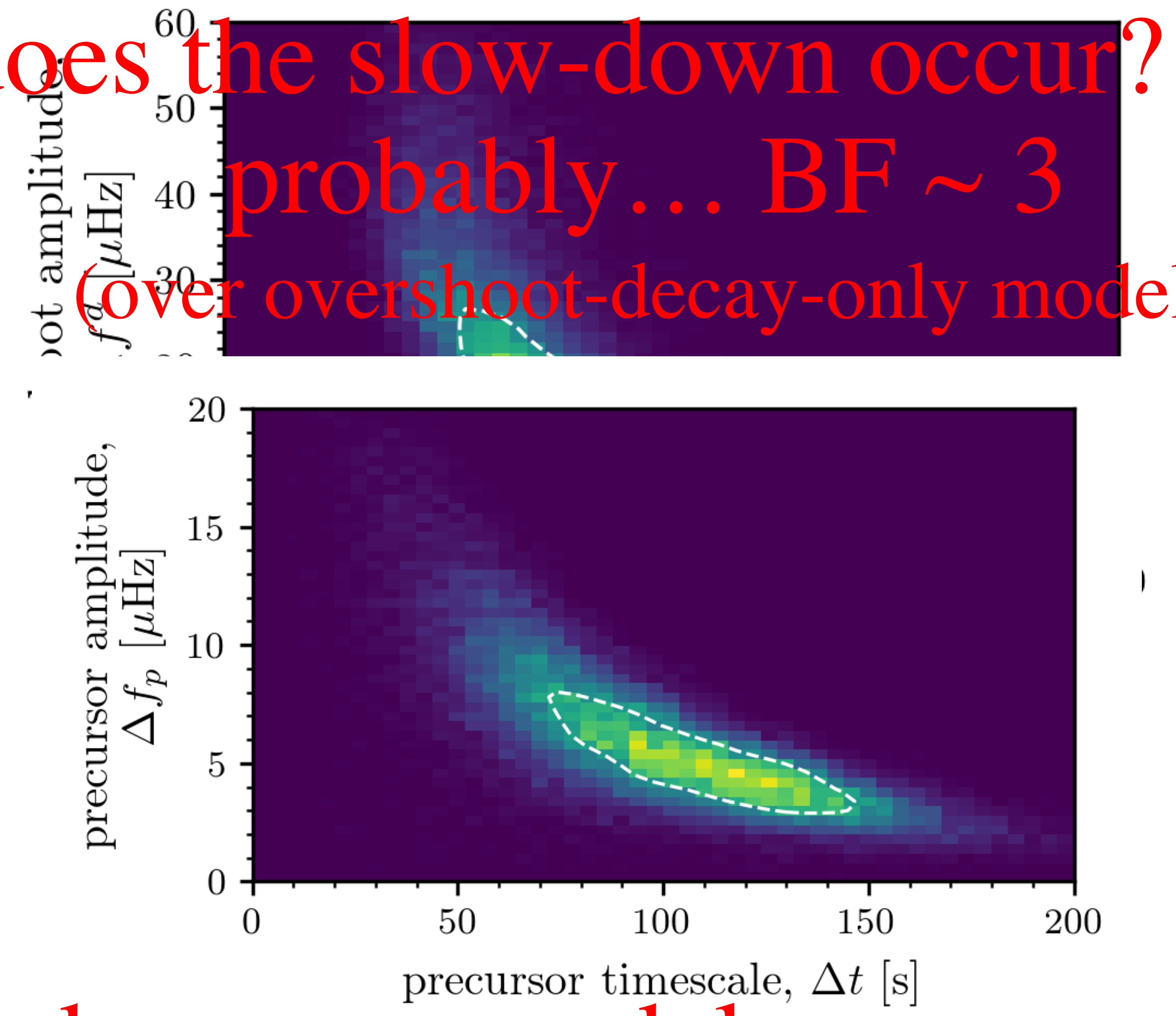
  yes! BF ~ 125

  (over simple step glitch)

- does the slow-down occur?

  probably… BF ~ 3

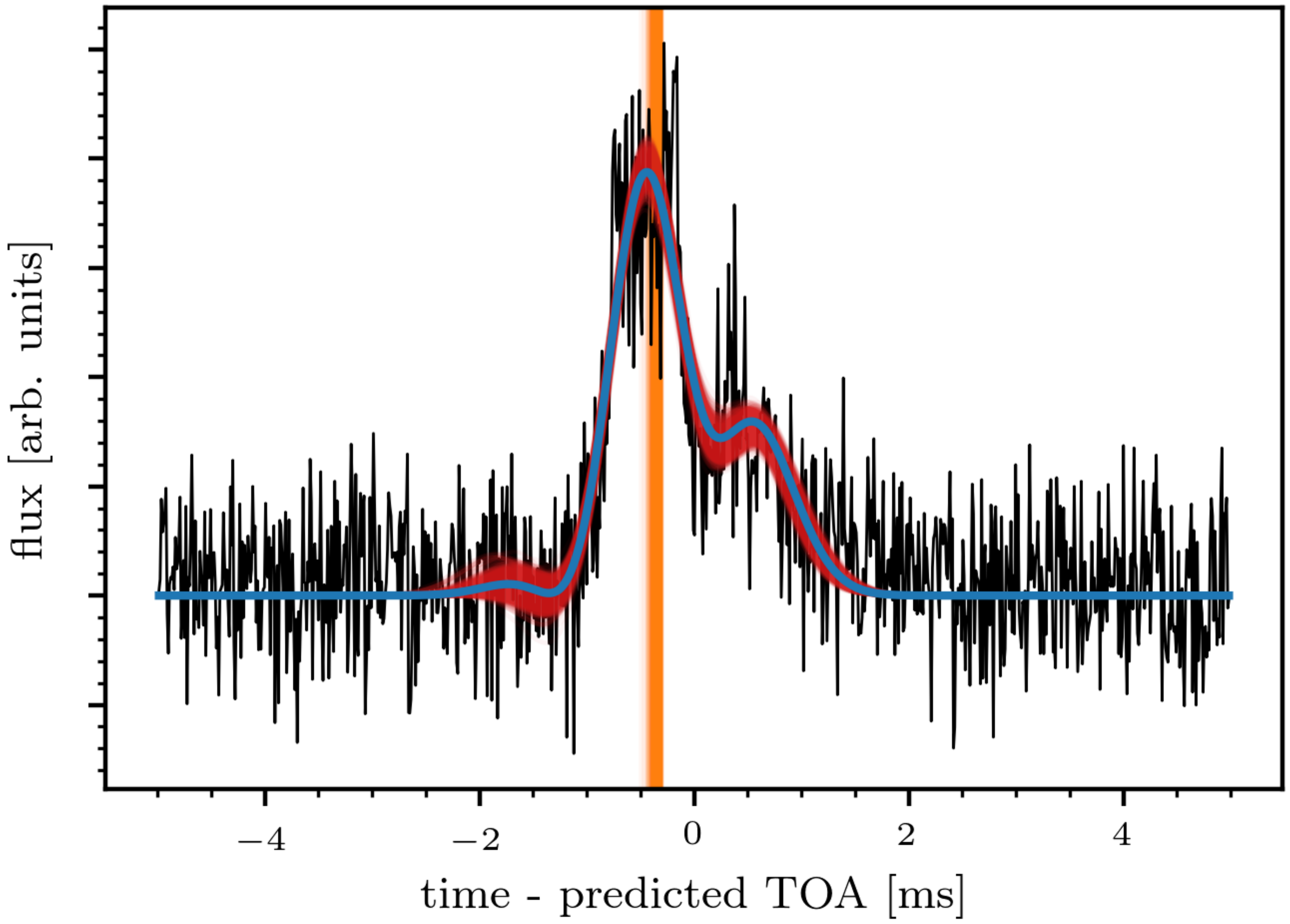
  (over overshoot-decay-only model)



word of caution: you're only as good as your models…

Ashton, PL, Graber, Palfreyman (submitted)

pulse-profile modelling



plot courtesy of Greg Ashton, via pyglit/Bilby

Greg Ashton has developed this;
speaking about it tomorrow!

- Pulsar timing uses pulse time-of-arrivals
- this is not the right way to do things
- Reverend Thomas Bayes would:
  - do full parameter estimation for each pulse
  - use a hierarchical model for the spin evolution, etc., of the pulsar

- hierarchical model includes things like
  - gravitational-wave background
  - pulsar red- and white-noise parameters
  - solar-system parameters
  - etc

“Bayesian inference is the future of gravitational-wave astronomy”

Matilda B. Bilby*

***still not a real Bilby**