Image: LIGO / Caltech / MIT / Sonoma State (Aurore Simonnet)

LIGO

An Introduction to Bayesian Parameter Estimation for Compact Binary Coalescences with bilby

Sylvia Biscoveanu

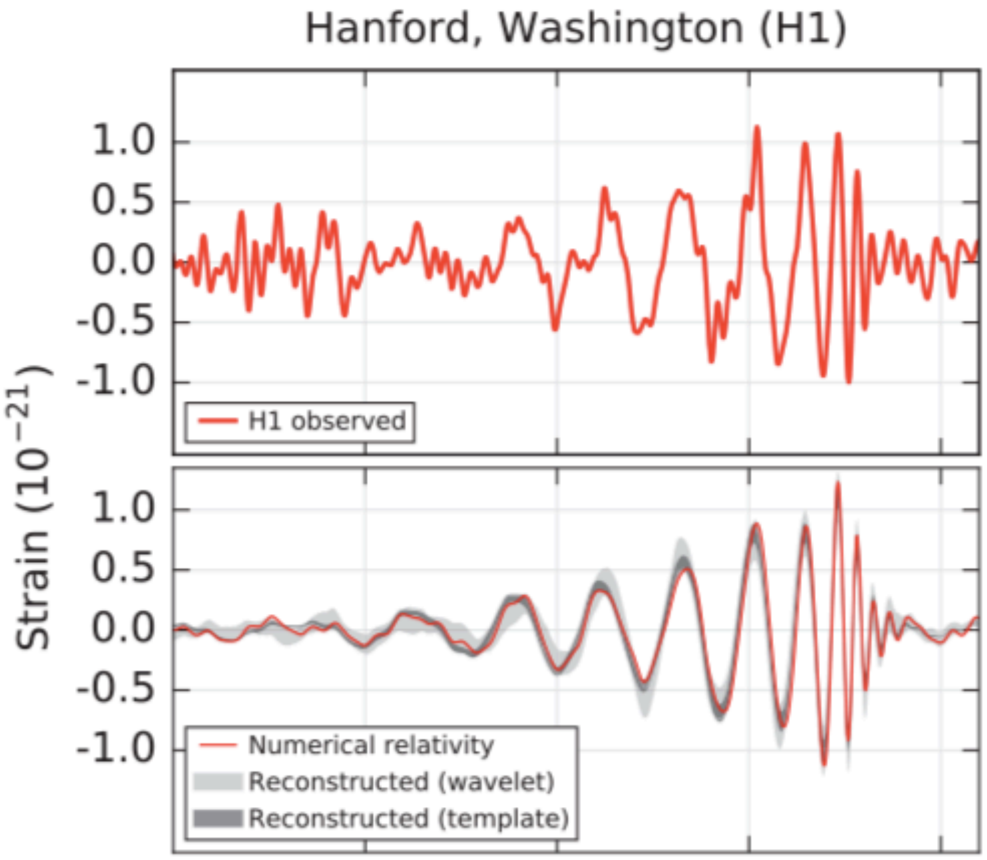
GWANW Student Workshop

June 28th 2021

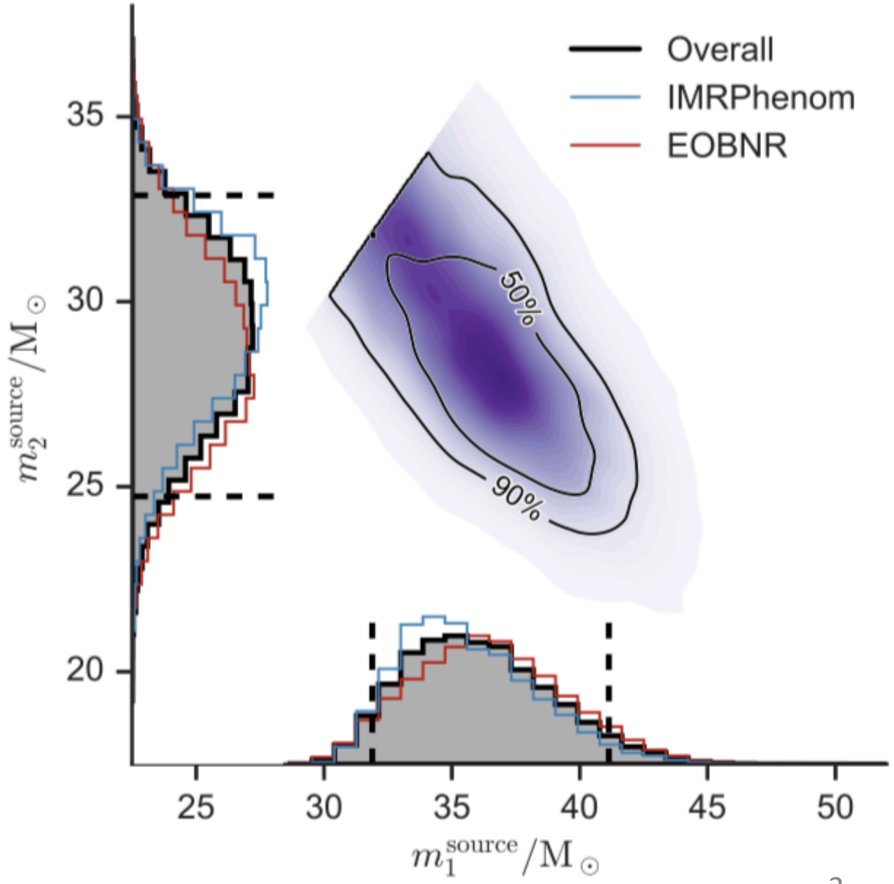
MIT Kavli Institute
for Astrophysics
and Space Research

Motivation

- How do we go from this… to this?



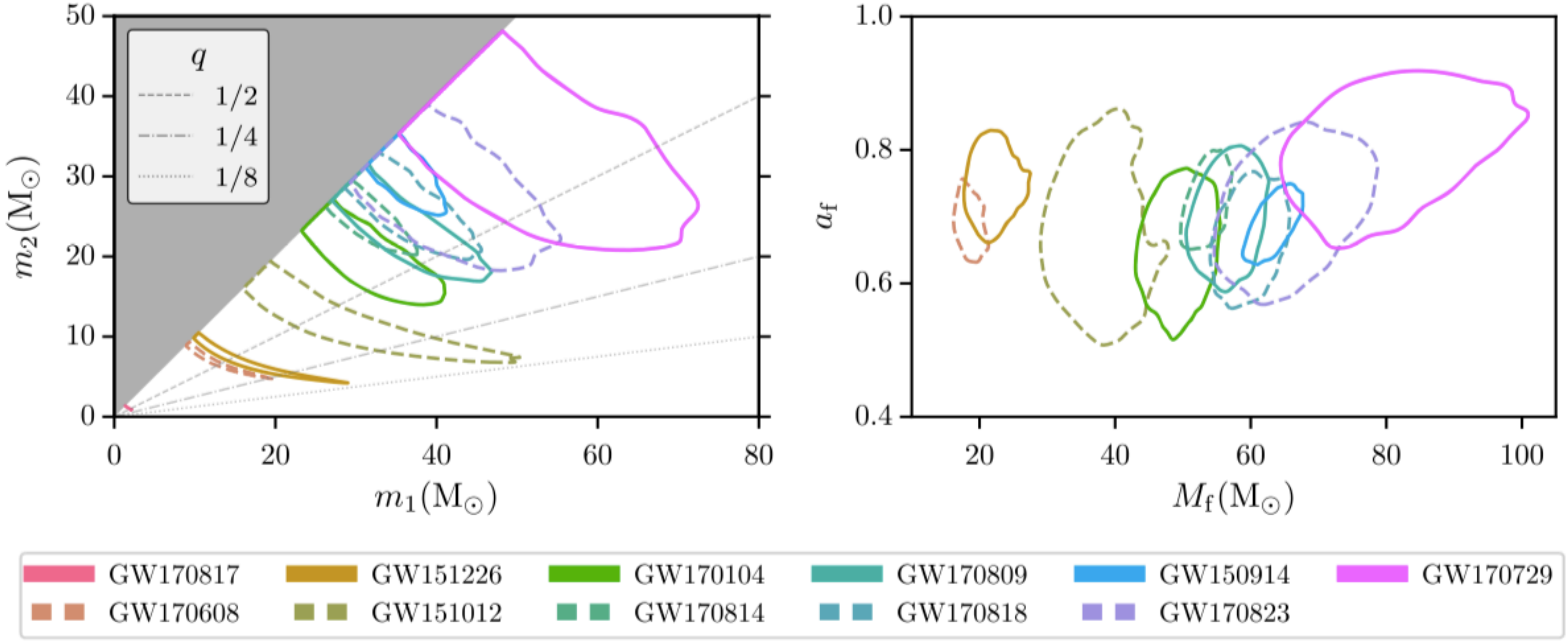
https://arxiv.org/abs/1602.03837



https://arxiv.org/pdf/1602.03840v2.pdf

Motivation

- Measure the properties of individual sources to learn about their evolutionary history



https://arxiv.org/pdf/1811.12907.pdf

Bayes' Theorem

$$\underbrace{p(\boldsymbol{\theta}|d, H)}_{\text{posterior}} = \frac{\overbrace{p(d|\boldsymbol{\theta}, H)}^{\text{likelihood}}\,\overbrace{p(\boldsymbol{\theta}|H)}^{\text{prior}}}{\underbrace{p(d|H)}_{\text{evidence}}}$$

posterior

parameters data model

likelihood prior

evidence

Bayes' Theorem Components

- Posterior – probability of the parameters θ given the data *d* and model *H*

$$p(\boldsymbol{\theta}|d, H)$$

- Likelihood – probability of the data *d* for parameters θ and model *H*

$$p(d|\boldsymbol{\theta}, H) \equiv \mathcal{L}(d|\boldsymbol{\theta}, H)$$

- Prior – initial probability of the parameters θ under model *H*

$$p(\boldsymbol{\theta}|H) \equiv \pi(\boldsymbol{\theta}|H)$$

Bayes' Theorem Components

- Evidence – normalization constant for the posterior, marginalized likelihood

$$p(d|H) \equiv \mathcal{Z}_H = \int \mathcal{L}(d|\boldsymbol{\theta}, H)\pi(\boldsymbol{\theta}|H)\ d\theta$$

Putting it all together:

$$p(\boldsymbol{\theta}|d, H) = \frac{\mathcal{L}(d|\boldsymbol{\theta}, H)\pi(\boldsymbol{\theta}|H)}{\mathcal{Z}_H}$$

The Gravitational-Wave Likelihood

- Under the assumption of Gaussian, stationary noise, the likelihood of observing data d given binary parameters θ is:

$$\mathcal{L}(d|\theta) \propto \exp\left(-\sum_k \frac{2|d_k - h_k(\theta)|^2}{T S_k}\right)$$

Segment duration

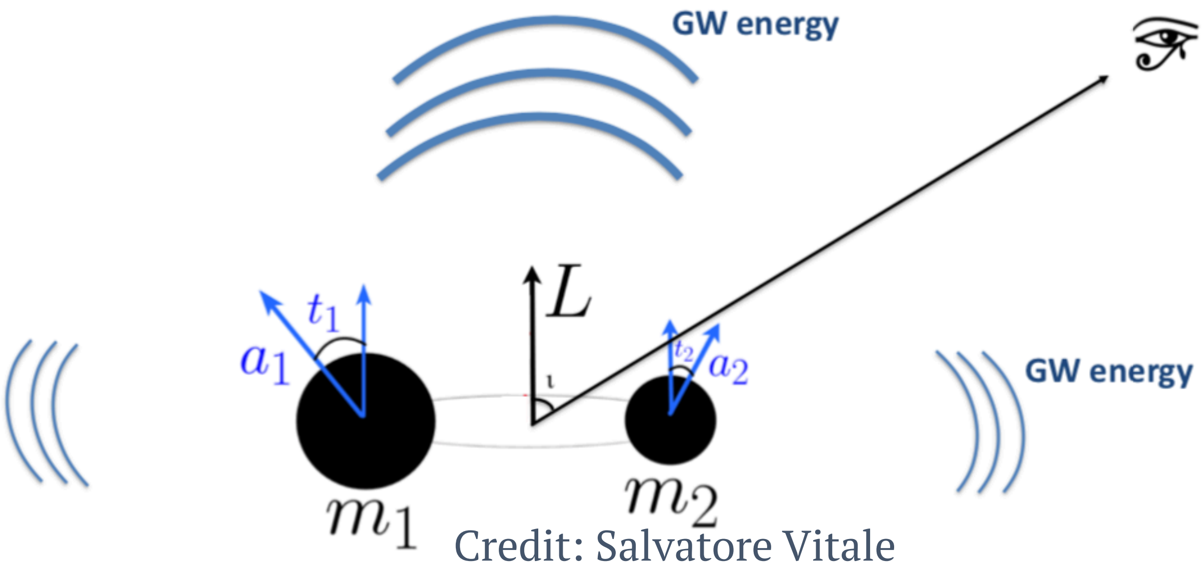
PSD

- $h(\theta)$ represents the gravitational waveform
- k subscript indicates frequency dependence
- Total likelihood is the product of the individual frequency bins

Binary parameters

- For compact binary coalescences, the astrophysical contribution is a waveform that depends on 17 parameters

Intrinsic:
Component masses
Component spins
(Tidal deformabilities)



Credit: Salvatore Vitale

Extrinsic:
Sky location
Distance
Inclination
Polarization
Reference phase
Time at coalescence

- Other models for other types of signals – sine gaussian wavelets, supernova waveforms, etc.

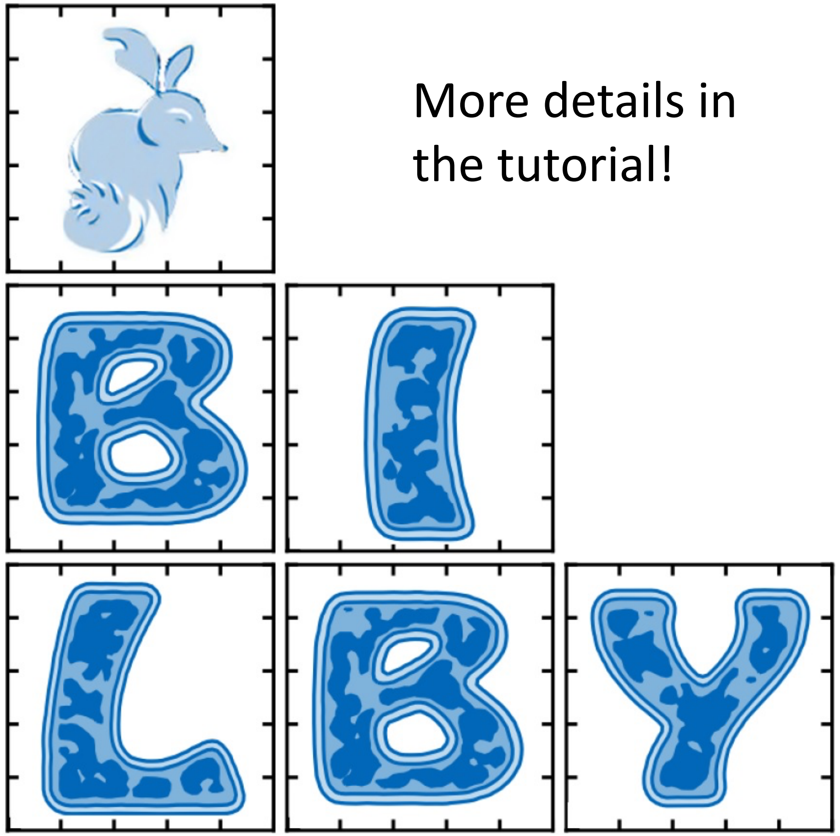
Priors

- Uniform in some parameterization of the mass
- Enforce $m_1 > m_2$
- Uniform in spin magnitudes
- Spin angles isotropic on the sphere
- Isotropic on the sky for right ascension and declination
- Uniform in luminosity volume ($\propto d_L^2$)

Sampling methods

- How do you actually obtain $p(\theta|d)$?
- Could evaluate the likelihood on a grid, but this isn't feasible with 17 parameters
- Instead use a stochastic sampler:
  - Markov Chain Monte Carlo (MCMC)
  - Nested sampling
- Obtain samples from the posterior probability distribution

Bilby



More details in the tutorial!

- The **B**ayesian **I**nference **Lib**rar**y** is a software package designed to enable parameter estimation for compact binary coalescences and more general problems
- Emphasis on modularity, transparency, and ease of use
- Wrapper for many different external samplers including dynesty, pymultinest, cpnest, emcee, ptemcee, and others
- Can analyze real data from LIGO and Virgo or simulated signals

Additional bilby resources

- Gitlab repo: https://git.ligo.org/lscsoft/bilby
- Documentation: https://lscsoft.docs.ligo.org/bilby/
- Slack workspace: bilby-code.slack.com
- GWTC-1 analysis: https://github.com/IsobelMarguarethe/Bilby-GWTC-1-Analysis-and-Verification/tree/v2.0
- Papers:
  - https://arxiv.org/abs/1811.02042
  - https://arxiv.org/abs/2006.00714

Backup

Measuring source properties from the waveform

$$\tilde{h}_+(f) = \frac{1}{2}\mathcal{A}_{\mathrm{GW}}(f)(1+\cos^2\iota)\cos\phi_{\mathrm{GW}}(f)$$

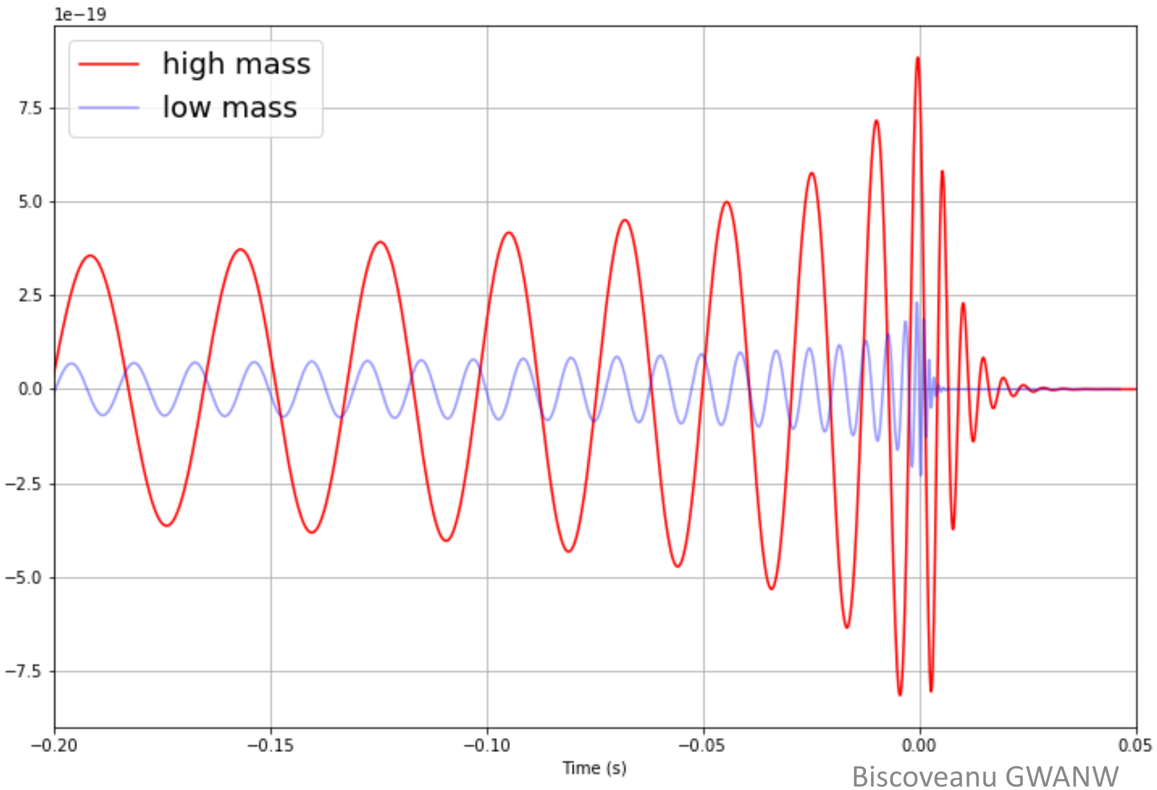
$$\tilde{h}_\times(f) = \mathcal{A}_{\mathrm{GW}}(f)\cos\iota\sin\phi_{\mathrm{GW}}(f)$$

$$\tilde{h}(f) = F_+\tilde{h}_+(f) + F_\times\tilde{h}_\times(f)$$

- Two polarizations
- Dependence on mass, spins, distance, etc. encoded in amplitude and phase
- Antenna patterns depend on the detector geometry and encode the effect of the extrinsic parameters

Effect of Mass

- Bigger mass → bigger amplitude
- Final mass measured from ringdown

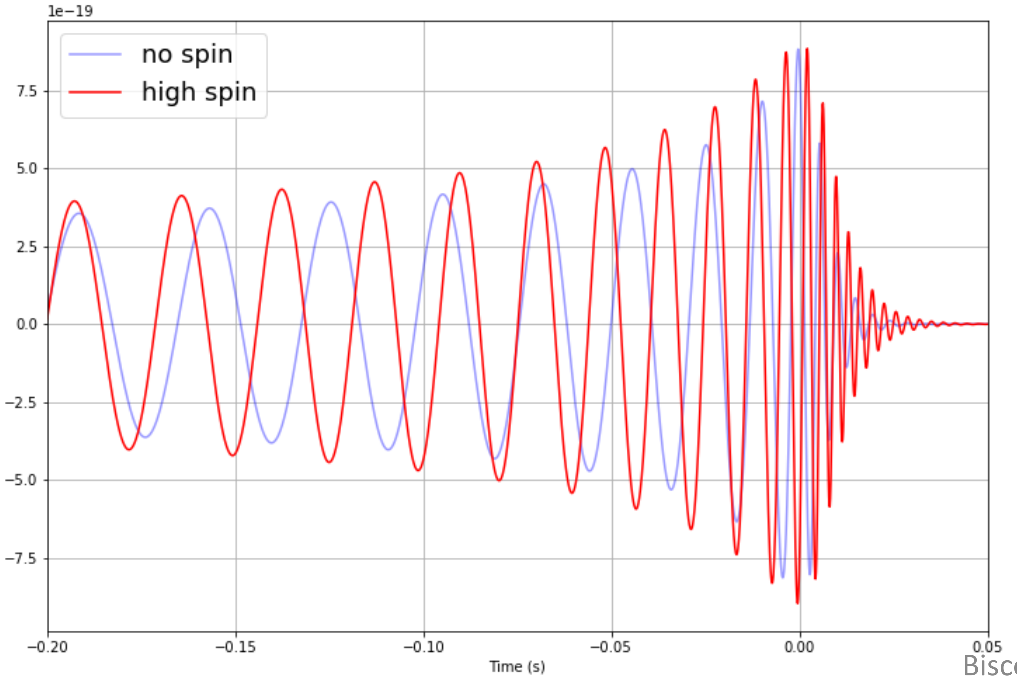


$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$

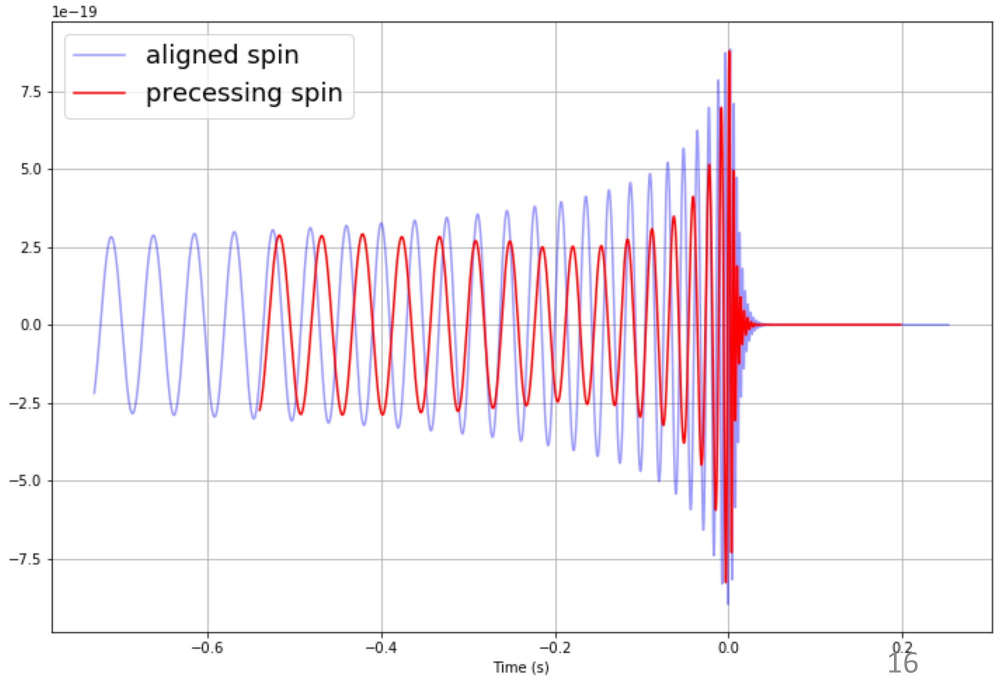
$$\mathcal{A}_{\mathrm{GW}} \propto \frac{\mathcal{M}^{5/6} f^{-7/6}}{d_L}$$

Effect of Spin

- More positive aligned spin → orbital hangup
- Takes longer for the system to merge

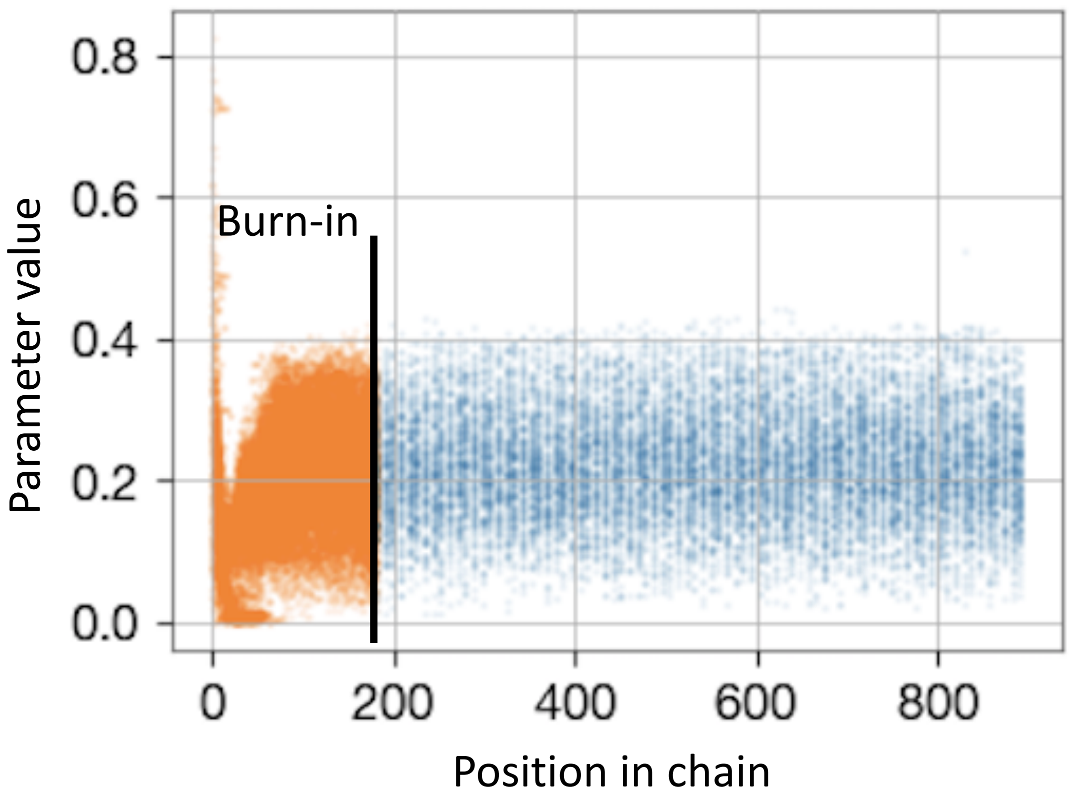


- Precessing spins → amplitude modulations
- Spins misaligned to orbital angular momentum

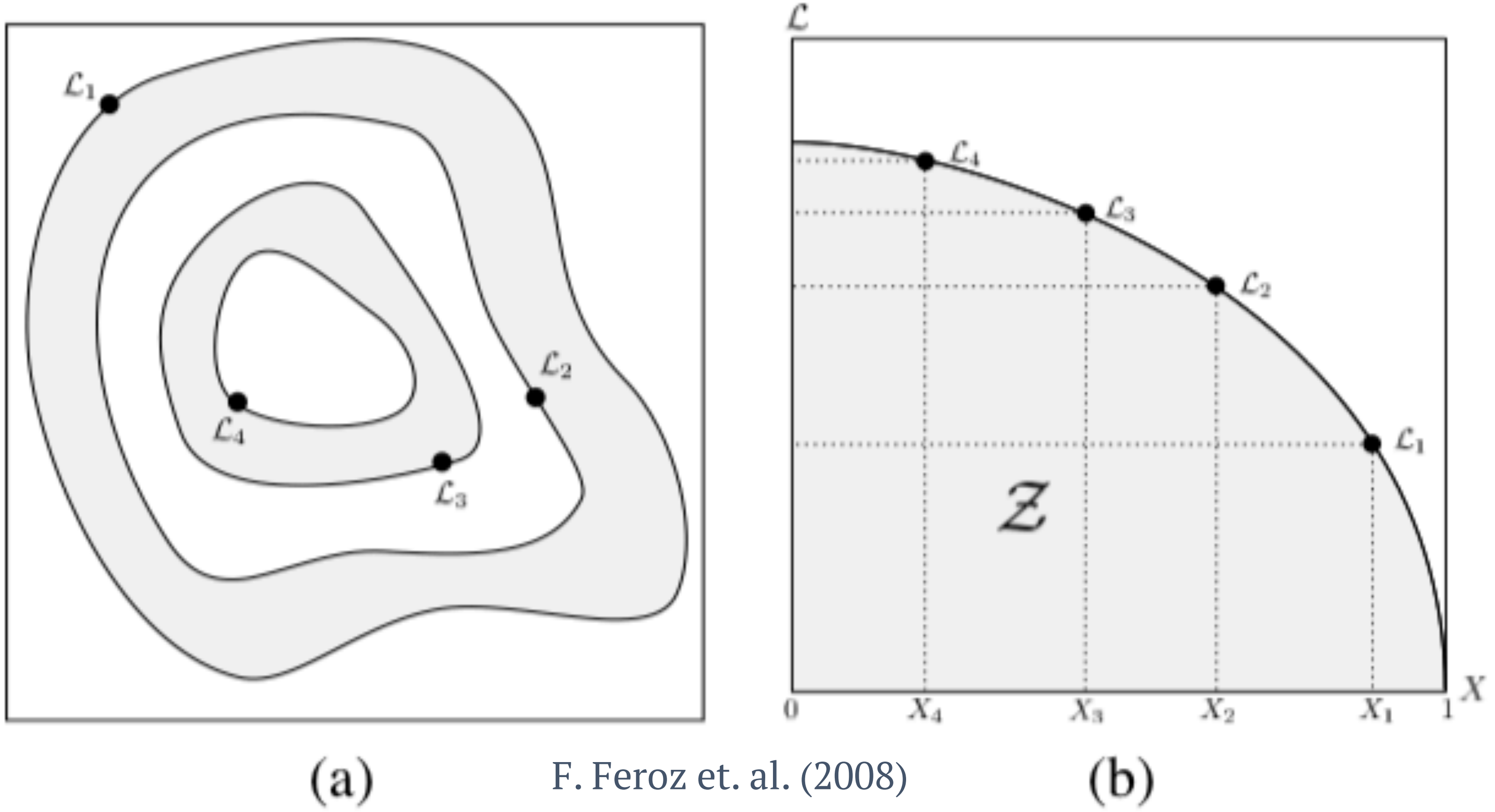


MCMC

- Particles undergo a random walk through the parameter space, where the probability of jumping to a new location is dictated by the proposal density function
- Determining a suitable proposal density function is the hard part of sampling – a simple example is a Gaussian centered on the current location
- Burn-in period before the walkers “forget” their starting positions
- Adjacent samples in a chain are correlated – chains need to be thinned by the integrated autocorrelation time



Nested Sampling



F. Feroz et. al. (2008)

- Sprinkle a set of live points over the prior space
- Replace the live point with the lowest likelihood with a point with a higher likelihood
- Evidence is the product of the likelihood at the discarded point and the difference in the prior volume between iterations
- Obtain samples from the prior in the process of calculating the evidence
- Proceed until a termination criterion is reached