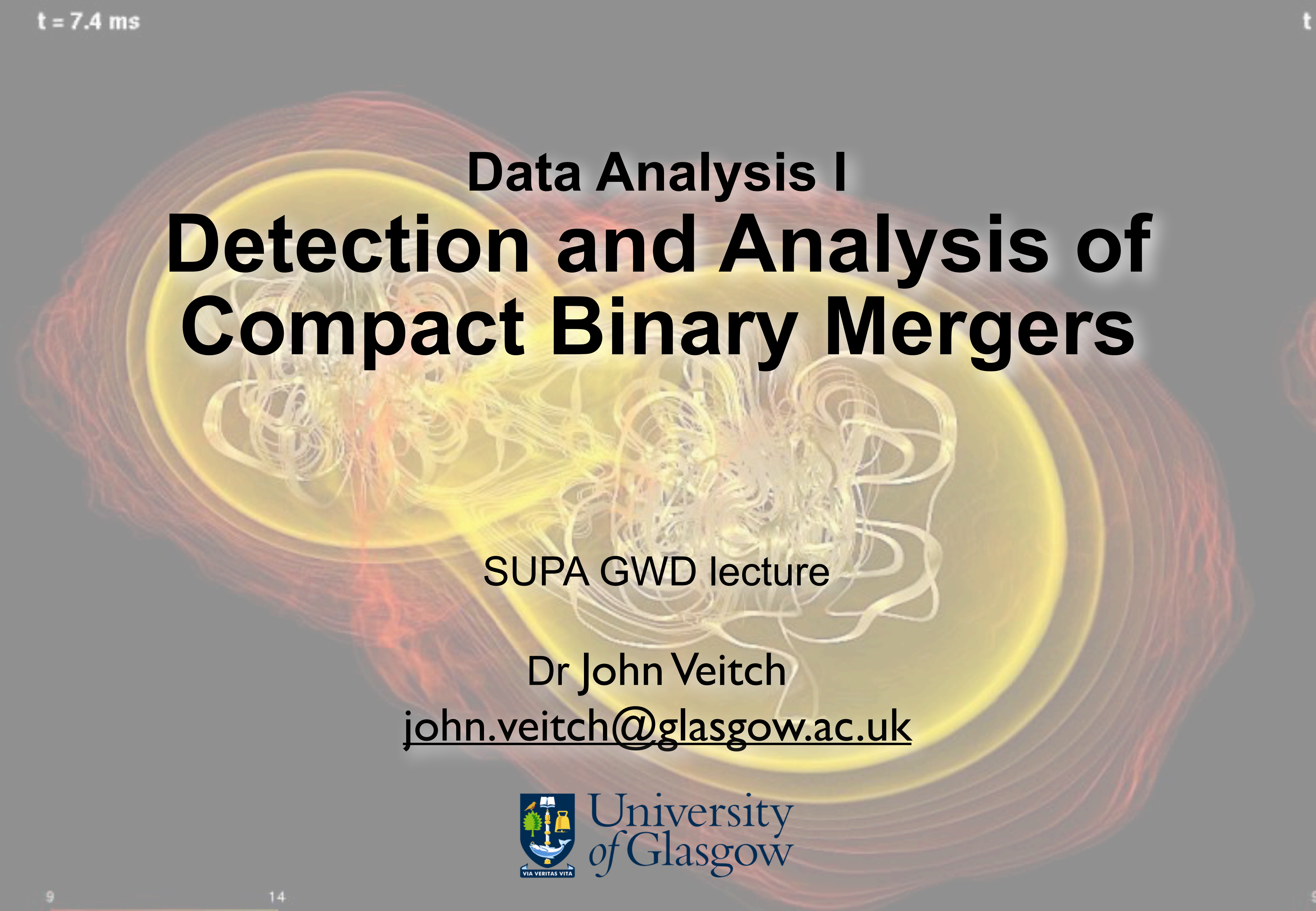


Data Analysis I

Detection and Analysis of Compact Binary Mergers

SUPA GWD lecture

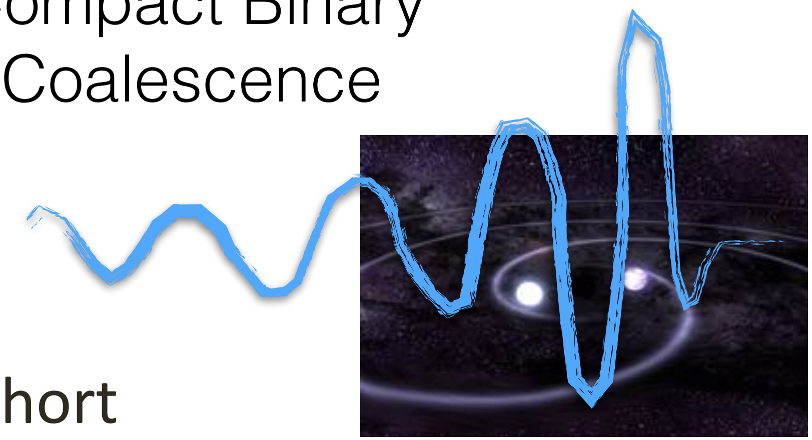
Dr John Veitch
john.veitch@glasgow.ac.uk
University of Glasgow

Gravitational wave source types

modelled

unmodelled

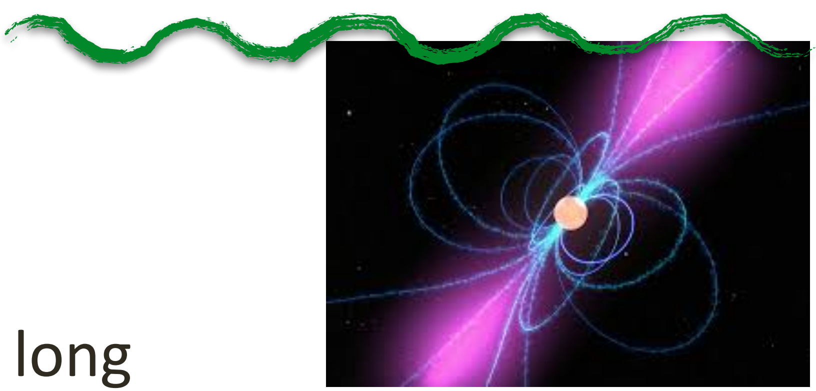
Compact Binary Coalescence



short

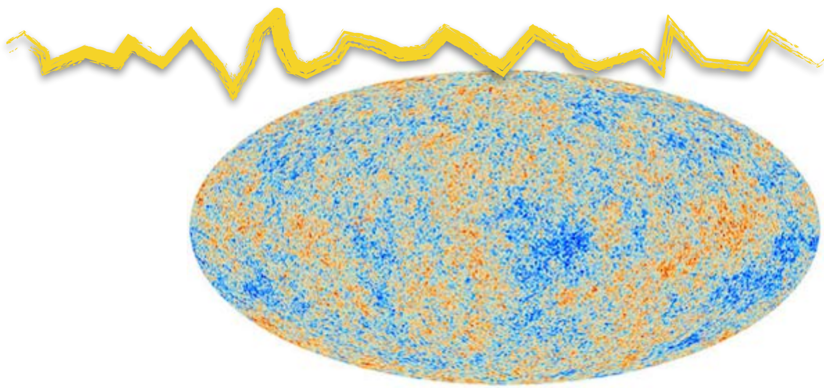
Burst

Continuous

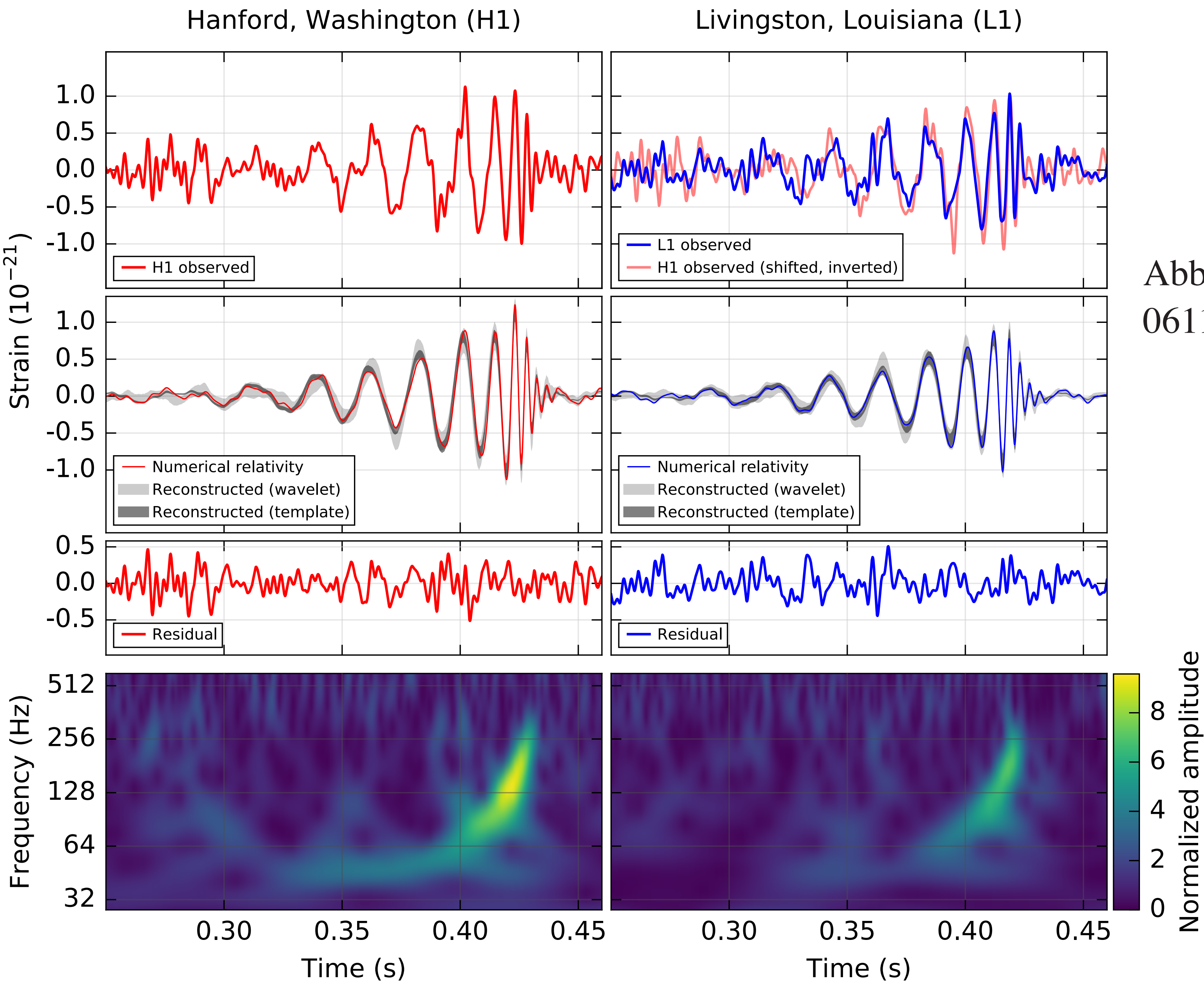


long

Stochastic



First detection



Abbot et al., PRL 116, 061102 (2016)

Basic considerations

- **Data analysis means learning about the world from data. Two basic questions for GW astronomy:**
  - ➡Is something there? [detection]
  - ➡If so, what are its characteristics? [parameter estimation]
- **To connect a theory to observations we need it to be able to describe the data**
  - ➡Data is noisy/uncertain - must be a statistical model
  - ➡Model may have unknown parameters
- **We want to answer questions like "Is the theory X true, given our observations?"**
  - ➡Use Statistical inference / Probability theory
  - ➡**P(X|Y)** means "Degree of belief that X is true, assuming Y"
  - ➡0 = False, 1 = True

Bayes' Theorem

- What we need is a method to update our beliefs given new information

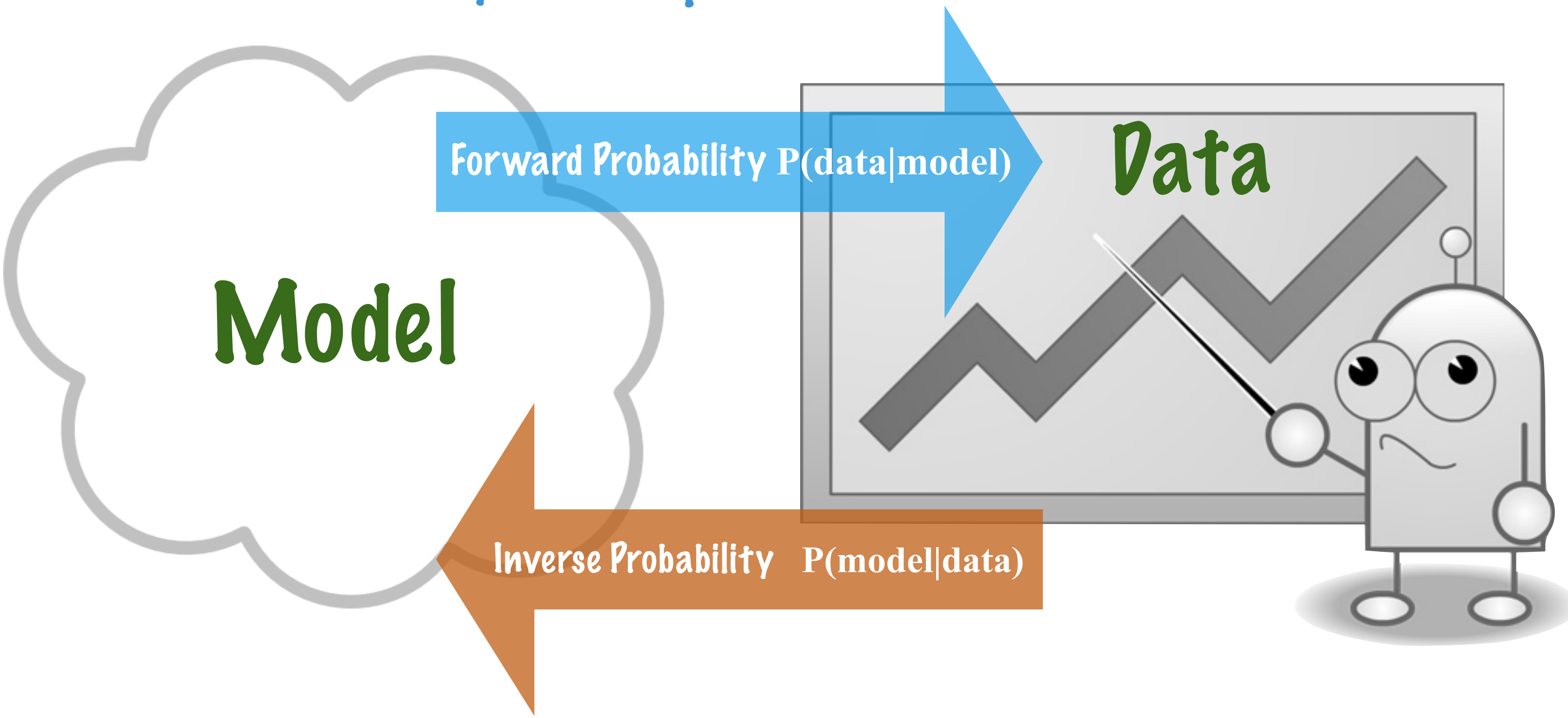
- A mathematical theory of learning that can be programmed into a computer

$$P(\text{theory}|\text{new data}) = \frac{P(\text{theory})P(\text{new data}|\text{theory})}{P(\text{new data})}$$

Forward and Inverse Probability

$$P(\text{theory}|\text{new data}) = \frac{P(\text{theory})P(\text{new data}|\text{theory})}{P(\text{new data})}$$

Forward Probability - make predictions about data from a model



Inverse Probability - Learn about a model from data

Bayesian model selection

- Recall Bayes' theorem where for desired model, M_A, and some observational data, D, we have

$$\underbrace{p(M_A|D,I)}^{\text{posterior}} = \frac{\overbrace{p(D|M_A,I)}^{\text{likelihood}} \times \overbrace{p(M_A|I)}^{\text{prior}}}{\underbrace{p(D|I)}_{\text{evidence}}}$$

- The posterior probability represents the state of our knowledge of the model ("the truth") in light of our observed data
- If we have a competing model or hypothesis, we use the ratio of the posterior probabilities for each model

$$\frac{p(M_A|D,I)}{p(M_B|D,I)} = \frac{p(M_A|I)}{p(M_B|I)} \times \frac{p(D|M_A,I)}{p(D|M_B,I)}$$

Bayesian model selection

- **This ratio of posterior probabilities is the odds ratio**

$$O_{AB} = \frac{p(M_A|D,I)}{p(M_B|D,I)} = \underbrace{\boxed{\frac{p(M_A|I)}{p(M_B|I)}}}_{\text{prior odds}} \times \underbrace{\boxed{\frac{p(D|M_A,I)}{p(D|M_B,I)}}}_{\text{Bayes factor}}$$

- **If $O_{AB} > 1$, M_A is preferred. If $O_{AB} < 1$, M_B is preferred**
- **If $O_{AB} = 1$, then there is insufficient information in the data to support either model**
- **There are a couple of points to note**

-Prior odds: If we believe that one model is more likely than the other, then that bias is reflected in this ratio.

-Bayes factor: The ratio of the likelihoods for each model, where we determine which model the data supports more

Detection: Is a signal present?

- **What are the odds of a signal being present "S" vs just noise "N", after observing data D?**
- $$\frac{P(S|D)}{P(S|N)} = \frac{P(S)}{P(N)} \times \frac{P(D|S)}{P(D|N)}$$
- **Use the likelihood ratio as a "detection statistic"**
- **Need likelihood functions (descriptions of the data)**
  - ➡ P(D|N): noise model
  - ➡ P(D|S): signal model
    - "**Modelled**" searches depend on theory to make strong predictions about the signal based on physical properties of the source: compact binaries, continuous waves
    - "**Unmodelled**" searches are less informed about the source (no good theory, or the source is stochastic) so rely on phenomenological parameters (e.g. signal energy, bandwidth, timespan?)

Noise Model

- GW detector noise is a series of **unknown fluctuations** *n* in the detector output *d* caused by the various noise sources. It is added to any true signal *h*
- $$d(t) = h(t) + n(t)$$
- We don't know precisely the values of *n(t)* so we describe it statistically: mean (zero), correlations, etc

➡We model the noise with a probability distribution function p(n|N) ~ "probability of seeing a specific noise n"

- If the autocorrelation function of the noise depends only on τ then we say it is a stationary noise process
- $$C(n(t), n(t+\tau)) = C_N(\tau)$$
- Power Spectral Density is Fourier transform of autocorrelation function

$$S_N(f) = 2\int_{-\infty}^{\infty} C_N(\tau)\exp(2\pi i f\tau)d\tau$$

Frequency Domain Analysis

- Noise likelihood knowing only mean=0 and correlation function is multi-variate Gaussian

$$p(n|N) \propto \exp\left(-\frac{1}{2}\iint_0^T n(t)C^{-1}(t,t')n(t')dtdt'\right) \propto \exp\left(-\frac{1}{2}\iint_0^T n(t)C_N^{-1}(\tau)n(t+\tau)dtd\tau\right)$$

- This is a convolution, use: $\mathscr{F}[a * b] = \mathscr{F}[a] \cdot \mathscr{F}[b]$
- Fourier frequency bins orthogonal so no correlations between frequencies. Likelihood simplifies to

$$\exp\left(-\int \frac{\tilde{n}(f)\tilde{n}^*(f)}{S_N(f)}df\right) = \exp\left(-\frac{1}{2} < \tilde{n}|\tilde{n} >\right)$$

- Define noise-weighted inner product

$$\langle a, b\rangle \equiv 2\int_0^\infty \frac{df}{S_h(f)}[\tilde{a}(f)\tilde{b}^*(f) + \tilde{a}^*(f)\tilde{b}(f)]$$

Signal Likelihood

- **Signal model predicts a gravitational waveform $h(t)$ or $\tilde{h}(f)$**
- $d = h + n$ **so likelihood of residuals is noise likelihood** $p(d|S) = \exp\left(-\frac{1}{2} < d-h|d-h > \right)$
- $\log L_S = -\frac{1}{2} < h|h > + < d|h > - \frac{1}{2} < d|d >$
- **For a fixed signal *h* the log-likelihood ratio between signal and noise models is**

$$\log(L_S/L_N) = -\frac{1}{2} < h|h > + < d|h >$$

Signal-to-Noise Ratio

- Consider a model that has an amplitude parameter $h = A\hat{h}$, where normalised template is $\hat{h} = \dfrac{h}{\sqrt{<h|h>}}$
- The likelihood is maximised at $\left.\dfrac{\partial \log L_S}{\partial A}\right|_{A_0} = 0$
- $\dfrac{\partial}{\partial A}\left[-\dfrac{1}{2}A^2 < \hat{h}|\hat{h} > + A < \hat{h}|d > \right] = 0$
- $A_0 = < \hat{h}|d > = < h|d > /\sqrt{<h|h>} = \rho$ Signal to Noise Ratio (SNR)
- The maximum log-likelihood ratio is $\log L_{max} = \dfrac{1}{2}\rho^2$
- Since this is monotonic in ρ, we can use the SNR as a detection statistic. Higher SNR gives higher posterior odds of signal!

Compact binary coalescence signal

- Compact binary systems are binaries which consist of neutron stars and/or black holes
- As the compact objects orbit, angular momentum is radiated away as gravitational radiation, leading to an inspiraling orbit
- The gravitational wave signal for the two polarisations takes the form

$$h_+(t) = A_{\rm GW}(t)(1+\cos^2\iota)\cos\phi_{\rm GW}(t)$$

$$h_\times(t) = -2A_{\rm GW}(t)\cos\iota\sin\phi_{\rm GW}(t)$$

- The observed signal is a function of the antenna patterns

$$h_{\rm det}(t) = F_+h_+(t) + F_\times h_\times(t)$$

Compact Binary Coalescence Signal

- The gravitational wave signal for the two polarisations takes the form

$$h_+(f) = \frac{1}{2}A_{GW}(f)(1+\cos^2\iota)\cos\phi_{GW}(f)$$

$$h_\times(f) = A_{GW}(f)\cos\iota\sin\phi_{GW}(f)$$

- The observed signal is a function of the antenna patterns

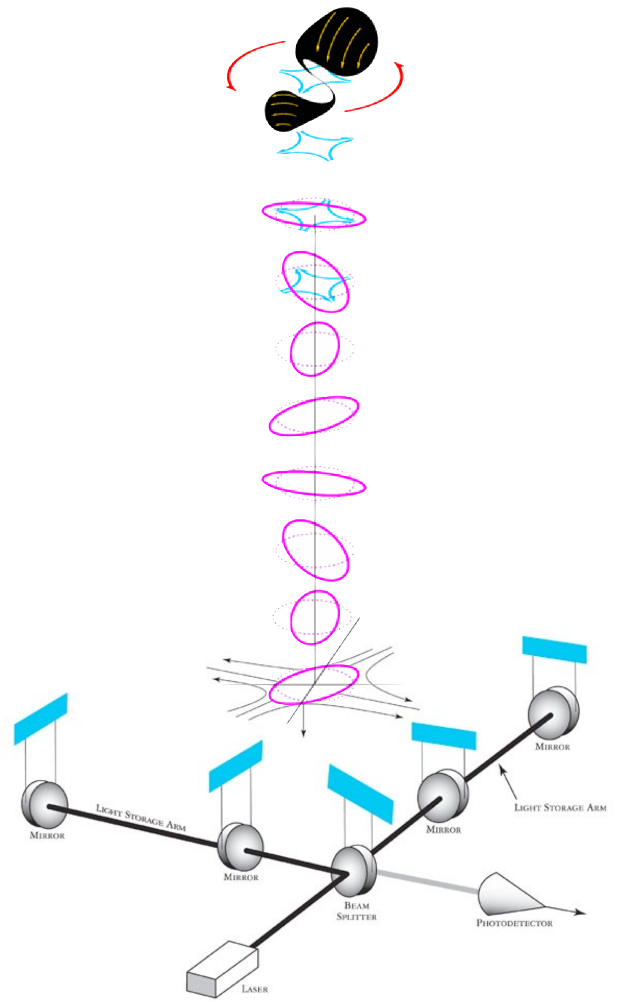
$$h_{\mathrm{det}}(t) = F_+h_+(t) + F_\times h_\times(t)$$

- Amplitude and phase from waveform model, e.g. PN

$$A_{GW}(f) \propto \frac{\mathcal{M}^{5/6}f^{-7/6}}{d_L}$$

$$\phi_{GW}(f,\mathcal{M},\eta) = \sum_k (\psi_k(\eta) + \psi_k^{(l)}(\eta))(\pi M f)^{(k-5)/3}$$

post-Newtonian series

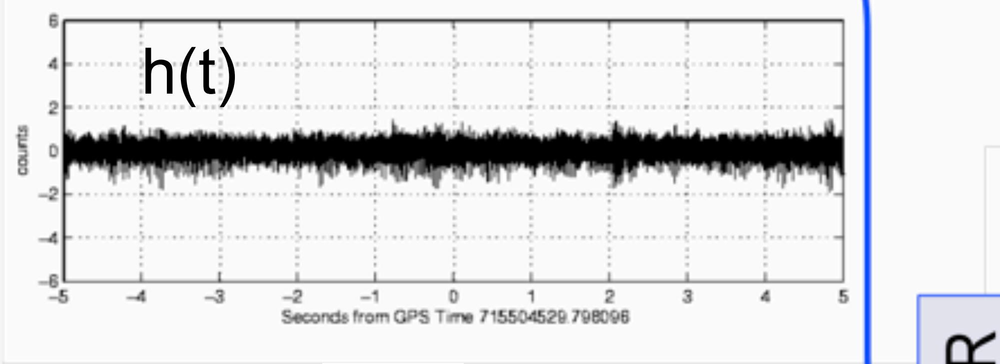


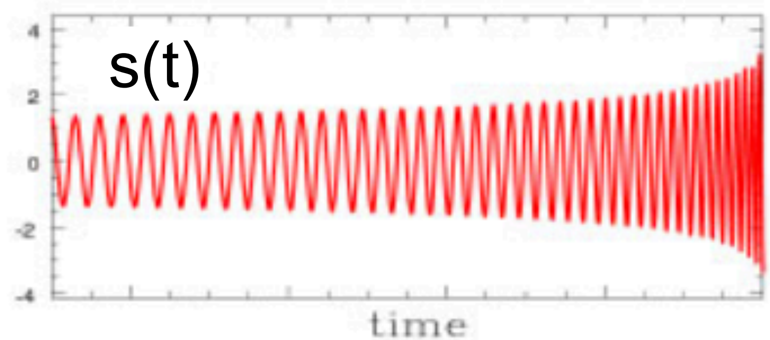
Matched Filtering

- Signal model is finite, and has some unknown reference time t_0 (e.g. time of coalescence)
- $h(t-t_0)$ becomes $\tilde{h}(f)e^{-2\pi i f t_0}$ in freq. domain
- We want to search over time to find the SNR time series
- $$\rho(t_0) = < \hat{h}(t-t_0) \,|\, d > \; = \int \frac{\tilde{d}^* \tilde{\hat{h}}(f) e^{-2\pi i f t_0}}{S_N(f)} df$$
- inverse Fourier transform of the integrand - use FFT to efficiently implement search.
- This is called the matched filter

Calculating Inspiral SNR

GW Channel
+ simulated inspiral

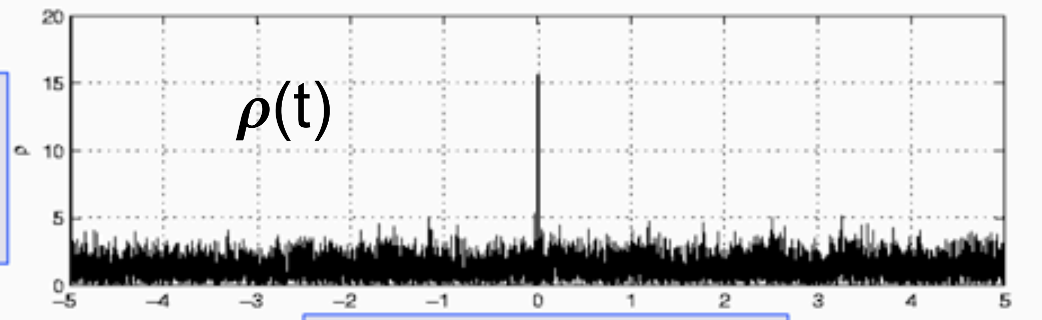




Signal template

SNR

Identify triggers as peaks of SNR



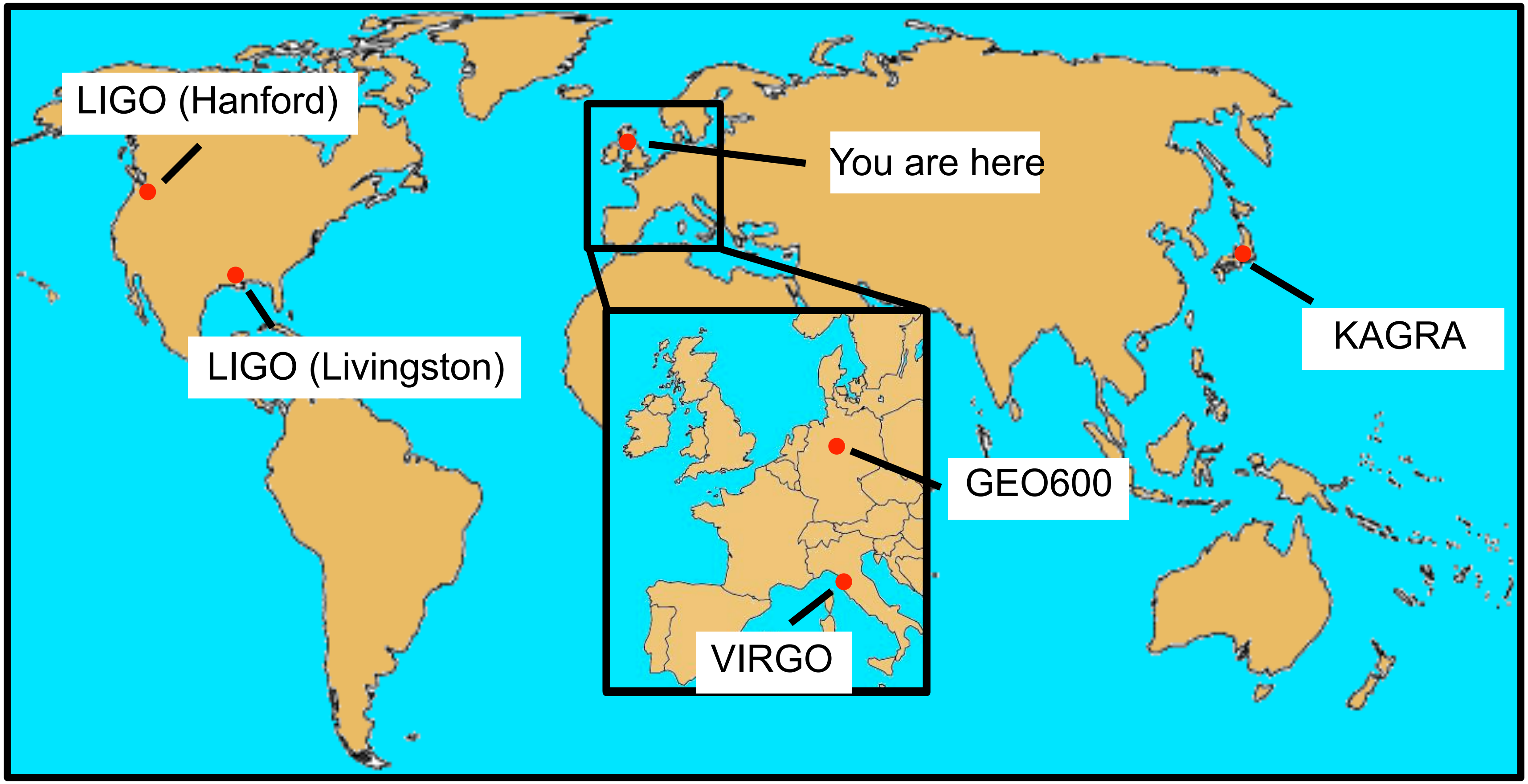
Coalescence Time

Proven to be the optimal filter for
a known signal in Gaussian noise.
Use ρ as a “ranking statistic”

Coincident Analyses

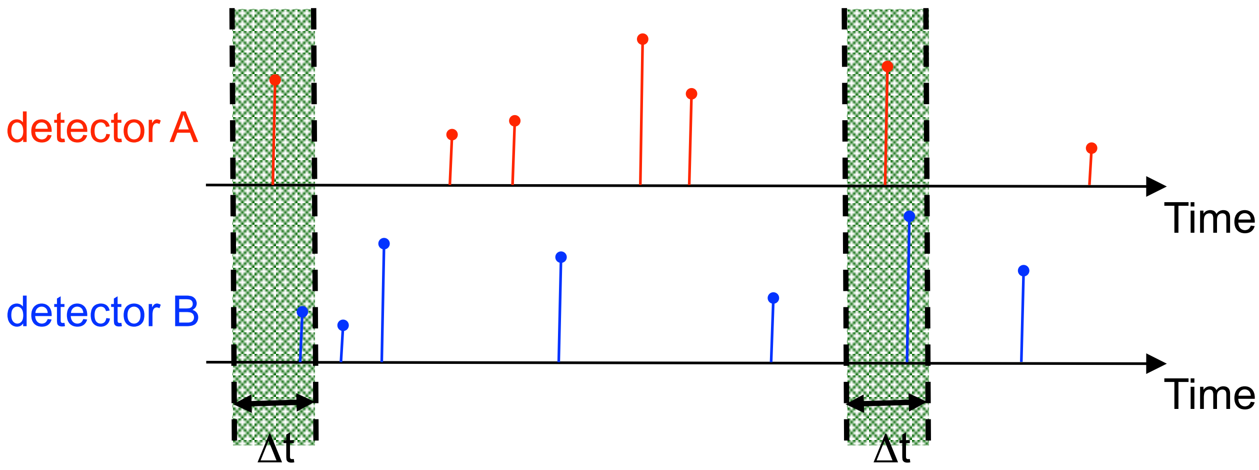
University of Glasgow

Interferometric detectors worldwide



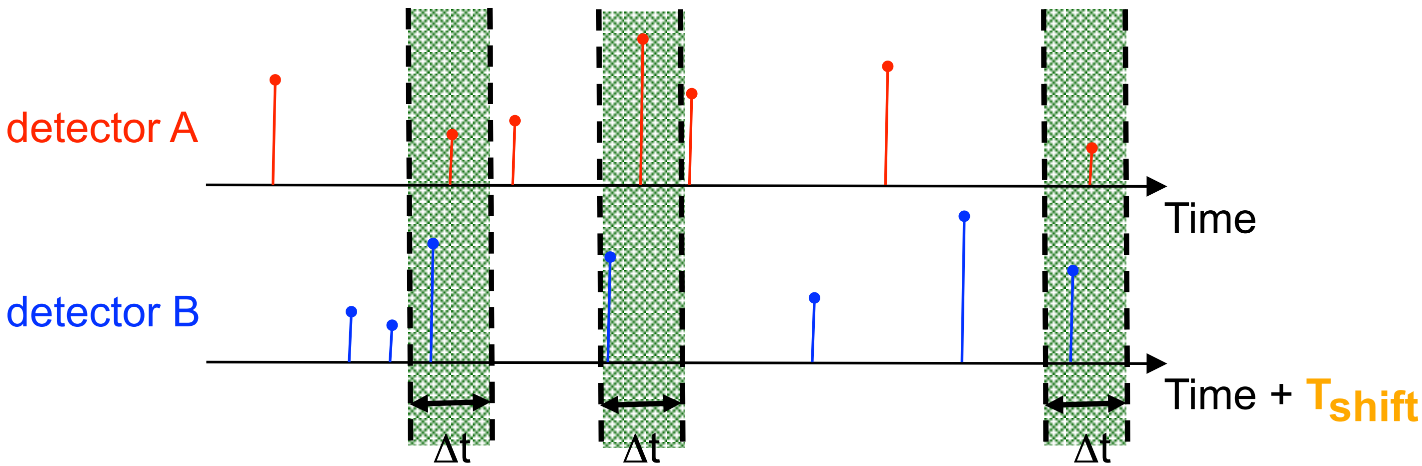
Coincidence analysis

- **Transient gravitational wave will** excite 2 widely-spaced detectors almost simultaneously (within light travel time between the detectors)
- **To identify possible signal,** count number of events that are coincident within a particular time window, Δt

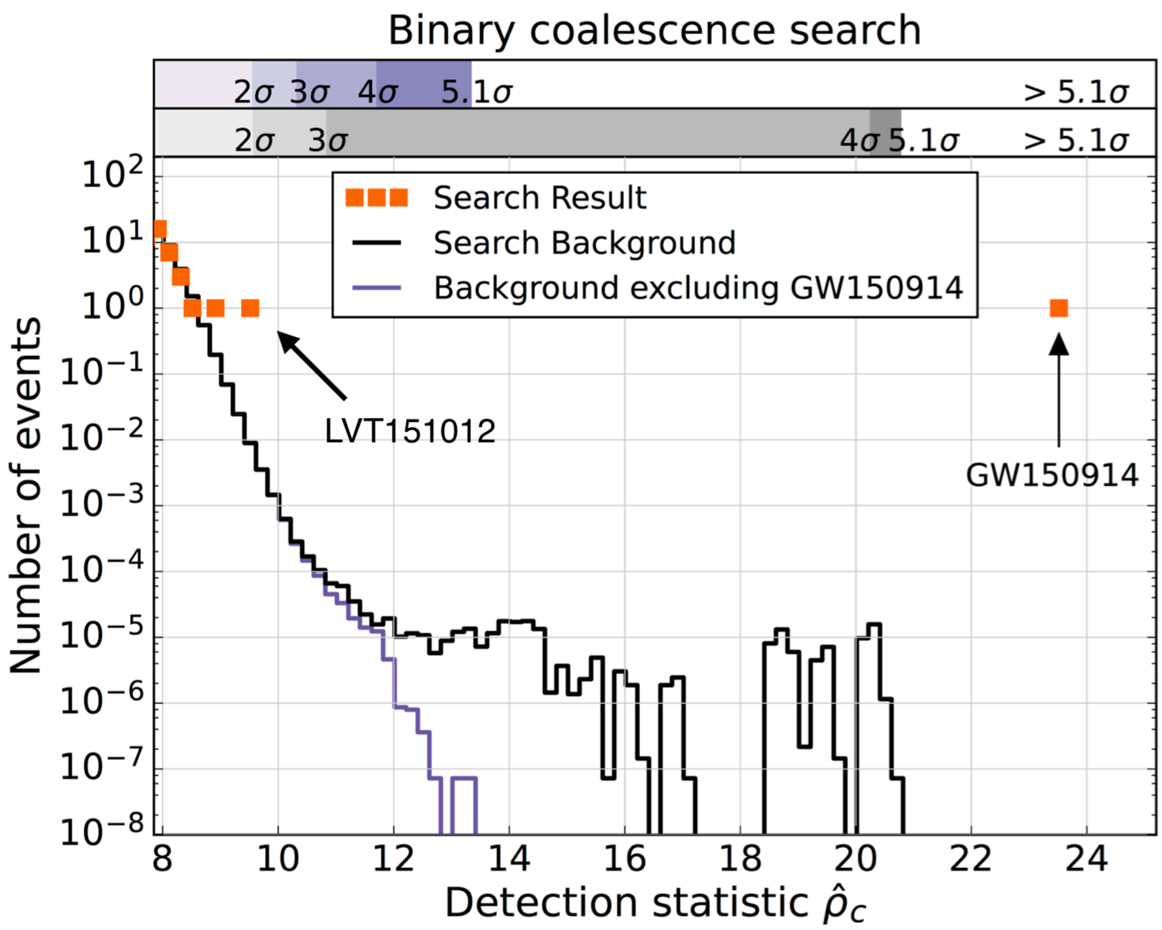


Coincidence analysis

- **Need to** estimate the rate of "accidental" coincidences (false alarms)
- **Add a time shift** (T_{shift}) to the data so that any coincidences from possible signals can no longer occur and count coincidences again



Estimating significance



The background is estimated by time-shifted data

GW150914 is louder than any background event and has an estimated significance of 1 in 203000 years

Parameter Estimation

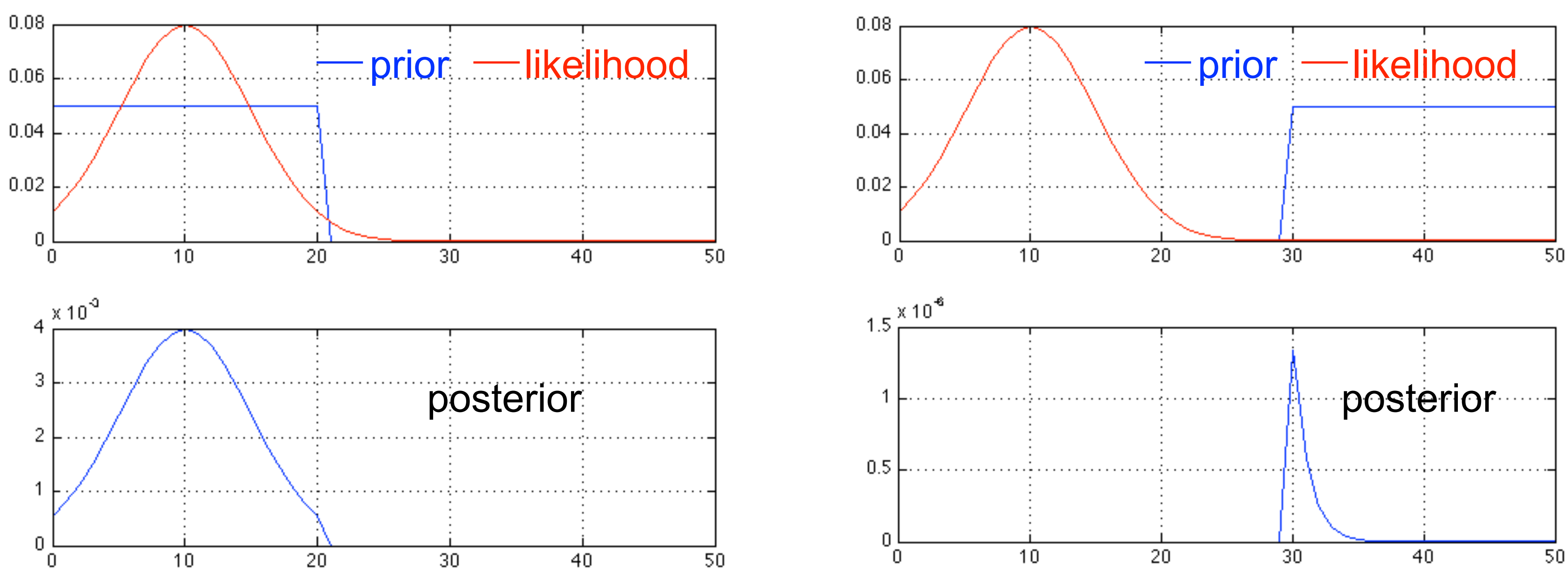
Bayesian inference

$$p(parameter|data, I) \propto p(data|parameter, I) \times p(parameter|I)$$

- **Here, the posterior probability represents our knowledge about the model given the data we have acquired**
- **The prior probability represents our state of knowledge about the model before any analysis of the data and this is modified by the acquisition of data through the likelihood function**
- **The posterior probability from one detector becomes the prior for the next - coherent analysis**

Revision: Bayesian inference

$$p(parameter|data, I) \propto p(data|parameter, I) \times p(parameter|I)$$



- **Inference about the model should be made using the posterior**
- **If we consider a broad, uniform prior, then the posterior is proportional to the likelihood**

What is there to measure?

- **Intrinsic Parameters**

-masses

➡ Chirp Mass $\mathcal{M} = \dfrac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$

-spins

- **Extrinsic Parameters**

-Inclination

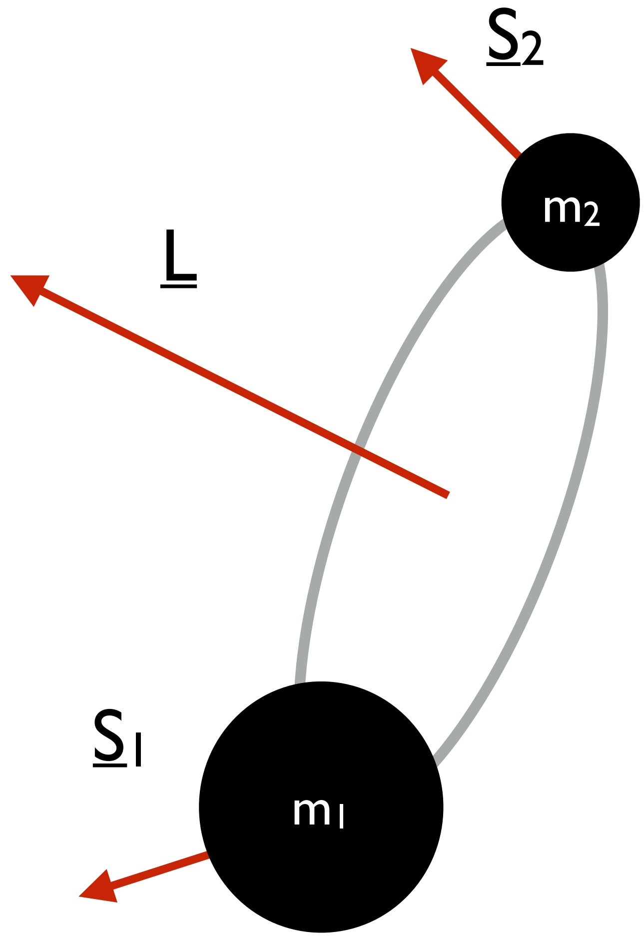
-Orientation

-Polarisation

-Sky position

-luminosity distance

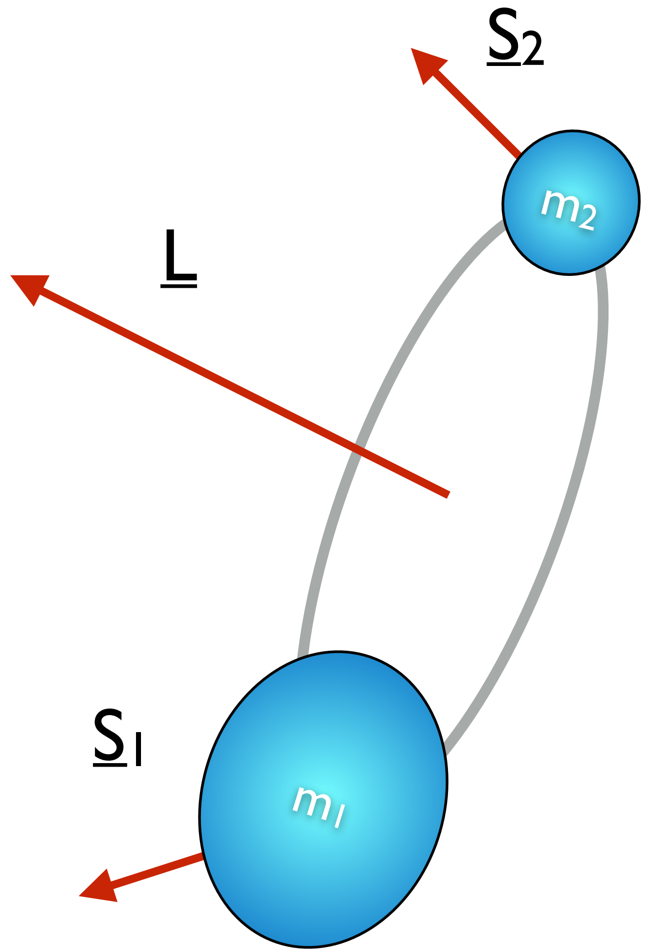
-time



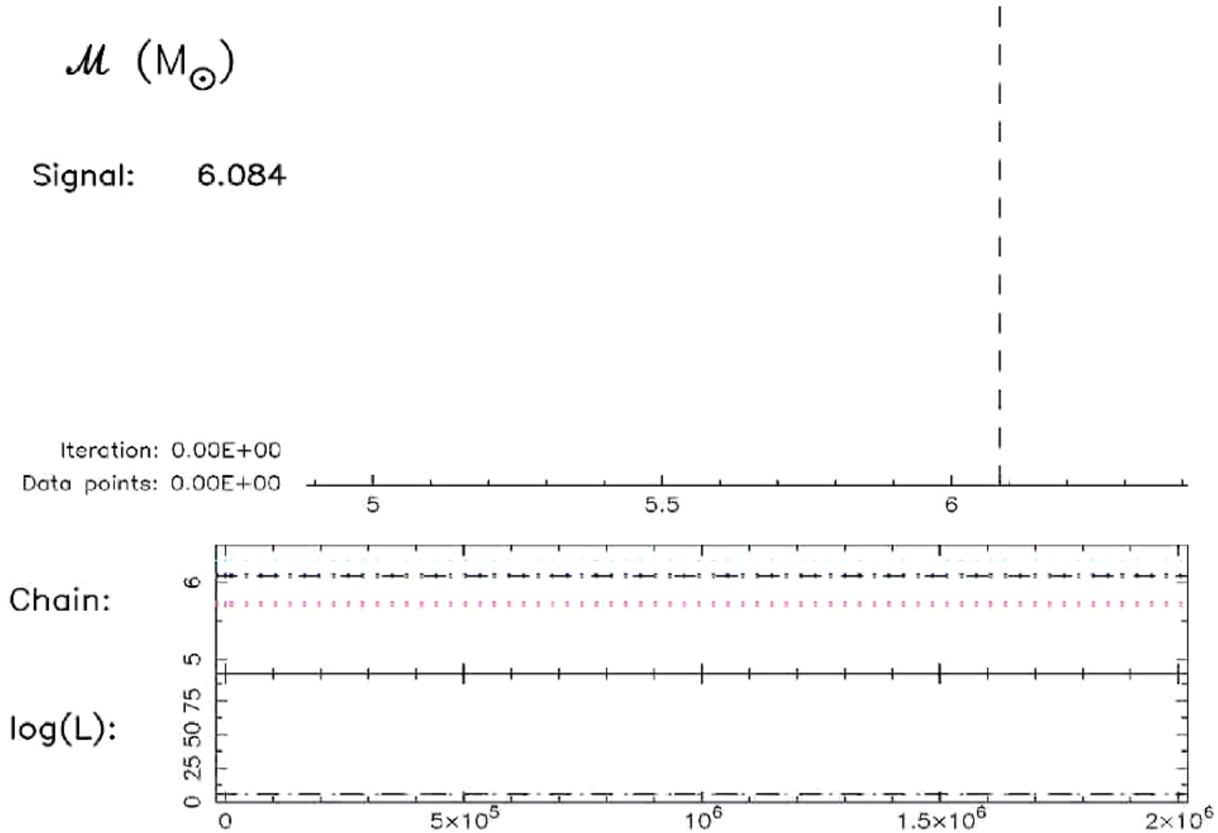
Subtler effects

- **NS Equation of state**
  - tidal deformation

- **Deviations from GR**
- **eccentricity**



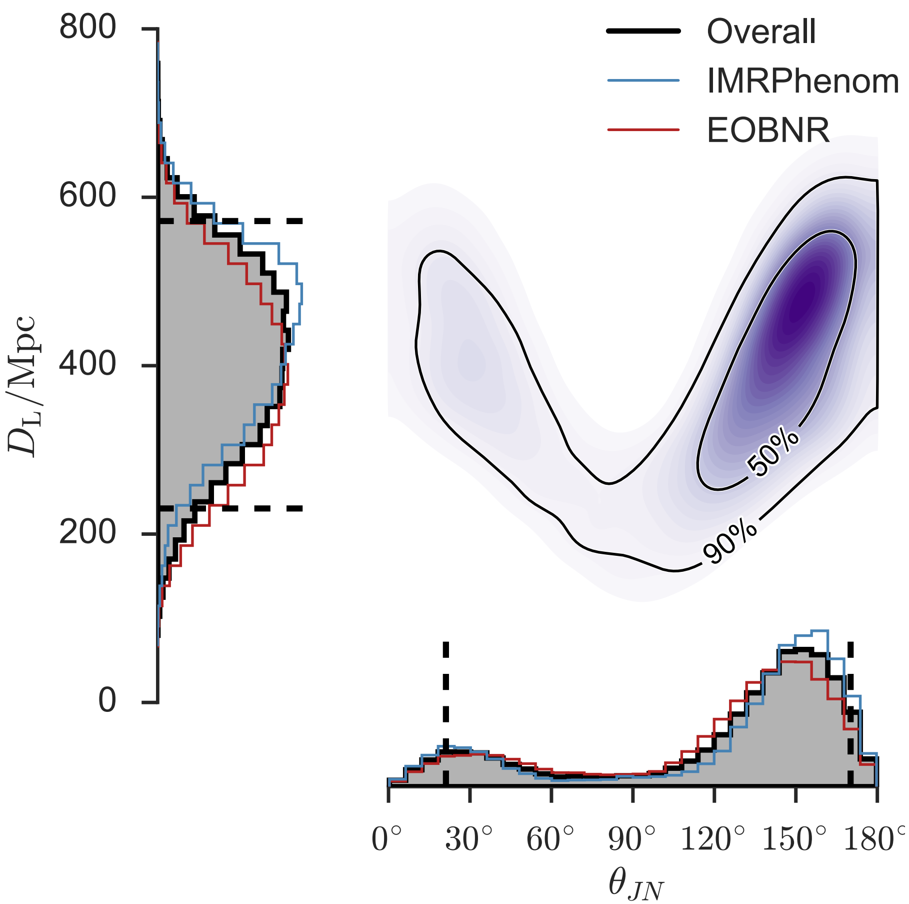
- >15 dimensional grid too large to work with
  - ➡Explore parameter space with random (stochastic) walk
  - ➡MCMC and Nested Sampling algorithms produce *samples* from the posterior distribution



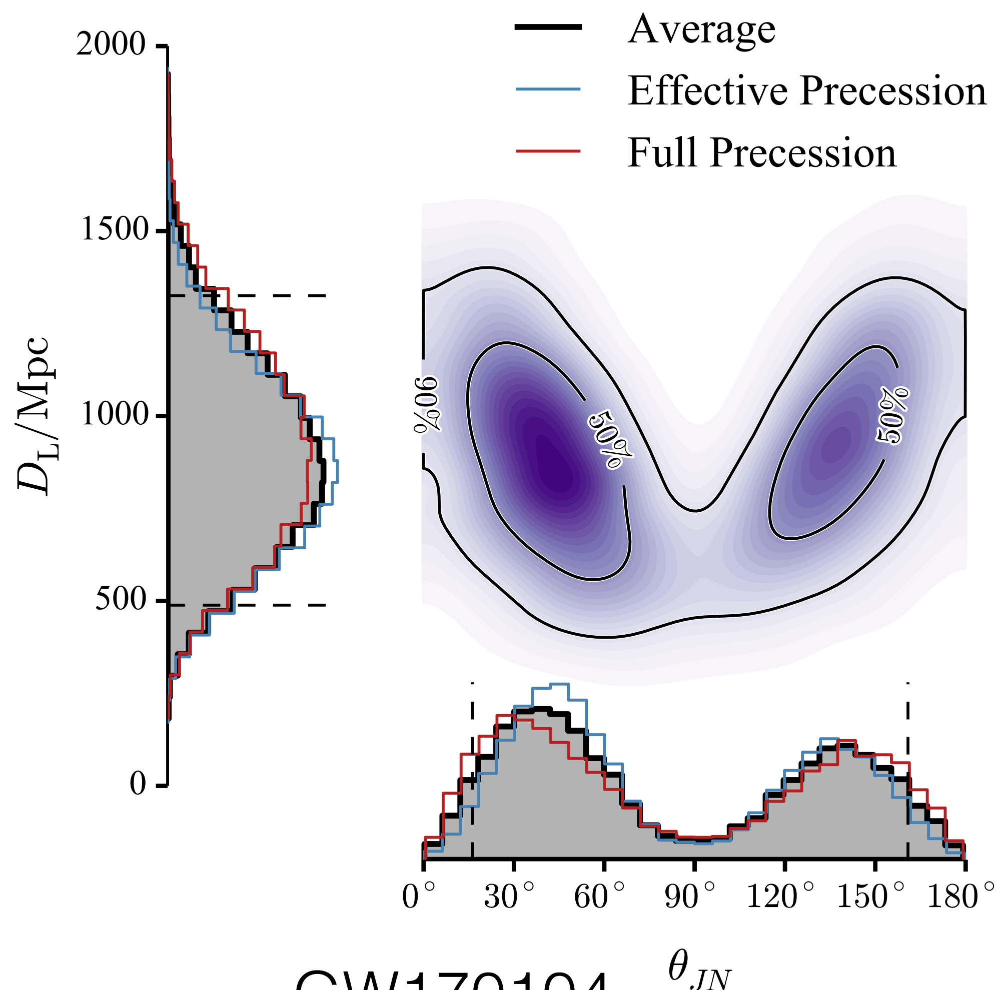
Parameter space degeneracies

- Certain groups of parameters are commonly difficult to distinguish
- We can say they are correlated if there is a bilinear dependence of the signal on the parameters
- For non-linear GW model we can say they are (partially) degenerate.
- Example: distance and inclination have characteristic V-shape joint posterior

Degeneracy - distance/inclination

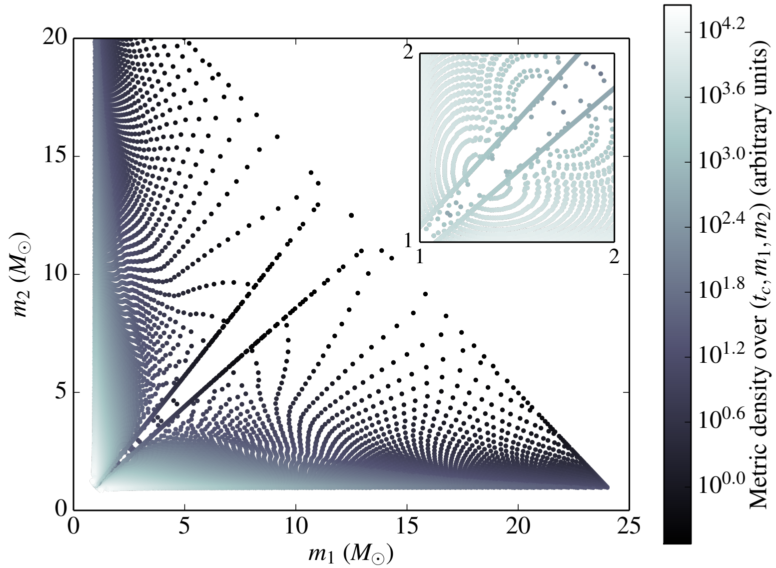


GW150914



GW170104

- If signal is loud, expect errors to be roughly Gaussian, and prior to be insignificant
- A quadratic expansion of log-likelihood about its maximum to determine covariance matrix = inverse of Fisher information matrix



$$\Gamma_{ij} = \left\langle \frac{\partial \hat{h}}{\partial \theta_i} \middle| \frac{\partial \hat{h}}{\partial \theta_j} \right\rangle$$

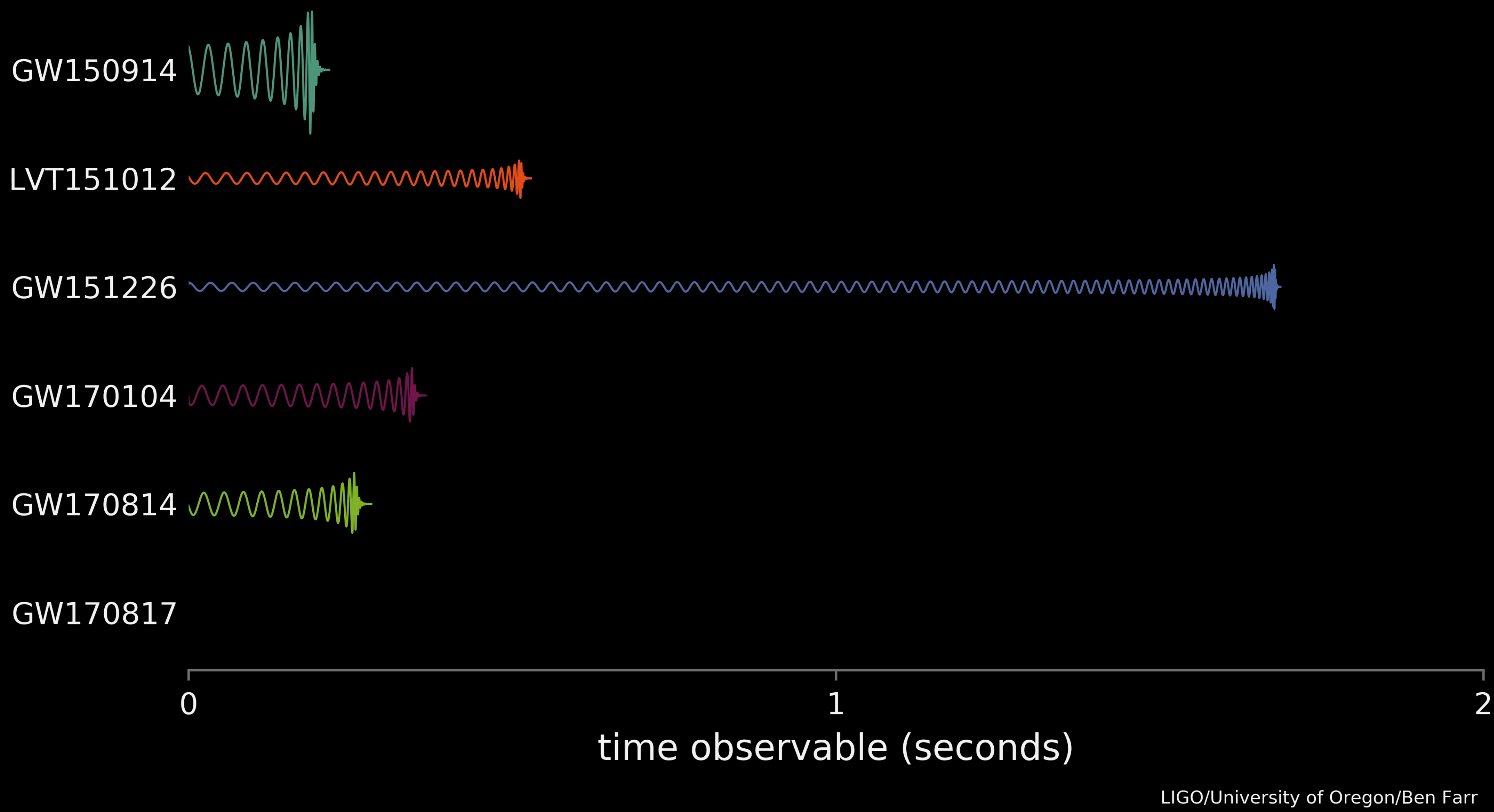
$$\log L(d|A,\vec{\theta}) \approx \frac{1}{2}\left[A^2_{ML}(1 - \Gamma_{ij}(\theta_{ML})\Delta\theta_i\Delta\theta_j) - \Delta A^2 + \ldots\right]$$

note: **in practice MCMC is used!**

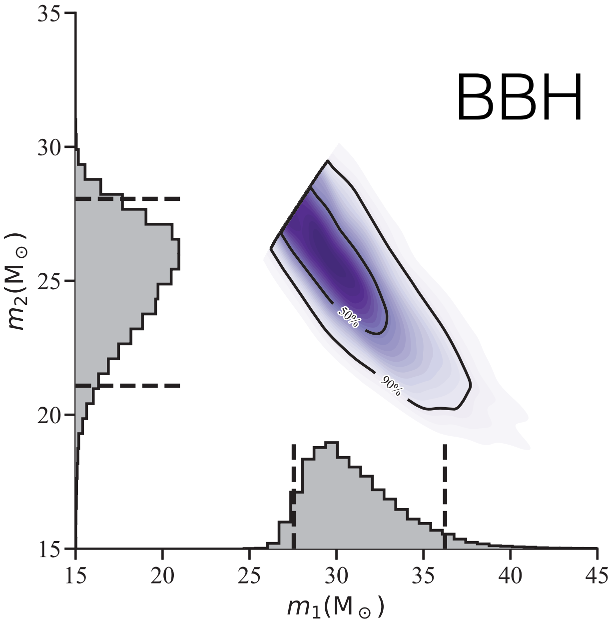
- Example calculation in notes [from Cutler & Flanagan 1994]

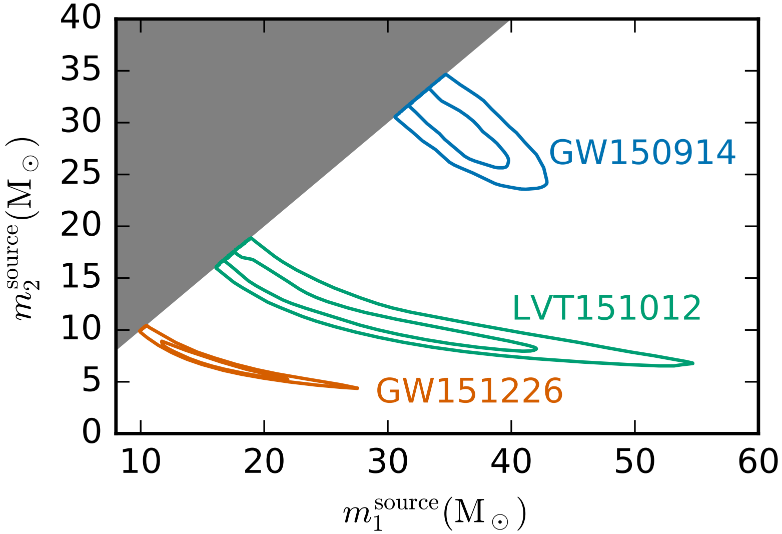
Mass precision

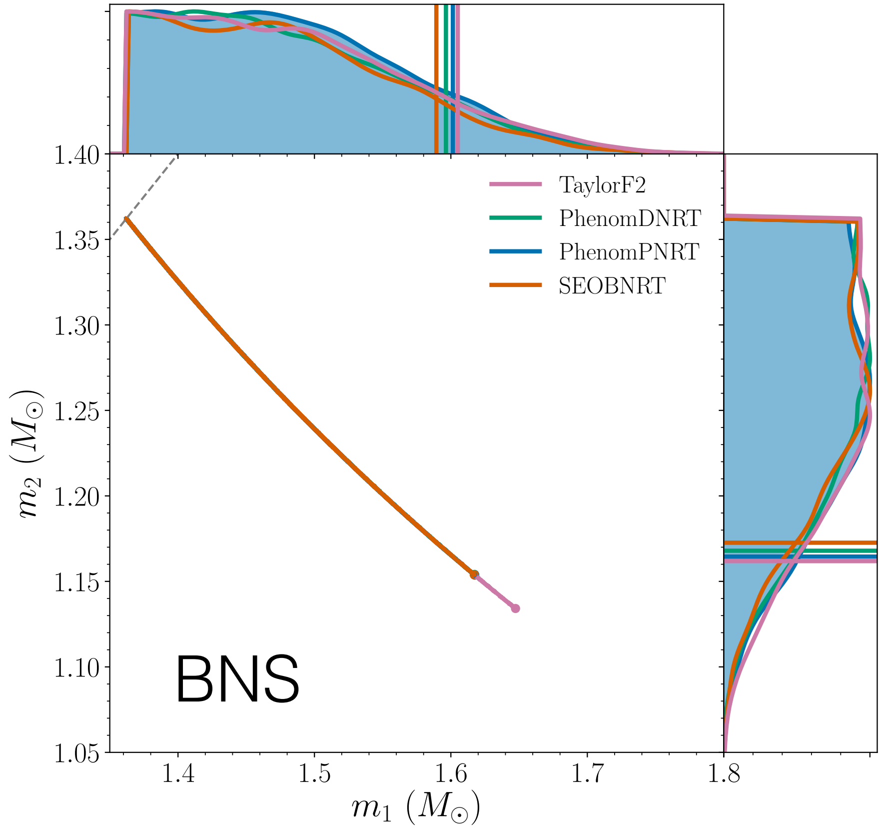
- For a simple (Newtonian) chirp, we have
- $\hat{h} = A \exp i\Phi(f)$, $\quad \Phi(f) = 2\pi f t_c - \phi_0 - \frac{\pi}{4} + \frac{3}{4}(8\pi \mathcal{M} f)^{-5/3}$
- Let us estimate the relative precision of the Chirp Mass $\mathcal{M}$ using Fisher Matrix formalism
- We have $\sigma_{\mathcal{M}}^2 = \Gamma^{-1}_{\mathcal{M}\mathcal{M}} = \left(\frac{\partial \hat{h}}{\mathcal{M}} \middle| \frac{\partial \hat{h}}{\mathcal{M}} \right)^{-1}$
- $\frac{\partial \hat{h}}{\partial \mathcal{M}} = \hat{h} \, . \, i \frac{\partial \Phi}{\partial \mathcal{M}} = -\frac{5i}{4\mathcal{M}} \left(8\pi \mathcal{M} f\right)^{-5/3} \hat{h}$
- So $\frac{\sigma_{\mathcal{M}}}{\mathcal{M}} \propto - \left((8\pi \mathcal{M} f)^{-5/3} \hat{h} \middle| (8\pi \mathcal{M} f)^{-5/3} \hat{h} \right)^{-1/2}$
- Since $\Phi(f) \sim (8\pi \mathcal{M} f)^{-5/3}$ is number of cycles, we have $\frac{\sigma_{\mathcal{M}}}{\mathcal{M}} \propto (\text{number of cycles})^{-1}$


GW150914
LVT151012
GW151226
GW170104
GW170814
GW170817
0
1
2
time observable (seconds)
LIGO/University of Oregon/Ben Farr

Mass accuracy



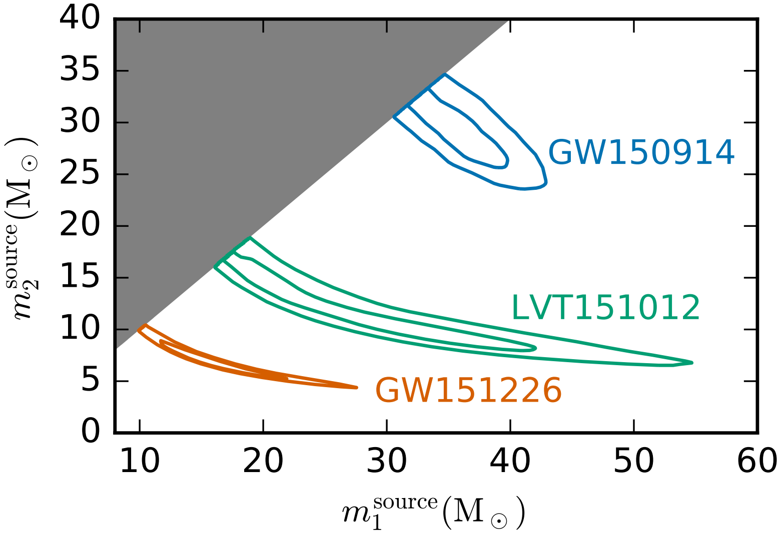


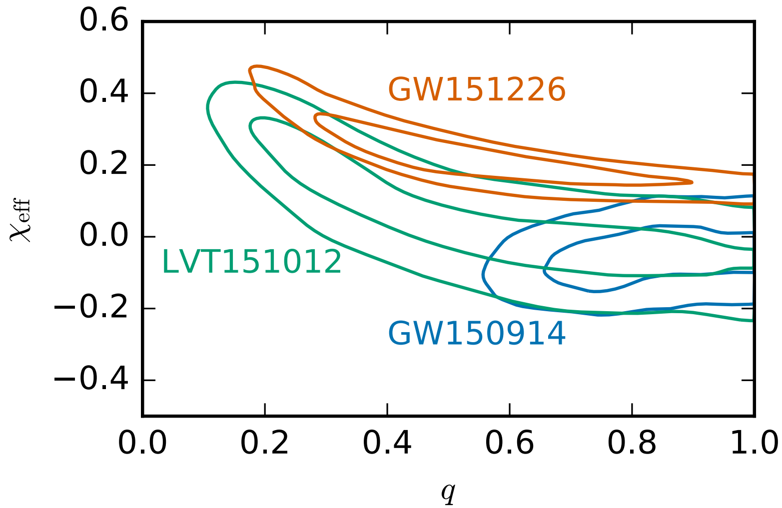


$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{M^{1/5}} \simeq \frac{c^3}{G}\left[\frac{5}{96}\pi^{-8/3} f^{-11/3} \dot{f}\right]^{3/5}$$

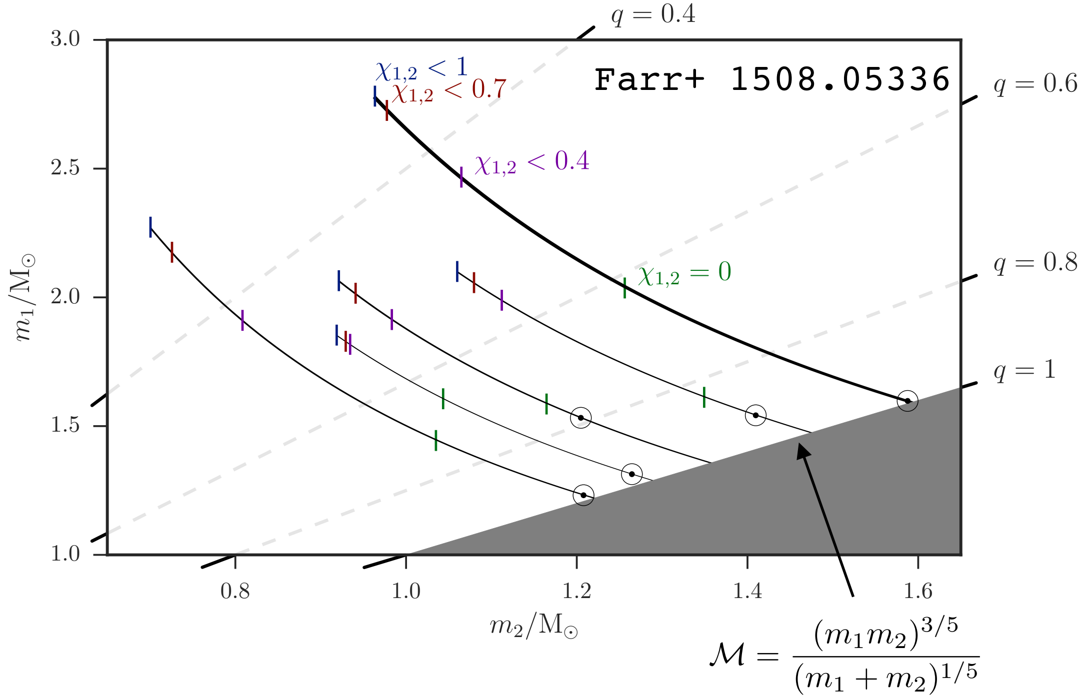
Mass and spin

Binary Black Holes





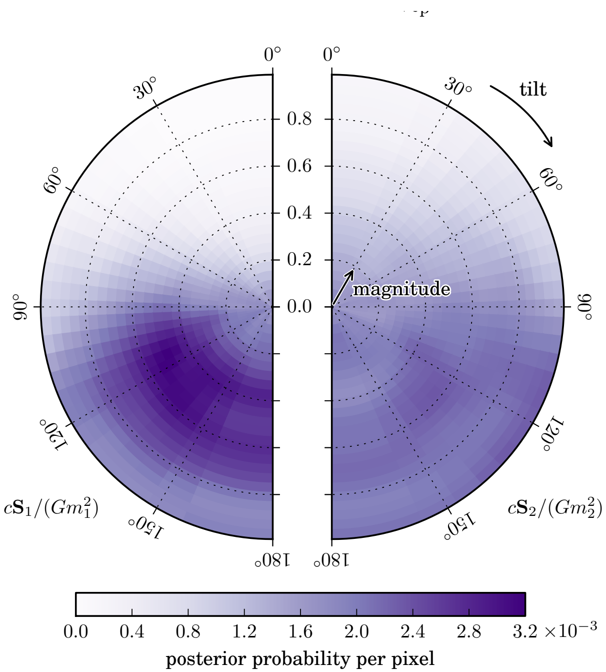
Binary Neutron Stars



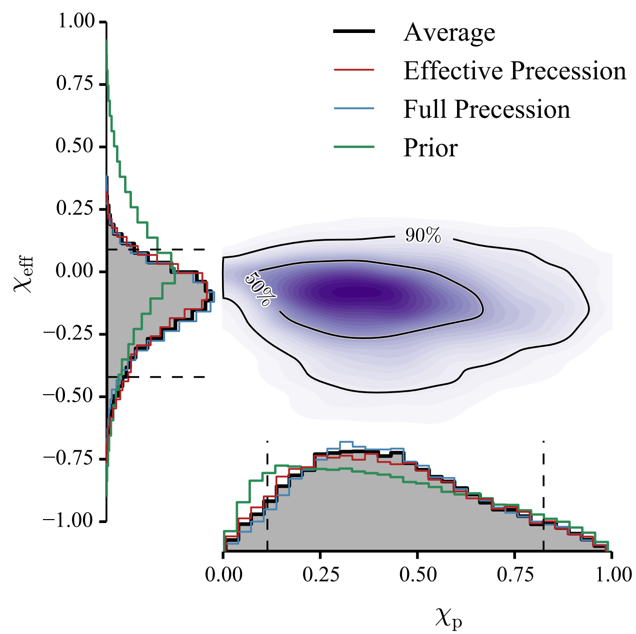
$$\chi_{\rm eff} = \frac{m_1 \chi_{1z} + m_2 \chi_{2z}}{m_1 + m_2}$$

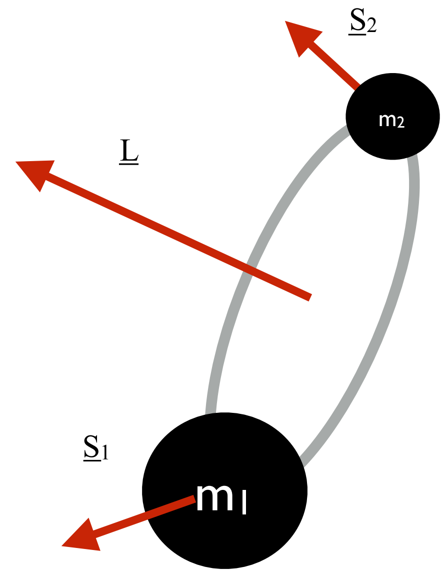
Effective spin enters PN expansion at 1.5PN
correlates with mass ratio

Spin effects



GW170104





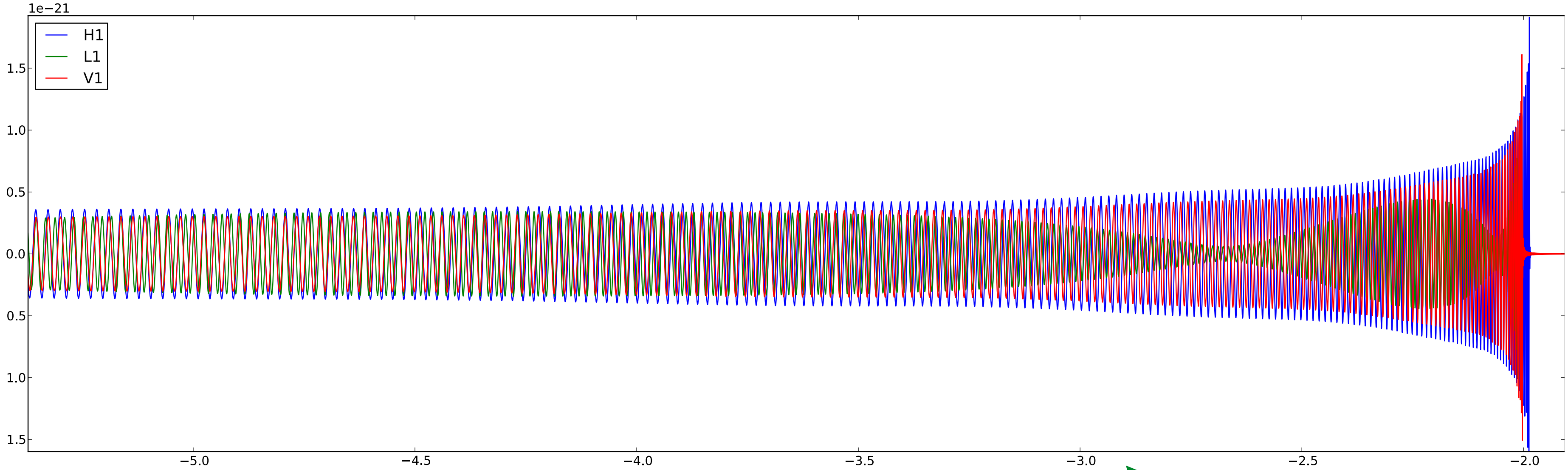
$$\chi_{\mathrm{eff}} = \frac{m_1\chi_{1z} + m_2\chi_{2z}}{m_1 + m_2}$$

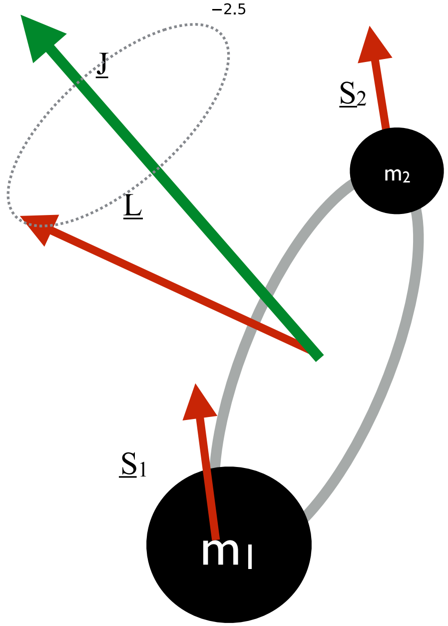
$$\chi_{\mathrm{p}} = \max\left(\chi_{1\perp}, \frac{3+4q}{4+3q} q\,\chi_{2\perp}\right)$$

Precession (spin-spin coupling) enters at 2PN order

- Poorly constrained by the signal at low SNRs

Precessing spins





- **Non-aligned spins contribute to the total angular momentum J differing from L. Spin-orbit interactions cause L to *precess.***
- **Precession of the orbital plane means that the + and x polarisations change over time w.r.t. fixed observer**

-Amplitude and frequency modulation of signal

-Better determination of F_+ and F_x!

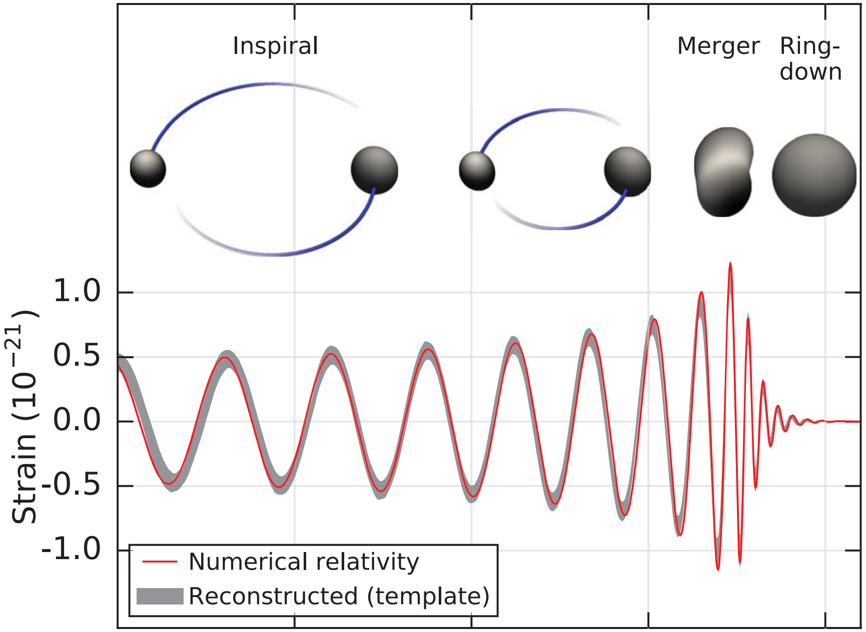
Merger-ringdown

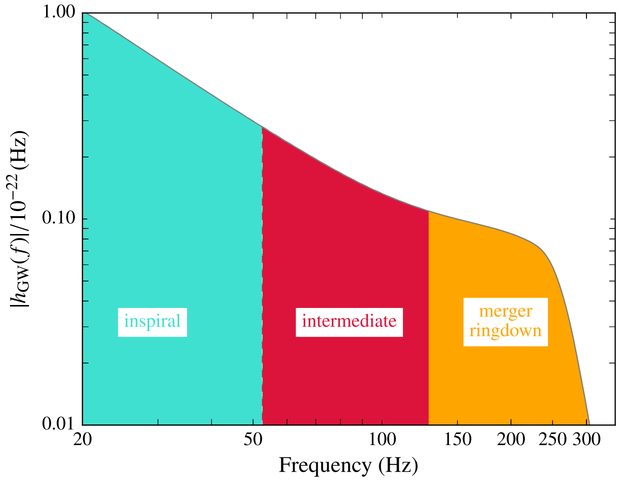
- **Modern waveforms model the entire merger: inspiral-merger-ringdown (IMR)**

-Effective-One-Body + Numerical relativity (EOBNR, SEOBNR)

-Phenomenological (IMRPhenom)

- **How does this affect parameter estimation?**

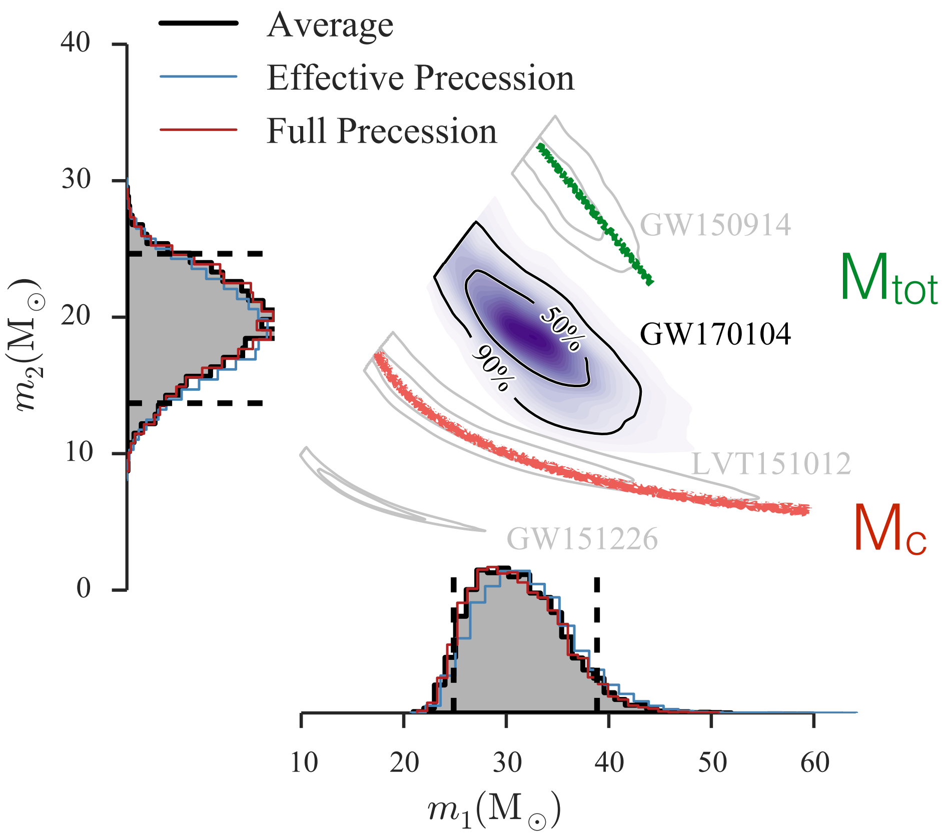




- **Ringdown modes of perturbed BH depends only on the Mass and Spin (no-hair theorem)**
- **Decompose signal into quasi-normal modes with frequency ω_{lm} and damping time τ_{lm}**
- **Spin of final BH dominated by angular momentum from the inspiral**
- **Final mass $M_f = M_1 + M_2 - M_{rad}$ so total mass determines ringdown properties**
- **When ringdown is prominent relative to inspiral, total mass (M_{tot}) not chirp mass (M_c) is relevant parameter**

$$h_+(t) = \frac{M}{D_L} \sum_{l,m>0} A_{l|m|} e^{-t/\tau_{lm}} Y_+^{lm}(\iota) \cos(\omega_{lm} t - m\phi),$$

$$h_\times(t) = -\frac{M}{D_L} \sum_{l,m>0} A_{l|m|} e^{-t/\tau_{lm}} Y_\times^{lm}(\iota) \sin(\omega_{lm} t - m\phi),$$



Summary

•Basic data analysis

- We use probabilistic models to represent uncertainty
- We have seen how detector noise is modeled

•Detection

- We shown the origin of the matched filtering algorithm
- Explained the *coincident* detection method
- Explained ranking statistic and timeslide estimation of false alarm rate

•Parameter Estimation

- We have looked at inferring model parameters
- We have looked at parameters for compact binaries
- We have explained precision and correlation in parameter estimation results.