Announcements

Please turn in Assignment 5.

Lecture today: HW will not be graded. But content from lecture and HW can show up on final exam.

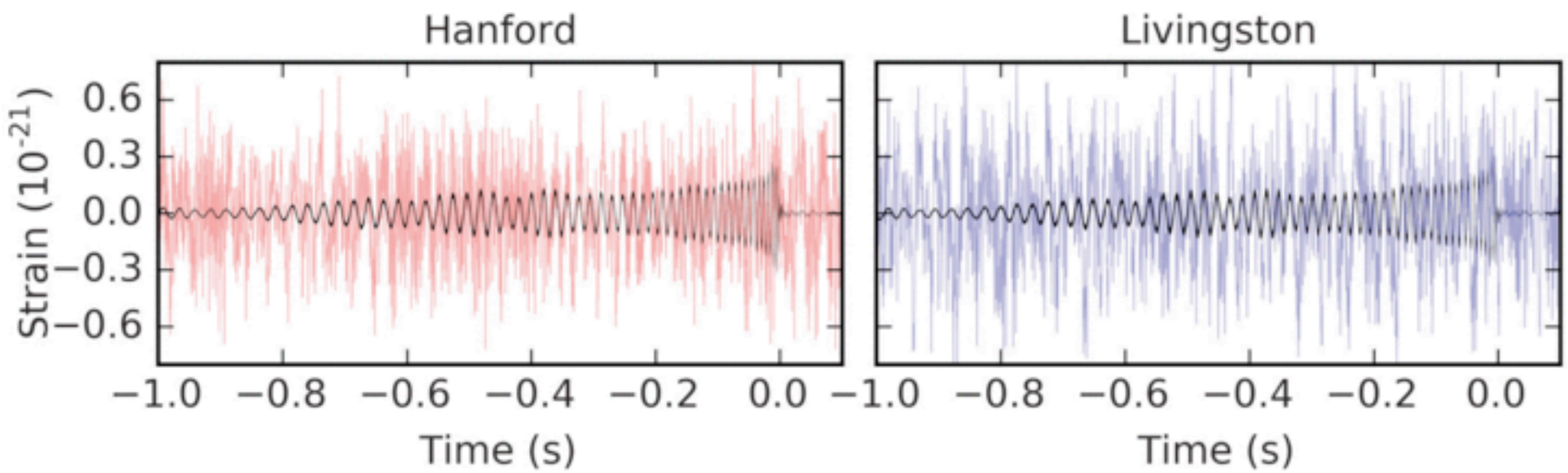
Final Exam: 13:00-16:00 Wednesday May 29, in H331
Note: no make-up of final exam except in cases of emergency or prior arrangement

Visualization Project due by email on May 28

Bayesian analyses for parameter estimation

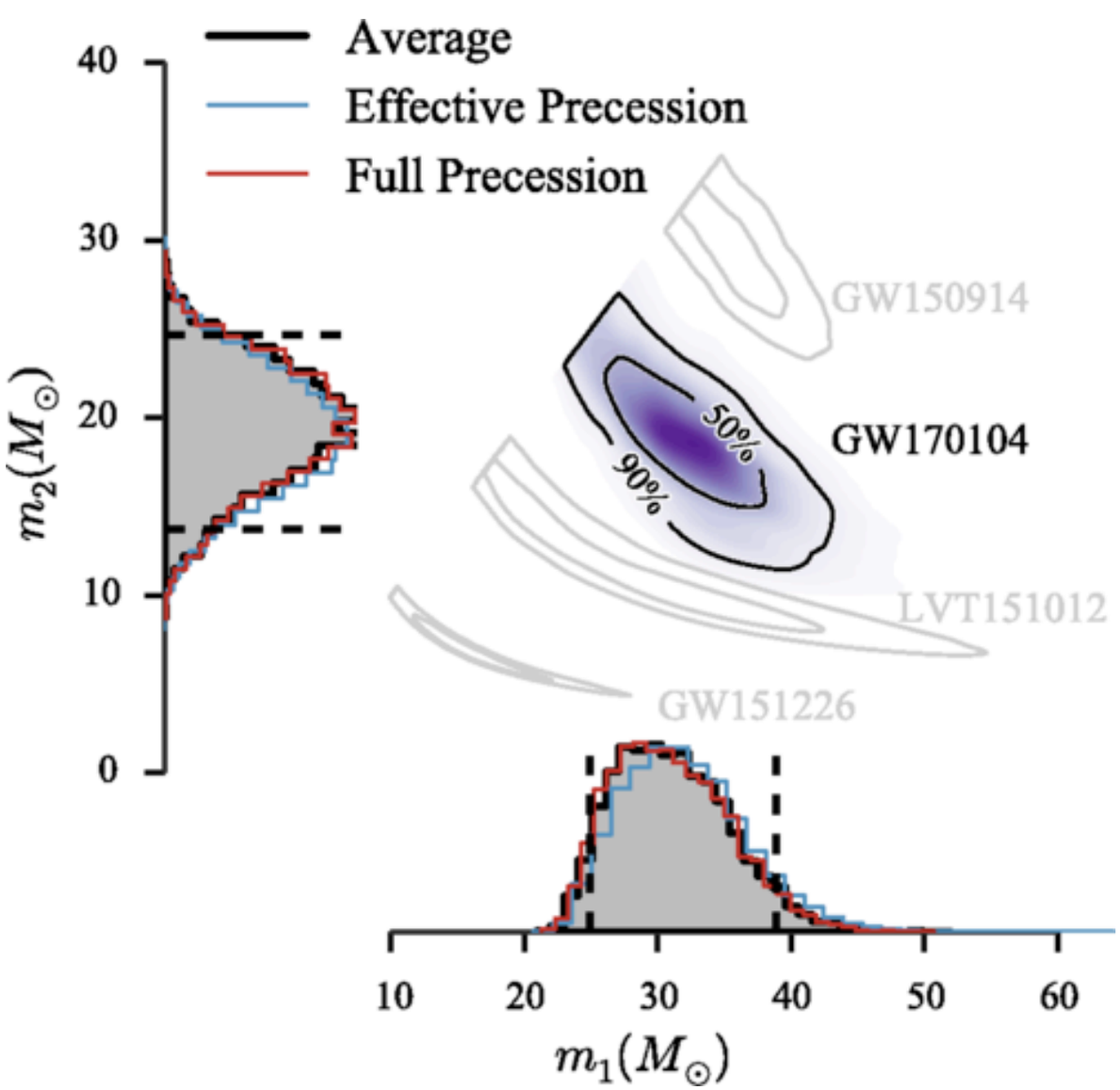
Lecture 5: Gravitational Waves MSc Course

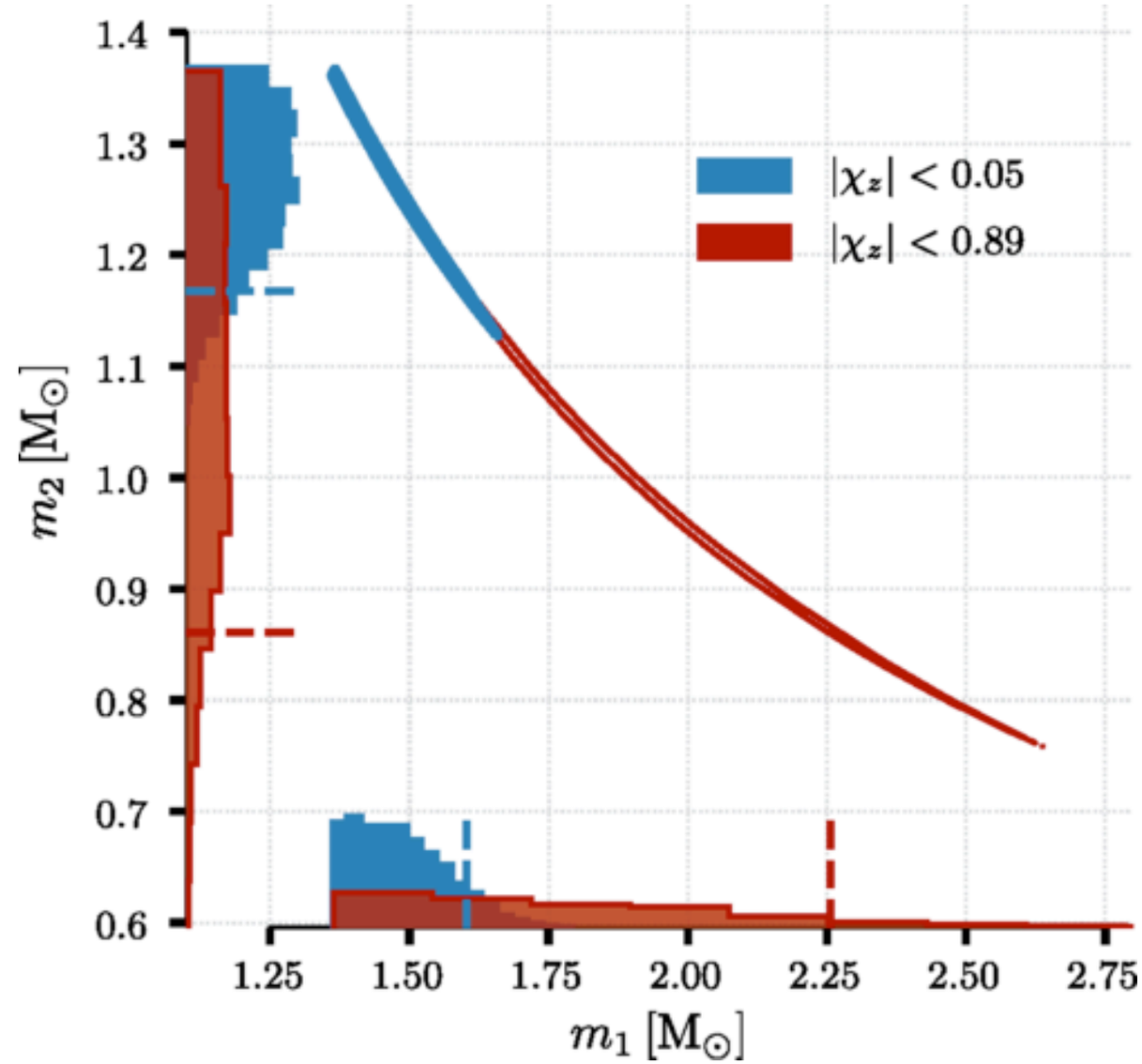
How do we go from detector data...



LVC, PRL 116, 241103 (2016)

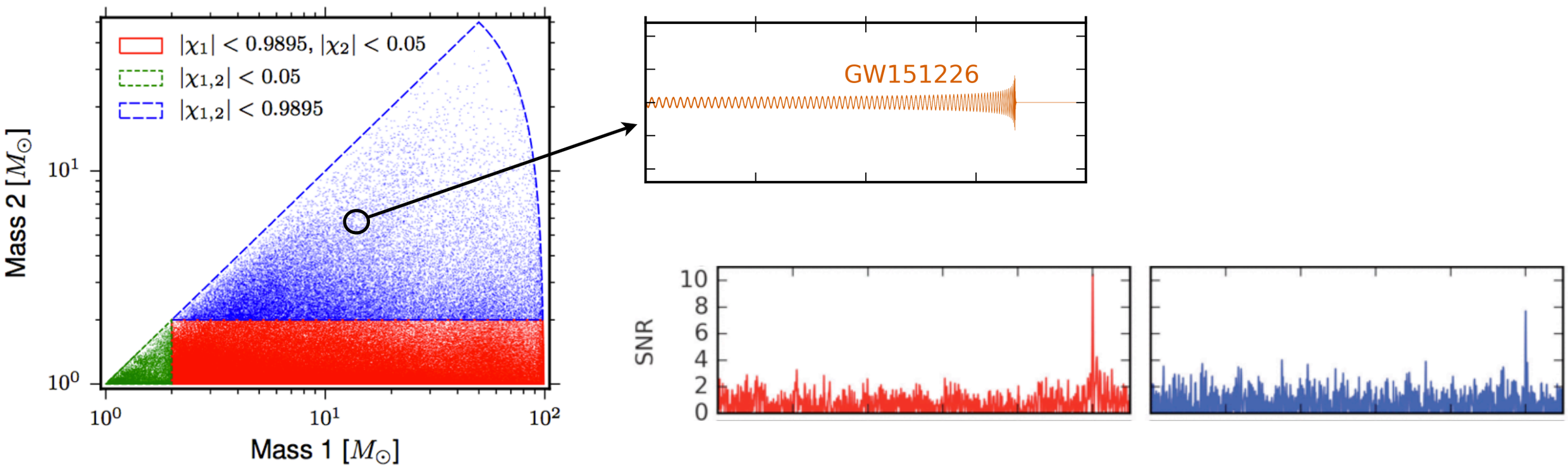
...to astrophysical parameters?





LVC, PRL 118, 221101 (2017)
LVC, PRL 119, 161101 (2017)

We've seen that we can apply the matched filtering technique with many different possible filters in a coarse template bank and extract possible events...



What can we conclude?

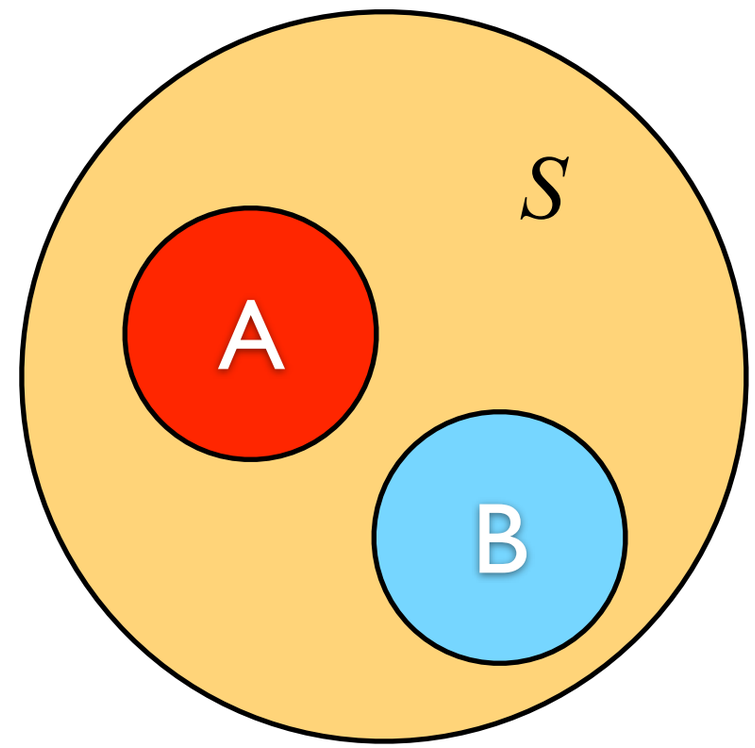
Can we claim detection?

If it is a detection, how can we reconstruct the properties of the source? And with what accuracy?

Probability

Consider set S with subsets A, B, ...

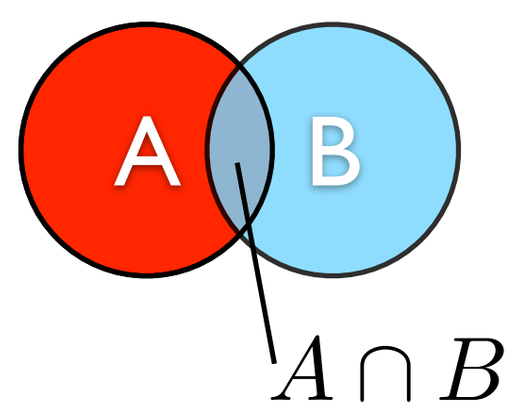
Probability is real-valued function that satisfies:



1. For every A in S, $P(A) \geq 0$.
2. For disjoint subsets ($A \cap B = 0$), $P(A \cup B) = P(A) + P(B)$.
3. $P(S) = 1$.

Conditional probability: probability of A given B

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



Frequentist versus Bayesian interpretation

Frequentist

1. A, B, ... are outcome of repeatable experiment

2. $P(A)$ is frequency of occurrence of A

3. $P(\text{data} \mid \text{hypothesis})$ or $P(\text{data} \mid \text{parameters})$ are probabilities of obtaining some data, given some hypothesis or given value of a parameter

4. Hypotheses are either correct or wrong and parameters have a true value. We do not talk about probabilities of hypotheses or parameters.

Frequentist versus Bayesian interpretation

Bayesian

1. A, B, ... are hypotheses, or theories, or parameters within a theory.

2. $P(A)$ is probability that A is true.

3. $P(\text{data} \mid \text{hypothesis})$ or $P(\text{data} \mid \text{parameters})$ are probabilities of obtaining some data, given some hypothesis or given value of a parameter.

4. Hypotheses and parameters are associated with probability distribution functions.

Bayes' Theorem

Given:

$$P(A \cap B) = P(A|B)P(B)$$

$$P(B \cap A) = P(B|A)P(A)$$

$$A \cap B = B \cap A$$

We can derive **Bayes' Theorem**:

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

A = hypothesis (or parameters or theory)
B = data

P(hypothesis|data) $\propto$ P(data|hypothesis) P(hypothesis)

More on conditional probability

It is customary to explicitly denote probabilities being conditional on “all background information we have”:

$$\mathrm{P}(A \mid I), \mathrm{P}(B \mid I), \ldots$$

All essential formulae are unaffected, for example:

$$P(A, B|I) = P(A|B, I)P(B|I)$$

$$P(A|B, I) = \frac{P(B|A, I)P(A|I)}{P(B|I)}$$

Marginalization

Consider sets B_k such that

- They are disjoint: $B_k \cap B_l = \emptyset$
- They are exhaustive: $\cup_k B_k$ is the Universe, or

$$\sum_k p(B_k|I) = 1$$

Then, $p(A|I) = \sum_k p(A, B_k|I)$

Marginalization Rule

Marginalization over continuous variable

Consider the proposition, "The continuous variable x has the value α."

Not a well-defined meaning of probability: $p(x = \alpha | I)$

Instead assign probabilities to finite intervals:

$$p(x_1 \leq x \leq x_2 | I) = \int_{x_1}^{x_2} \mathrm{pdf}(x) dx$$

where $\mathrm{pdf}()$ is the probability density function.

$$\int_{x_{\mathrm{min}}}^{x_{\mathrm{max}}} \mathrm{pdf}(x) dx = 1$$

Marginalization for continuous variables: $$p(A) = \int_{x_{\mathrm{min}}}^{x_{\mathrm{max}}} \mathrm{pdf}(A, x) dx$$

More on Bayes' Theorem

Initial Understanding + New Observation = Updated Understanding

Likelihood function

Prior probability

Posterior probability

$$p\left(h'|d\right) = \frac{p\left(d|h'\right)p\left(h'\right)}{p\left(d\right)}$$

Evidence

More on Bayes' Theorem

An experiment is performed, data d is collected.

We are measuring parameter θ.

Consider a model H that allows us to calculate the probability of getting data d if parameter θ is known.

Posterior probability of θ:

$$p(\theta|d, H, I) = \frac{p(d|\theta, H, I)p(\theta|H, I)}{p(d|H, I)}$$

The evidence doesn't depend on θ so ignore for now:

$$p(\theta|d, H, I) \propto p(d|\theta, H, I)p(\theta|H, I)$$

More parameters

Can extend to more parameters: **joint posterior**

$$p(\theta_1, \ldots, \theta_N | d, H, I)$$

If we want posterior distribution just for variable θ_1,

$$p(\theta_1 | d, H, I)$$

then we marginalize

$$p(\theta_1 | d, H, I) = \int_{\theta_2^{\min}}^{\theta_2^{\max}} \ldots \int_{\theta_N^{\min}}^{\theta_N^{\max}} p(\theta_1, \ldots, \theta_N | d, H, I) d\theta_2 \ldots d\theta_N$$

The likelihood function

Likelihood function

$$p\left(h'|d\right) = \frac{p\left(d|h'\right)p\left(h'\right)}{p\left(d\right)}$$

Probability of data given hypothesis - a true “frequentist” probability

In GW science, the likelihood function is the noise model.

The likelihood function: the data

If the detector noise is stationary and Gaussian:

$$\langle \tilde{n}^*(f)\tilde{n}(f')\rangle = \delta(f-f')\frac{1}{2}S_n(f)$$

Gaussian probability distribution for noise:

$$p(n_0) = \mathcal{N}\exp\left\{-\frac{1}{2}\int_{-\infty}^{\infty} df\,\frac{|\tilde{n}_0(f)|^2}{(1/2)\,S_n(f)}\right\}$$

$$= \mathcal{N}\exp\left\{-\frac{(n_0|n_0)}{2}\right\}$$

Output of detector: $s(t) = h(t;\theta_t) + n_0(t)$, $n_0 = s - h(\theta_t)$

Plug into $p(n_0)$ to get: $\Lambda(s|\theta_t) = \mathcal{N}\exp\left\{-\frac{1}{2}(s-h(\theta_t)|s-h(\theta_t))\right\}$

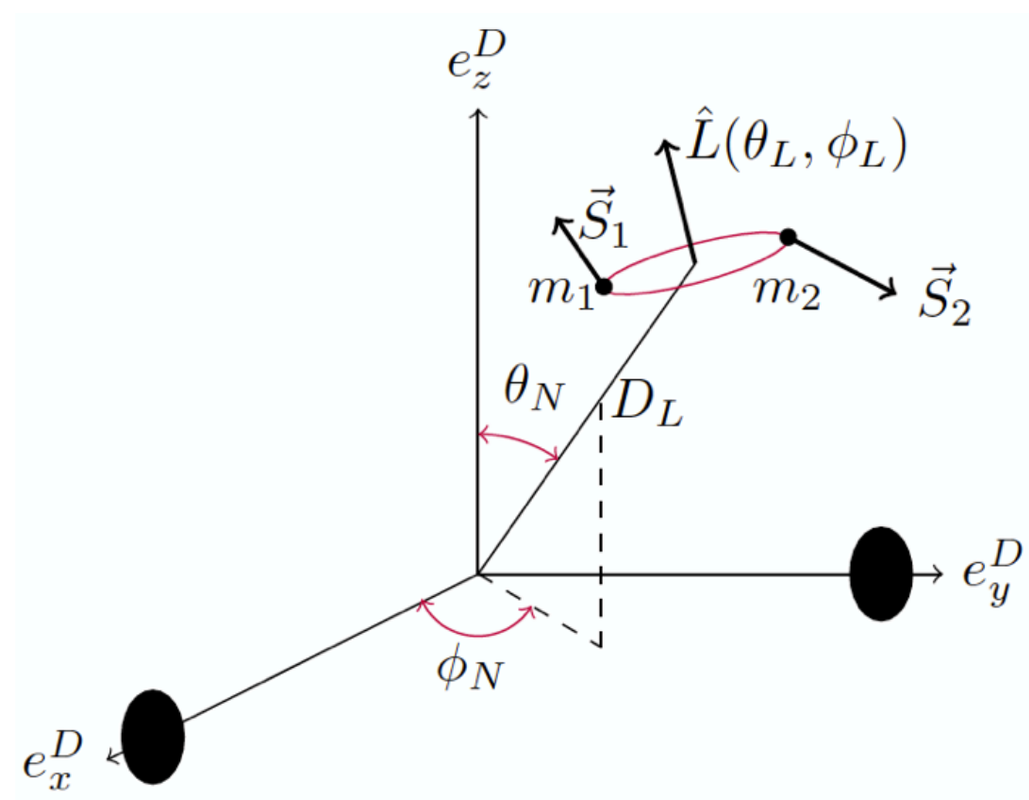
The likelihood function: the data

$$h_t \equiv h(\theta_t)$$

$$\Lambda(s|\theta_t) = \mathcal{N}\exp\left\{(h_t|s) - \frac{1}{2}(h_t|h_t) - \frac{1}{2}(s|s)\right\}$$

In this form, information might not be very manageable.

For binary coalescence there could be more than 15 parameters θ^i



The prior probability

Prior probability

$$p\left(h'|d\right) = \frac{p\left(d|h'\right)p\left(h'\right)}{p\left(d\right)}$$

Probability of hypothesis; makes no sense in frequentist interpretation.

But for a Bayesian, one can make assumptions to include a prior; can be subjective.

Thus, prior choices can influence results.

Can be seen as the “degree of belief” that the hypothesis is true **before** a measurement is made.

The prior probability

$$p^{(0)}(\theta_t)$$

Examples in GW science:

* Known distributions in space

$p^{(0)}(r)dr \sim r^2 dr$ for isotropic sources

$p^{(0)}(r)dr \sim r dr$ for sources in the Galaxy

* Known mass distribution of neutron stars ~1.35 $M_\odot$

The posterior probability

Posterior probability

$$p\left(h'|d\right) = \frac{p\left(d|h'\right)p\left(h'\right)}{p\left(d\right)}$$

Can be seen as the “degree of belief” that the hypothesis is true **after** a measurement is made.

$$p(\theta_t|s) = \mathcal{N}p^{(0)}(\theta_t)\exp\left\{(h_t|s) - \frac{1}{2}(h_t|h_t)\right\}$$

The evidence

$$p\left(h'|d\right) = \frac{p\left(d|h'\right)p\left(h'\right)}{p\left(d\right)}$$

Evidence

The evidence is unimportant for parameter estimation (but not model selection).

It is basically a normalization factor for parameter estimation.

Notice that it doesn't depend on the parameter being measured.

The evidence: model selection

$$p(h'|d, M) = \frac{p(d|h', M)p(h'|M)}{p(d|M)}$$

M: any overall assumption or model (e.g. the signal is a GW, the binary black hole is spin-precessing, the binary components are neutron stars)

Odds Ratio: Compare competing models, for example "GW170817 was a BNS" vs "GW170817 was a BBH":

$$\mathcal{O}_{ij} = \frac{p(M_i|d)}{p(M_j|d)}$$

$$= \frac{p(M_i)p(d|M_i)}{p(M_j)p(d|M_j)}$$

What is the most probable value of the parameters, θ_t ?

A rule for assigning the most probable value is called an estimator. Choices of estimators include:

1. Maximum likelihood estimator

2. Maximum posterior probability

3. Bayes estimator

1. Maximum likelihood estimator

Define $\hat{\theta}$ as value which maximizes probability distribution:

$$p(\theta_t|s) = \mathcal{N} p^{(0)}(\theta_t) \exp\left\{ (h_t|s) - \frac{1}{2}(h_t|h_t) \right\}$$

Let prior be flat. Then problem is to maximize the likelihood $\Lambda(s|\theta_t)$

Generally simpler to maximize $\log \Lambda$.

$$\log \Lambda(s|\theta_t) = (h_t|s) - \frac{1}{2}(h_t|h_t)$$

$$\frac{\partial}{\partial \theta_t^i} \left[(h_t|s) - \frac{1}{2}(h_t|h_t) \right] = 0$$

2. Maximum posterior probability

Allows us to include prior information.

Then we maximize the full posterior probability:

$$p(\theta_t|s) = \mathcal{N} p^{(0)}(\theta_t) \exp\left\{ (h_t|s) - \frac{1}{2}(h_t|h_t) \right\}$$

Non-trivial priors can lead to conceptual issues.
For example, if $(\bar{\theta}_1, \bar{\theta}_2)$ is the maximum of the distribution function

$$p(\theta_1, \theta_2|s),$$

it is no longer true that $\bar{\theta}_1$ is the maximum of the reduced distribution function

$$\tilde{p}(\theta_1|s) = \int d\theta_2 \; p(\theta_1, \theta_2|s),$$ integrating out θ_2

3. Bayes' Estimator

Neither 1) nor 2) minimizes the error on the parameter estimation.

Most probable values of parameters defined by

$$\hat{\theta}^i_B(s) \equiv \int d\theta \; \theta^i \, p(\theta|s)$$

Errors on parameters defined by matrix:

$$\Sigma^{ij}_B = \int d\theta \left[\theta^i - \hat{\theta}^i_B(s)\right] \left[\theta^j - \hat{\theta}^j_B(s)\right] p(\theta|s)$$

Independent of whether we integrate out a variable, minimizes parameter estimation error, but has a high computational cost.

Confidence versus Credibility

Frequentist relies on confidence interval (CI).

Bayesian approach relies on a credible region (CR)

For example, consider an experimental apparatus that provides values distributed as Gaussian around true value x_t with standard deviation sigma σ

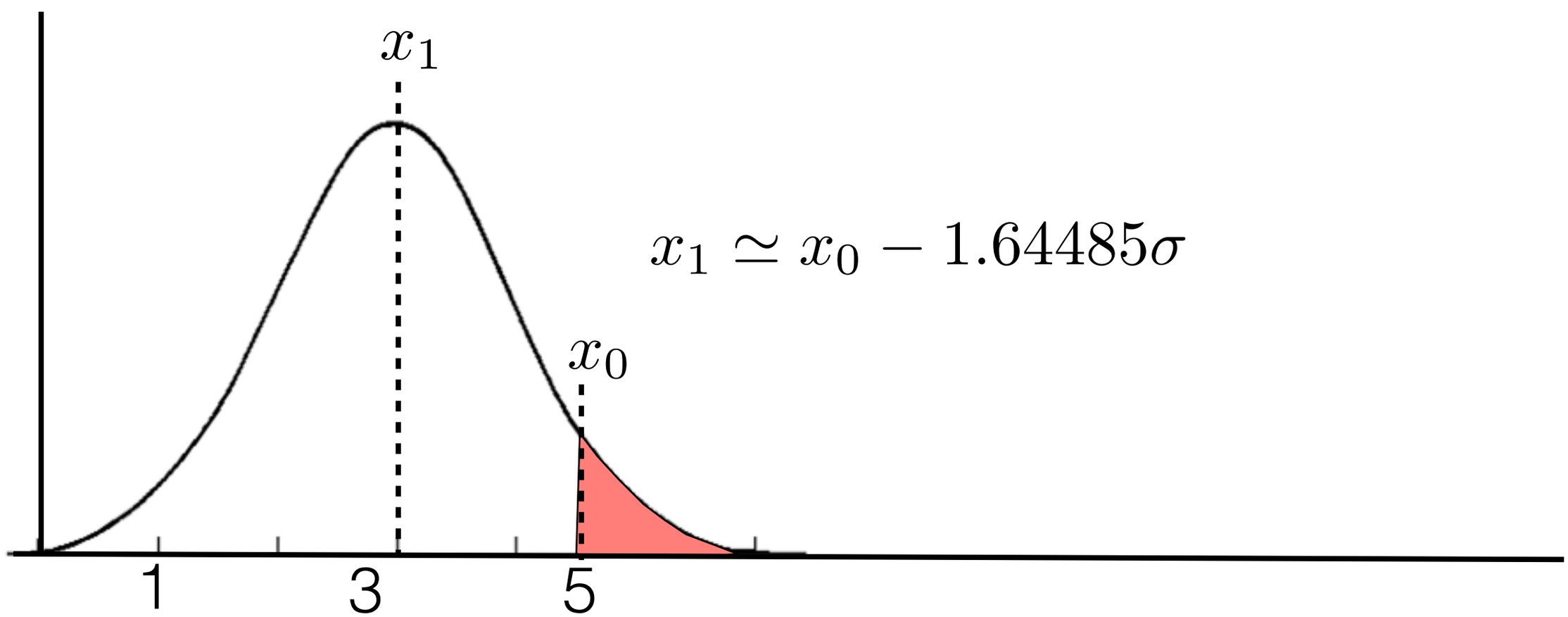
$$P(x|x_t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{ -\frac{(x - x_t)^2}{2\sigma^2} \right\}$$

One repetition of experiment yields value $x_0 = 5$.

Frequentist confidence interval

Use Neyman's construction for 90% confidence level.

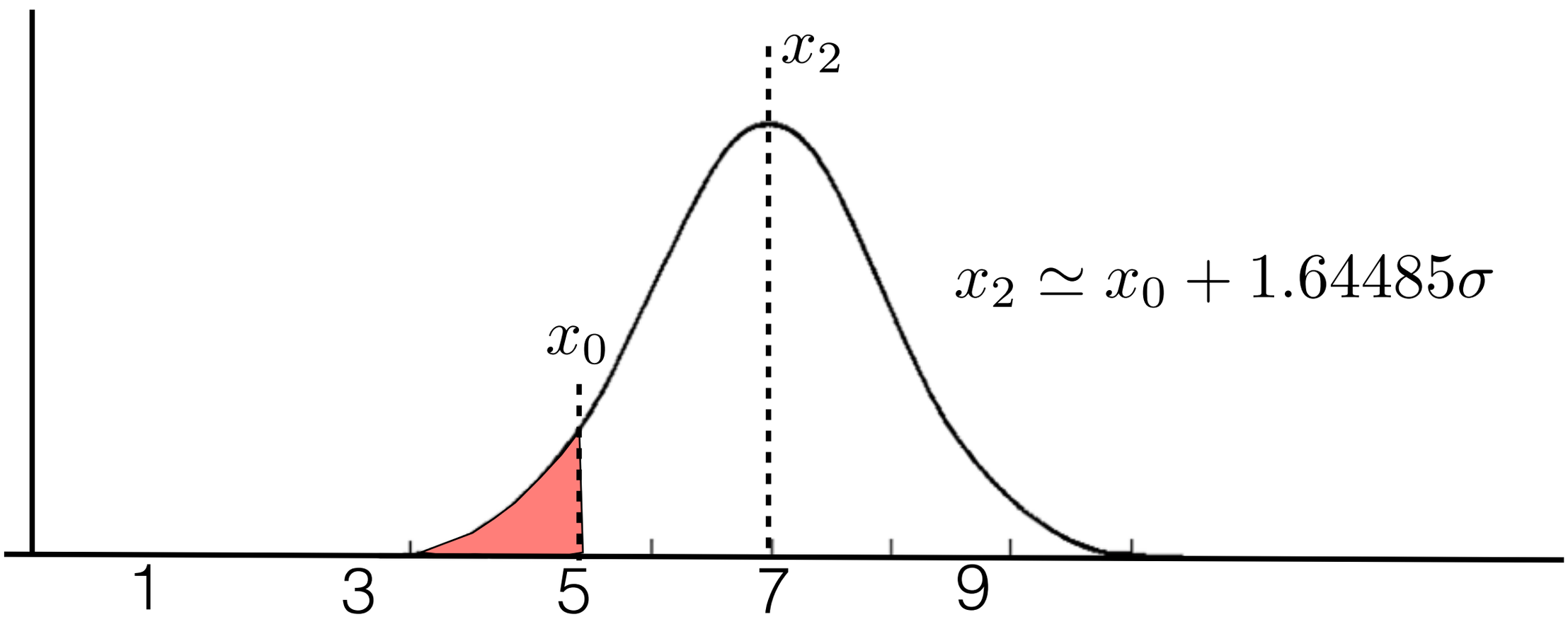
1. Find value $x_1 < x_0$ such that 5% of area under $P(x|x_1)$ is at $x > x_0$.



Frequentist confidence interval

Use Neyman's construction for 90% confidence level.

1. Find value $x_2 > x_0$ such that 5% of area under $P(x|x_2)$ is at $x < x_0$.



Bayesian credible region

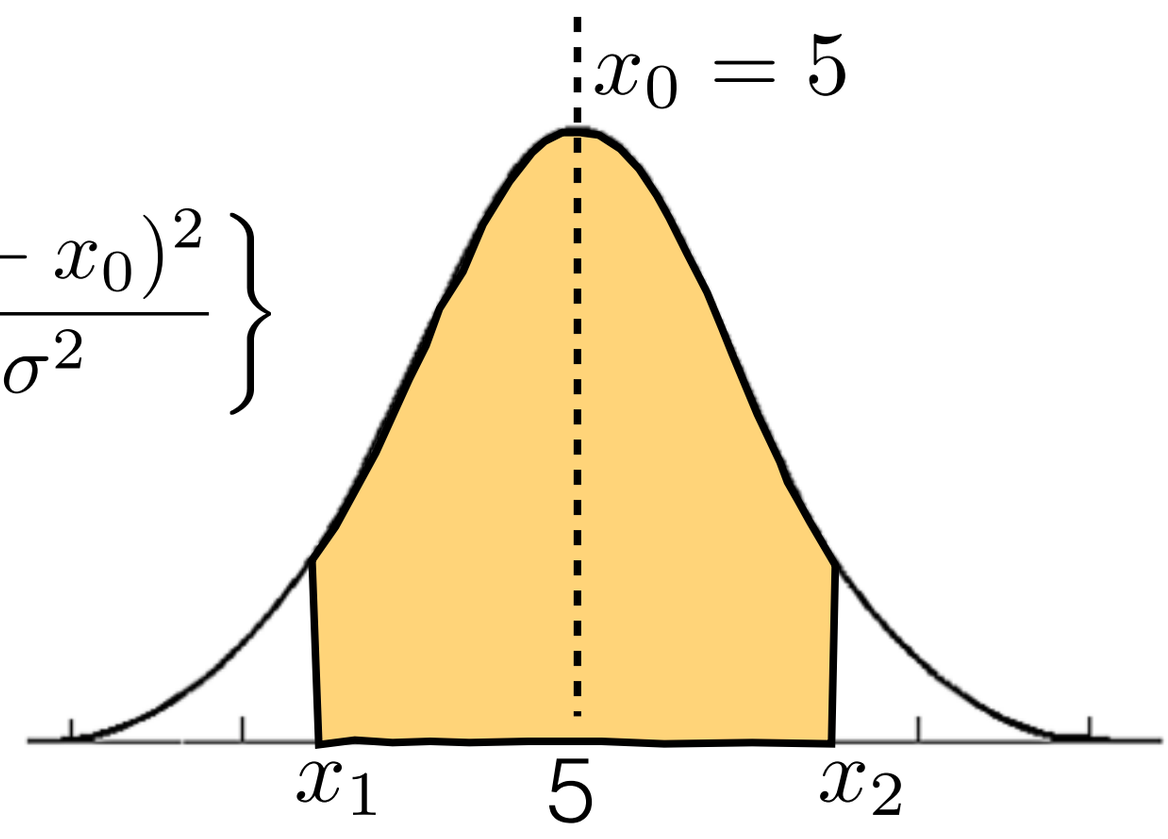
Bayesian approach: construct probability distribution for true value x_t from the likelihood function P(data|hypothesis)

$$\Lambda(x_0|x_t) = P(x_0|x_t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x_0 - x_t)^2}{2\sigma^2}\right\}$$

P(hypothesis|data) $\propto$ P(data|hypothesis) P(hypothesis)

Flat prior so P(hypothesis|data) $\propto$ P(data|hypothesis)

$$P(x_t|x_0) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x_t - x_0)^2}{2\sigma^2}\right\}$$



Confidence versus Credibility

For Gaussian distributions, the Frequentist and Bayesian definitions give the same result for x_1 and x_2 but the interpretation is different.

Frequentist: In the limit of a large number of repetitions, 90% of the confidence regions obtained by the different repetitions of the experiment will cover the true value of x_t.

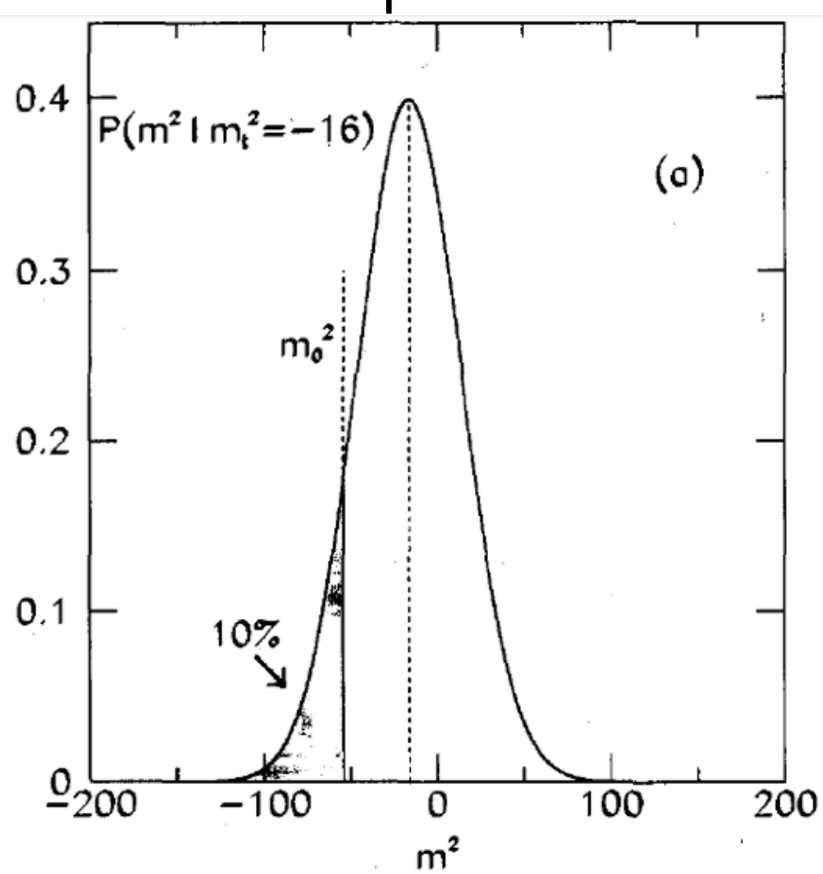
Bayesian: 90% confidence interval is interval which subtends an area equal to 90% the total area of the p.d.f. of the true value x_t.

Confidence versus Credibility

Consider variable with bounded domain like a mass or rate. We can accommodate the physical constraint with a prior.

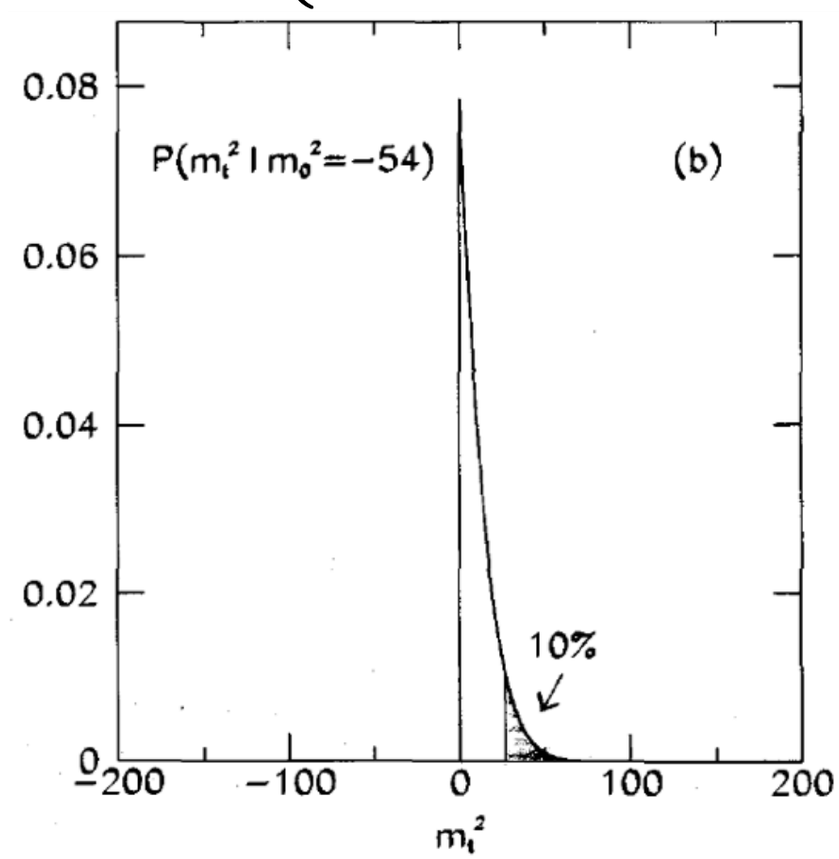
Example: square of mass of electron neutrino

No prior



$$m^2 = (-54 \pm 30)\mathrm{eV}^2$$

$$P(m^2) = \begin{cases} 0 & m^2 < 0 \\ \text{uniform} & m^2 \geq 0 \end{cases}$$



$$m^2 < 26.6\mathrm{eV}^2$$

FD Cousins (1995)

Matched filtering statistics

Hypothesis: GW signal $h(t;\theta)$ is present.

We've found the most probable value of parameters θ

Question: What is the statistical significance of an event found at a given level of signal-to-noise ratio?

Matched filtering statistics

In any detector, we have two kinds of noise:

1. Well-behaved Gaussian noise

$\sim e^{-x^2/2}$ drops very fast for large values of argument x.

Can eliminate Gaussian noise by setting a large threshold for signal-to-noise ratio.

Matched filtering statistics

In any detector, we have two kinds of noise:

2. Non-Gaussian noise

Typically characterized by long tails at large values of signal-to-noise ratio, decays as power law.

Cannot be eliminated with large threshold.

Best way to eliminate is to require coincidence.

Matched filtering statistics

Question: What is the statistical significance of obtaining a given signal-to-noise ratio assuming only Gaussian noise is present.

$$\rho = \frac{\hat{s}}{N} \qquad\qquad S/N = \langle \rho \rangle$$

$$\hat{s} = \int_{-\infty}^{\infty} dt \, [h(t) + n(t)] \, K(t)$$

In **absence of GW signal**, the probability distribution of ρ is

$$p(\rho | h = 0) d\rho = \frac{1}{\sqrt{2\pi}} e^{-\rho^2/2} d\rho$$

Matched filtering statistics

If there is a GW with signal-to-noise ratio $\bar{\rho}$ in the output, then the full signal-to-noise ratio will be

$$\rho = \bar{\rho} + \hat{n}/N$$

$\rho - \bar{\rho} = \hat{n}/N$ is Gaussian variable with zero mean and unit variance

$$p(\rho|\bar{\rho})d\rho = \frac{1}{\sqrt{2\pi}} e^{-(\rho-\bar{\rho})^2/2} d\rho$$

Let $R \equiv \rho^2$.

$$P(R|\bar{R})dR = p(\rho|\bar{\rho})d\rho + p(-\rho|\bar{\rho})d\rho$$

$$P(R|\bar{R})dR = \frac{1}{\sqrt{2\pi R}} e^{-(\bar{R}+R)/2} \cosh\left[\sqrt{R\bar{R}}\right] dR$$

Matched filtering statistics

False Alarm Probability

$$p_{\mathrm{FA}} = \int_{R_t}^{\infty} dR \; P(R|\bar{R} = 0)$$
$$= 2\,\mathrm{erfc}(\rho_t/\sqrt{2})$$

False Dismissal Probability - probability of losing a real GW signal.

$$p_{\mathrm{FD}} = \int_{0}^{R_t} dR \; P(R|\bar{R})$$

A Case Study - Binary Black Hole Merger GW150914

Detected on September 14, 2015 @ 09:50:45 UTC

False alarm probability $< 2 \times 10^{-7}$

Parameter estimation started:
* coherent across the LIGO network
* used waveform models that include full richness of physics with black hole spins
* covers full parameter space with fine sampling

A Case Study - Binary Black Hole Merger GW150914

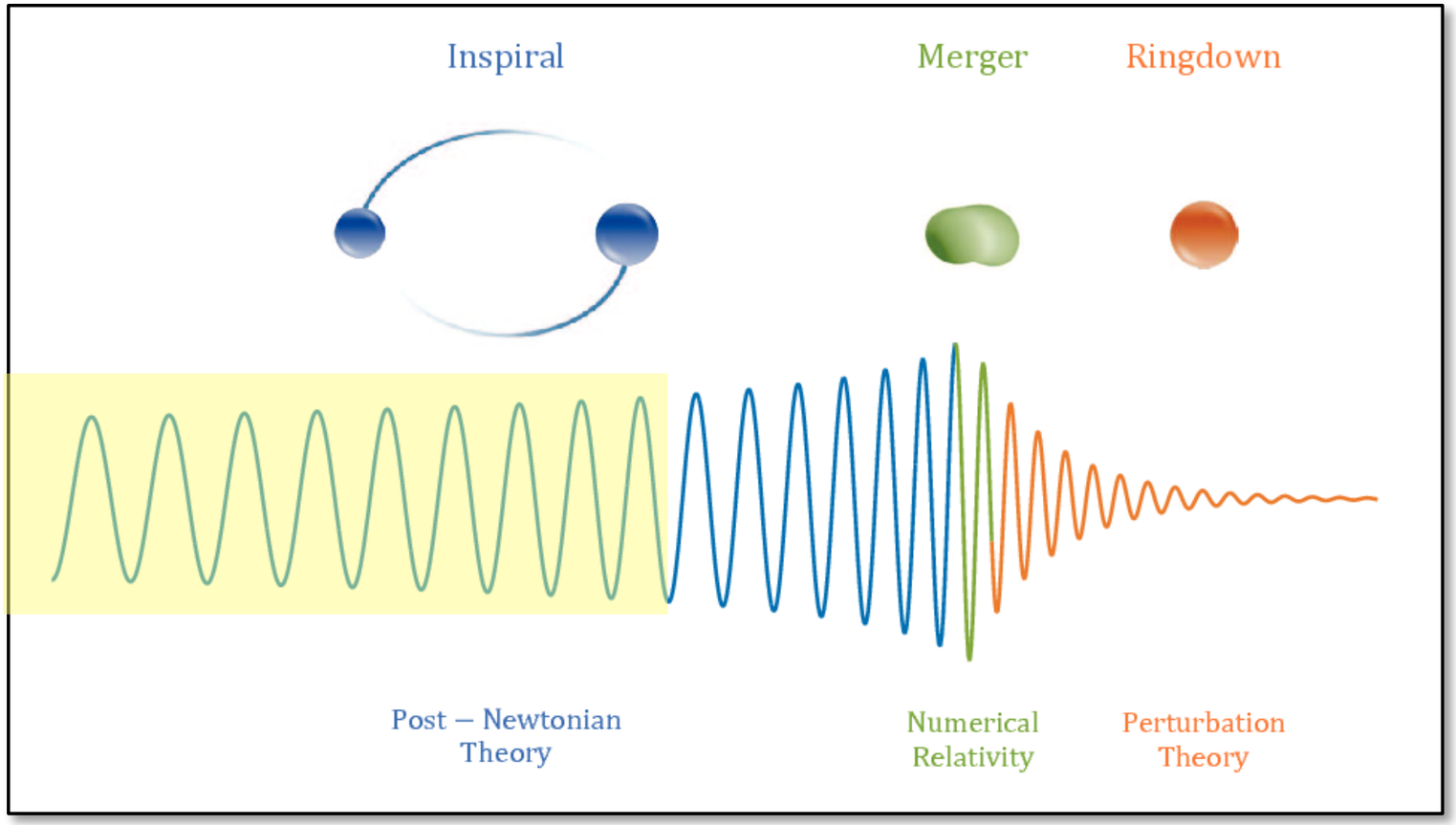
8 intrinsic parameters

mass 1
mass 2
spin1x
spin1y
spin1z
spin2x
spin2y
spin2z

9 extrinsic parameters

luminosity distance
right ascension
declination
binary orbital inclination
binary polarization angle
coalescence time
coalescence phase
eccentricity magnitude
periapsis

Radiation reaction circularizes orbits for signals in LIGO/Virgo band so ignore this



During inspiral, phase evolution $\phi_{\rm GW}(t; m_{1,2}, \mathbf{S}_{1,2})$ can be computed with PN-theory in powers of v/c.

leading order

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

$$\simeq \frac{c^3}{G}\left[\frac{5}{96}\pi^{-8/3} f^{-11/3}\dot{f}\right]^{3/5}$$

higher order

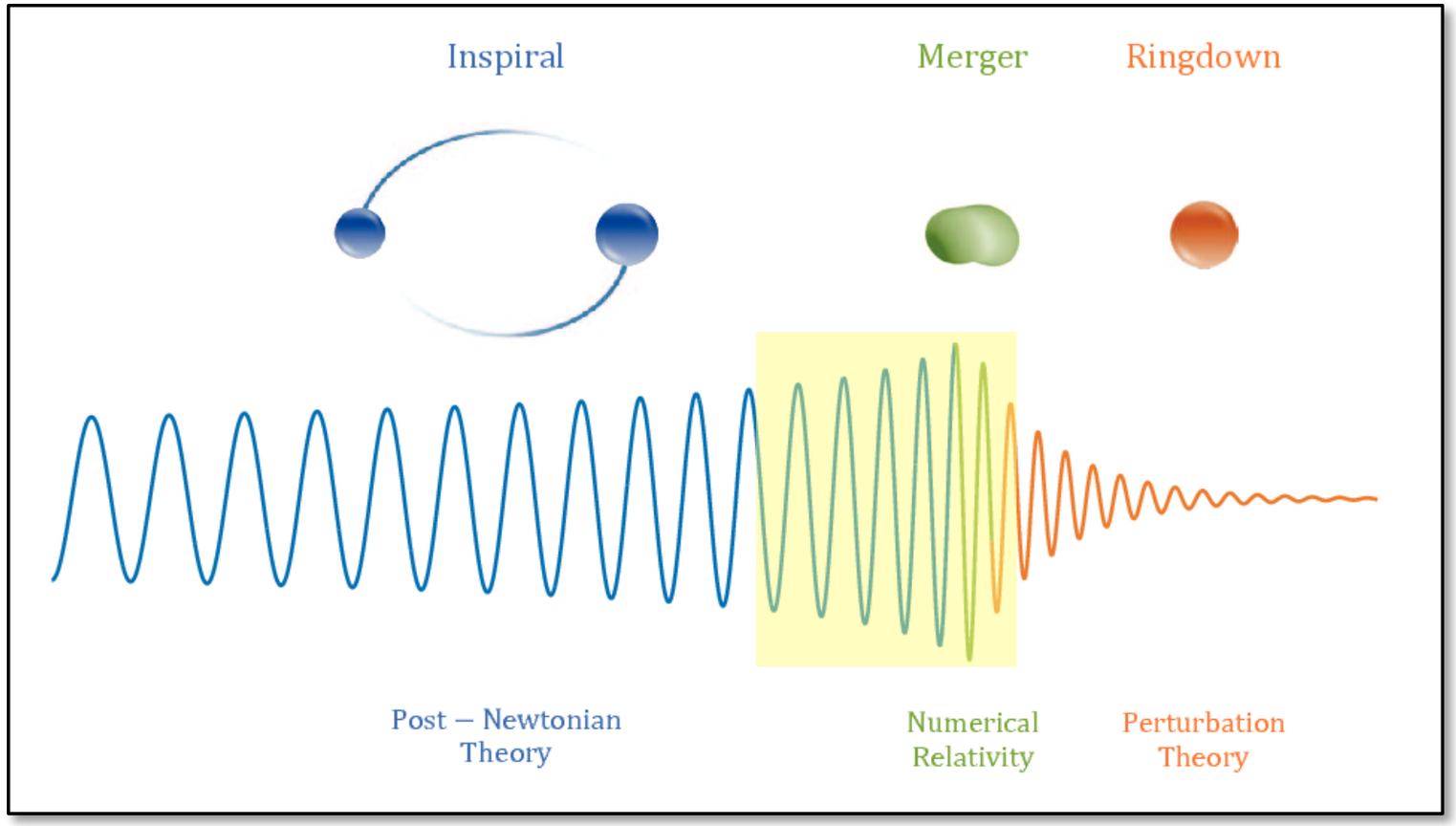
$$q = \frac{m_2}{m_1} \leq 1$$

$$\mathbf{S}_{1,2} \parallel \mathbf{L}$$

even higher order

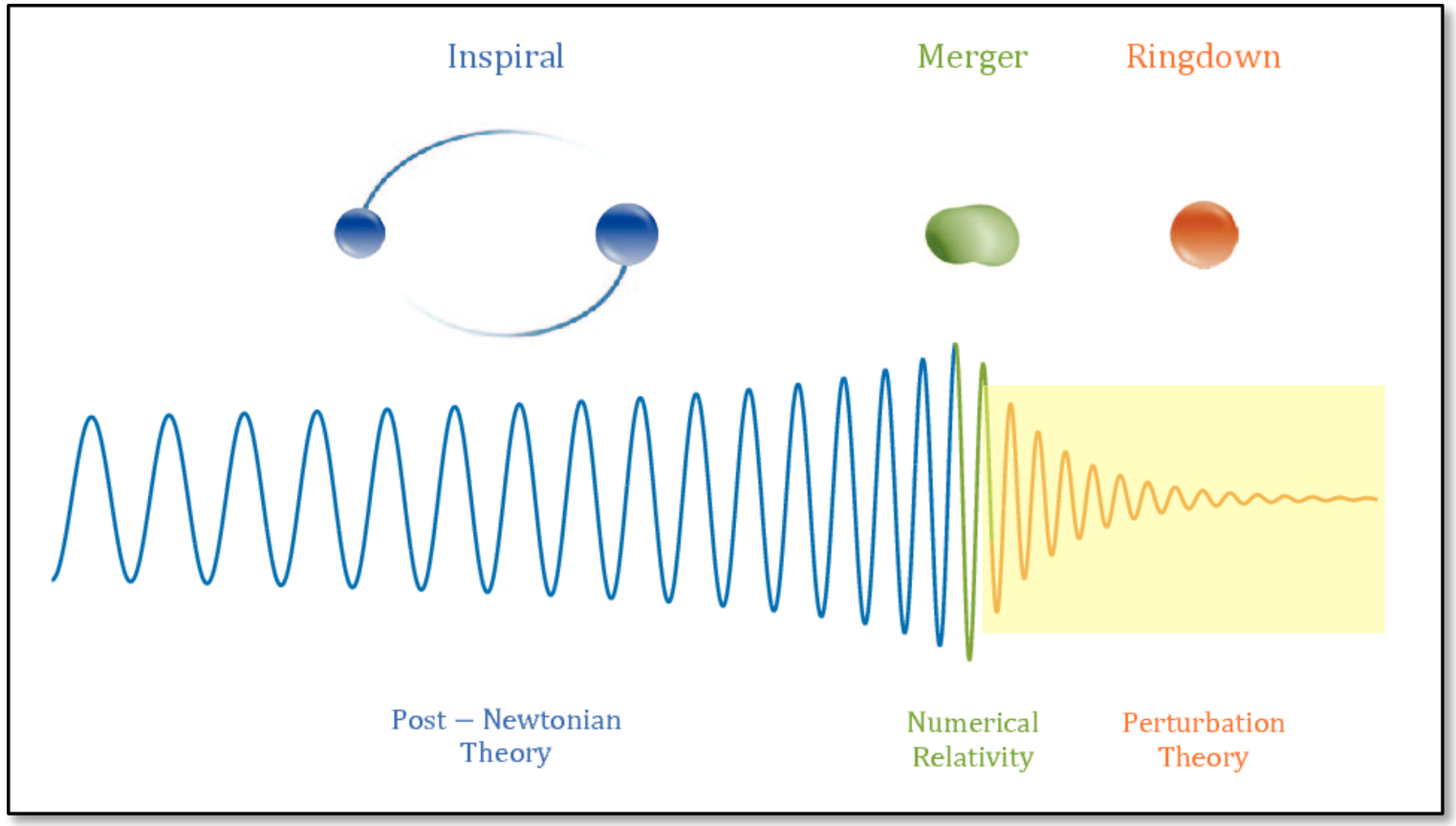
$$S_{1x}, S_{1y}, S_{1z}$$

$$S_{2x}, S_{2y}, S_{2z}$$



Numerical relativity needed for binary evolution in late inspiral and merger.

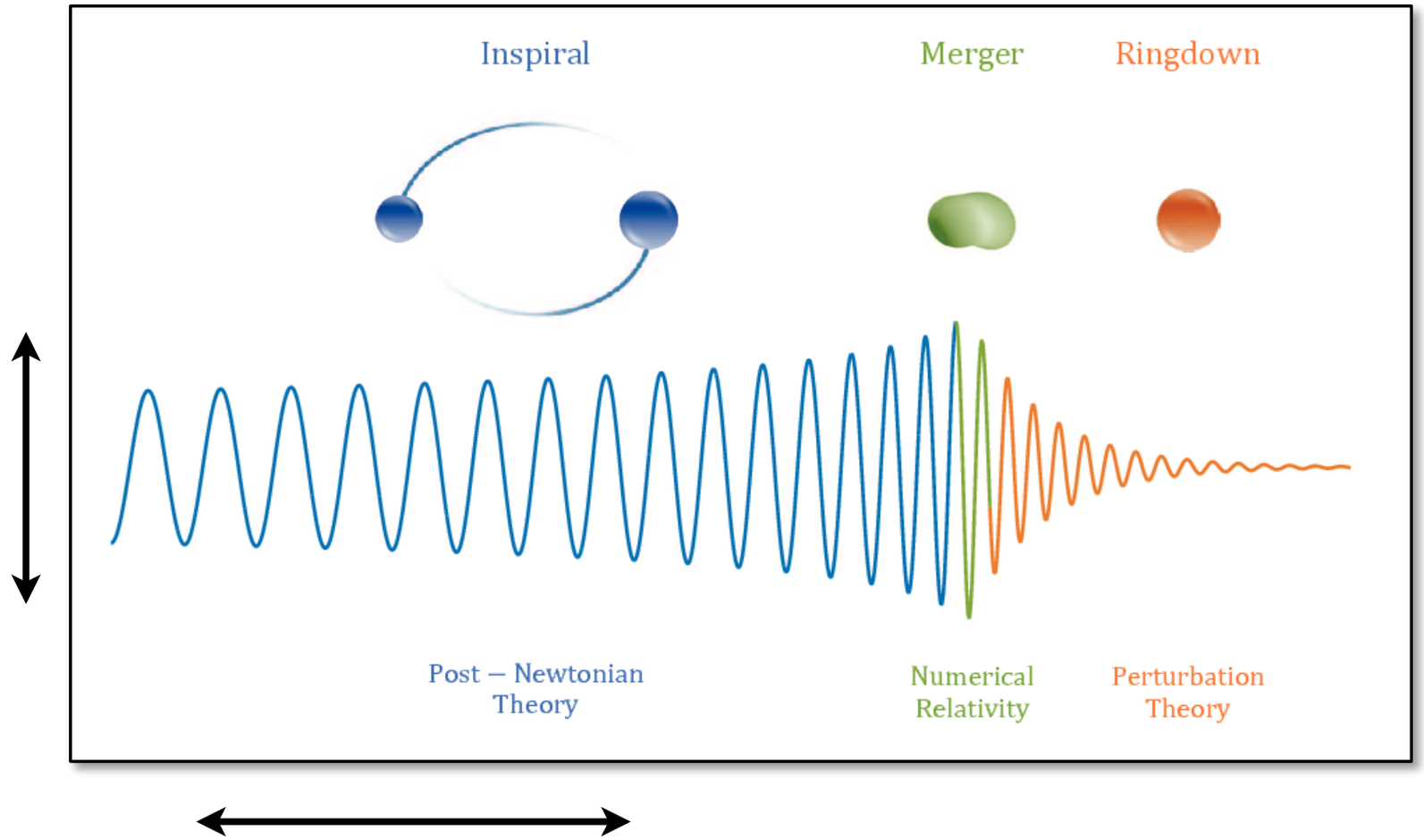
$$M_{\text{total}} = m_1 + m_2$$



Details of ringdown

Final dimensionless spin magnitude $a_f = \frac{c|\mathbf{S_f}|}{Gm_f^2} \leq 1$

Final mass M_f



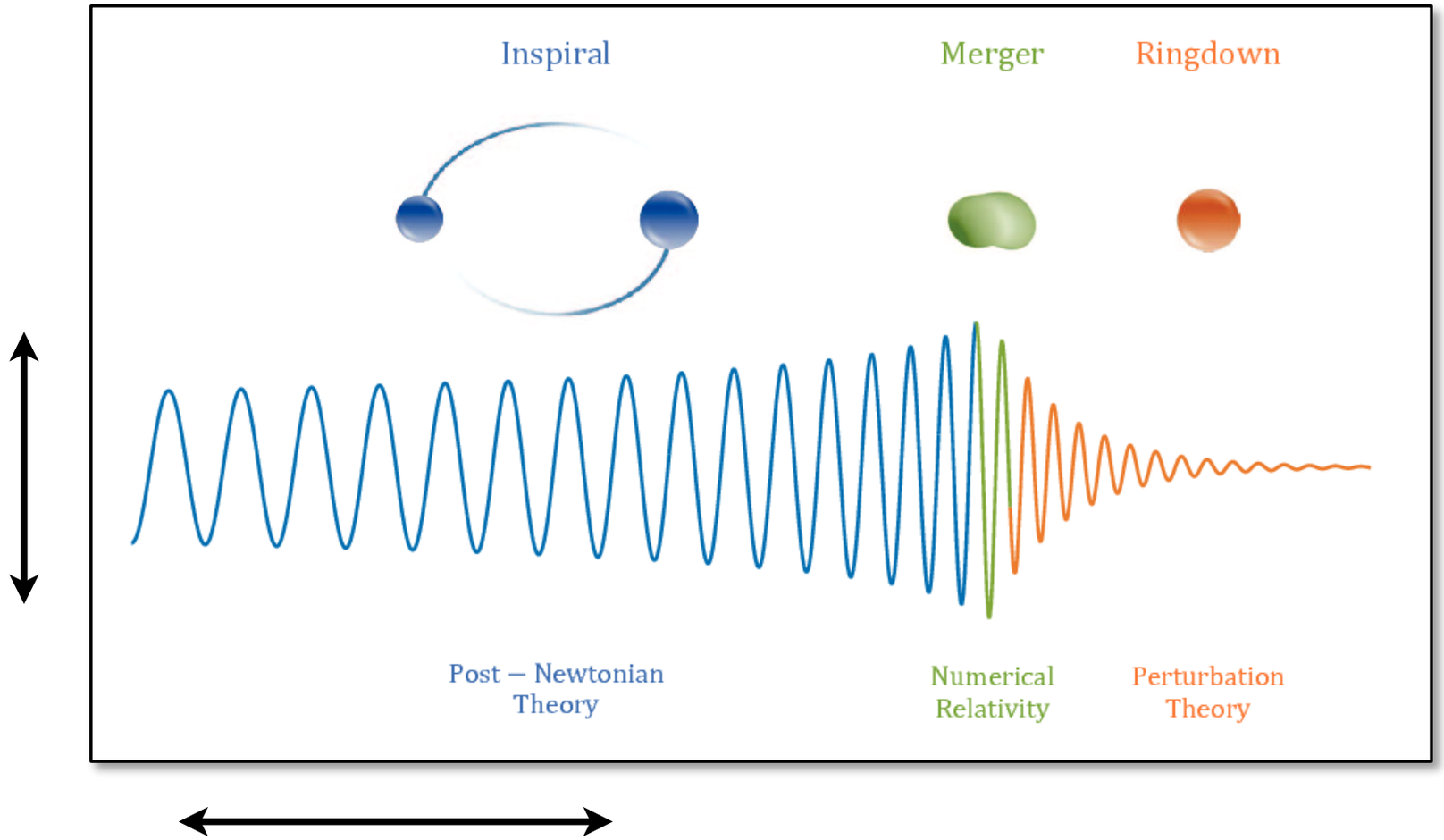
Amplitude

$$A_{\mathrm{GW}} \propto \frac{1}{D_L}$$

Observed frequency redshifted by a factor of $(1+z)$

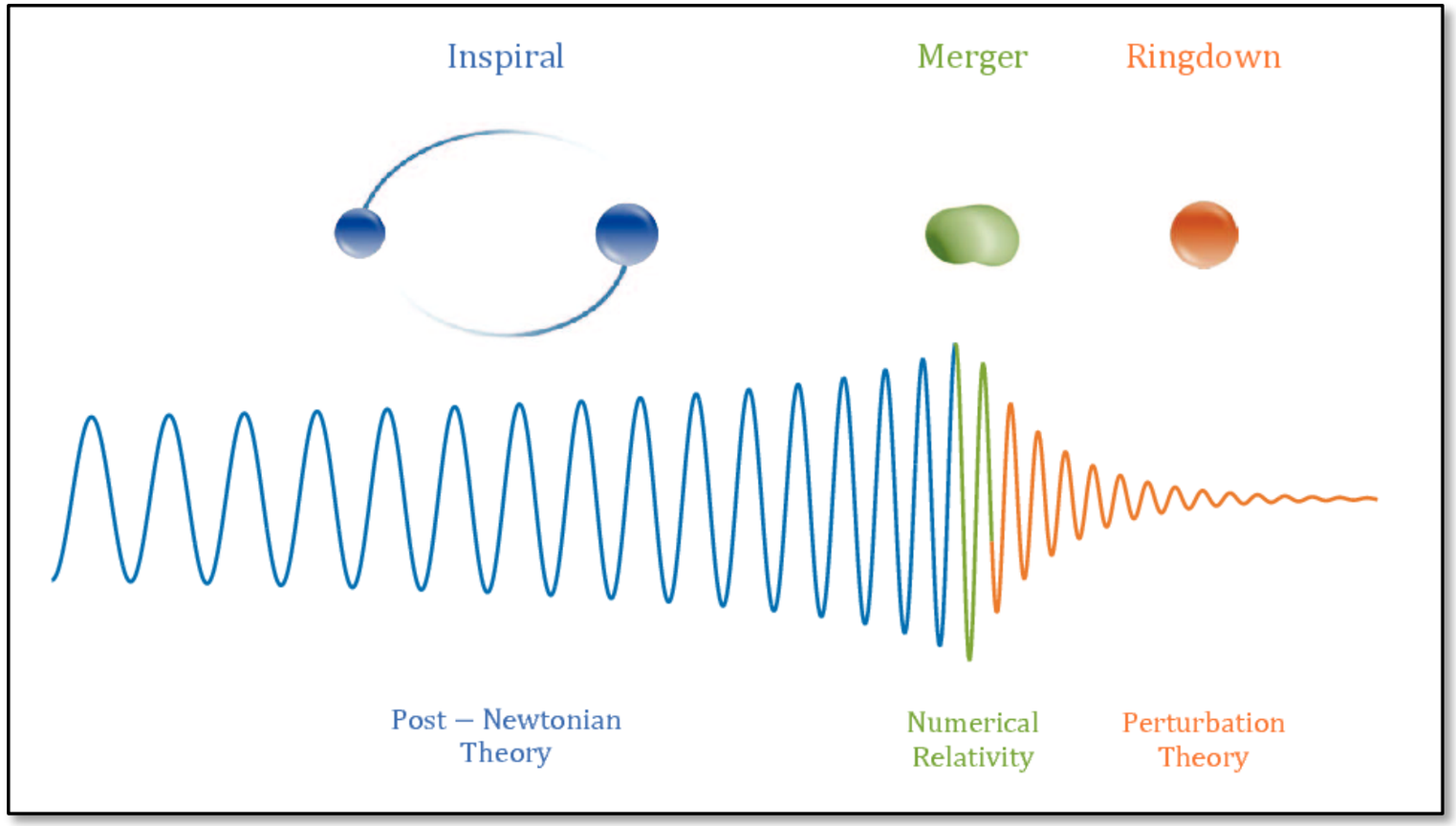
z - cosmological redshift

Indistinguishable from rescaling of masses $m = (1+z)m^{\mathrm{source}}$



For systems with minimal precession, all the following change the overall amplitude and phase but not signal morphology:

$$D_L, \alpha, \delta, \iota, \psi, t_c, \phi_c$$



Amplitude and phase modulations

ψ , ι become time-dependent; binary's orbital plane precesses around direction of total angular momentum:

$$\mathbf{J} = \mathbf{L} + \mathbf{S_1} + \mathbf{S_2}$$

Depends on viewing angle

A Case Study - Binary Black Hole Merger GW150914

Evaluation of multidimensional integrals using two independent stochastic sampling engines based on:

* Markov-chain Monte Carlo
* Nested Sampling

Hypothesis: GW from compact binary coalescence

Use model waveforms for inspiral/merger of two black holes

Results are posterior PDFs for parameters describing the signal and the model evidence

A Case Study - Binary Black Hole Merger GW150914

Priors

t_c in ± 0.1 s

$m_{1,2} \in [10, 80] M_\odot$

ϕ_c in $[0, 2\pi]$

$a_{1,2} \in [0, 1]$

$f \in [20, 1024]$ Hz

uniform in volume

isotropically oriented

Precessing model: isotropic spin orientation

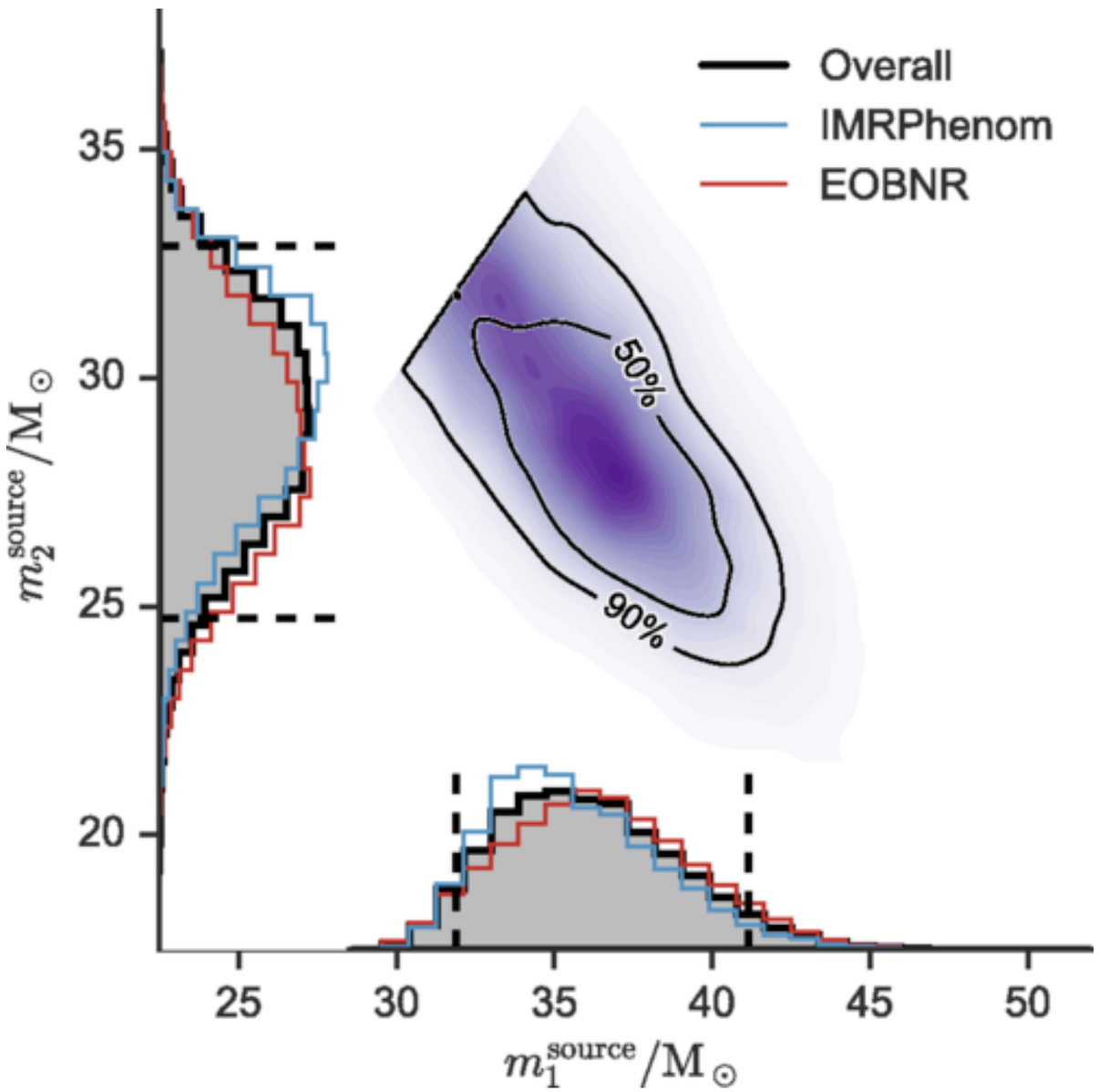
Aligned-spin model: uniform distribution [-1, 1]

	EOBNR	IMRPhenom	Overall
Detector-frame total mass $M/M_\odot$	$70.3^{+5.3}_{-4.8}$	$70.9^{+4.0}_{-3.9}$	$70.6^{+4.6\pm0.5}_{-4.5\pm1.3}$
Detector-frame chirp mass $\mathcal{M}/M_\odot$	$30.2^{+2.5}_{-1.9}$	$30.6^{+1.8}_{-1.8}$	$30.4^{+2.1\pm0.2}_{-1.9\pm0.5}$
Detector-frame primary mass $m_1/M_\odot$	$39.4^{+5.5}_{-4.9}$	$38.5^{+5.6}_{-3.6}$	$38.9^{+5.6\pm0.6}_{-4.3\pm0.4}$
Detector-frame secondary mass $m_2/M_\odot$	$30.9^{+4.8}_{-4.4}$	$32.2^{+3.6}_{-4.8}$	$31.6^{+4.2\pm0.1}_{-4.7\pm0.9}$
Detector-frame final mass $M_f/M_\odot$	$67.1^{+4.6}_{-4.4}$	$67.6^{+3.6}_{-3.5}$	$67.4^{+4.1\pm0.4}_{-4.0\pm1.2}$
Source-frame total mass $M^{\rm source}/M_\odot$	$65.0^{+5.0}_{-4.4}$	$65.0^{+4.0}_{-3.6}$	$65.0^{+4.5\pm0.8}_{-4.0\pm0.7}$
Source-frame chirp mass $\mathcal{M}^{\rm source}/M_\odot$	$27.9^{+2.3}_{-1.8}$	$28.1^{+1.7}_{-1.6}$	$28.0^{+2.0\pm0.3}_{-1.7\pm0.3}$
Source-frame primary mass $m_1^{\rm source}/M_\odot$	$36.3^{+5.3}_{-4.5}$	$35.3^{+5.2}_{-3.4}$	$35.8^{+5.3\pm0.9}_{-3.9\pm0.1}$
Source-frame secondary mass $m_2^{\rm source}/M_\odot$	$28.6^{+4.4}_{-4.2}$	$29.6^{+3.3}_{-4.3}$	$29.1^{+3.8\pm0.1}_{-4.3\pm0.7}$
Source-frame final mass $M_f^{\rm source}/M_\odot$	$62.0^{+4.4}_{-4.0}$	$62.0^{+3.7}_{-3.3}$	$62.0^{+4.1\pm0.7}_{-3.7\pm0.6}$
Mass ratio q	$0.79^{+0.18}_{-0.19}$	$0.84^{+0.14}_{-0.20}$	$0.82^{+0.17\pm0.01}_{-0.20\pm0.03}$
Effective inspiral spin parameter $\chi_{\rm eff}$	$-0.09^{+0.19}_{-0.17}$	$-0.05^{+0.13}_{-0.15}$	$-0.07^{+0.16\pm0.01}_{-0.17\pm0.05}$
Dimensionless primary spin magnitude a_1	$0.32^{+0.45}_{-0.28}$	$0.32^{+0.53}_{-0.29}$	$0.32^{+0.49\pm0.06}_{-0.29\pm0.01}$
Dimensionless secondary spin magnitude a_2	$0.57^{+0.40}_{-0.51}$	$0.34^{+0.54}_{-0.31}$	$0.44^{+0.50\pm0.08}_{-0.40\pm0.02}$
Final spin a_f	$0.67^{+0.06}_{-0.08}$	$0.66^{+0.04}_{-0.06}$	$0.67^{+0.05\pm0.01}_{-0.07\pm0.02}$
Luminosity distance D_L/Mpc	390^{+170}_{-180}	440^{+150}_{-180}	$410^{+160\pm20}_{-180\pm40}$
Source redshift z	$0.083^{+0.033}_{-0.036}$	$0.093^{+0.029}_{-0.036}$	$0.088^{+0.032\pm0.005}_{-0.037\pm0.008}$
Upper bound on primary spin magnitude a_1	0.65	0.74	0.69 ± 0.08
Upper bound on secondary spin magnitude a_2	0.93	0.78	0.89 ± 0.13
Lower bound on mass ratio q	0.64	0.68	0.66 ± 0.03
Log Bayes factor $\ln \mathcal{B}_{s/n}$	288.7 ± 0.2	290.3 ± 0.1	...

Parameter estimates are broadly consistent across two models.

Log Bayes factors are comparable so we cannot prefer one model over the other.

A Case Study - Binary Black Hole Merger GW150914



$$m_1^{\rm source} = 36^{+5}_{-4} M_\odot$$

$$m_2^{\rm source} = 29^{+4}_{-4} M_\odot$$

$$0.66 \leq q \leq 1$$

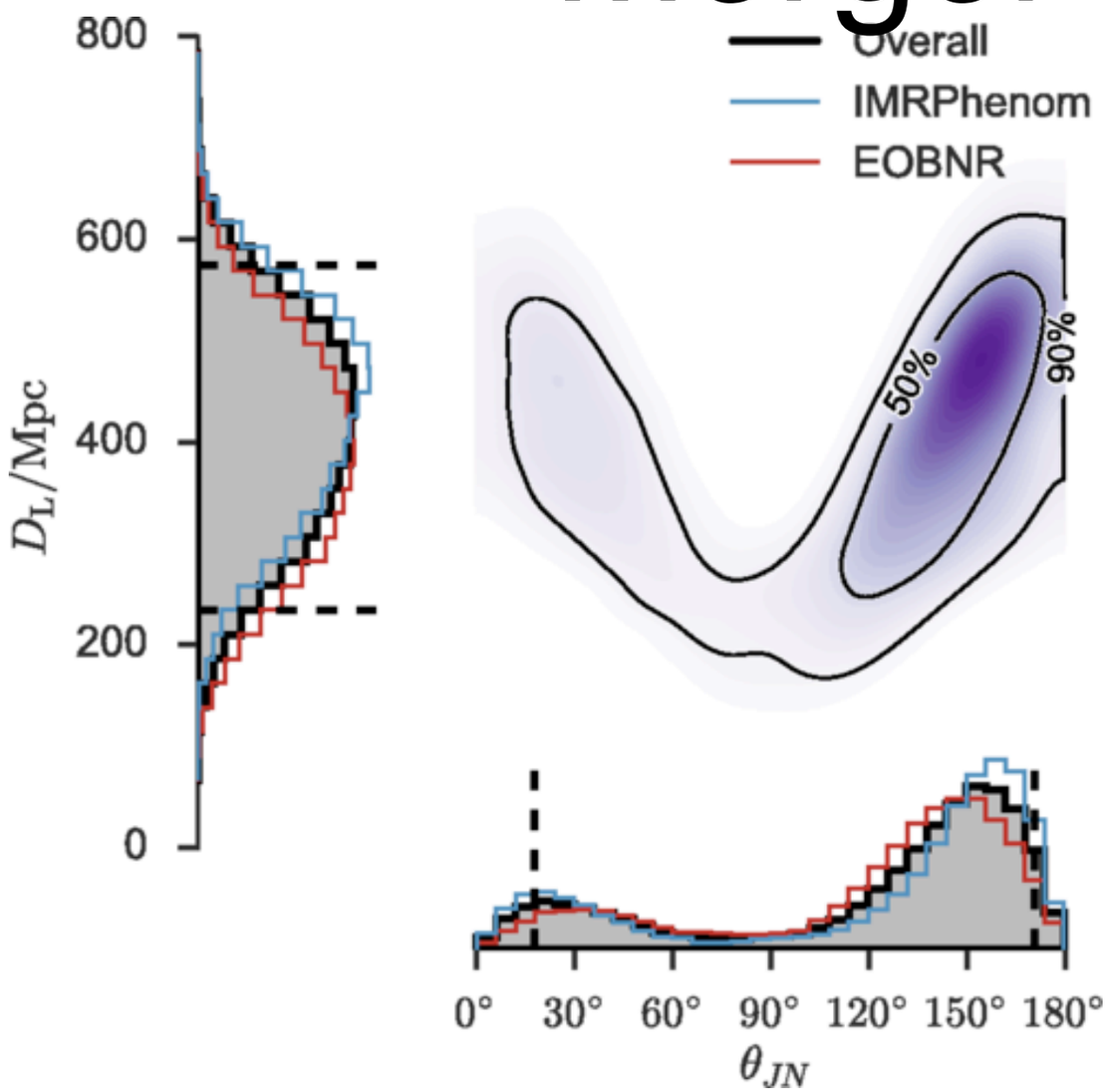
with 90% probability

Conservative upper limit for mass of stable NS is $3M_\odot$

Could consider exotic alternatives.

LVC, PRL 116, 241102 (2016)

A Case Study - Binary Black Hole Merger GW150914



$$D_L = 410^{+160}_{-180}\mathrm{Mpc}$$

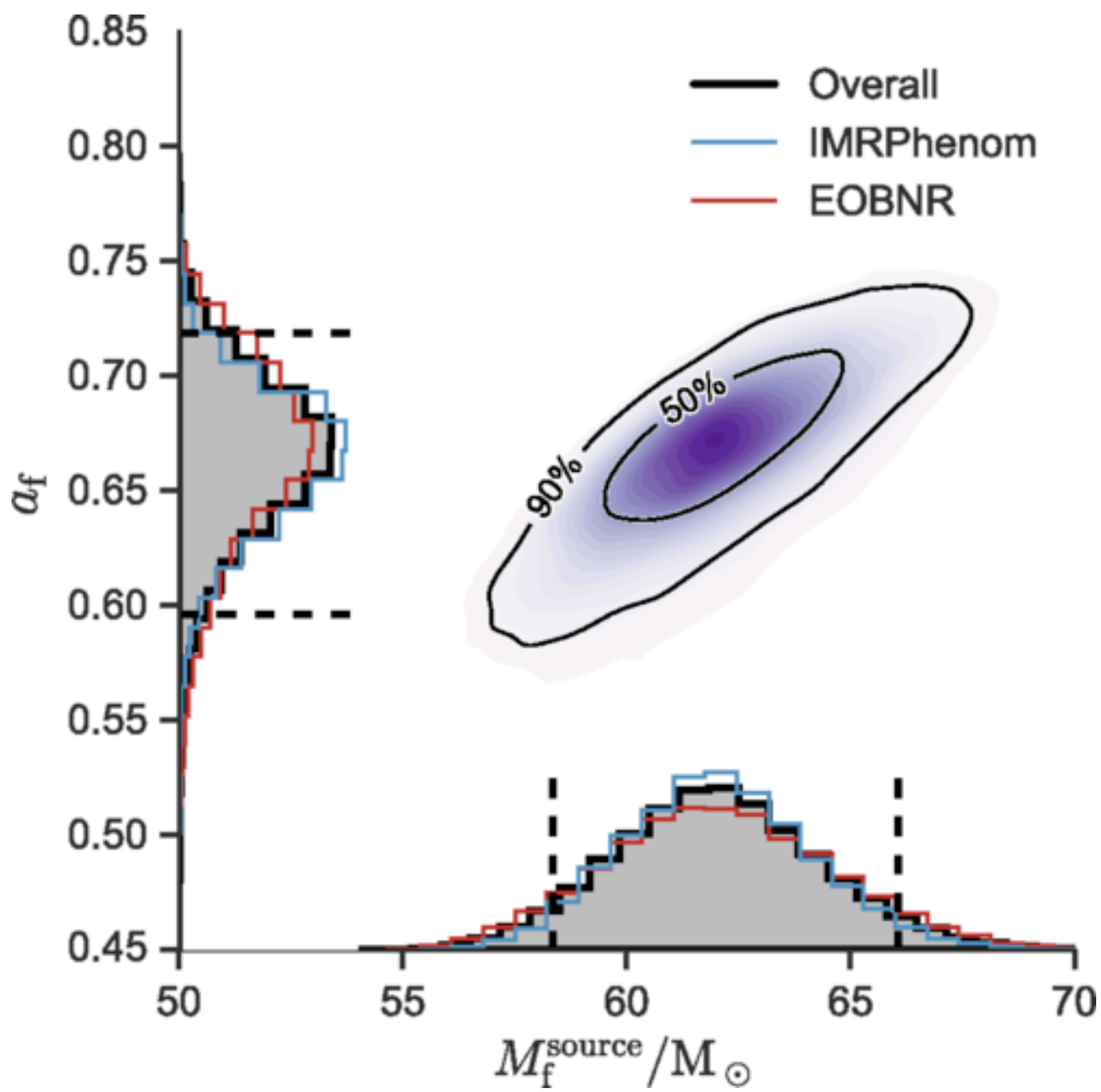
Assuming flat ΛCDM cosmology, the inferred luminosity distance corresponds to redshift:

$$z = 0.09^{+0.03}_{-0.04}$$

D_L correlated with the inclination of orbital plane with respect to line of sight

Orientation of total orbital angular momentum misaligned to line of sight is disfavored

A Case Study - Binary Black Hole Merger GW150914



Does not use waveform models but rather fitting formula calibrated to NR simulations.

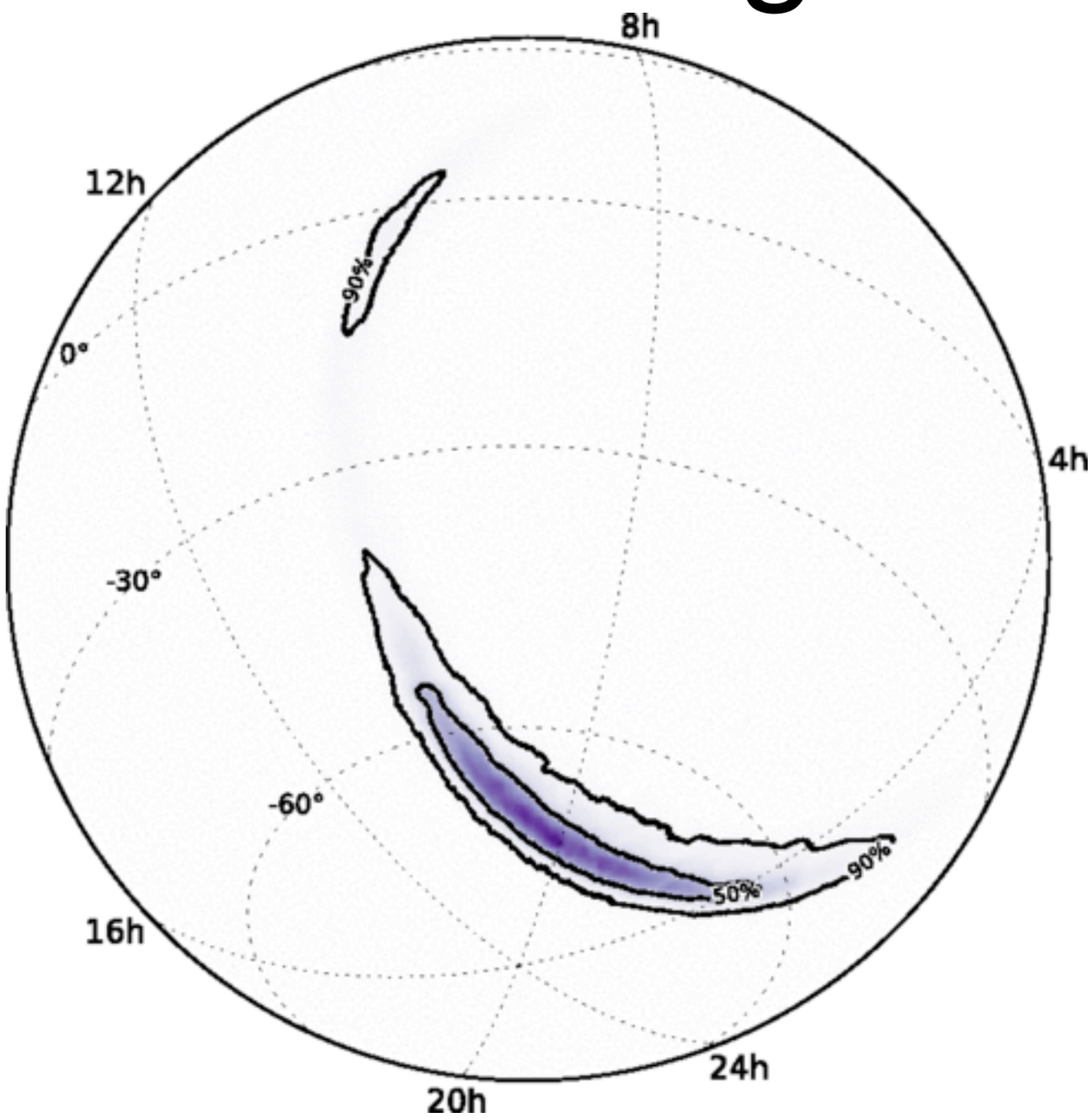
$$M_f^{\rm source} = 62^{+4}_{-4} M_\odot$$

$$a_f = 0.67^{+0.05}_{-0.07}$$

Final spin is precisely determined.

$$E_{\rm rad} = M^{\rm source} - M_f^{\rm source} = 3.0^{+0.5}_{-0.4} M_\odot c^2$$

A Case Study - Binary Black Hole Merger GW150914

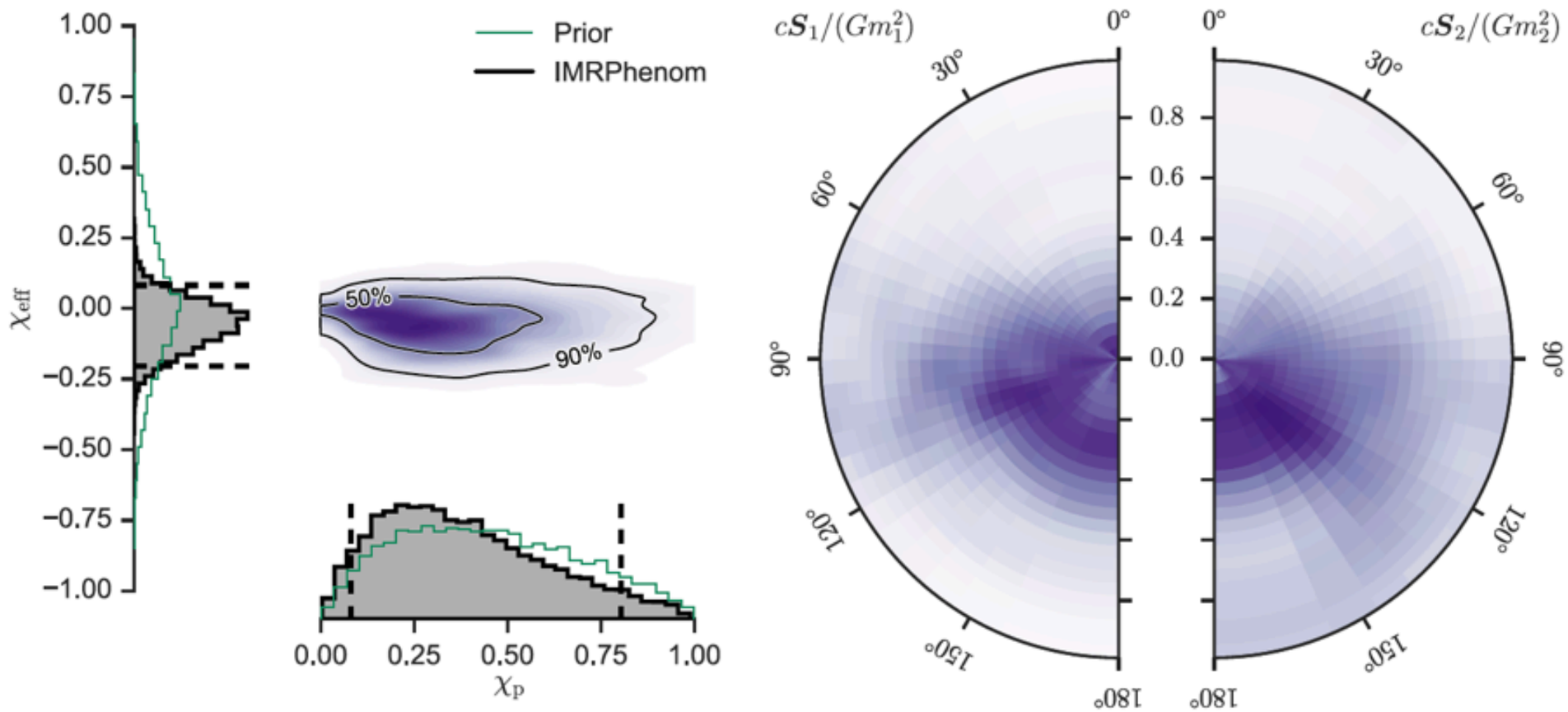


Network of GW detectors needed to reconstruct location of GW in sky via time of arrival and amplitude and phase consistency

$$\Delta t_{\mathrm{HL}} = 6.9^{+0.5}_{-0.4}\mathrm{ms}$$

$610\,\mathrm{deg}^2$ (90% probability)

A Case Study - Binary Black Hole Merger GW150914

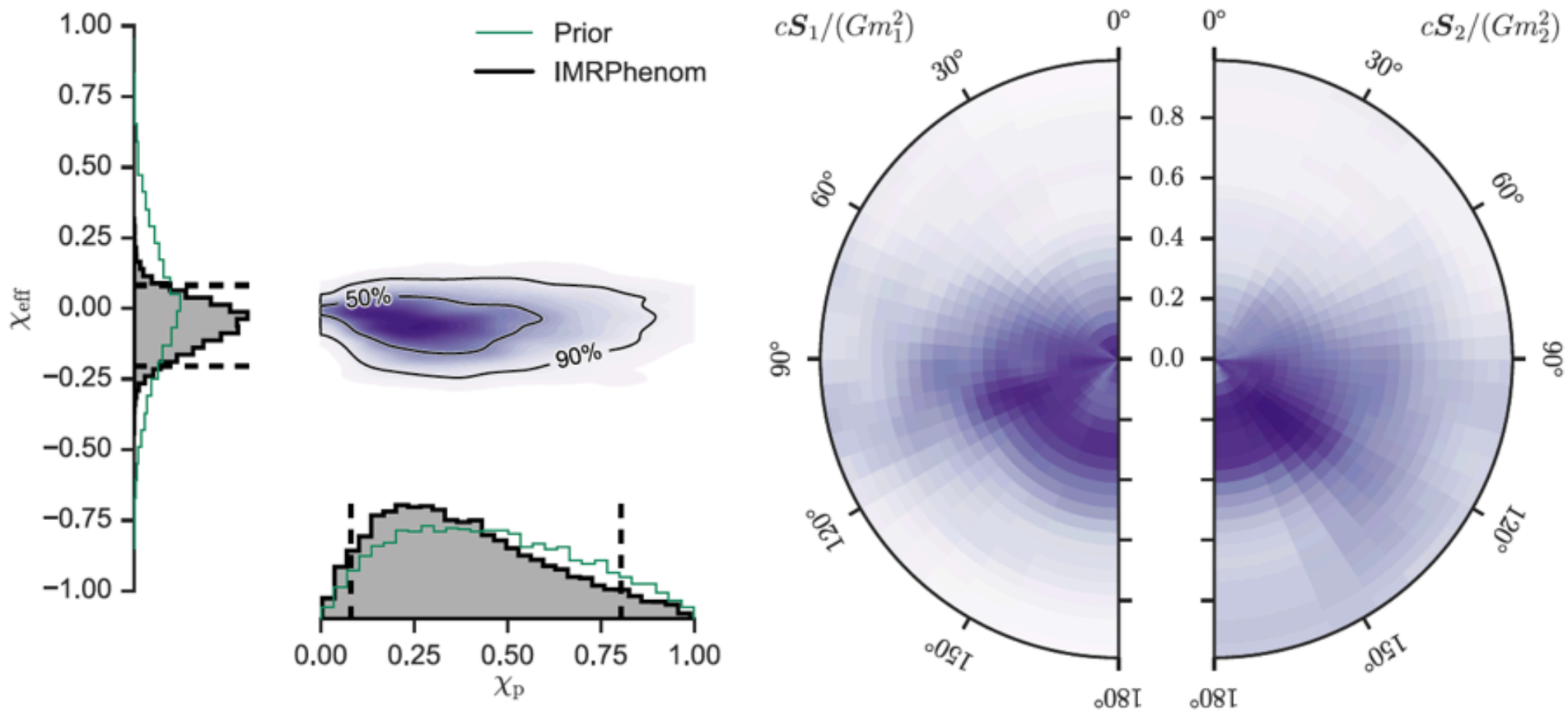


Spin projections along direction of orbital angular momentum affect inspiral rate of binary.

$a_1 < 0.7$ (at 90% probability)

$a_2 < 0.9$ (at 90% probability)

A Case Study - Binary Black Hole Merger GW150914

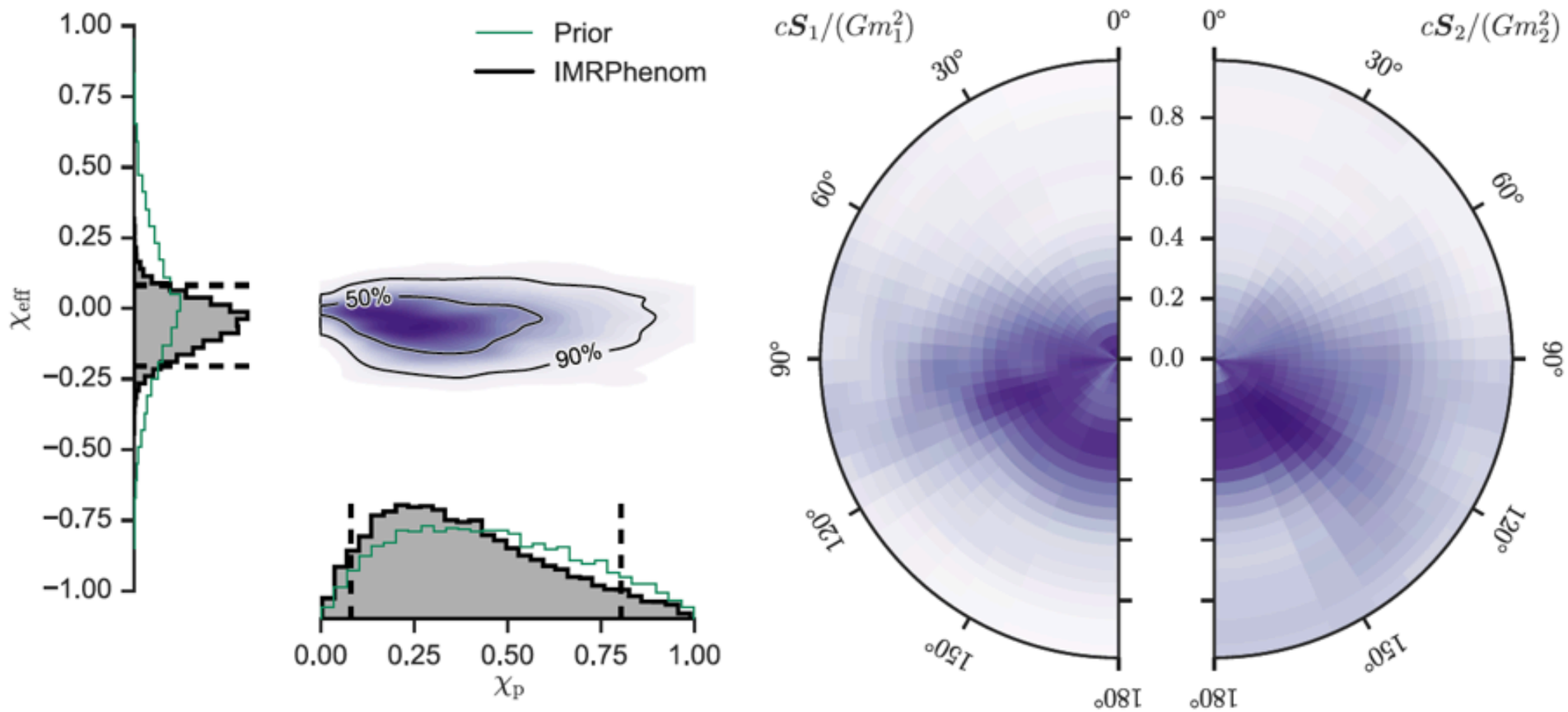


Difficult to untangle full degrees of freedom but several one dimensional parameterizations have been defined.

$$\chi_{\rm eff} = \frac{c}{GM}\left(\frac{\mathbf{S_1}}{m_1} + \frac{\mathbf{S_2}}{m_2}\right) \cdot \frac{\mathbf{L}}{|\mathbf{L}|}$$

$$\chi_{\rm eff} = -0.07^{+0.16}_{-0.17}$$

A Case Study - Binary Black Hole Merger GW150914



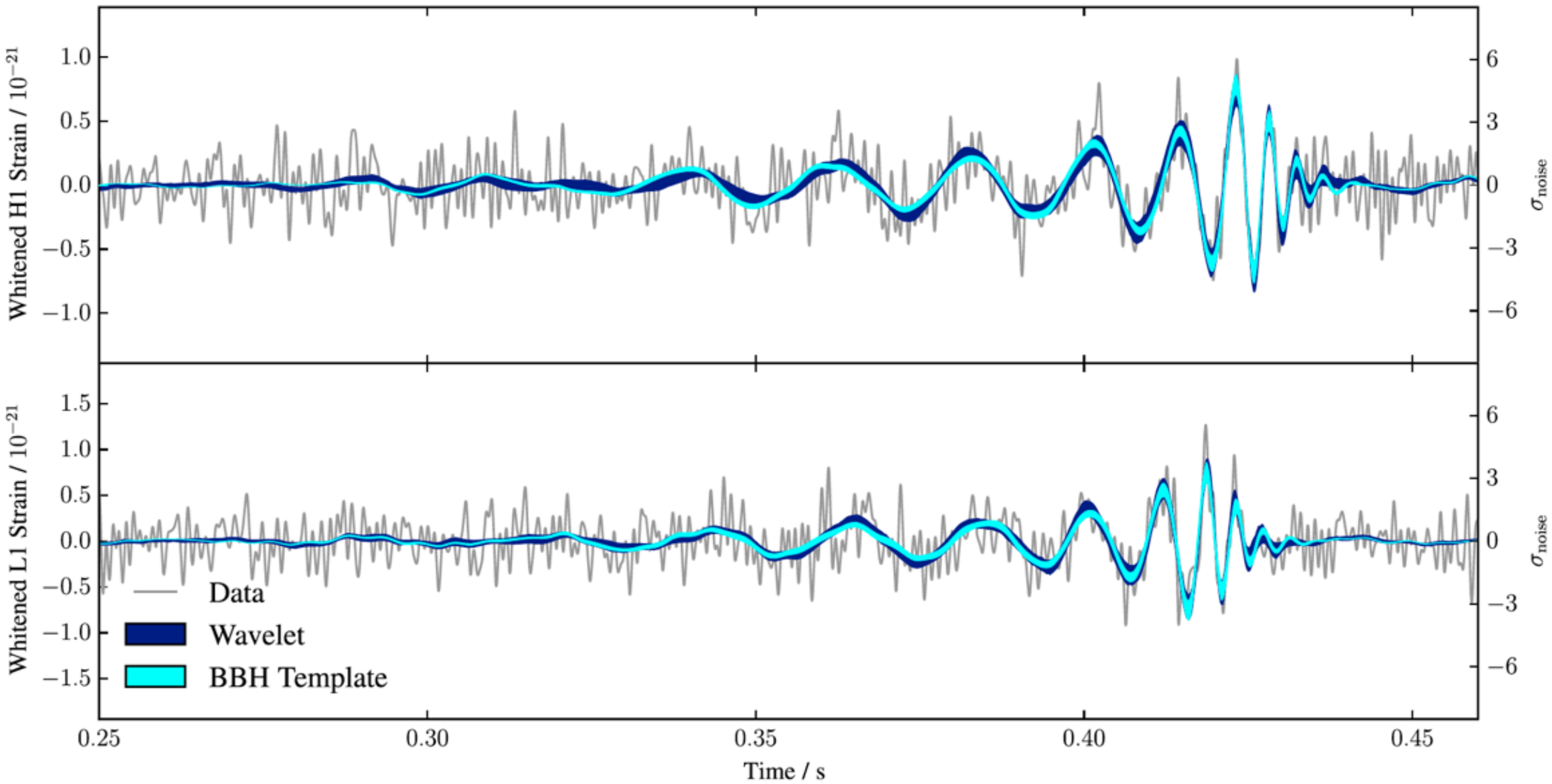
Difficult to untangle full degrees of freedom but several one dimensional parameterizations have been defined.

Effective precession spin parameter

$$\chi_p = \frac{c}{B_1 G m_1^2} \max(B_1 S_{1\perp}, B_2 S_{2\perp}) > 0 \qquad B_{1,2}(q)$$

χ_p consistent with prior.

Minimal assumption analysis: not necessarily derived from binary system



Remarkable agreement between the actual data and the reconstructed waveform under two model assumptions.

* ~10 cycles during inspiral phase from 30Hz
* merger
* ringdown

A Case Study - Binary Black Hole Merger GW150914

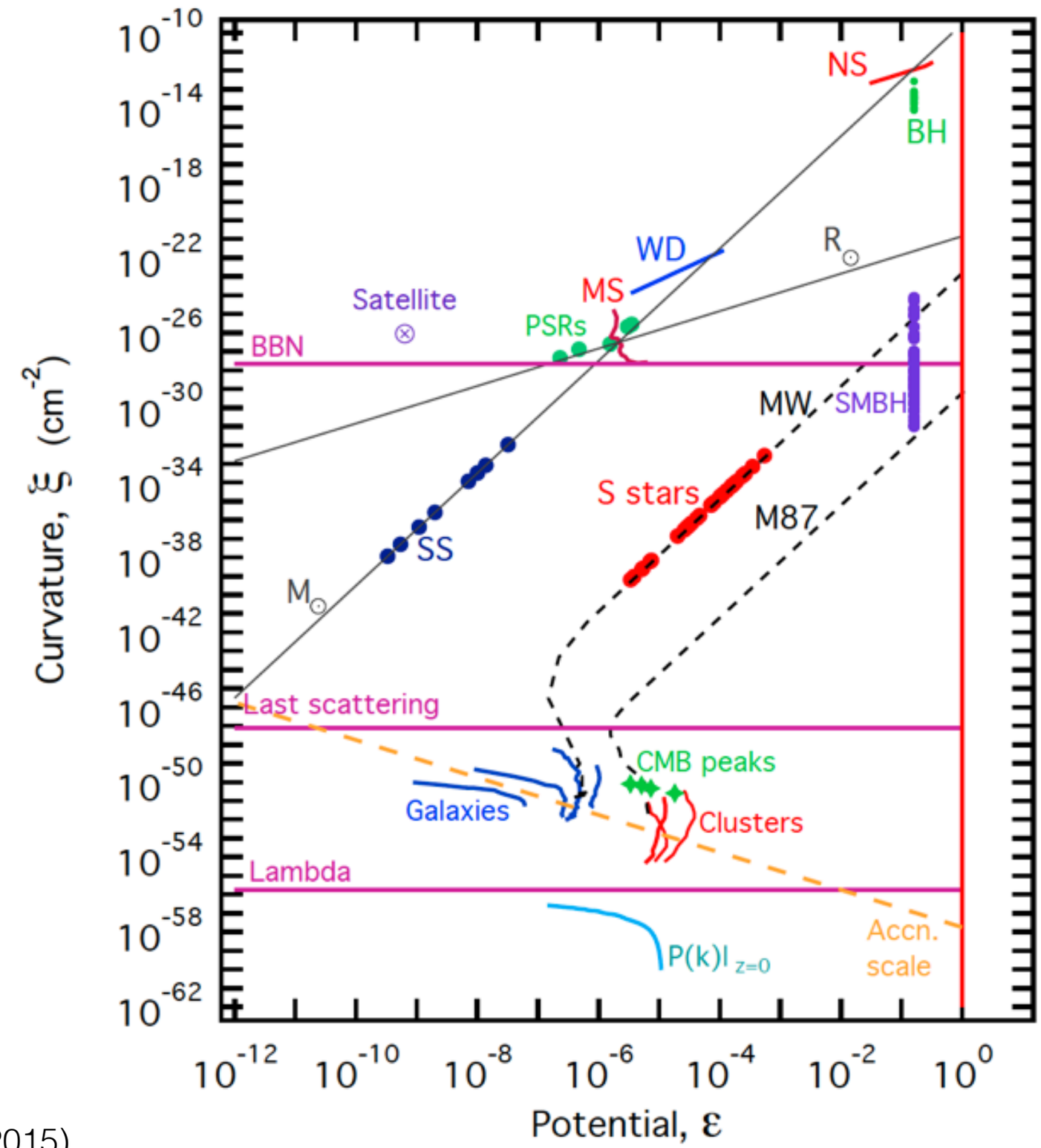
Heaviest stellar mass BHs known to date

First stellar-mass binary BH

Binary black holes do form and merge within Hubble time

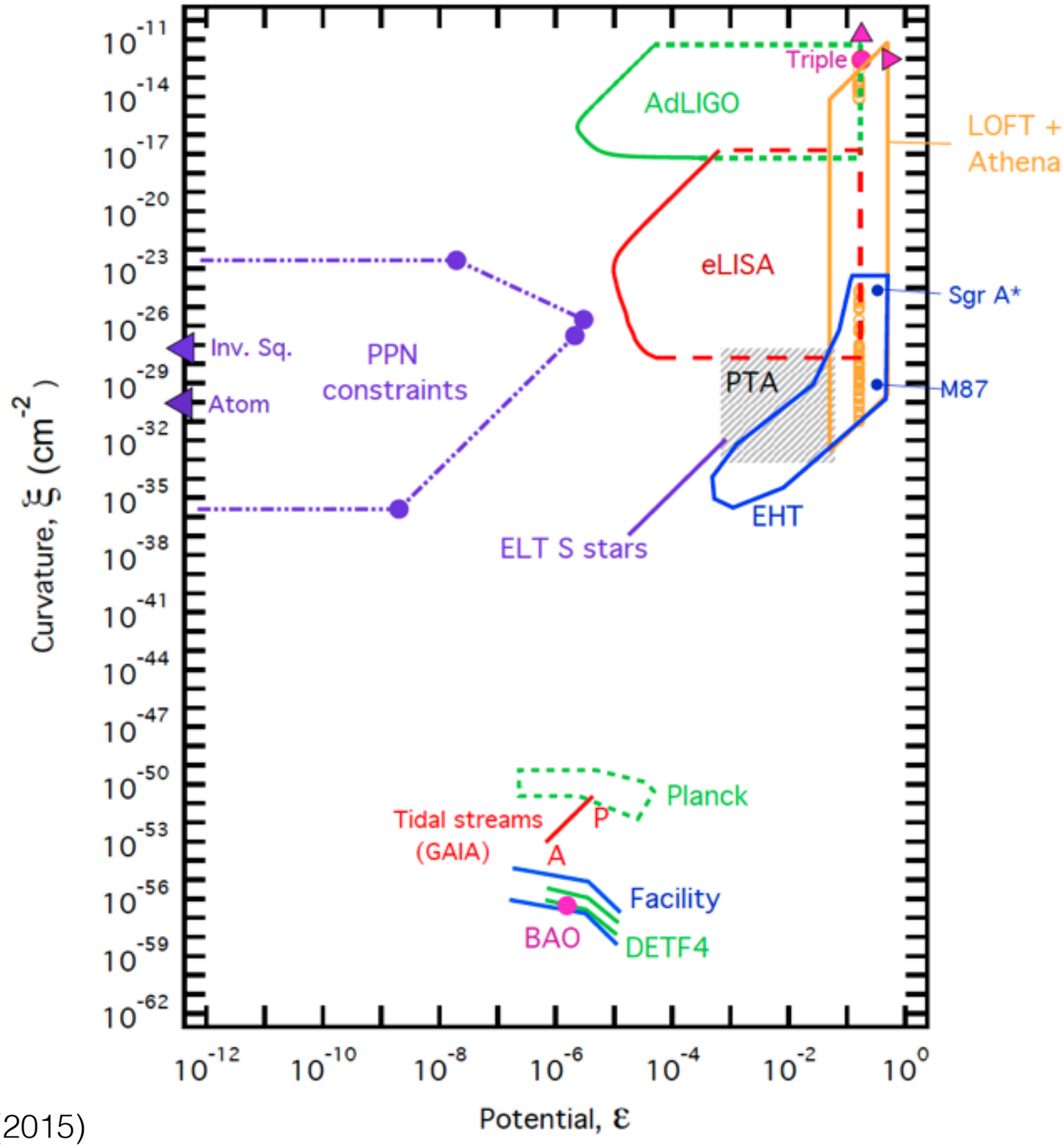
First BH spin constraints independent of x-ray spectra observations

Strong field tests of General Relativity



Baker, et al ApJ 802, 63 (2015)

Strong field tests of General Relativity

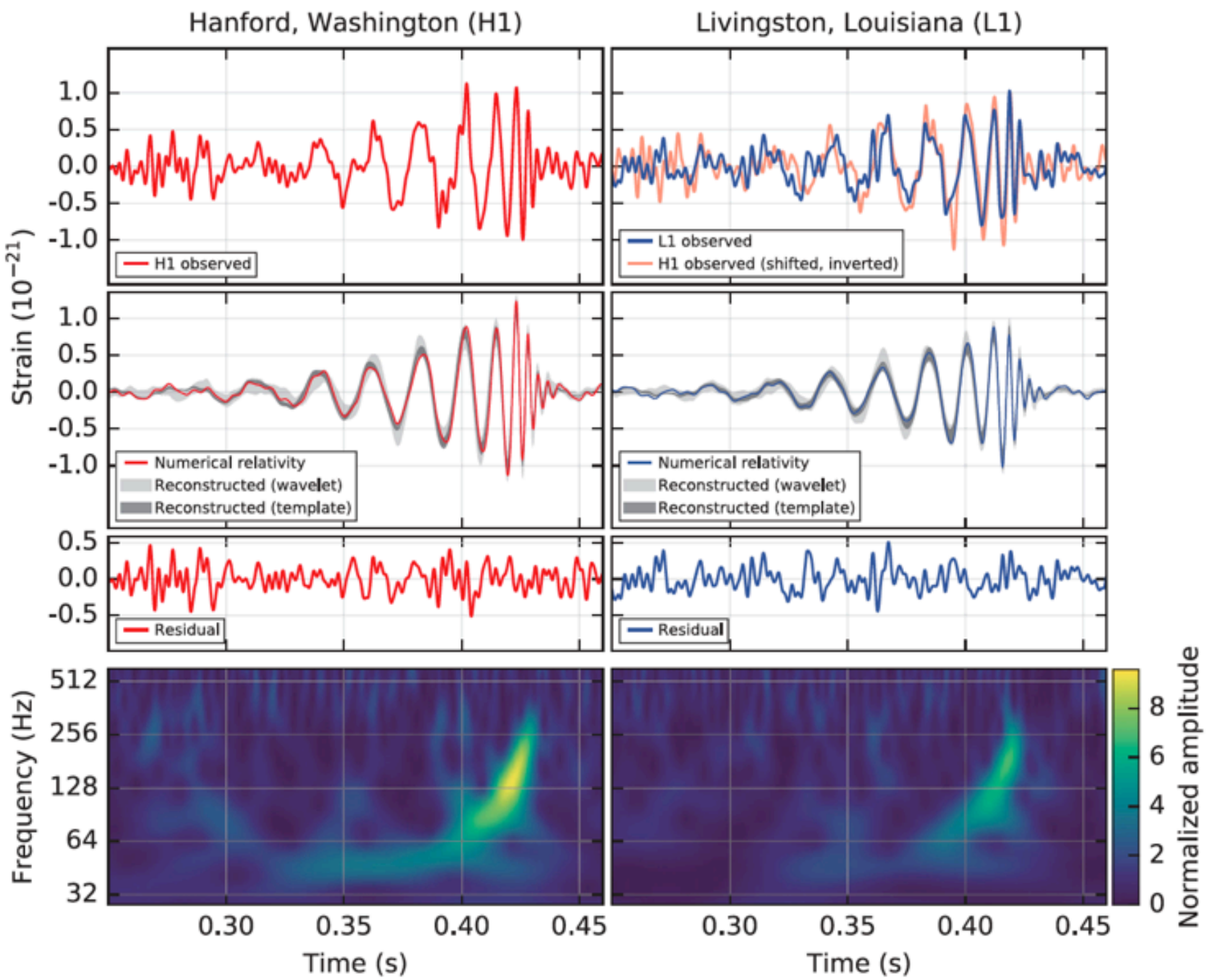


Baker, et al ApJ 802, 63 (2015)

Residual strain

Remove most probable GR waveform from data.

Calibrated against waveforms from direct numerical integration of Einstein equations.



Analysis reveals that GW150914 residual favors instrumental noise over the presence of coherent signal or glitches.

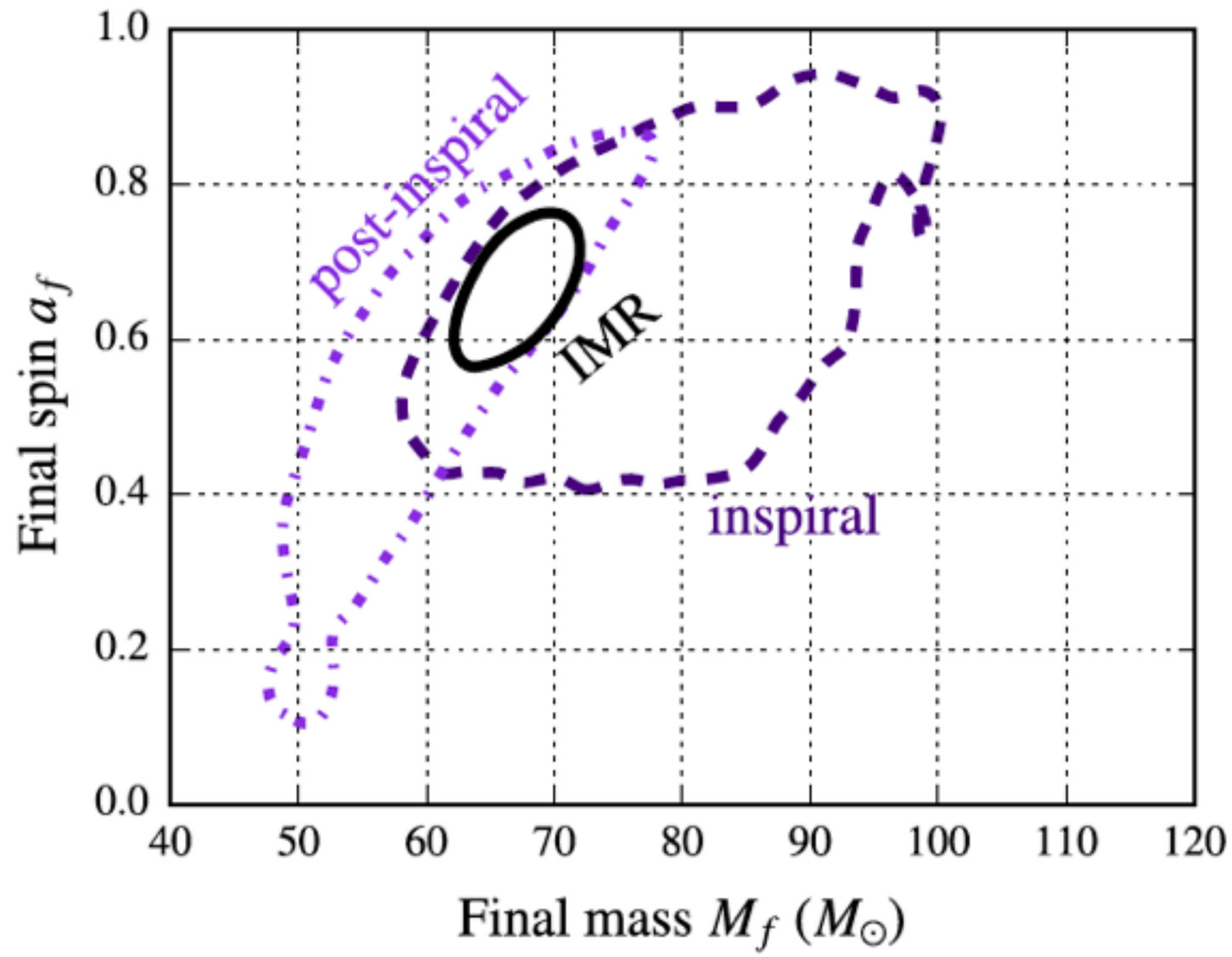
Inspiral-merger-ringdown consistency test

Mass and spin parameters predicted from binary inspiral

versus

Mass and spin inferred from post-inspiral signal

Numerical relativity provides fitting formulas for relations between the binary's components and final masses and spins.



Analysis reveals that GW150914 inspiral and post-inspiral have significant region of overlap.

Parameterized deviations from GR

Look for possible departures from GR, parameterized by set of testing coefficients.

Orbital phase during inspiral is function of ever increasing orbital speed:

$$\Phi(v) = \left(\frac{v}{c}\right)^{-5} \left[\varphi_{0\mathrm{PN}} + \varphi_{0.5\mathrm{PN}} \left(\frac{v}{c}\right) + \varphi_{1\mathrm{PN}} \left(\frac{v}{c}\right)^{2} + \ldots + \varphi_{2.5\mathrm{PN}^{(l)}} \log\left(\frac{v}{c}\right) \left(\frac{v}{c}\right)^{5} + \ldots + \varphi_{3.5\mathrm{PN}} \left(\frac{v}{c}\right)^{7}\right]$$

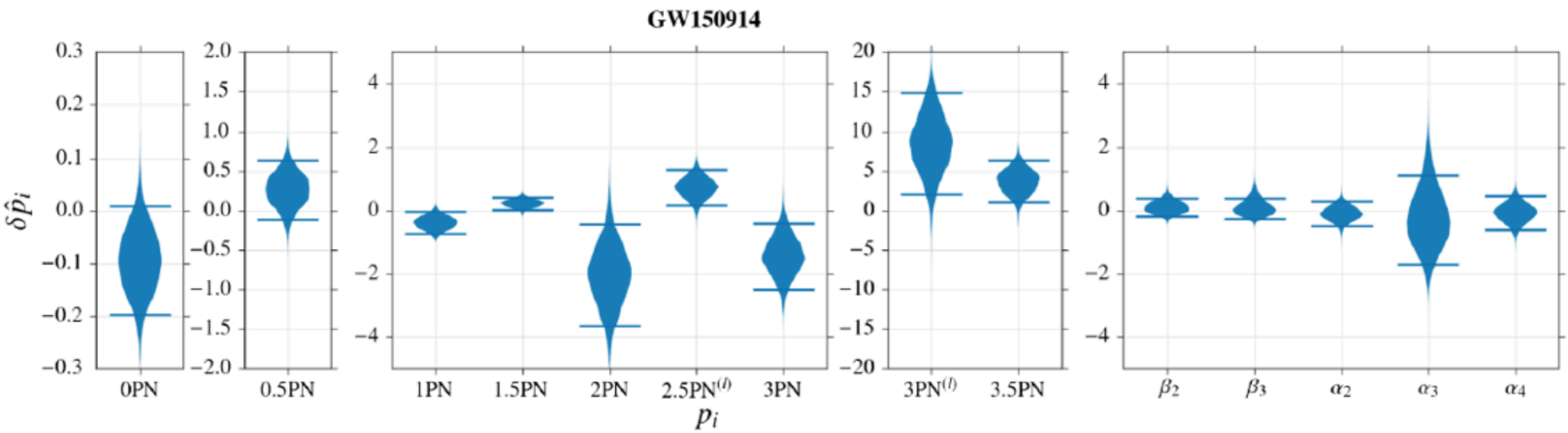
In GR, these have known functions.

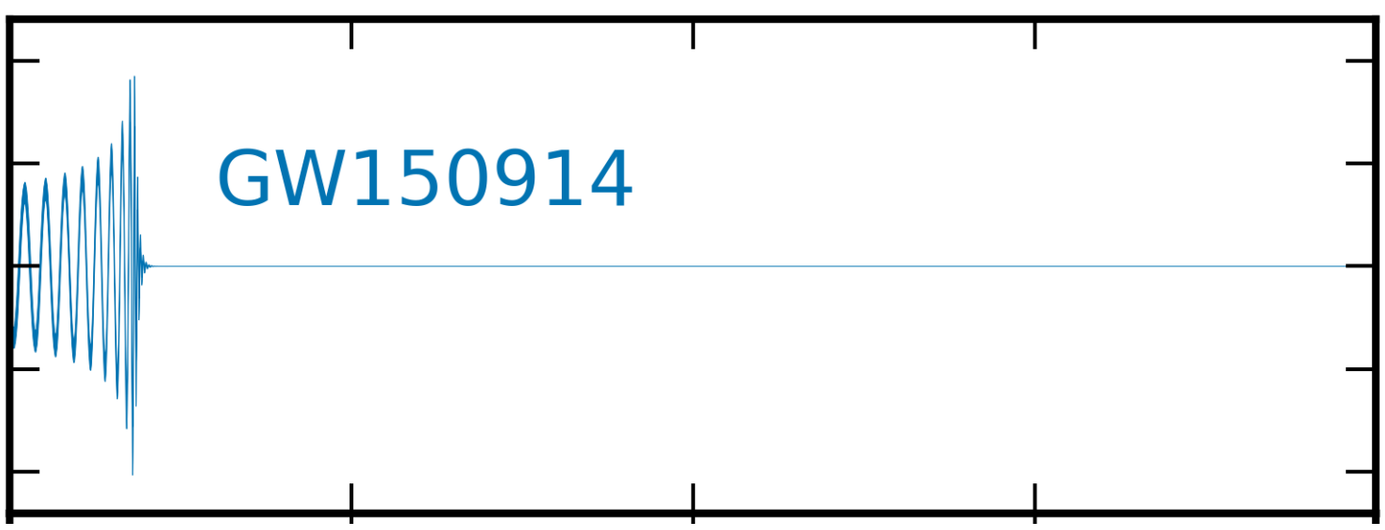
Orbital phase between inspiral and merger-ringdown parameterized by β_j.

Orbital phase of merger-ringdown parameterized by α_j.

Parameterized deviations from GR

Look for possible departures from GR, parameterized by set of testing coefficients.

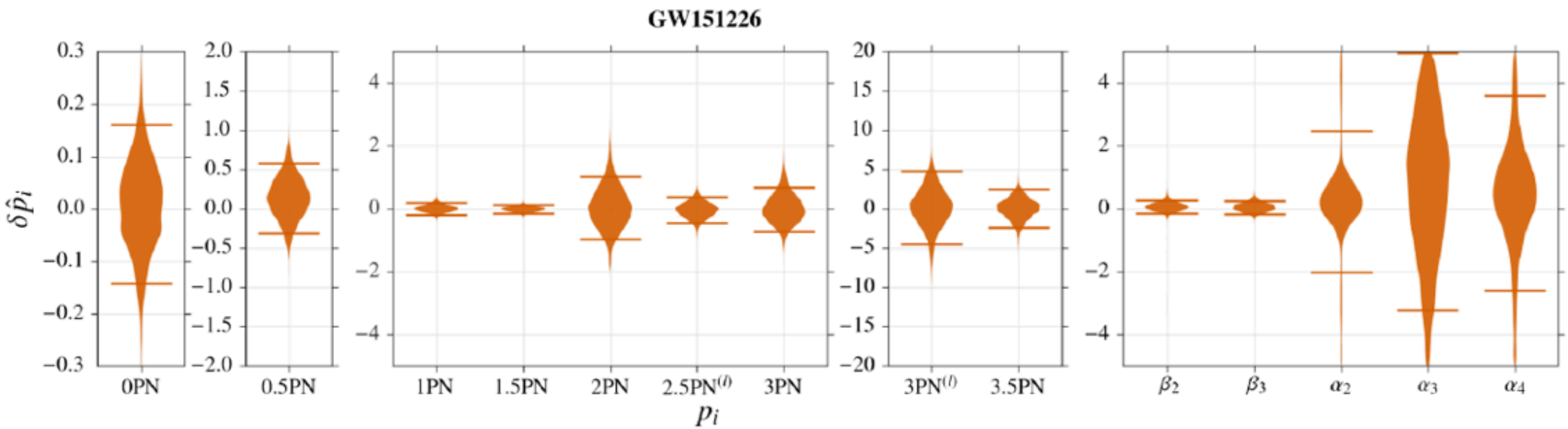


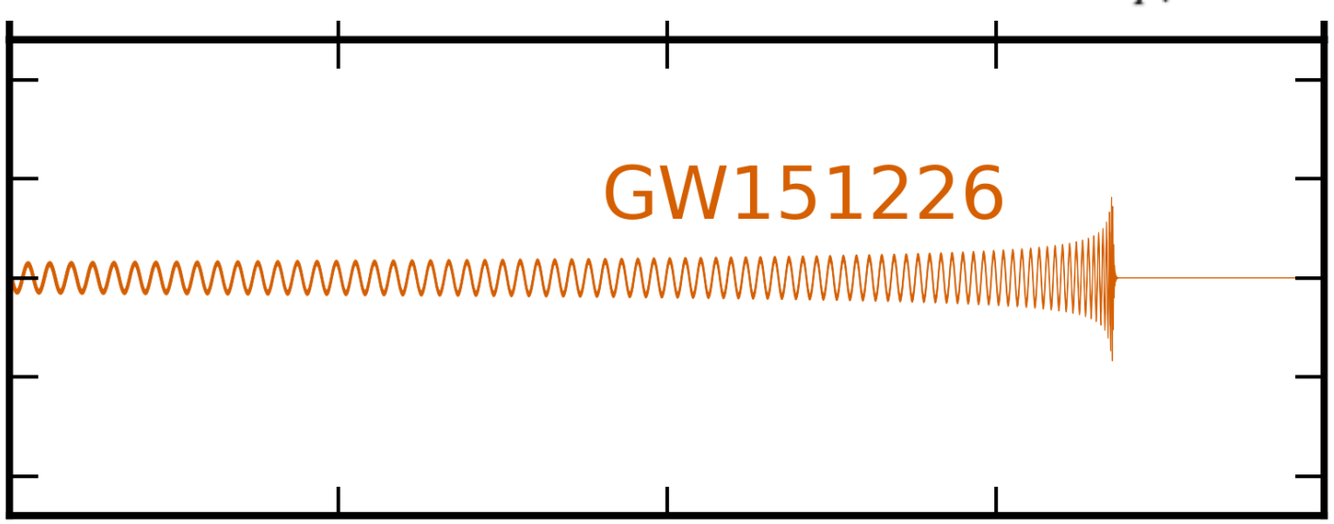


GW150914 provided probe of late inspiral and merger.

Parameterized deviations from GR

Look for possible departures from GR, parameterized by set of testing coefficients.

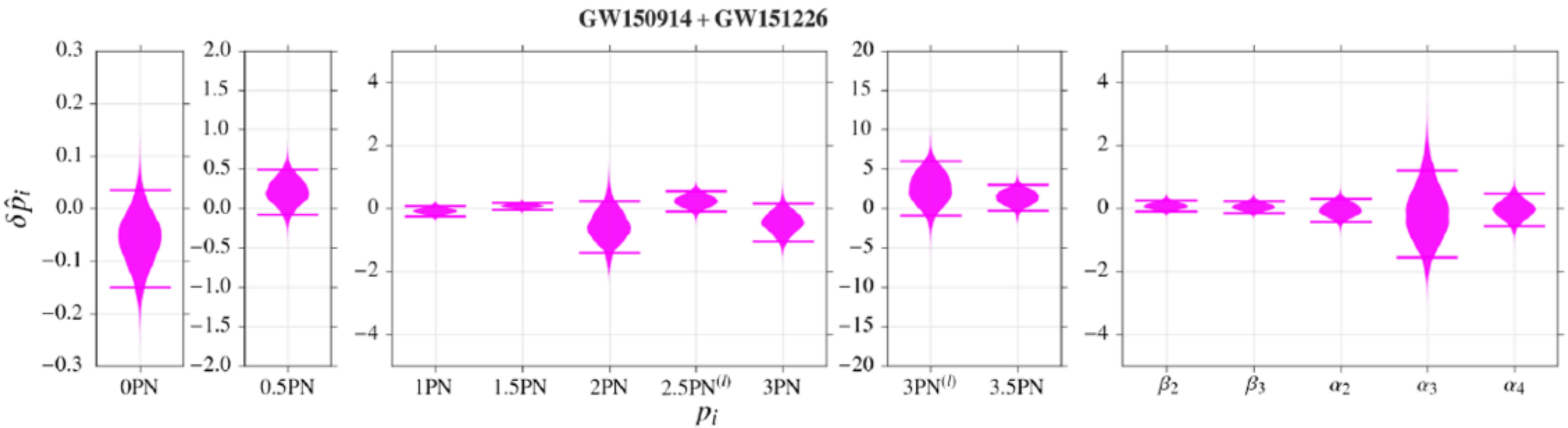




GW151226 provided opportunity to probe PN inspiral with many more waveform cycles.

Parameterized deviations from GR

Look for possible departures from GR, parameterized by set of testing coefficients.



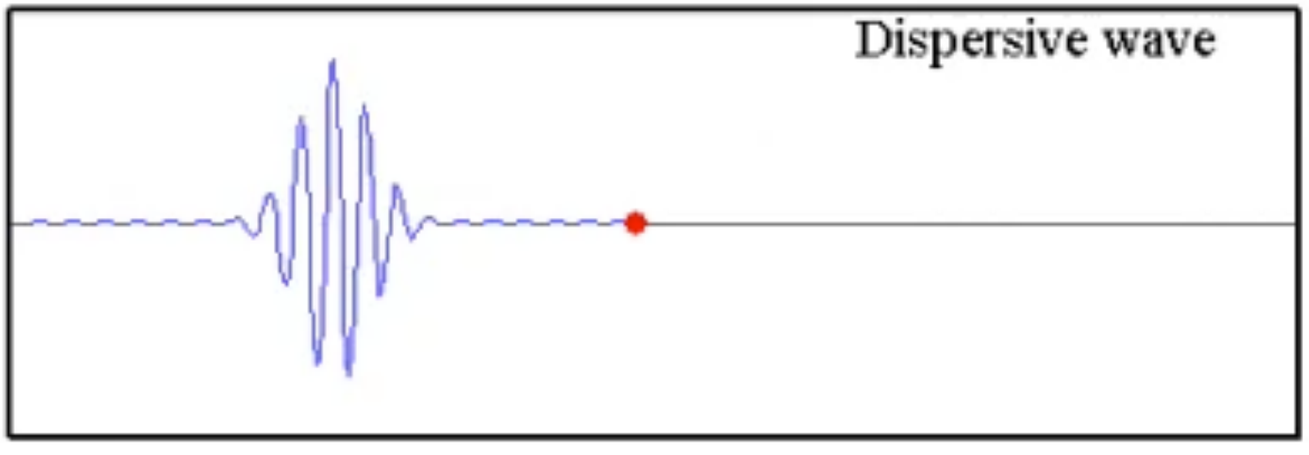
Posterior distributions for deviations can be combined to yield stronger constraints.

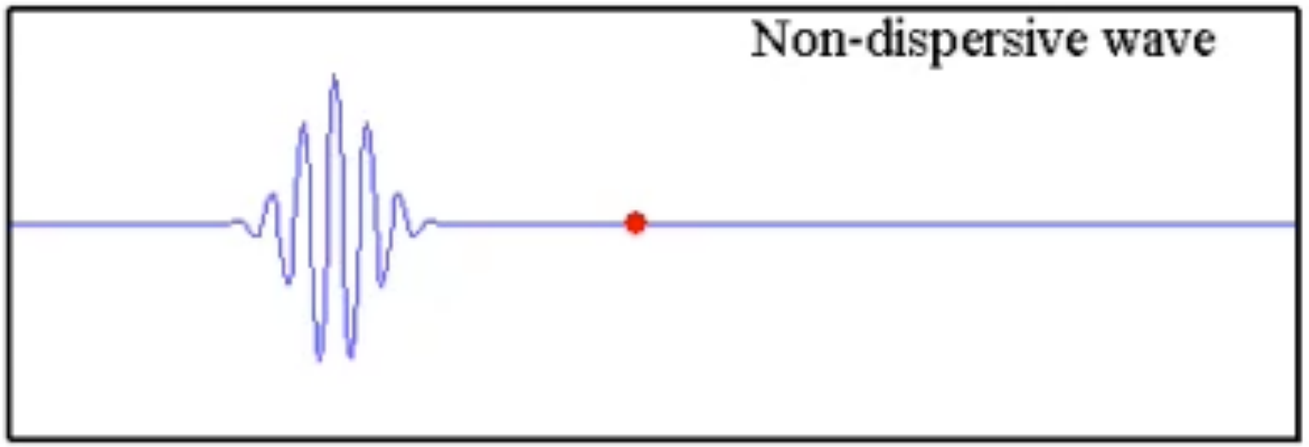
No evidence for disagreement with predictions of GR.
Accuracies will improve with $\sqrt{N}$.

Constraints on Lorentz violations

In GR, GWs are nondispersive.

But modifications to the dispersion relation can arise in theories that include violations of local Lorentz invariance.





isvr

Thus, modified propagation of GWs can be mapped to Lorentz violation.

Constraints on Lorentz violations

Consider modified dispersion relation of the form:

$$E^2 = p^2c^2 + Ap^\alpha c^\alpha$$

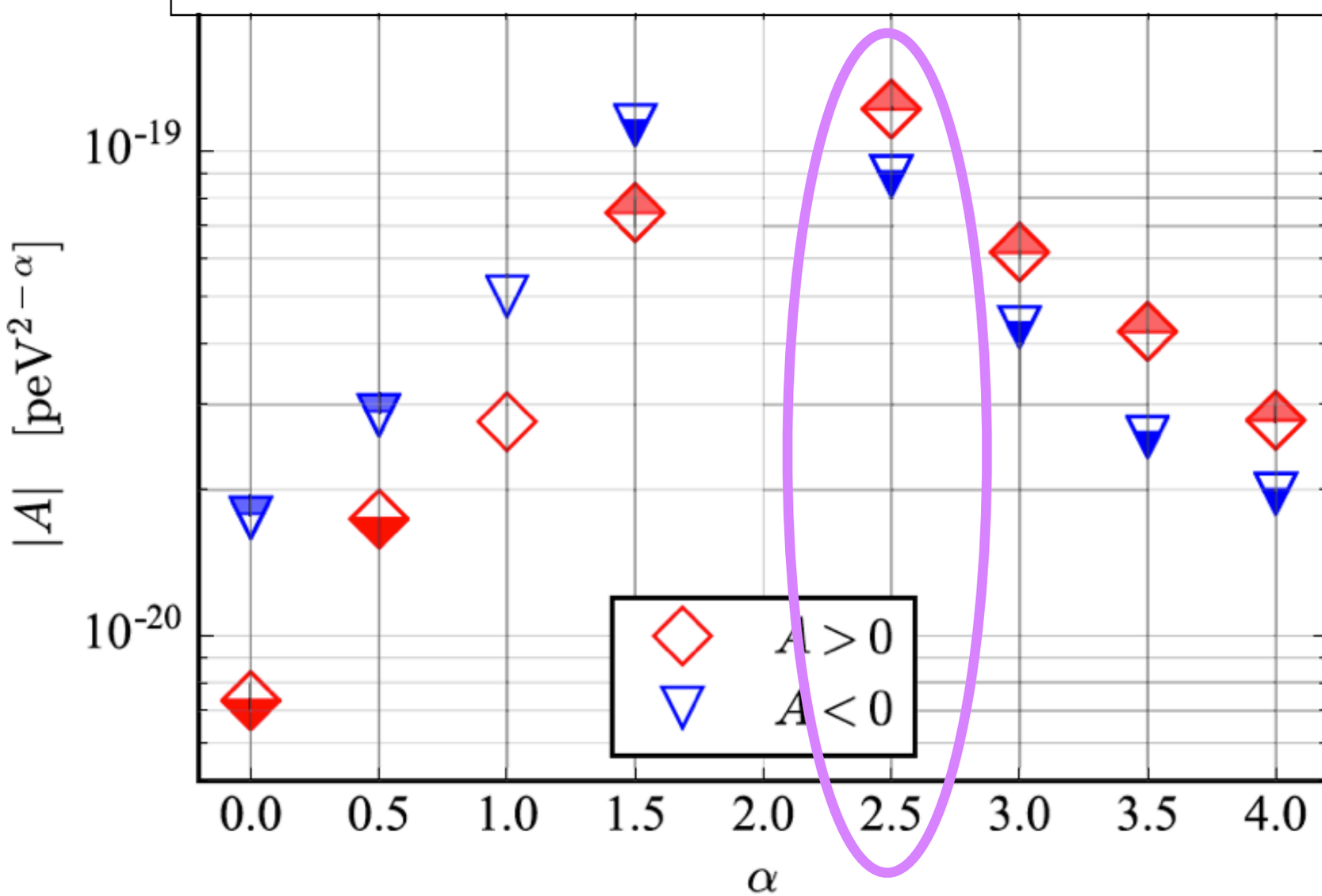
A - amplitude of dispersion. GR predicts A=0.

Several modified theories of gravity predict specific values of $\alpha \geq 0$.

Dispersion occurs during propagation of GW toward Earth. GW170104 provides the best constraint since it was the furthest signal so far. Redshift ~0.2.

Constraints on Lorentz violations

Combined posterior of GW150914, GW151226, and GW170104



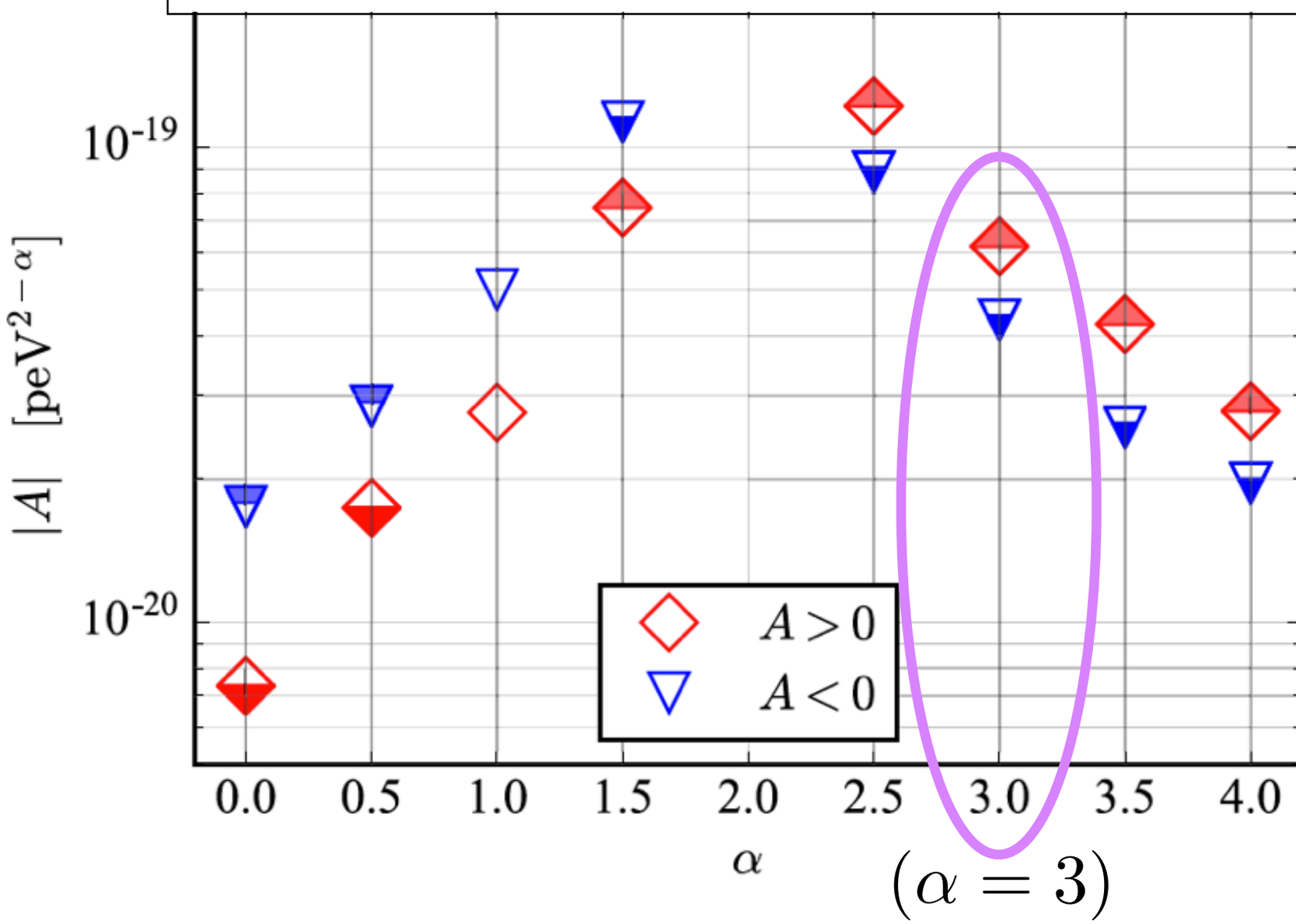
$(\alpha = 2.5)$

multifractal spacetime

LVC, PRL 118, 221101 (2017).

Constraints on Lorentz violations

Combined posterior of GW150914, GW151226, and GW170104

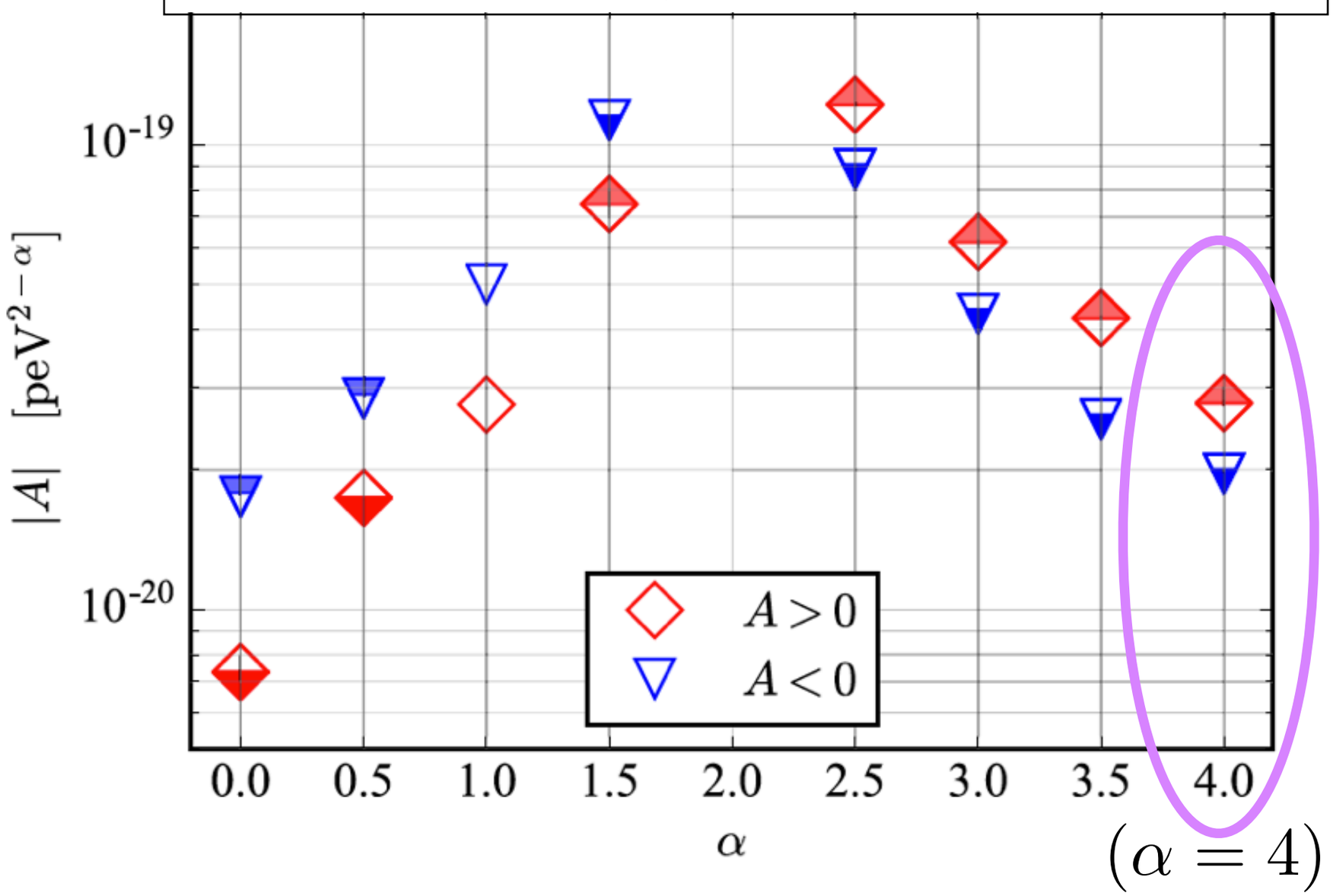


doubly special relativity

LVC, PRL 118, 221101 (2017).

Constraints on Lorentz violations

Combined posterior of GW150914, GW151226, and GW170104

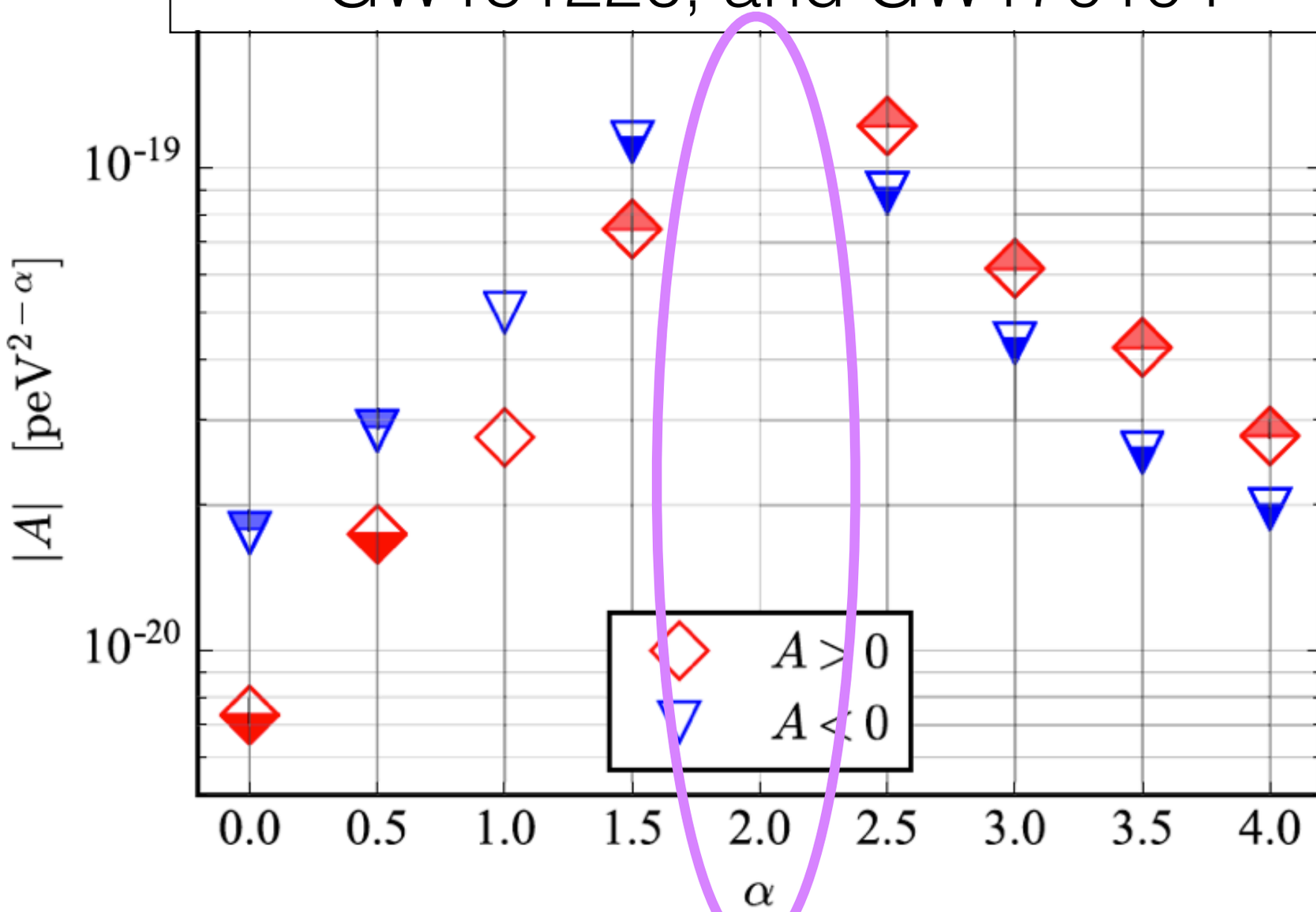


extra-dimensional theories

LVC, PRL 118, 221101 (2017).

Constraints on Lorentz violations

Combined posterior of GW150914, GW151226, and GW170104

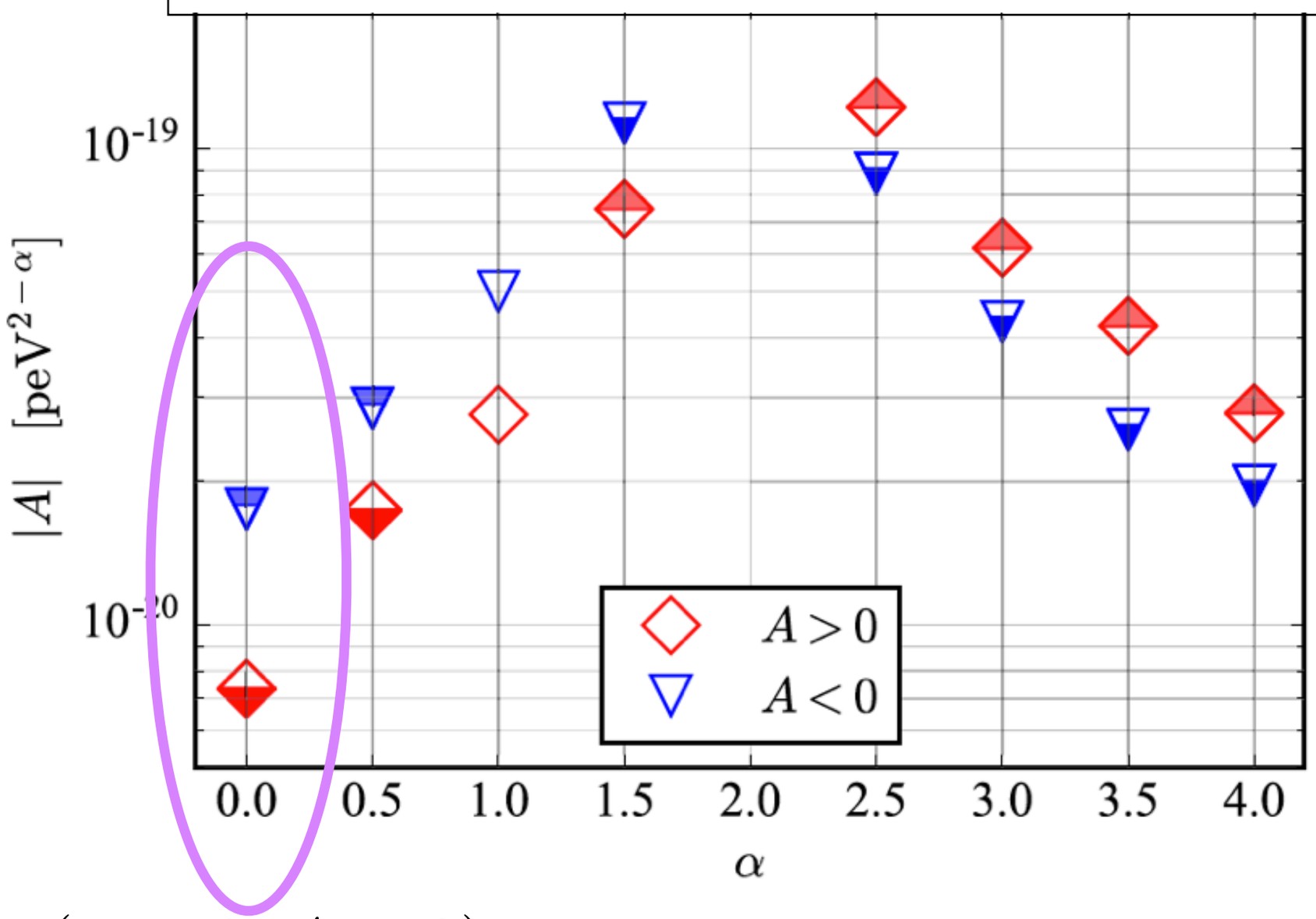


$(\alpha = 2)$ degenerate with arrival time of signal

LVC, PRL 118, 221101 (2017).

Constraints on Lorentz violations

Combined posterior of GW150914, GW151226, and GW170104



$(\alpha = 0, A > 0)$ massive-graviton theories

LVC, PRL 118, 221101 (2017).

Massive graviton

$(\alpha = 0, A > 0)$ - reparameterized to derive lower bound on graviton Compton wavelength

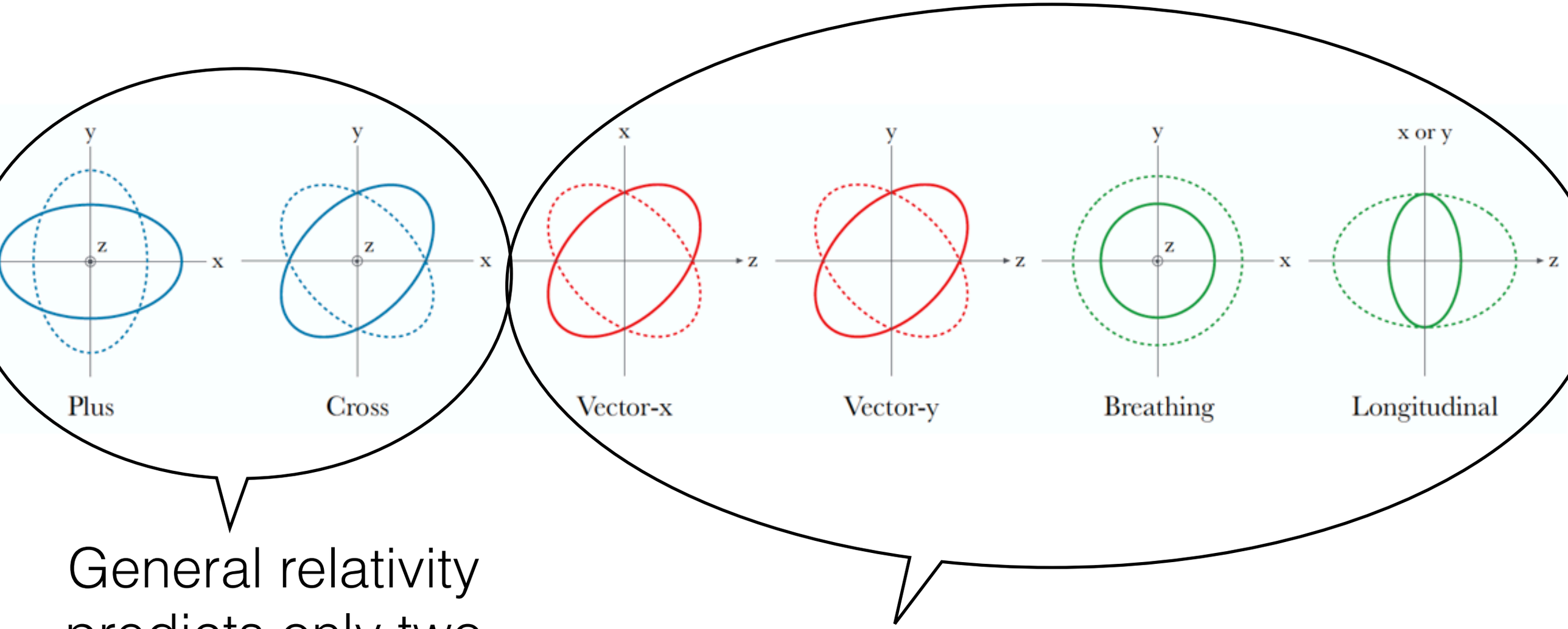
Finite Compton wavelength ⇒ nonzero mass

$$\lambda_g > 1.6 \times 10^{13}\,\mathrm{km}$$

$$m_g \leq 7.7 \times 10^{-23}\,\mathrm{eV}/c^2$$

LVC, PRL 118, 221101 (2017).

Gravitational-wave Polarizations



General relativity predicts only two tensor GW polarizations.

Alternate theories allow for up to four additional vector and scalar modes.

In principle, full generic metric theories predict any combination of tensor, vector or scalar polarizations.

Gravitational-wave Polarizations

Two LIGO detectors are almost aligned so they can't really give us information on other polarizations.

With Virgo, we get a little more information.

Consider models where polarization states are pure tensor, pure vector, or pure scalar.

Bayes' factors for GW170814 (triple BBH):

$$\frac{P(\theta|\text{tensor})}{P(\theta|\text{vector})} = 200 \qquad \frac{P(\theta|\text{tensor})}{P(\theta|\text{scalar})} = 1000$$

Network of at least six detectors is required to determine the polarization content of GW transient.

LVC, PRL 119, 141101 (2017).