Selection Effects in Gravitational-Wave Observations

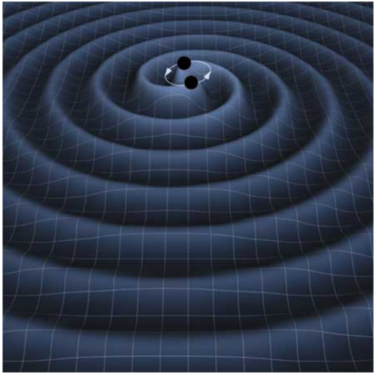
Kipp Cannon

Theories of Astrophysical Big Bangs 2025
2025-2-17

Overview

- Two problems my students are working on:
  1. **IMAFUKU Hayato**: Testing for non-GR polarization states in gravitational-wave observations of compact object collisions.
     - Master's thesis, University of Tokyo, 2025.
     - Imafuku, H., *et al.*, "Statistical biases in parameterized searches for gravitational-wave polarizations", arXiv:2501.16788 [gr-qc].
  2. **HARADA Reiko**: Inferring the Hubble parameter from gravitational-wave observations of compact object collisions.

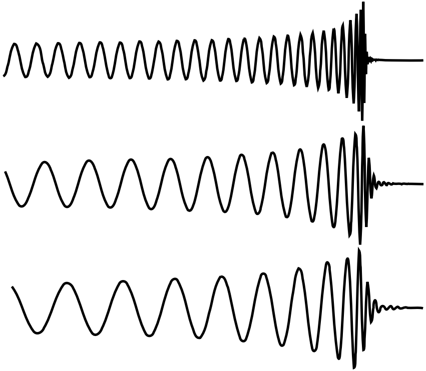
Gravitational Waves From Compact Object Collisions



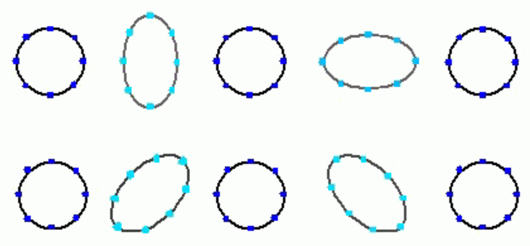
Key Features:

- Very few degrees of freedom compared to electromagnetic sources.
- Amplitude and frequency evolution determined by masses and spins.
- Amplitude inversely proportional to distance.
- Amplitude also depends on geometry factors
- GR predicts two polarization states usually called “plus” and “cross” or “left-handed circular” and “right-handed circular”.

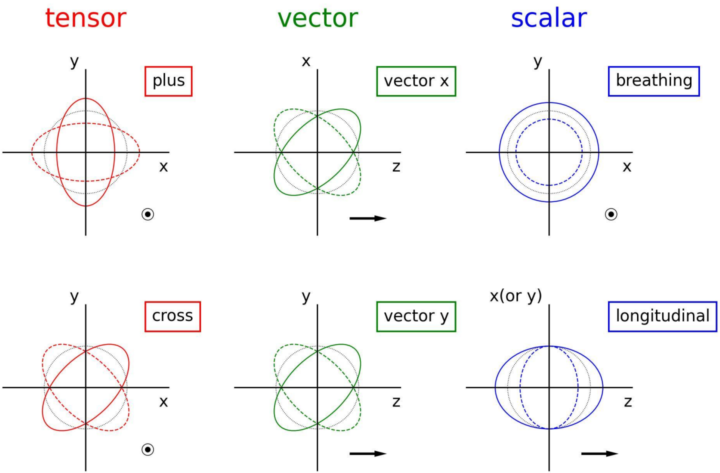
Gravitational Waves From Compact Object Collisions



Testing for non-GR Polarization States



Testing for non-GR Polarization States



Testing for non-GR Polarization States

How to test for non-GR polarizations?

- Each GW detector observes a single polarization component.
- 3 or more GW detectors allows one to test for a 3rd degree of freedom:
  - **Is the 3rd detector's output a linear combination of the other two?**
- Not possible with today's network: LIGO detectors too closely aligned.

Testing for non-GR Polarization States

- If the source radiates energy into more than 2 DOF, then loses energy faster:
  - **Amplitude and frequency evolution differ from GR's prediction.**
- Cannot directly observe all polarizations, but can at least test for the presence of more than 2.
- Takeda, H., *et al.*, "Search for scalar-tensor mixed polarization modes of gravitational waves", Phys. Rev. D, 105(8), 084019, 2022.
- Examined GW170814 (first signal observed with 3 detectors) and GW170817 (highest SNR signal observed up to that time).
- NOTE: examined "good" signals, not all signals.

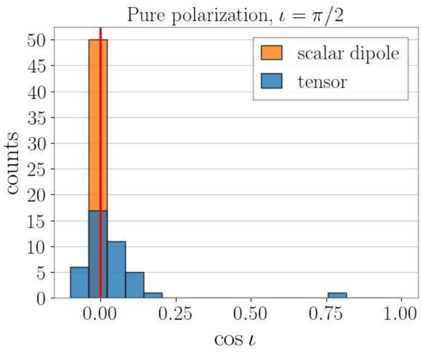
Testing for non-GR Polarization States

From Imafuku, H., *et al.*, "Statistical biases in parameterized searches for gravitational-wave polarizations", arXiv:2501.16788 [gr-qc].

- Earlier work by Takeda, *et al.*, set independent constraints on scalar polarization amplitude for each GW signal.
- Wanted to combine results into a single overall constraint on the scalar component radiated by compact object mergers.
- Were curious about the compatibility of the inferred properties of a merger when observed from different locations.

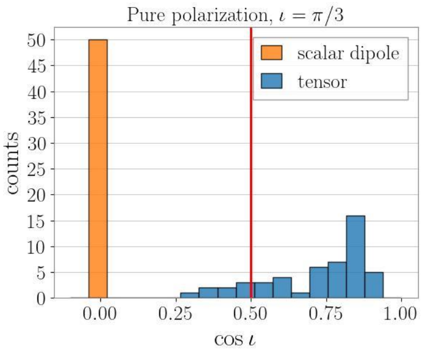
Testing for non-GR Polarization States

From Imafuku, H., *et al.*, "Statistical biases in parameterized searches for gravitational-wave polarizations", arXiv:2501.16788 [gr-qc].



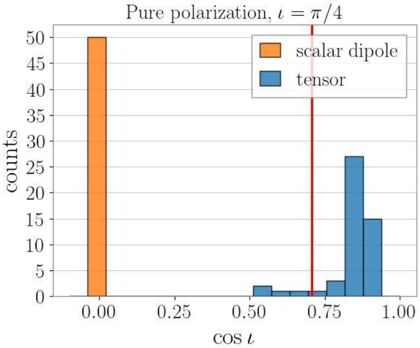
Testing for non-GR Polarization States

From Imafuku, H., *et al.*, “Statistical biases in parameterized searches for gravitational-wave polarizations”, arXiv:2501.16788 [gr-qc].



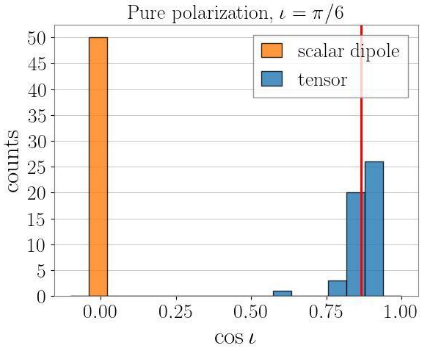
Testing for non-GR Polarization States

From Imafuku, H., *et al.*, "Statistical biases in parameterized searches for gravitational-wave polarizations", arXiv:2501.16788 [gr-qc].



Testing for non-GR Polarization States

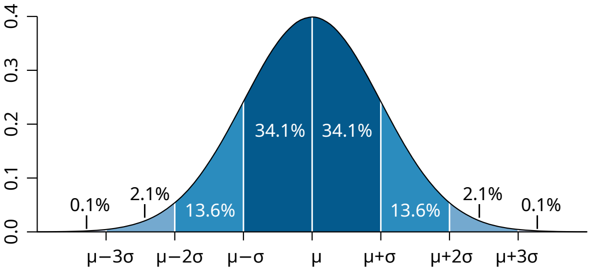
From Imafuku, H., *et al.*, "Statistical biases in parameterized searches for gravitational-wave polarizations", arXiv:2501.16788 [gr-qc].



Testing for non-GR Polarization States

Constructing and Interpreting a “P-P Diagram”

- Imagine we have performed a Bayesian estimation of a parameter and obtained a posterior PDF like

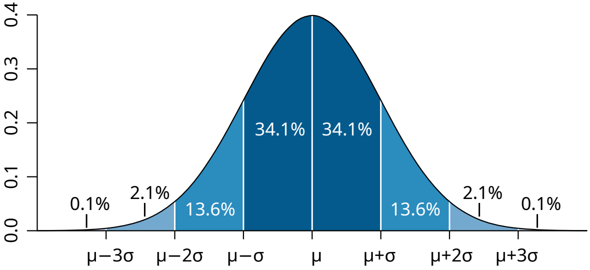


- The area under the curve in some interval tells us the probability with which the true value of the parameter occurs within that interval.

Testing for non-GR Polarization States

Constructing and Interpreting a “P-P Diagram”

- But is our posterior PDF correct!? What does that mean, how do we know?



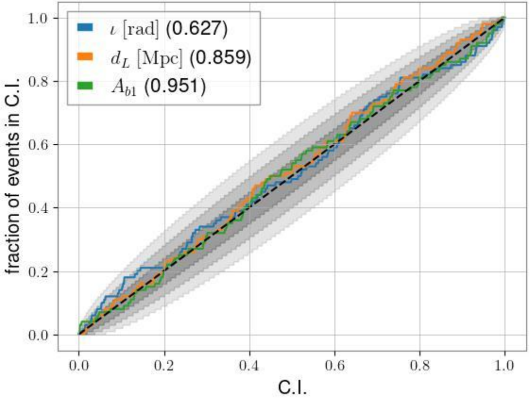
- Obviously, it’s “correct” if the true value really is found within any given interval with the given probability.

Testing for non-GR Polarization States

Constructing and Interpreting a "P-P Diagram"

- We can quantify this by constructing test cases:
  1. Generate a simulated data set (containing a signal whose parameters we know).
  2. Process it as usual, and obtain a posterior PDF.
  3. Choose a variety of intervals, doesn't matter how, and for each record:
     - The area under the curve.
     - a Yes/No answer to the question: is the true value in that interval?
  4. Repeat until we are tired
- Finally plot the fraction of simulations that were found inside the chosen interval as a function of the area under the PDF in that interval.
- If the posterior PDF's coverage is correct, the result is an $x = y$ diagonal line.

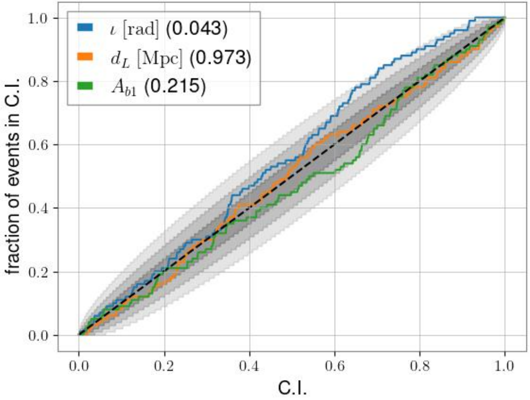
Testing for non-GR Polarization States



- Simulated signals generated by drawing from astrophysical distribution.
- Orbit inclination, distance, and scalar polarization amplitude inferred for all simulations.
- Prior in Bayes' theorem equals astrophysical distribution.

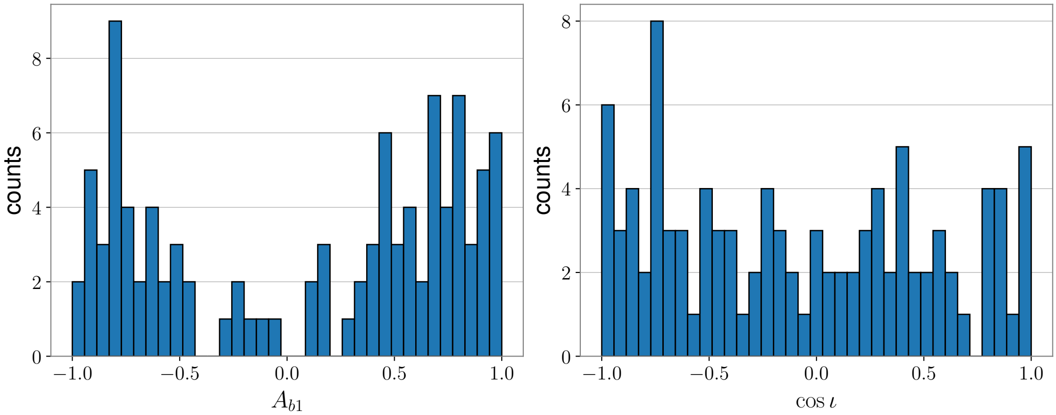
From Imafuku, H., "On Constraining Scalar-Tensor Polarization of Gravitational Waves via Observations of Compact Object Collisions", Master's thesis, University of Tokyo, 2025.

Testing for non-GR Polarization States

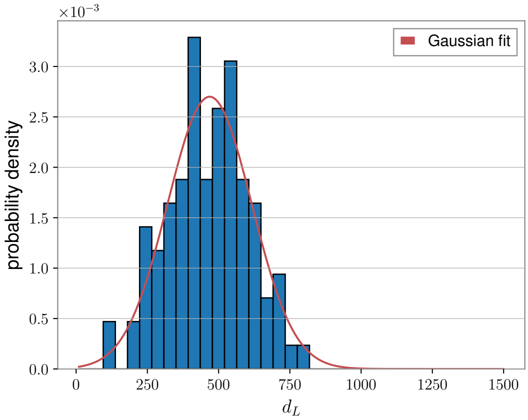


- Simulated signals generated by drawing from astrophysical distribution.
- Select 10% loudest (highest signal-to-noise ratio) simulations, and use only those for inferring parameters.
- Prior in Bayes' theorem still equals astrophysical distribution.

Testing for non-GR Polarization States

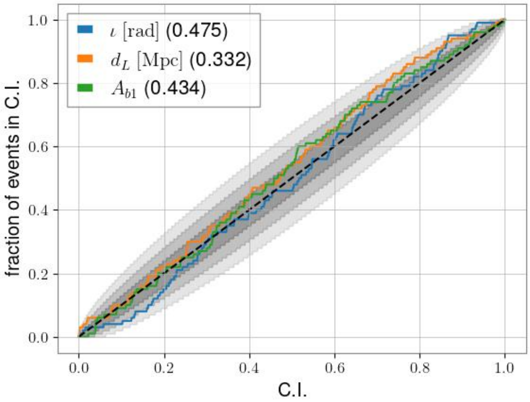


Testing for non-GR Polarization States



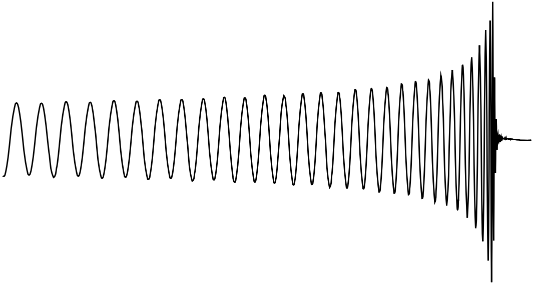
- Intrinsic, astrophysical, distribution is $\propto d_L^2$.
- Selecting loudest subset from a collection of detected signals significantly alters the distribution of d_L.

Testing for non-GR Polarization States



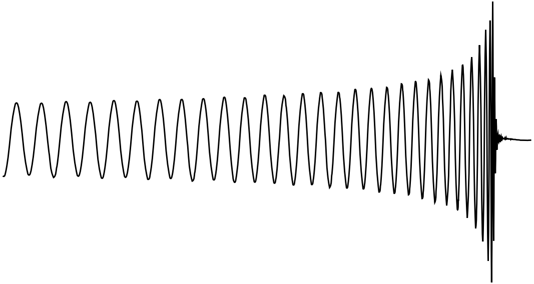
- Simulated signals generated by drawing from astrophysical distribution.
- Select 10% loudest (highest signal-to-noise ratio) simulations, and use only those for inferring parameters.
- Prior in Bayes' theorem modified so that distance prior models selection process (other parameters' priors unchanged).

Measuring Hubble Parameter from Compact Object Mergers



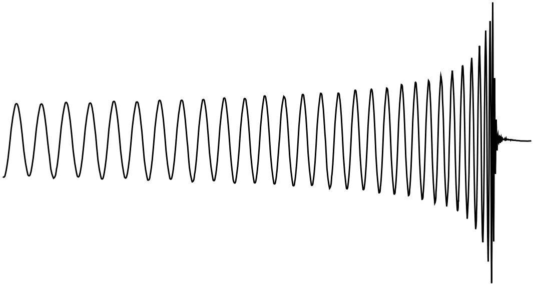
- The GWs radiated by a compact object collision are (almost) completely determined by the masses, spins, orbital parameters. Source composition is not very important.
- The amplitude and frequency evolution both fixed by these same parameters.

Measuring Hubble Parameter from Compact Object Mergers



- As the wave travels to Earth it is Doppler shifted by the expansion of the universe, and the flux density is diminished as $1/r^2$, altering the frequency and amplitude.
- By determining the correction required to put the signal back the way GR says it should be, we obtain a direct measurement of the distance/red-shift relationship — the Hubble parameter.

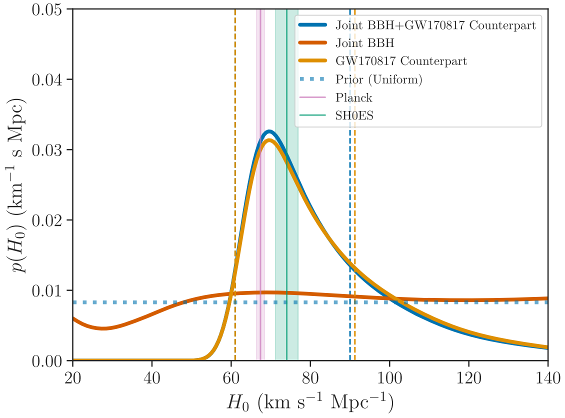
Measuring Hubble Parameter from Compact Object Mergers



- In practice, our inability to precisely determine the properties of the source means we don't really know what the signal is supposed to look like.
- We don't exactly know sky location, orbit inclination, we don't know the geometry factors upon which the amplitude depends.

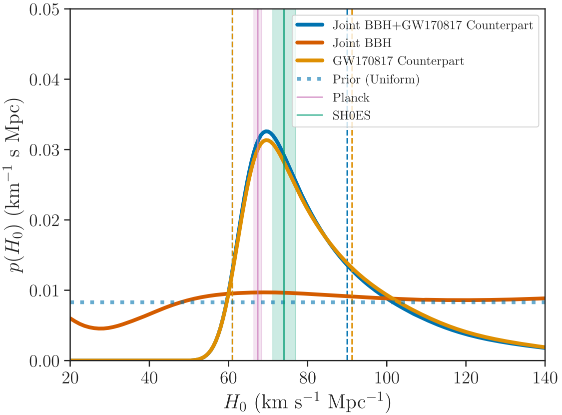
Measuring Hubble Parameter from Compact Object Mergers

- Each compact object merger observation provides a posterior PDF for the Hubble parameter. Observations are independent, therefore joint posterior is product of individual posteriors.
- Requires full Bayesian parameter estimation (PE) analysis for each signal = posterior PDF for each source's masses, spins, orbital parameters, location on sky.



Measuring Hubble Parameter from Compact Object Mergers

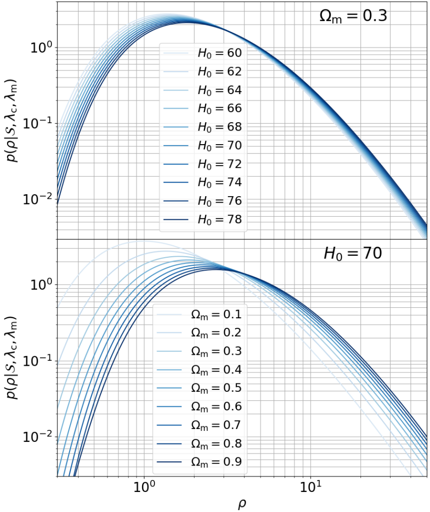
- PE requires computer time, people time: only a subset of signals is used.
- Selecting "good" signals introduces the same selection bias effects observed above.
- Also new risk of a human-induced bias in the final answer.
- If the analysis produces an unpopular H_0, does the selection criteria get adjusted and people try again?



Measuring Hubble Parameter from Compact Object Mergers

- Harada-san's approach: only use information available directly from the detection pipeline.
- By removing the requirement for full parameter estimation analysis, the need to set a selection criterion is removed. Full GW signal catalogue is used.
- Maybe less informative, but removes the need for parameter estimation analyses, removes the need to choose good signals. We use the entire contents of the GW catalogue.

Measuring Hubble Parameter from Compact Object Mergers



- For BBHs (no EM counterpart) mostly the Hubble parameter is being estimated from the SNR distribution of the observed signals, which is information the detection pipeline is perfectly good at supplying directly.
- It's a *tiny* effect in today's detectors.
- Selection of "good" events can easily overwhelm the effect.

Conclusions

Public Service Announcement:

- **Do not assume that your statistical inference results are correct.**
- Especially don't assume that because the code ran without crashing then it must be correct!
- *Test your parameter estimation.* Confirm your posteriors have the correct coverage. Demonstrated that your P-P diagram is diagonal.
- In my experience, I have never, ever, seen Bayesian parameter estimation work properly the first try. There have always been effects that were assumed would be small but proved to be more significant than anticipated.