Quantitative measurement of model differences for CBCs

Qian Hu, John Veitch @ University of Glasgow

Mar 14 2022, LVK meeting

LIGO-G2200415

Overview

❑ A new approach to evaluate GW waveform accuracy

- By looking into difference between two waveform models
- Free from the unknown true waveform

❑ Applied to…

- GWTC-3 and GWTC-2.1 PE samples: How was IMRPhenomXPHM and SEOBNRv4PHM's performance? E.g., did they generate faulty waveforms for the extreme-mass-ratio event GW191219_163120?
- The relation between waveform difference and posterior difference
- Simulations: Good and bad regions in the parameter space & future accuracy requirements

Assessment of one waveform model

- Can detectors distinguish it from the real one?

- “Accurate enough”: the detector can not distinguish it from the real waveform
- Construct such a waveform family for plus polarization: (Lindblom+, Phys. Rev. D **78**, 124020, 2008)

$$H_1^+(\lambda) = (1-\lambda)h_0^+ + \lambda h_1^+ = h_0^+ + \lambda\delta h_1^+, \quad 0 < \lambda < 1,$$ h0: real waveform, h1: model waveform

- Distinguishing waveforms <=> measuring λ

$$\sigma_\lambda^{-2} = \left(\frac{\partial h^+}{\partial\lambda} \mid \frac{\partial h^+}{\partial\lambda}\right) = (\delta h_1^+ \mid \delta h_1^+).$$

$$(a \mid b) = 4\int_0^{+\infty} \frac{a^*(f)b(f)}{S_n(f)}df$$

- If the error of measuring λ is greater than its domain of definition (also the parametric distance between real and model waveforms), the detector can not distinguish

$$\|\delta h_1^+\|^2 = (\delta h_1^+ \mid \delta h_1^+) < 1.$$

- It shows: waveform error should lie within a unit ball in the inner-product space
- Note: when calculating the inner product, we need to minimize it over an arbitrary phase φ_0 and a time shift t_0, in order to eliminate the non-physical difference between models

Assessment of one waveform model

- A single polarization vs detector response

$$\|\delta h_1^+\|^2 = (\delta h_1^+ \mid \delta h_1^+) < 1.$$

- We can extend it to the detector response: the radius of the ball is weighted

$$h_0 = F_+ h_0^+ + F_\times h_0^\times$$
$$h_1 = F_+ h_1^+ + F_\times h_1^\times$$

$$\begin{aligned}\|\delta h_1\| &= \|F_+ \delta h_1^+ + F_\times \delta h_1^\times\| \\ &\leq \|F_+ \delta h_1^+\| + \|F_\times \delta h_1^\times\| \\ &< |F_+| + |F_\times|.\end{aligned}$$

Triangle inequality in the inner-product space

- To evaluate the waveform itself: use one polarization (or $h_+ - ih_\times$ etc)
- To evaluate the hypothetical signals we used in data analysis, use detector response: the waveform errors are weighted by antenna response functions when projecting waveforms to the detector

Assessment of waveform pair

- Eliminate the unknown real waveform

$$\|\delta h_1^+\|^2 = (\delta h_1^+ \mid \delta h_1^+) < 1. \qquad \delta h_1^+ = h_1^+ - h_0^+$$

- The calculation of δh_1^+ needs the real waveform, which we don't know
- Use Numerical Relativity (NR) simulations as real waveform, but the number of NR simulations is limited
- Introduce another waveform model h_2, pair it with h_1

$$\begin{aligned} \Delta^+ &= \delta h_1^+ - \delta h_2^+ \\ &= (h_1^+ - h_0^+) - (h_2^+ - h_0^+) \\ &= h_1^+ - h_2^+. \end{aligned}$$

Real waveform is cancelled!

- Assume two waveforms are both accurate enough, we have

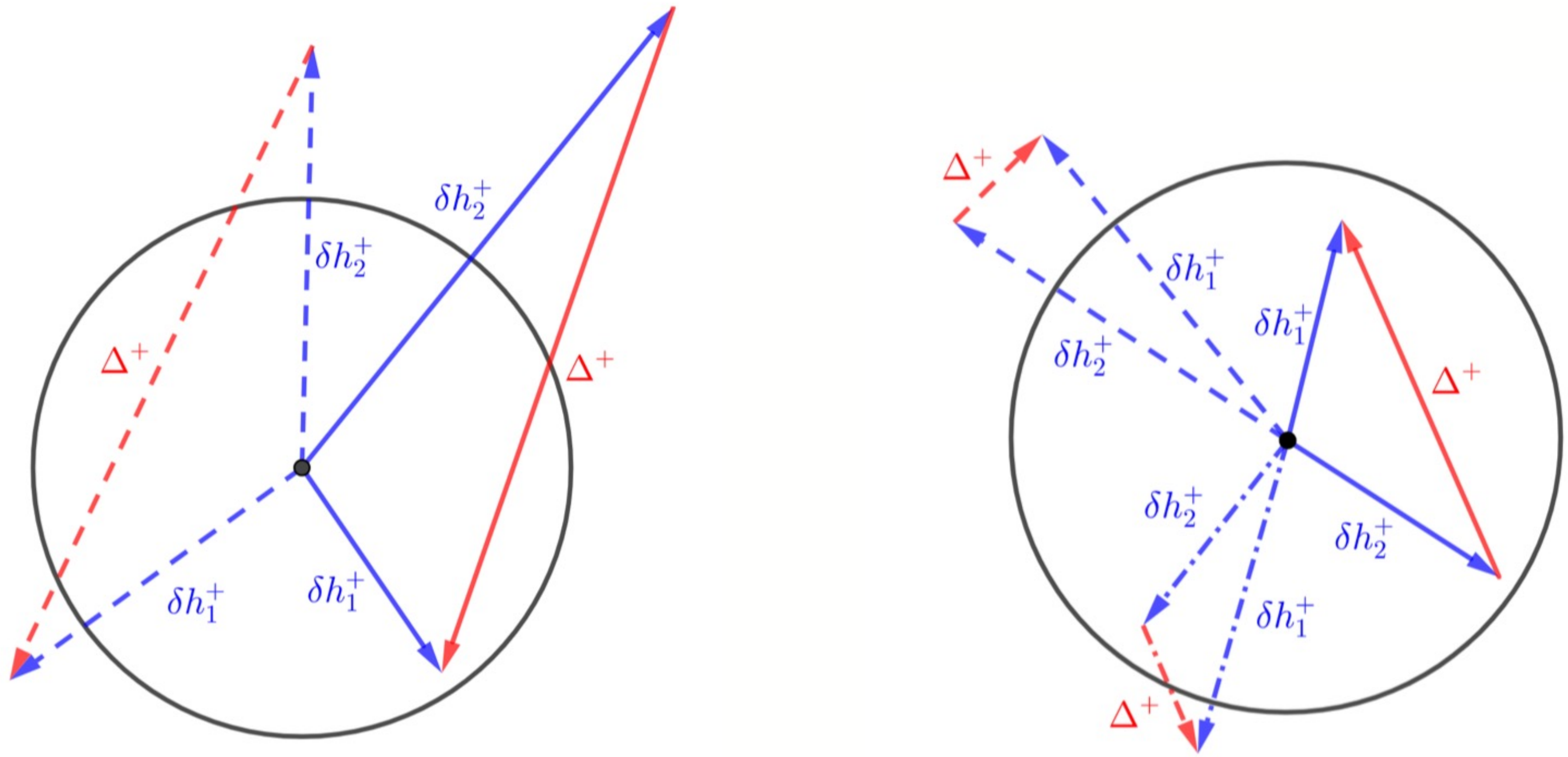
$$\|\Delta^+\| \leq \|\delta h_1^+\| + \|\delta h_2^+\| < 2.$$

- If we find $|| \Delta^+ || > 2$, at least one of the waveforms is not accurate enough. It's a necessary condition of "a pair of waveform models are both accurate".

Assessment of waveform pair

- An illustration of all possible cases

- If we find $|| \Delta^+ || > 2$, at least one of the waveforms is not accurate enough
- Even though we have got $|| \Delta^+ ||$, we don't know the real situation (possibilities are plotted in different line styles.)



Case I: $\boldsymbol{\Delta^+ > 2}$ **Case II:** $\boldsymbol{\Delta^+ < 2}$

Assessment of waveform pair

$$\|\Delta^{+}\| \leq \|\delta h_1^{+}\| + \|\delta h_2^{+}\| < 2.$$

- Extend to detector response:

$$\|\Delta\| \leq \|\delta h_1\| + \|\delta h_2\| < 2(|F_{+}| + |F_{\times}|).$$

- Extend to detector network:

$$\mathbf{C} = (\mathbf{D}|\mathbf{B}) \Rightarrow C_{jk} = \sum_{p=1}^{n} (D_{jp} \mid B_{pk})$$

$$\|\boldsymbol{\Delta}_{\mathrm{det}}\| = (\delta\mathbf{h}^{\mathbf{T}}|\delta\mathbf{h}) = \sum_k (\delta h^{(k)}|\delta h^{(k)}) = \sum_k \left(\Delta^{(k)}\right)^2 < 2\sum_k (|F_{+}^{(k)}| + |F_{\times}^{(k)}|).$$

- To sum up:

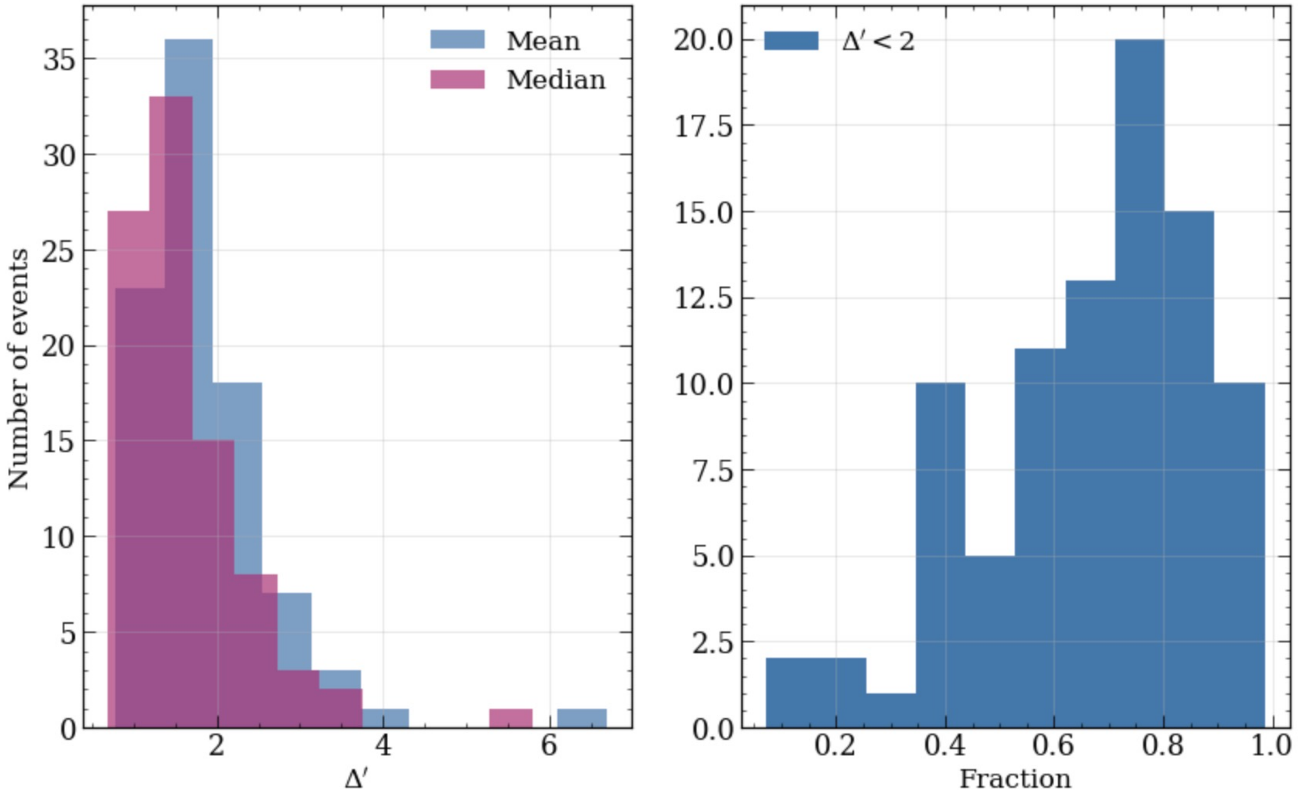
$$\Delta'^{(k)} = \frac{\Delta^{(k)}}{|F_{+}^{(k)}| + |F_{\times}^{(k)}|}, \quad \boldsymbol{\Delta}'_{\mathrm{det}} = \frac{\boldsymbol{\Delta}_{\mathrm{det}}}{\sum_k (|F_{+}^{(k)}| + |F_{\times}^{(k)}|)}$$

They should be less than 2 if both models are accurate!

Applying to PE samples

- Overview: histograms

- For each event, calculate Δ'_{net} for the mixed posterior samples from IMRPhenomXPHM & SEOBNRv4PHM
- Calculate mean, median of Δ'_{net} for each event (left panel)
- Calculate fraction of $\Delta'_{net} < 2$ samples for each event (right panel)
- There are several events having "worse" performance compared to the others

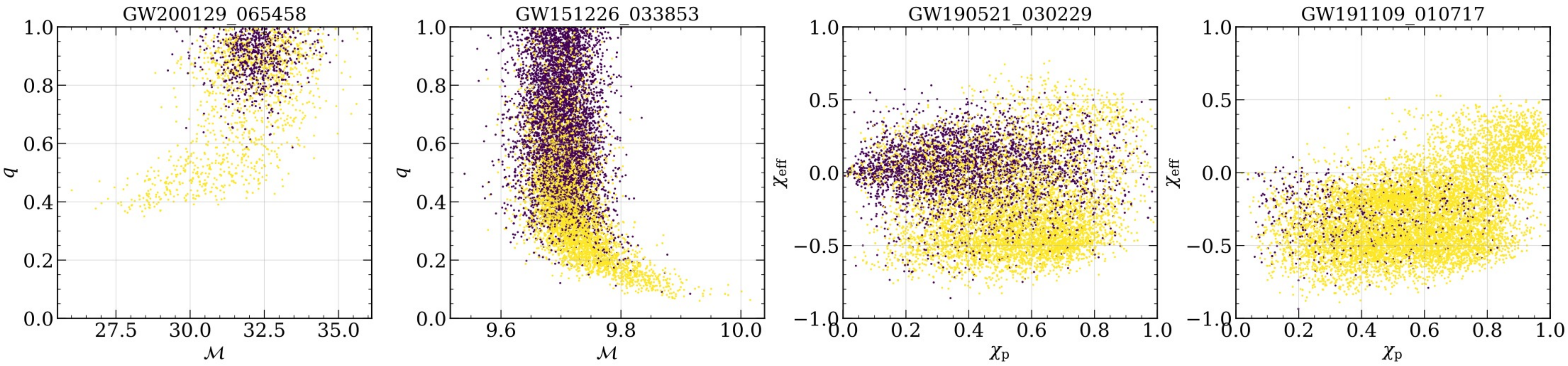


$$\Delta'_{\mathrm{det}} = \frac{\Delta_{\mathrm{det}}}{\sum_k (|F_+^{(k)}| + |F_\times^{(k)}|)}$$

Applying to PE samples

- Overview: distribution in mass and spins

- Yellow points: $\Delta'_{net} > 2$ samples
- Purple points: $\Delta'_{net} < 2$ samples
- Accuracy becomes worse when mass ratio decreases or spins increase



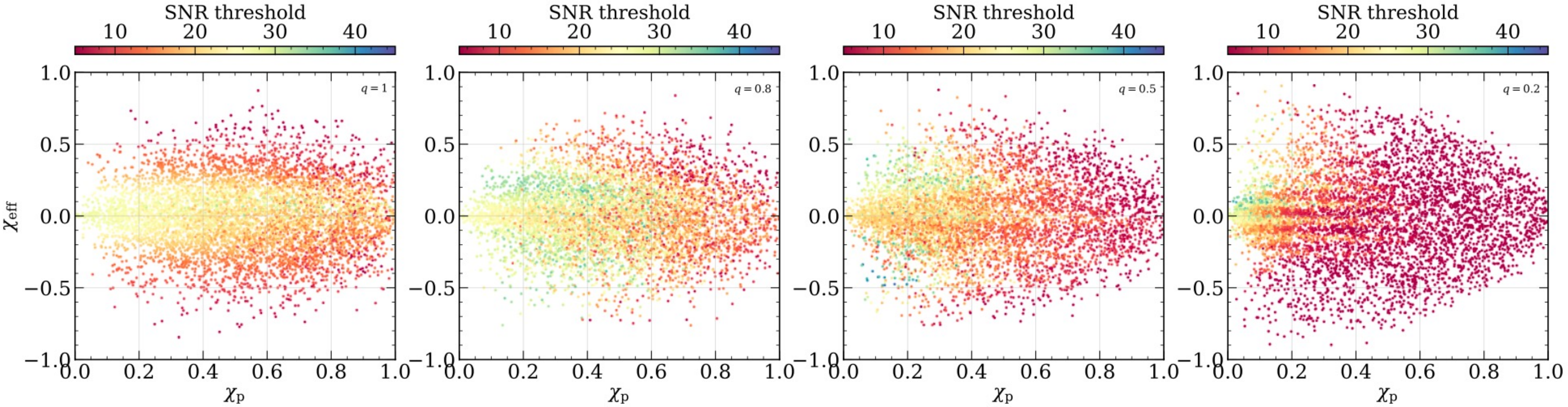
chirp mass - mass ratio

precession spin - effective spin

BBH Simulations

- IMRPhenomXPHM and SEOBNRv4PHM

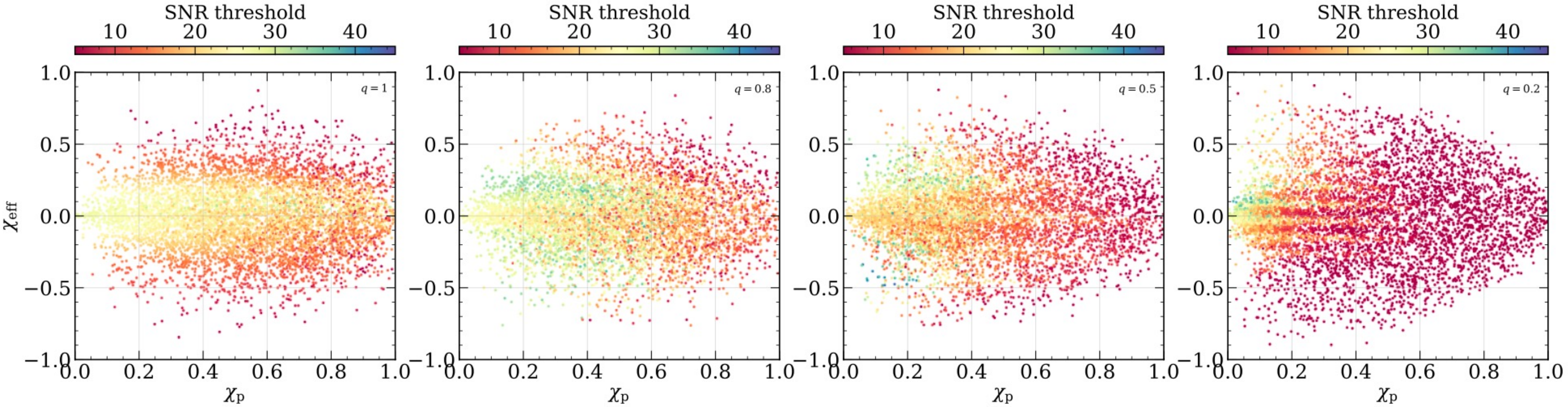
- $m_1 = 30M_{\odot}, q = 1, 0.8, 0.5, 0.2$
- Spins are randomly generated (isotropic, uniform between 0 and 1)
- SNR threshold: SNR when waveform difference reaches upper limit 2



BBH Simulations

- IMRPhenomXPHM and SEOBNRv4PHM

- Waveform accuracy deteriorates as spin goes up or mass ratio goes down
- In some cases, SNR threshold drops below 5
- Using $\Delta \propto SNR$, for 3rd-gen detectors (SNR 30~1000), the model mismatch from true waveform should be improved by 3-4+ orders of magnitude (consistent with Pürrer+, Phys. Rev. Research **2**, 023151)

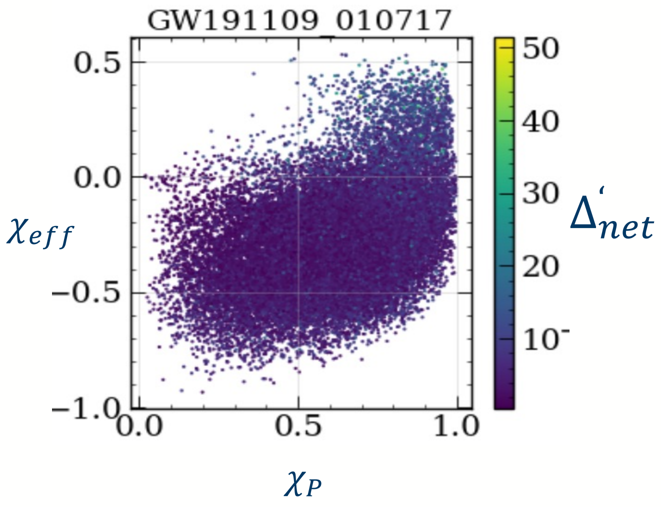


Applying to O3b PE samples

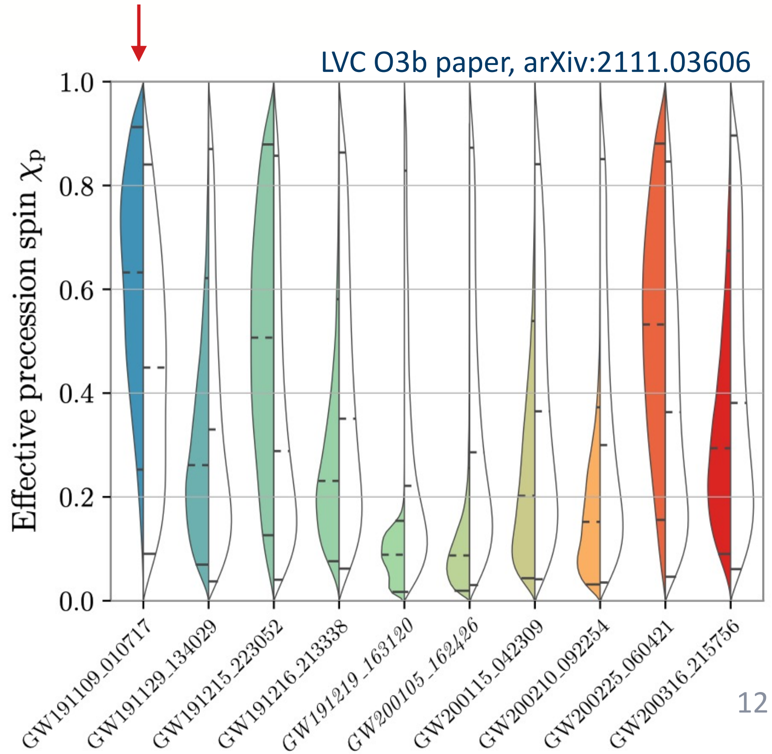
- Two “bad” events in O3b: GW191109 and GW200129

GW191109

- High mass BBH (Mchirp ~ 50Msun)
- Has the smallest (negative) χ_{eff} in O3b catalog, “where waveform differences may be expected” – O3b paper, also arXiv:2010.05830, arXiv:2106.06492
- Also has large χ_P
- Large difference in waveforms means the two waveforms can not be both accurate

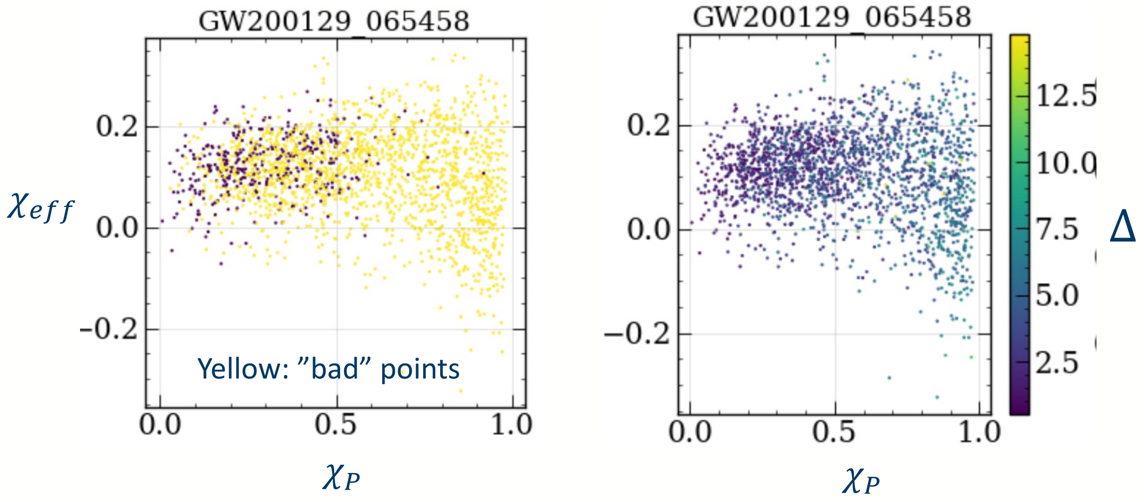


LVC O3b paper, arXiv:2111.03606



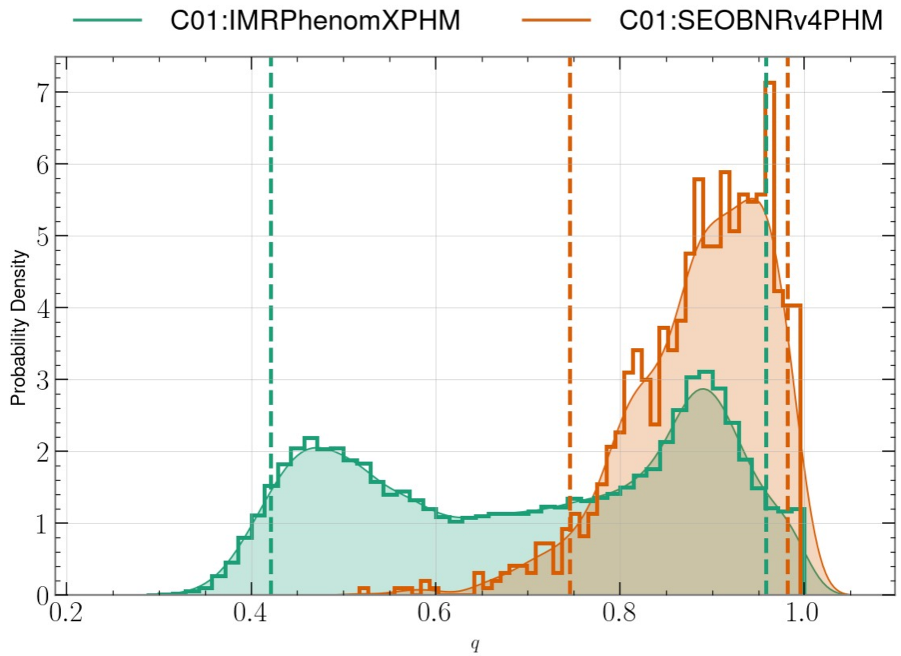
Applying to O3b PE samples

- Two “bad” events in O3b: GW191109 and GW200129



GW200129

- The highest SNR in O3b Catalog (SNR~26.8)
- The highest inferred χ_P
- PE results showed difference between two waveform models

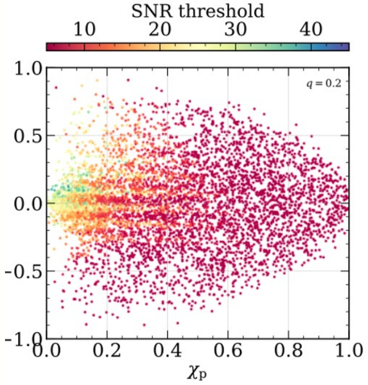


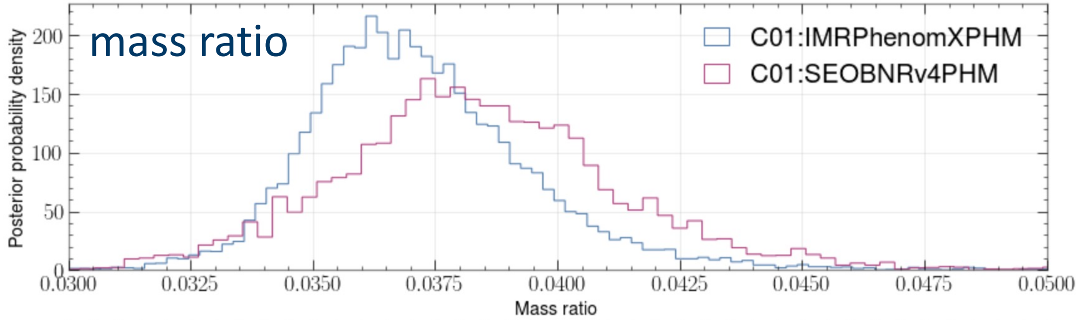
Applying to O3b PE samples

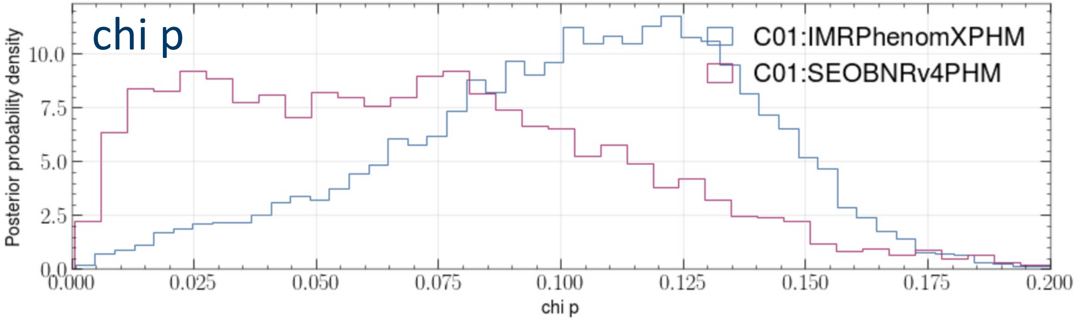
- “Extreme”-mass-ratio event GW191219

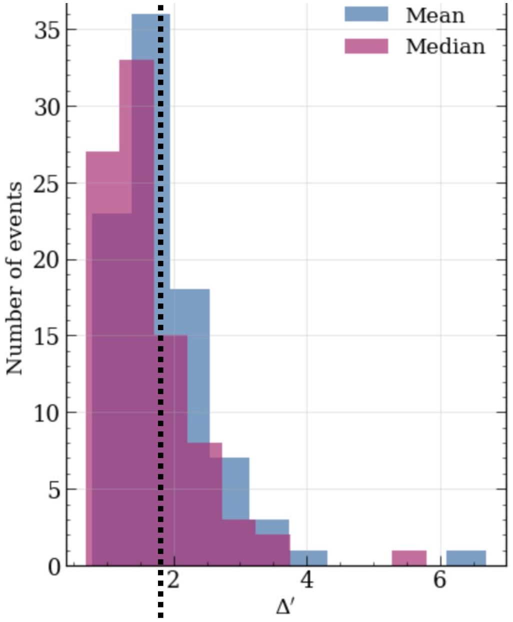
GW191219

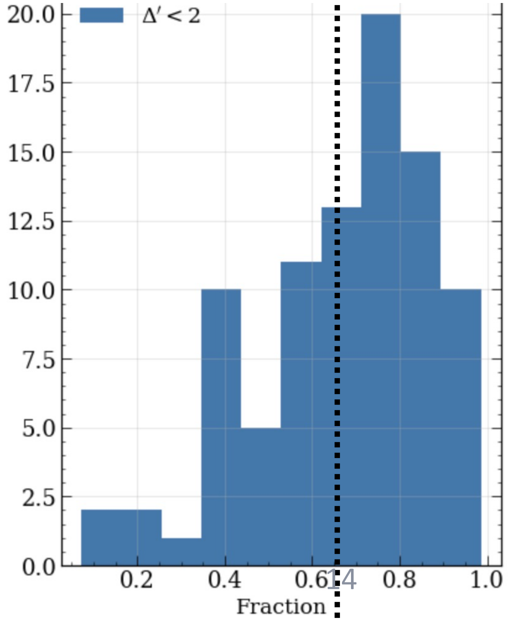
- The lowest mass ratio to-date, out of the waveform calibration range
- The smallest χ_P in O3b, $\chi_{eff} = 0.00^{+0.07}_{-0.09}$
- Mean value of $\Delta^{'}_{net}$: 1.77 (< 2)
- Fraction of $\Delta^{'}_{net}$ <2 samples: 0.62
- Waveform performance is “not too bad” compared to other events
- Spin is more problematic than mass ratio





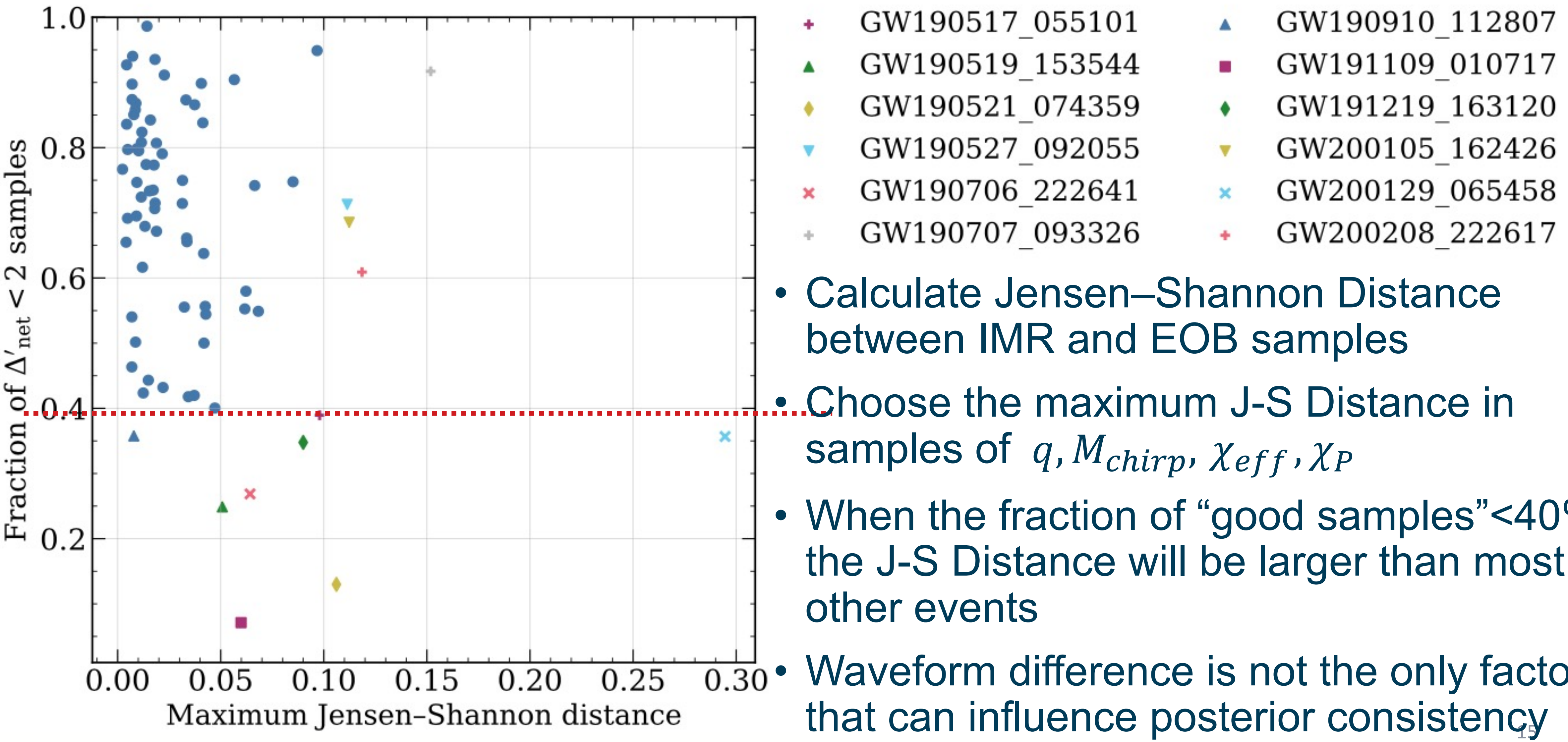






Δ'_{net} vs posterior inconsistency

University of Glasgow | IGR



- Calculate Jensen–Shannon Distance between IMR and EOB samples
- Choose the maximum J-S Distance in samples of $q, M_{chirp}, \chi_{eff}, \chi_P$
- When the fraction of “good samples”<40%, the J-S Distance will be larger than most other events
- Waveform difference is not the only factor that can influence posterior consistency

Summary

- A waveform accuracy evaluation approach, free from NR simulations
  - Key idea: if two waveforms have significant difference, they can not be accurate at the same time
  - Drawback: can not determine which one is inaccurate, or both inaccurate
  - Easy to apply, can be used as quantitative check in PE workflow

- BBH Real events & simulations
  - Only part of PE samples can pass our assessment; they are in the "well-behaved" regions of parameter space (low spin and equal mass)
  - Waveform difference has correlation with posterior sample consistency
  - Future 3rd-gen detectors: accuracy need to be improved 3+ orders of magnitude

Assessment of waveform pair

- Normalization & Relations with overlap

- $\Delta^+ = (h_1^+ - h_2^+ | h_1^+ - h_2^+)$, is proportional to the amplitude of GWs. Louder events tend to have larger Δ^+. We want to eliminate the impact of SNR and investigate waveform model's intrinsic performance in some specific parameter regions.
- Normalize Δ^+ with SNR (geometric mean of SNRs of two waveforms, i.e. $\sqrt{\rho_1 \rho_2}$.)

$$\|\Delta^+_{\mathrm{SNR}=1}\|^2 = \frac{(h_1^+ - h_2^+ | h_1^+ - h_2^+)}{\sqrt{(h_1^+|h_1^+)(h_2^+|h_2^+)}}$$

$$\|\Delta^+_{\mathrm{SNR}=\rho_0}\| = \rho_0 \|\Delta^+_{\mathrm{SNR}=1}\|$$

- Compared to overlap which is widely-used in the waveform community

$$\mathcal{O}(h_1, h_2) = \Re \frac{(h_1|h_2)}{\sqrt{(h_1|h_1)(h_2|h_2)}},$$

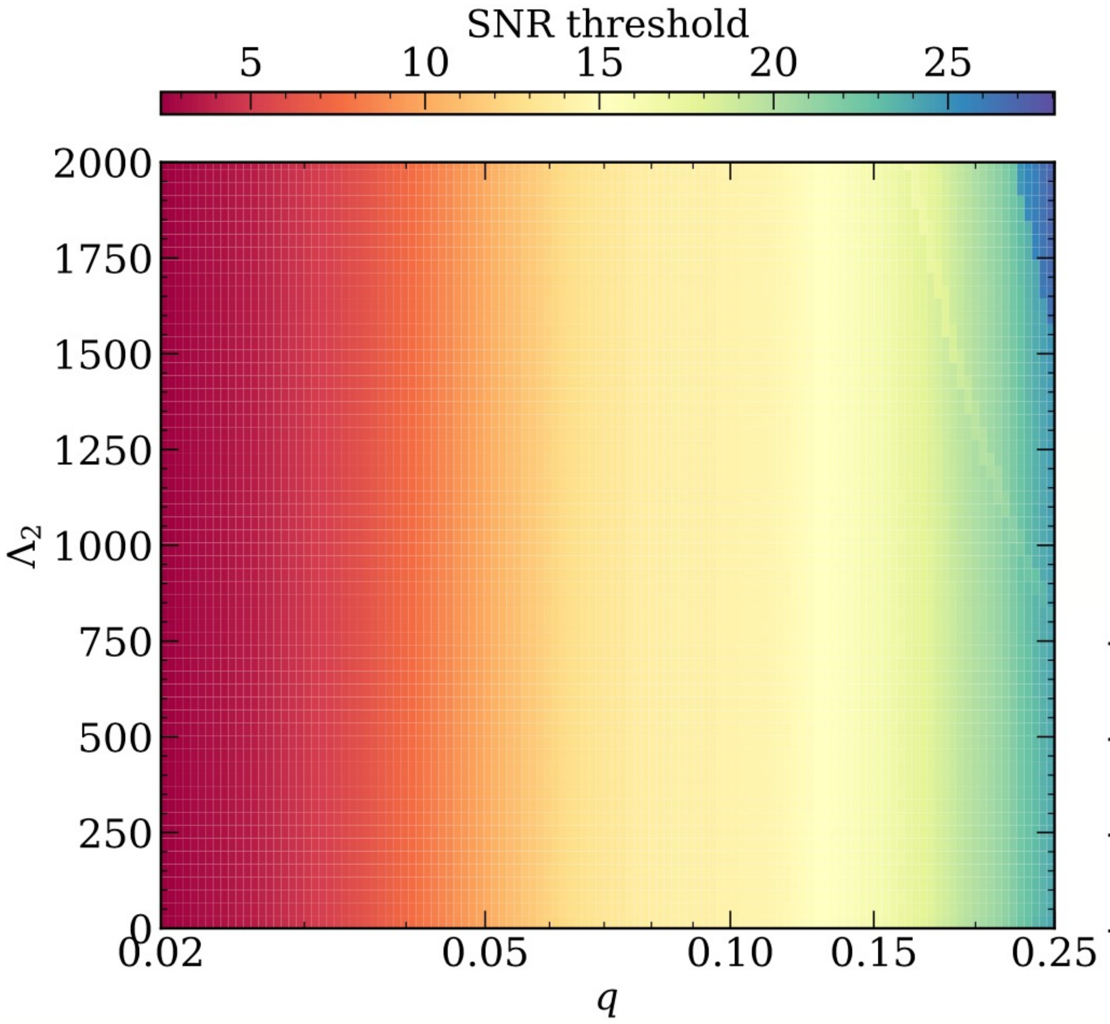
$$\|\Delta^+_{\mathrm{SNR}=1}\|^2 = \frac{\rho_1^+}{\rho_2^+} + \frac{\rho_2^+}{\rho_1^+} - 2\mathcal{O}(h_1^+, h_2^+).$$

- Δ^+ analysis is consistent with overlap method. But Δ^+ has a clear upper limit 2.

NSBH simulations

IMRPhenomNSBH and SEOBNRv4_ROM_NRTidalv2_NSBH

- $m_2 = 1.4 M_{\odot}, q \in [0.02, 0.25], \Lambda_2 \in [0,2000]$
- We assume zero-spin, as both models are calibrated with non-spin simulations

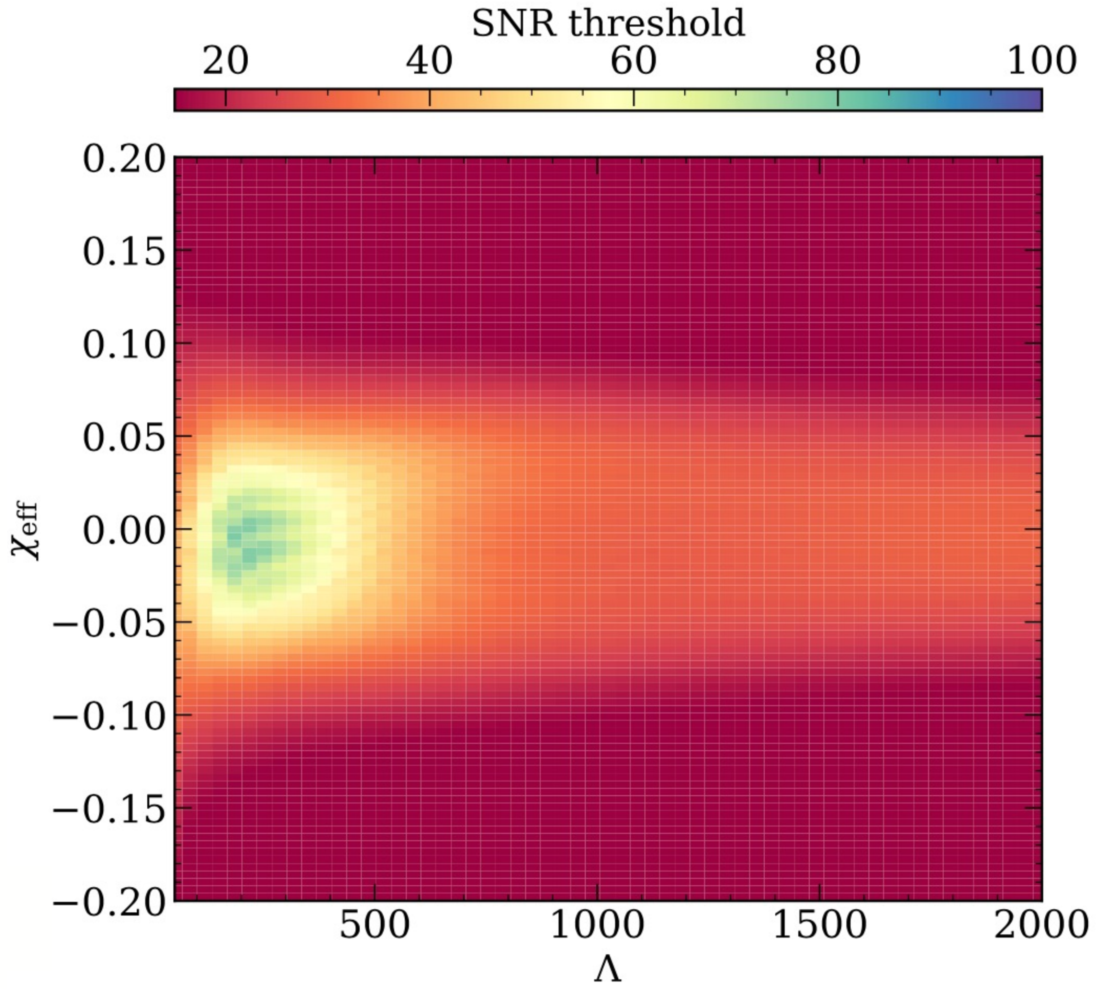


- Compared to matter effects, mass ratio has more impacts on waveform accuracy
- Waveform accuracy should also be improved for future high SNR observations , or when more complex physical effects are included (spins, higher modes or eccentricity etc)

BNS simulations

IMRPhenomPv2_NRTidalv2 and SEOBNRv4T_surrogate

- $m_1 = m_2 = 1.4 M_\odot, S_1 = S_2, \Lambda_1 = \Lambda_2$
- Aligned spin $|S_1| < 0.2,\ \Lambda_1 \in [0,2000]$



- Two waveform models agree with each other quite well in $\Lambda < 500, |S| < 0.05$, this is the region that coincides with our current knowledge of neutron star
- Waveform accuracy should be improved for future high SNR observations, or when more complex physical effects are included (high spin scenario, precession effects etc)