Waveform Systematics and Tests of General Relativity: present and future

Sylvain Marsat (L2IT, Toulouse)







Waveform systematics, environment, tests of GR

Environmental effects

Modelled:
- accretion, dynamical friction
- line-of-sight acceleration
- strong lensing
- microlensing

Unmodelled ?
- stochastic accretion
- stochastic lensing

Waveform systematics

Waveform inaccuracy:
- amplitude/phase errors with respect to true GR

Missing physics:
- higher modes
- spin precession
- eccentricity
- memory, post-merger kick...

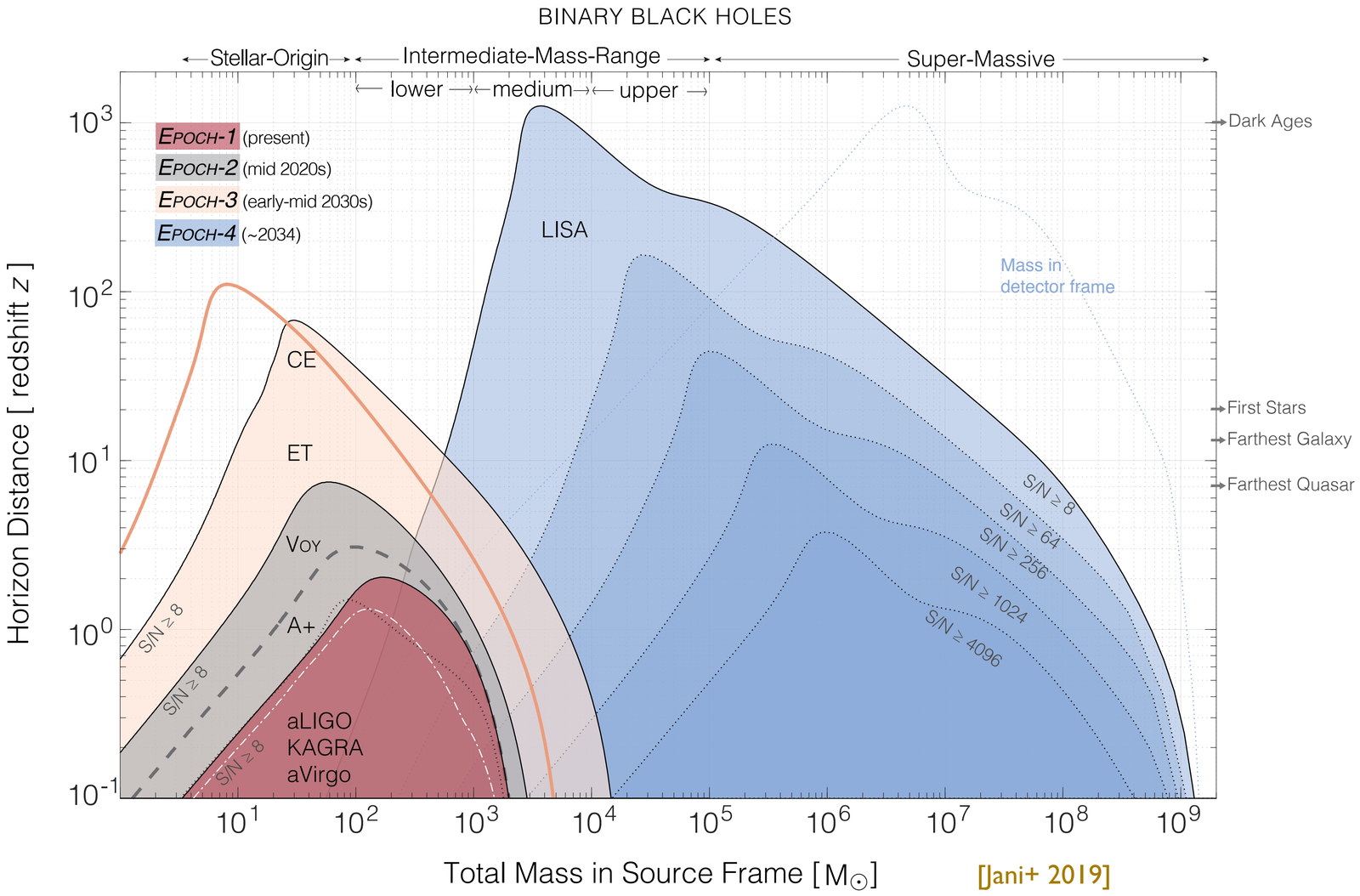
Tests of GR

Individual sources:
- consistency tests
- theory-agnostic tests
- theory-consistent tests

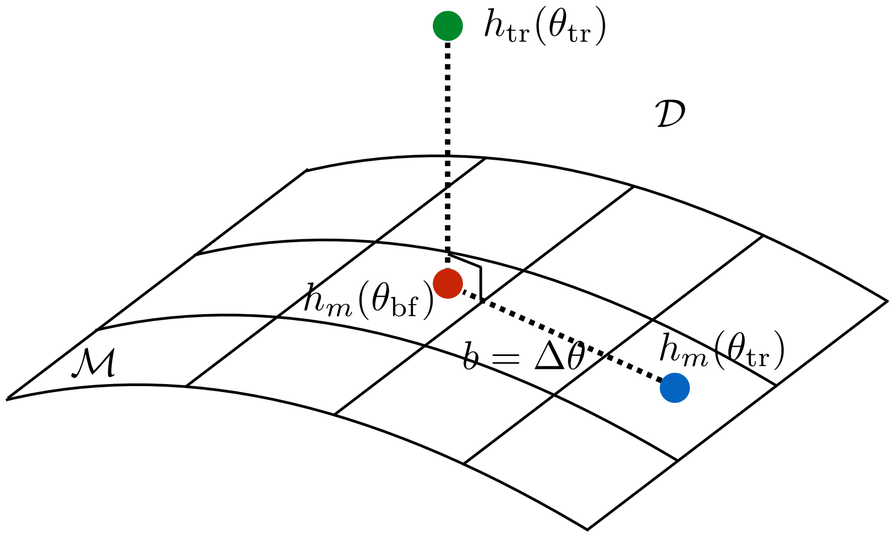
Collective inference

For the LVK context, review by [Gupta+ 2024]

A glimpse of the future of GW astronomy



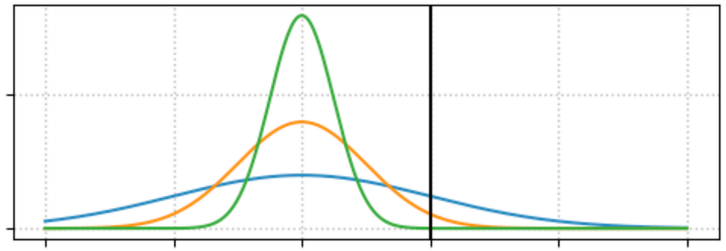
Waveform systematics and parameter estimation



Systematic biases:

Ignoring the effect of the noise, bias given by the **best-fit** parameters on the model signal manifold: $\Delta\theta = \theta_{\rm bf} - \theta_{\rm tr}$

- the **bias** is SNR-independent (optimization problem), but requires to explore the full parameter space [**expensive**]
- the statistical errors scale with SNR

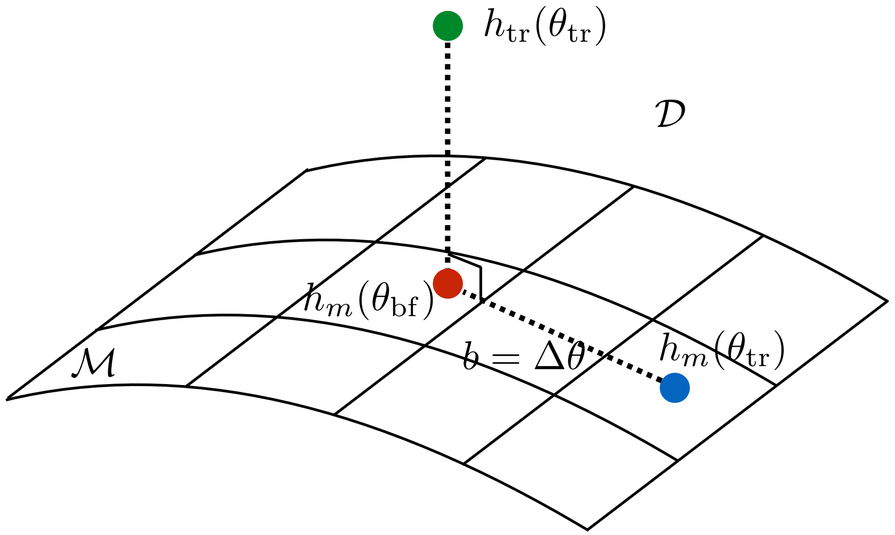


Different questions for systematics:

Are current waveform models accurate enough for GW science ? How accurate do they need to be ?

- Are waveforms good enough for prospective science ? Are e.g. Fisher errors accurate with approx. waveforms ? More forgiving.
- Will the biases be statistically significant ?
- Will biases affect hierarchical analyses (population, cosmology) ? How will they stack (coherent/incoherent) ?
- Will residuals affect the rest of the data (LISA global fit) ?
- How will tests of GR be affected ?
- Computational aspects: tradeoff accuracy/speed for low-latency ?

Waveform systematics and parameter estimation



Mismatch (unfaithfulness):

Mismatch, optimization over time/phase/polarization:

$$\mathrm{MM} = 1 - \max_{t,\varphi,\psi,\ldots} \frac{(h_m|h_{\rm tr})}{\sqrt{(h_m|h_m)}\sqrt{(h_{\rm tr}|h_{\rm tr})}}$$

- Computed locally [**fast**]
- SNR-independent
- Different versions: single-detector optimized over sky, combining $h_+, h_\times$

Indistinguishability criterion:

[Lindblom&al 2008]
[Chatziioannou&al 2019]
[Toubiana-Gair 2024]

$$\ln \mathcal{L}(\theta) = -\frac{1}{2}(h(\theta) - h_{\rm tr}|h(\theta) - h_{\rm tr})$$

$$\ln \mathcal{L}(\theta_{\rm bf}) \sim \ln \mathcal{L}(\theta_{1-\sigma})$$

$$\mathrm{MM} < \frac{D}{2}\frac{1}{\mathrm{SNR}^2}$$

- Constant D: dimension, approximate
- Scaling SNR^2 robust

Linearized biases (Cutler-Vallisneri):

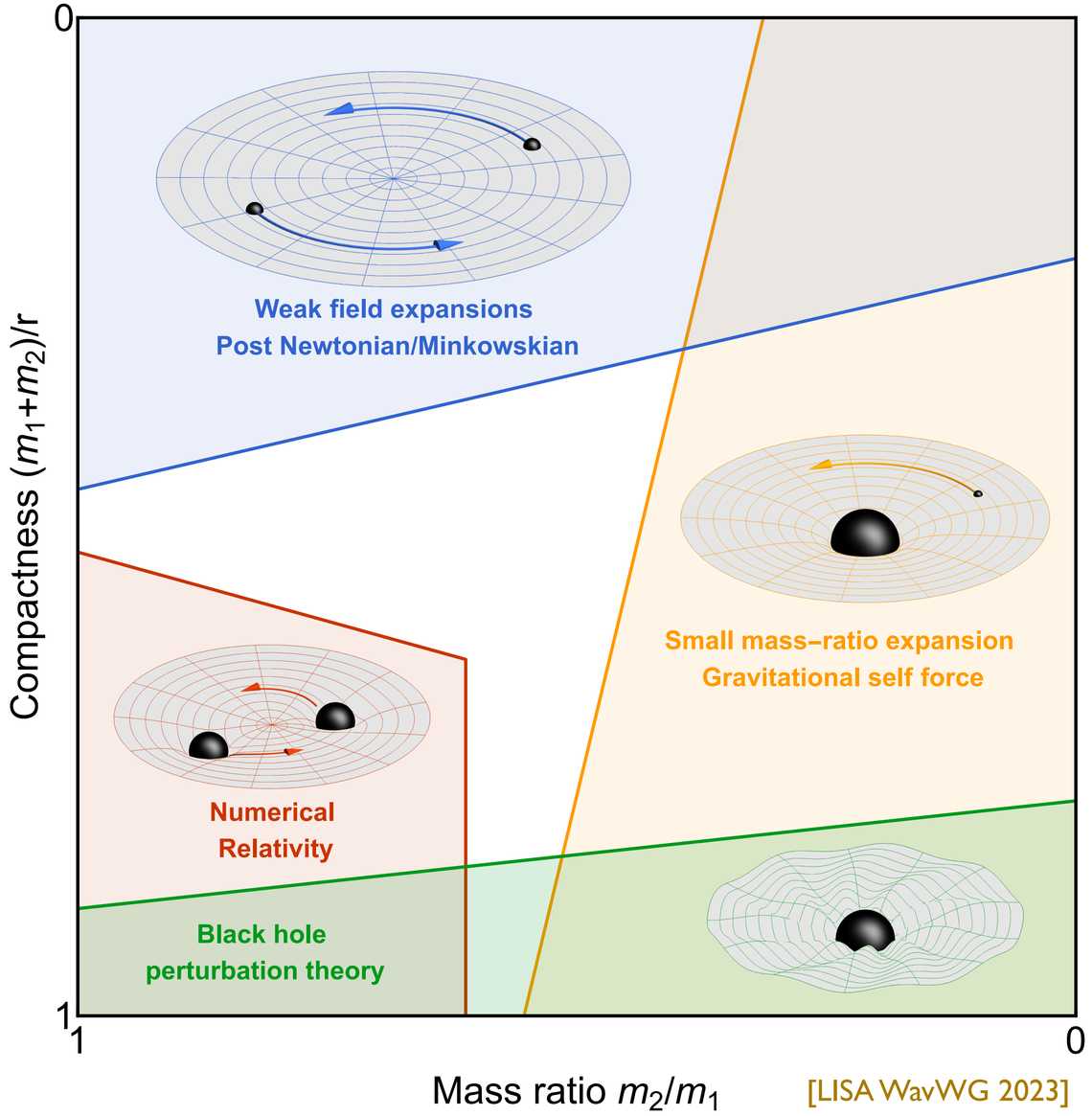
[Flanagan-Hughes 1997]
[Cutler-Vallisneri 2007]

In the linear signal approximation, estimation of bias [**fast**]:

$$F_{ij} = (\partial_i h|\partial_j h)$$
$$\Delta\theta_i = F^{-1}_{ij}(\partial_j h|\delta h)$$

- Bias assessment: take best waveform and pretend it to be exact
- Can we assess biases with efficient tools ? Do we need full simulations ?

Overview of waveform models

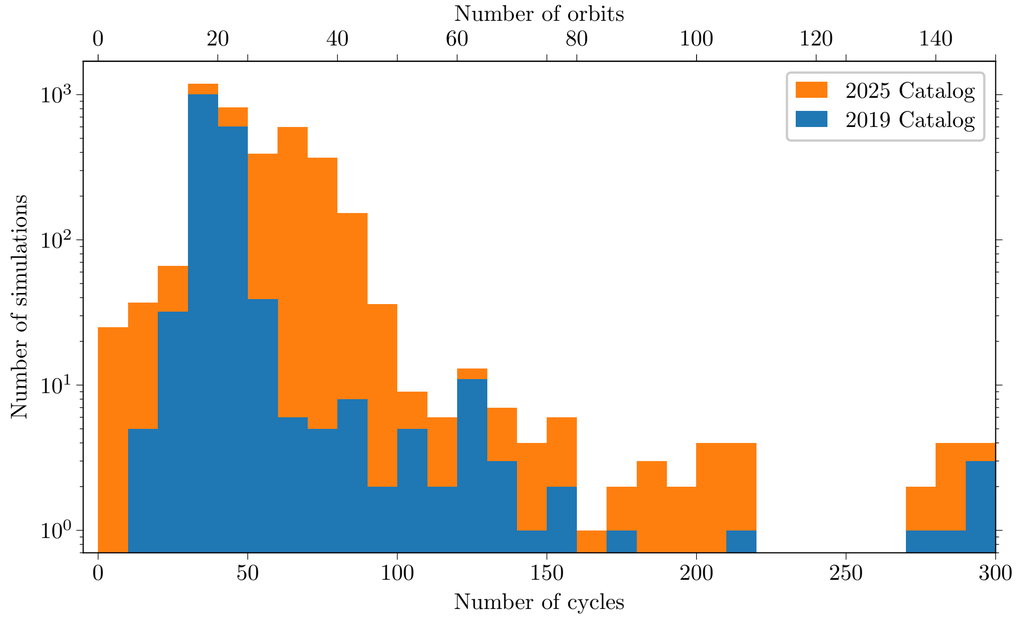


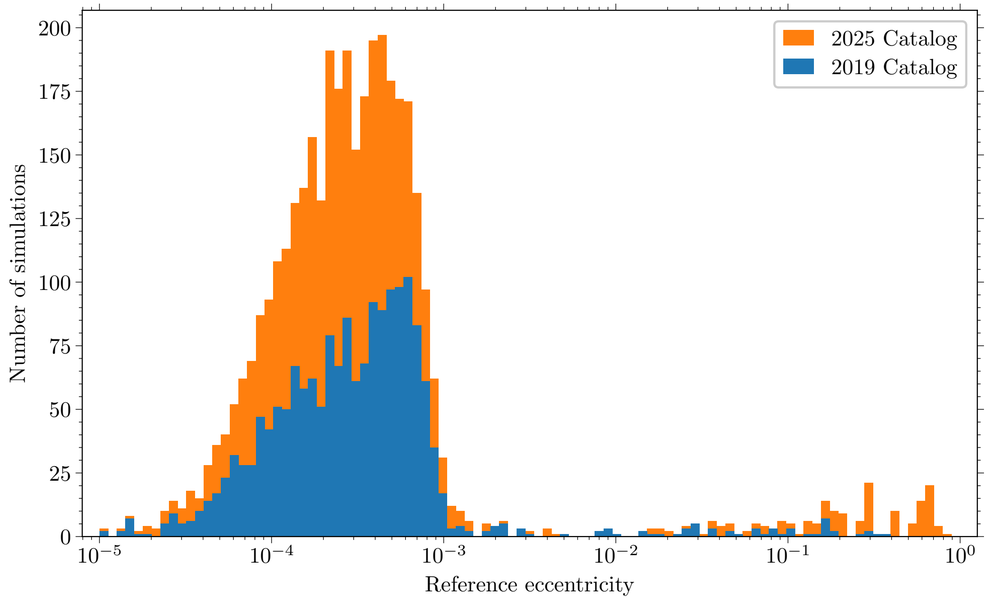
[LISA WavWG 2023]

Overview of waveform models: NR coverage

[SXS catalog 2025]







Overview of waveform models: Phenom / SEOB / NR Surrogates

Phenom:

`IMRPhenomXPNR` [Hamilton+ 2025]
- Fourier-domain, precessing

`IMRPhenomTPHM` [Estelles+ 2021]
- Time-domain, precessing
- GPU acceleration

`IMRPhenomTEHM` [Planas+ 2025]
- Time-domain, eccentric

EOB:

`SEOBNRv5PHM` [Ramos-Buades+ 2023]
- Time-domain, precessing

`SEOBNRv5EHM` [Gamboa+ 2024]
- Eccentric, spin-aligned

`TEOBResumS-Dali` [Nagar+ 2024]
- Spin-precessing, Eccentric

+ `_ROM`
- accelerated reduced order models (spin aligned)

NR surrogates:

`NRHybSur3dq8` [Varma+ 2018]
- SXS NR simulations hybridized with long EOB inspirals (covers ~6 months for $M = 10^5 M_\odot$)
- Surrogate interpolant, time-domain

`NRSur7dq4` [Varma+ 2019]
- SXS NR simulations, short
- 7d, fully precessing, with mode asymmetries

+ high-q surrogates
- BH pert. + NR

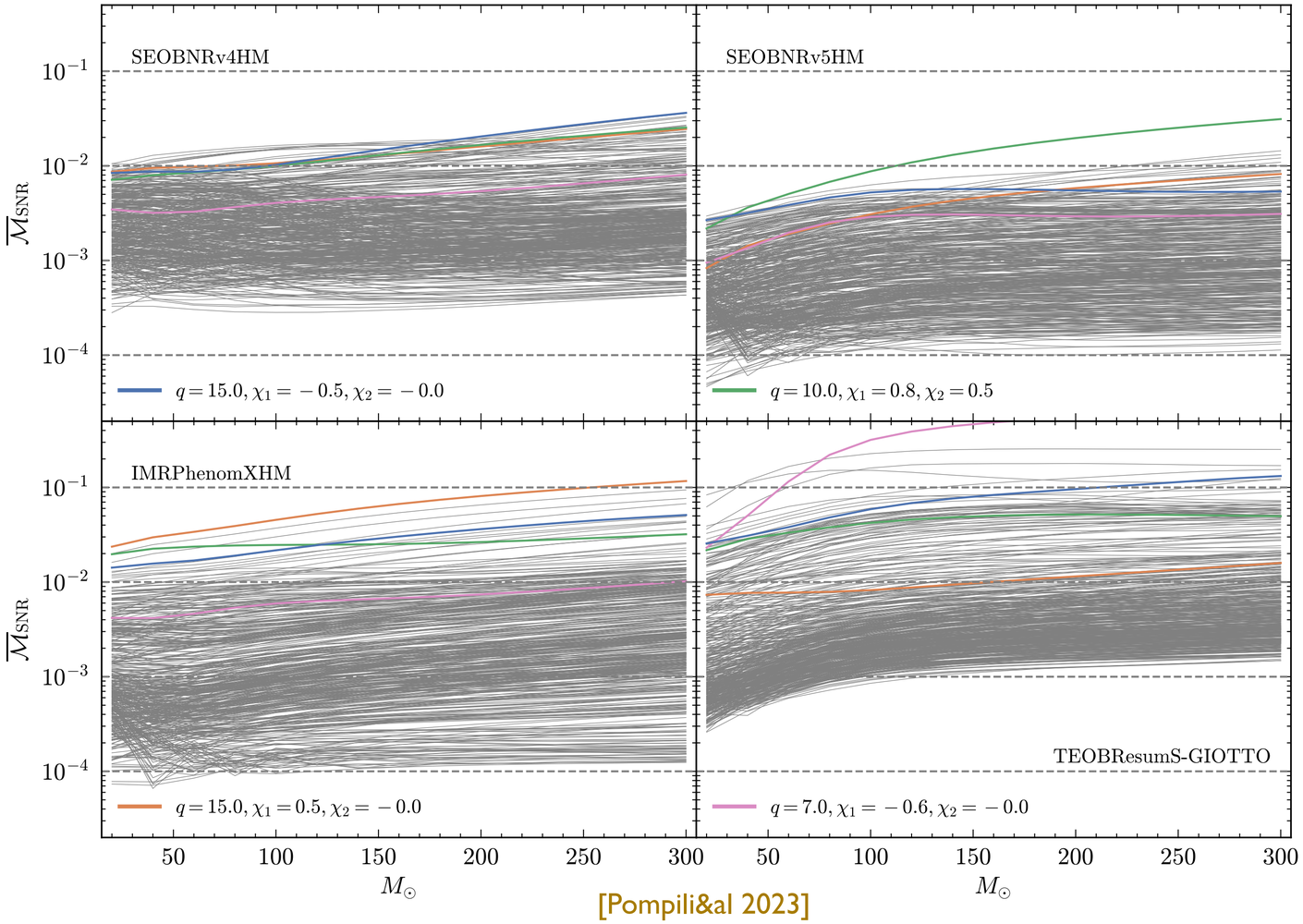
+ eccentric surrogates
- non-spinning

Packages:

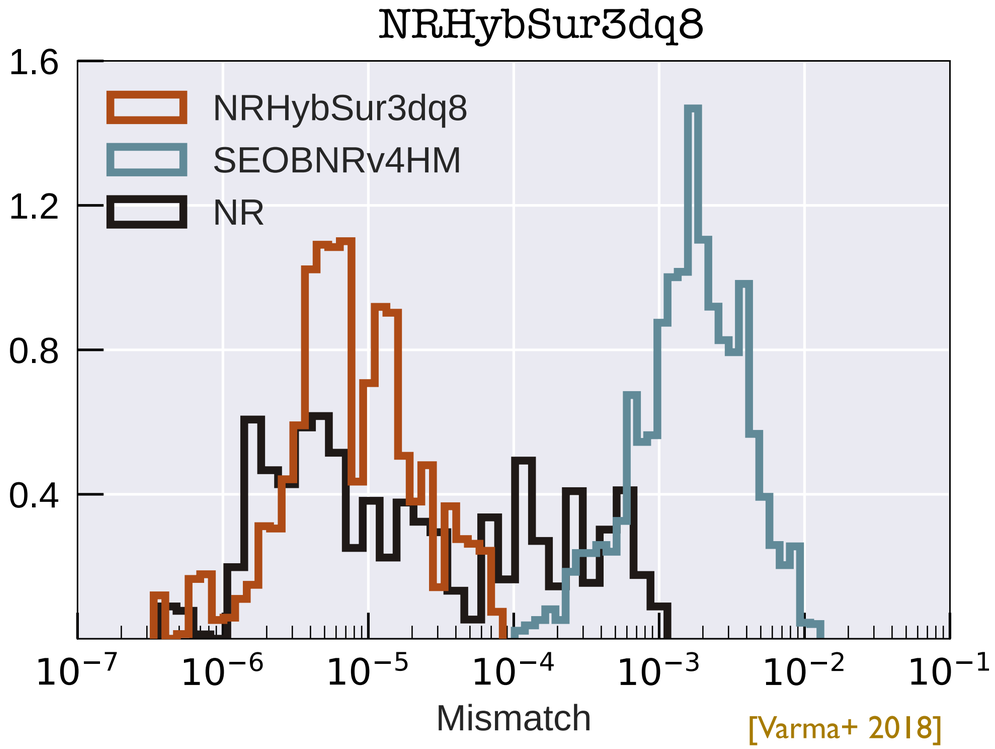
Moving towards modular codes:
- `phenomxpy`
- `pyseobnr`
- `gwsurrogate`

+ recent move to GPU implementations (Jax)

Overview of waveform models: mismatches for the circularized non-precessing case



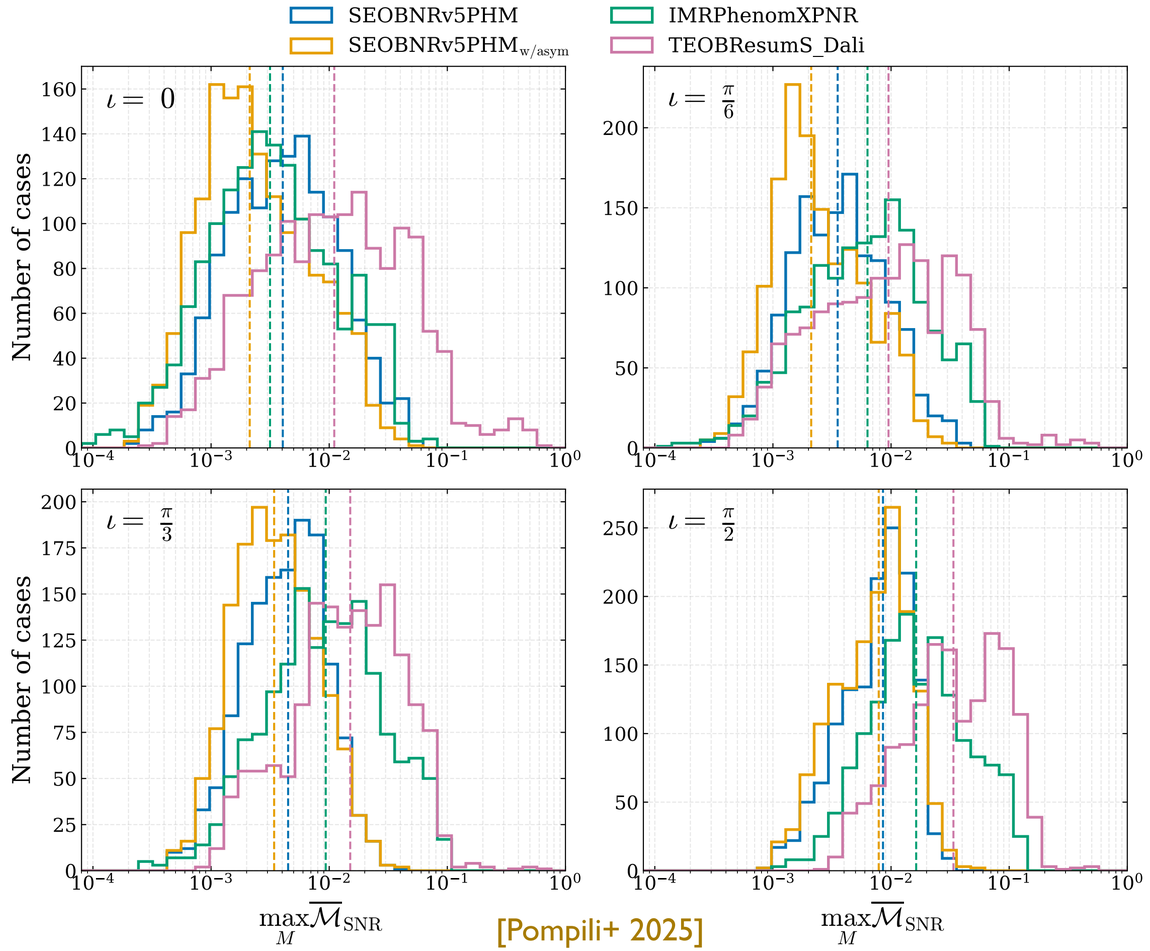
[Pompili&al 2023]



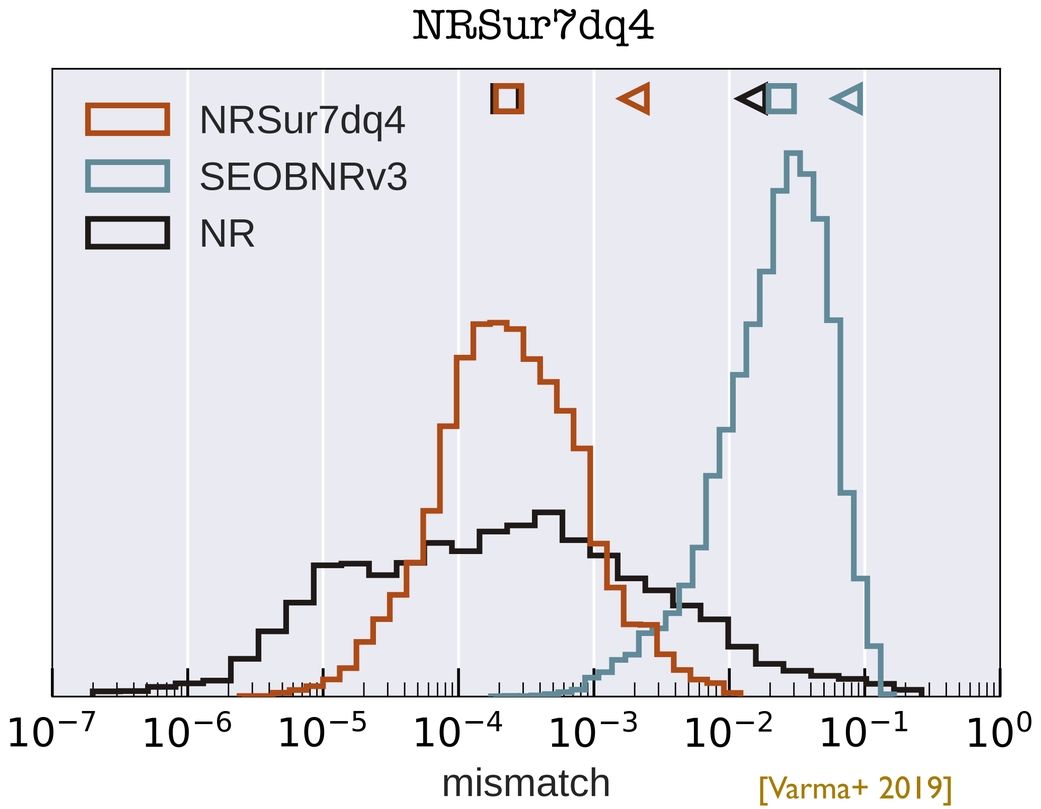
[Varma+ 2018]

Aligned spin case: mismatch with NR ~ 10^{-4} – 10^{-2}

Overview of waveform models: mismatches for precession



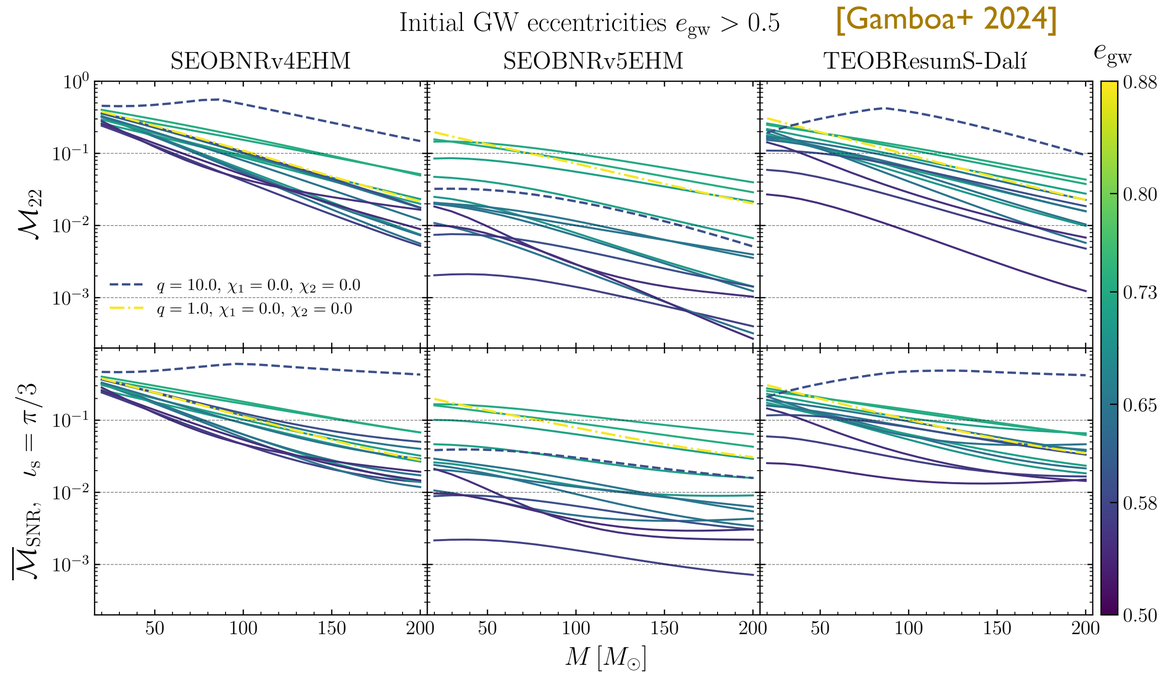
[Pompili+ 2025]



[Varma+ 2019]

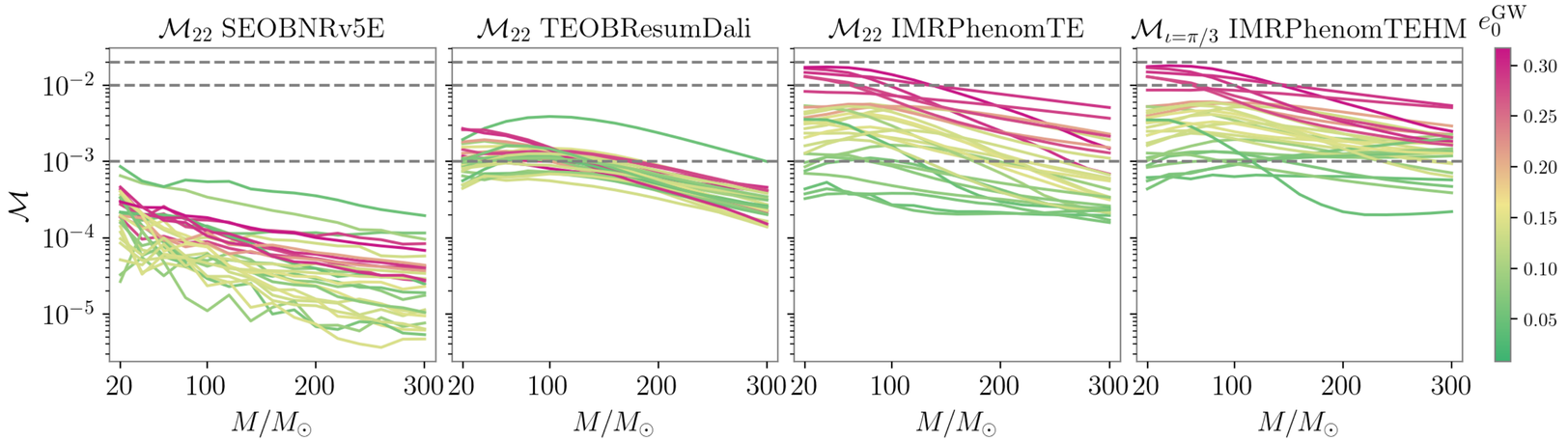
Precessing spin case: mismatch with NR ~ 10^{-3} – few 10^{-2}

Overview of waveform models: mismatches for eccentricity



[Gamboa+ 2024]

Eccentric case: mismatch with NR ~ $10^{-2} - 10^{-1}$



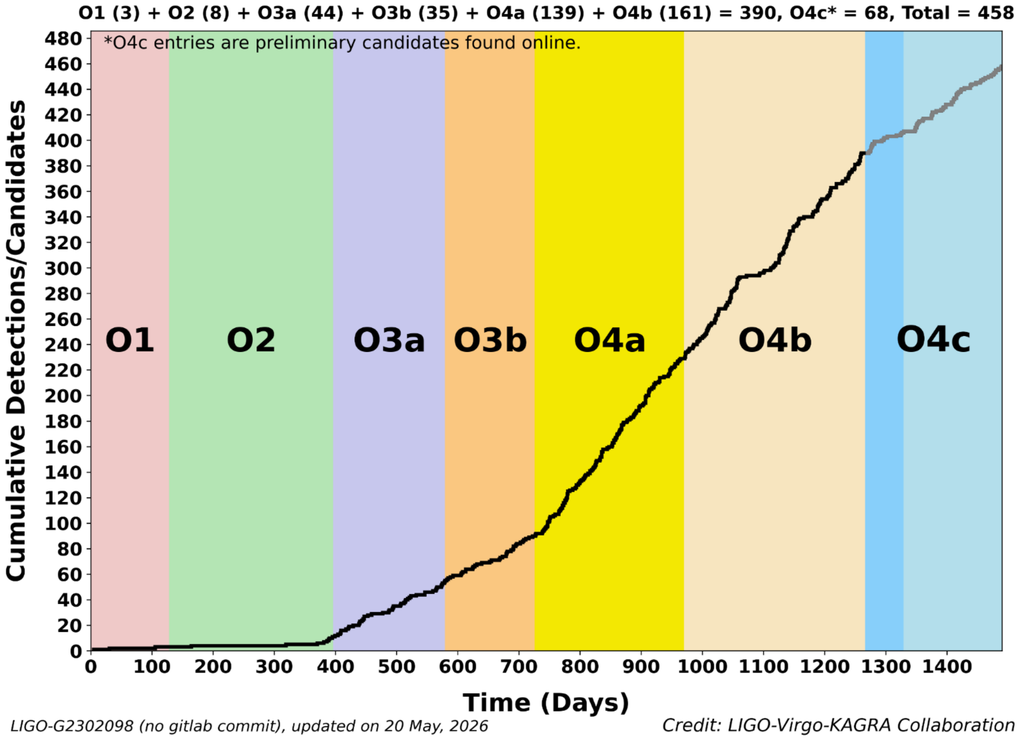
[Planas+ 2025]

See also:

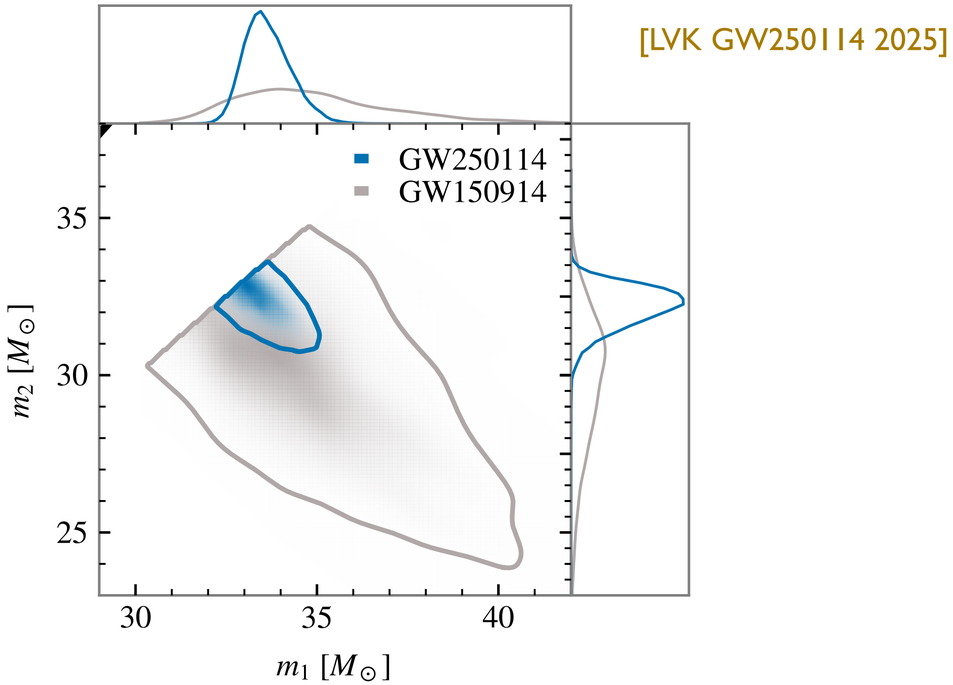
NRSurE_q4NoSpin_22

[Nee+ 2026]

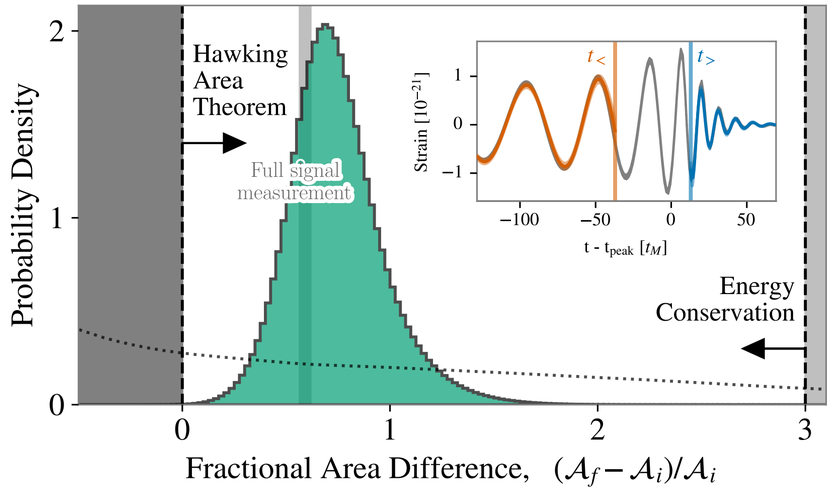
LVK status: waveform systematics



[LVK GWTC-5 2026]



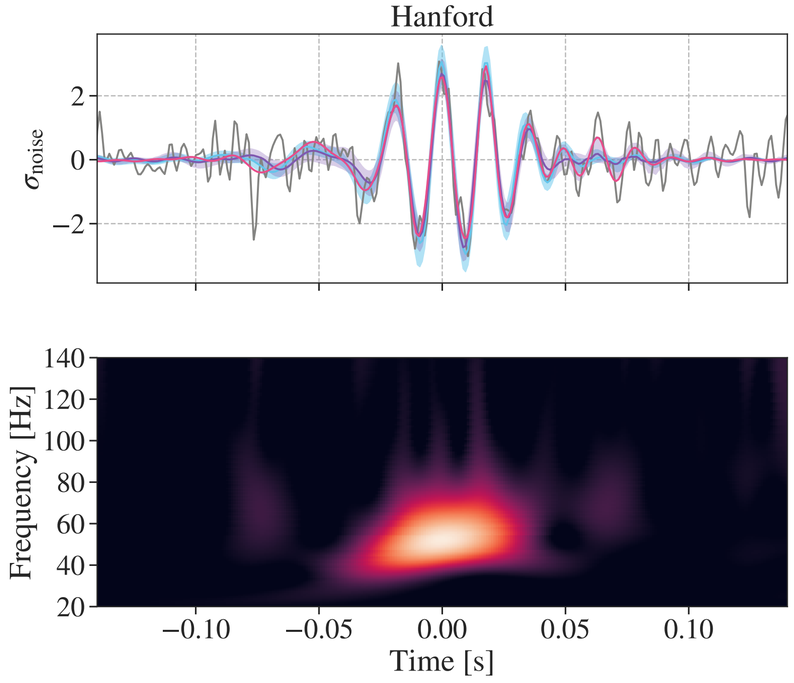
[LVK GW250114 2025]

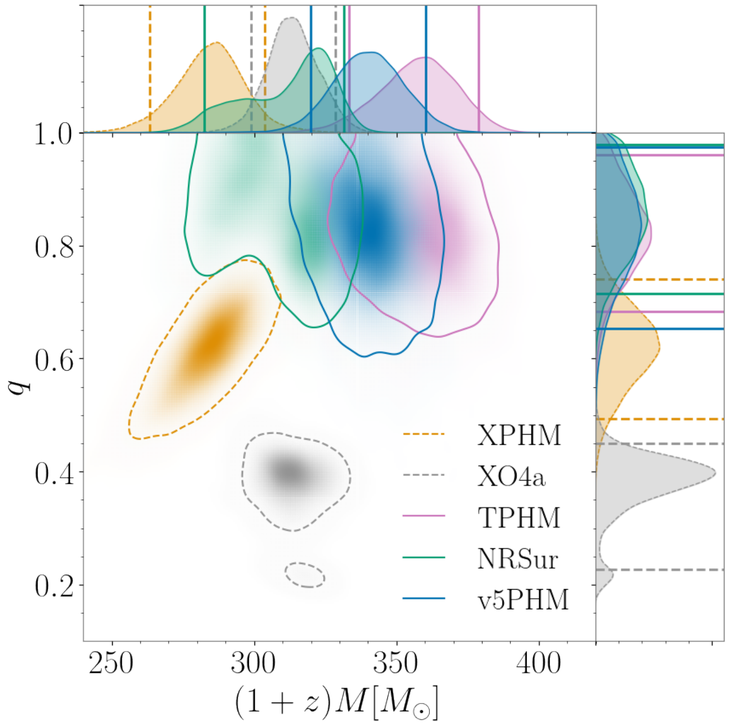


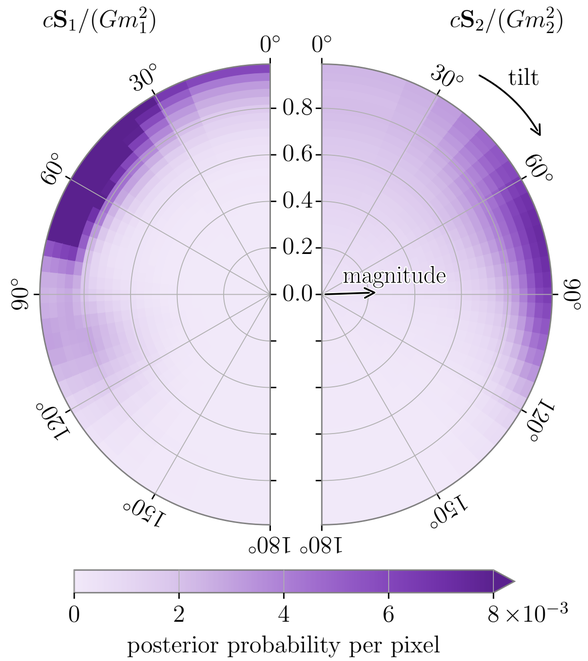
- GWTC-5: close to 400 compact binary coalescences
- Tests of GR now performed on 'interesting subset' of events
- Exceptional events: example of GW250114, 'clone' of GW150914, SNR 24 → 80, test of the area theorem

LVK: waveform systematics for GW231123

[LVK GW231123 2025]

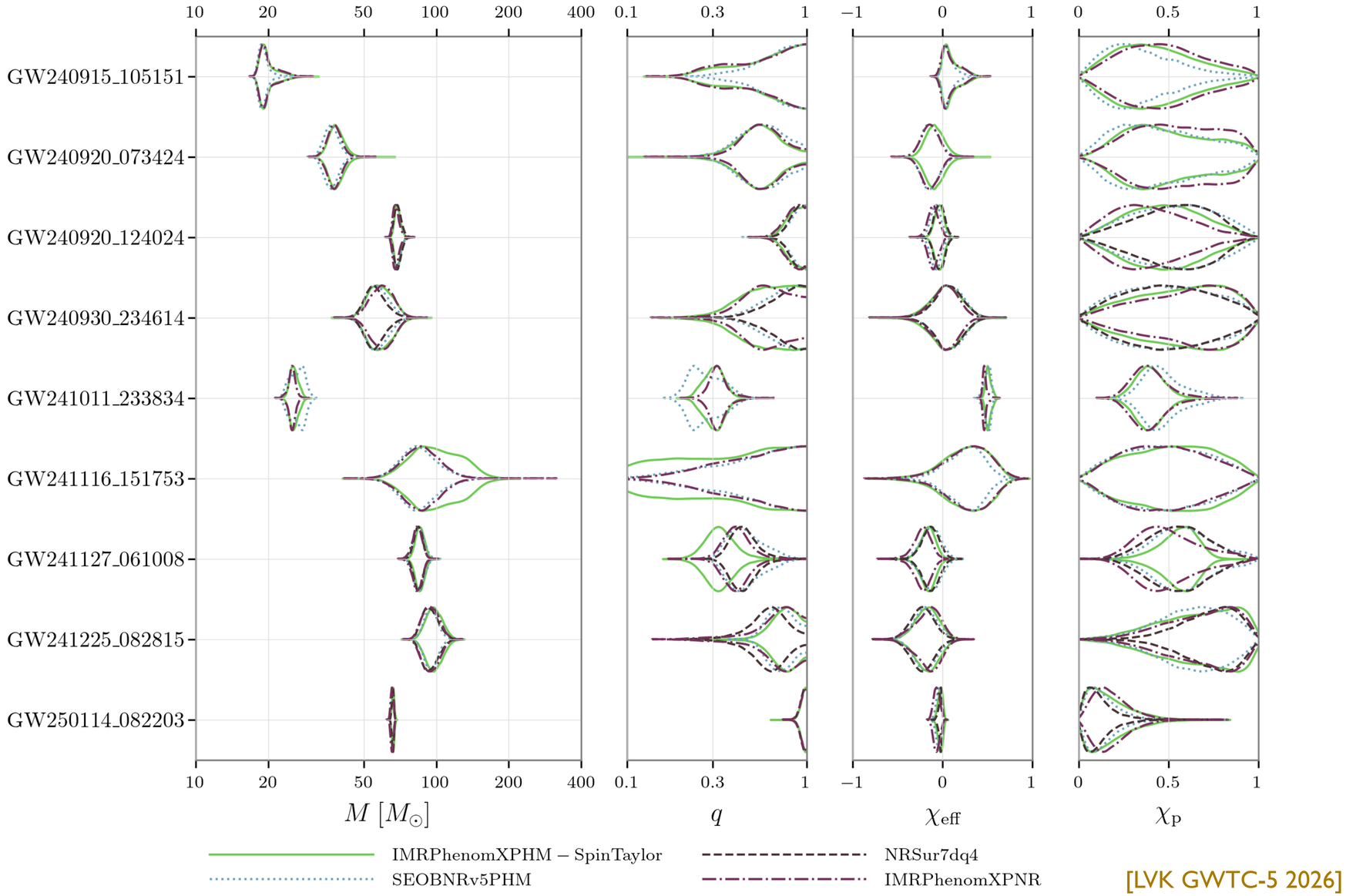






- GW231123: high-SNR, very high-mass event (similarities with GW190521)
- Signal of just a few cycles around merger
- Very high precessing spins recovered in PE
- Significant differences between waveform models for this event
- Numerous re-analyses: eccentricity, strong lensing, lensing in the wave optics regime

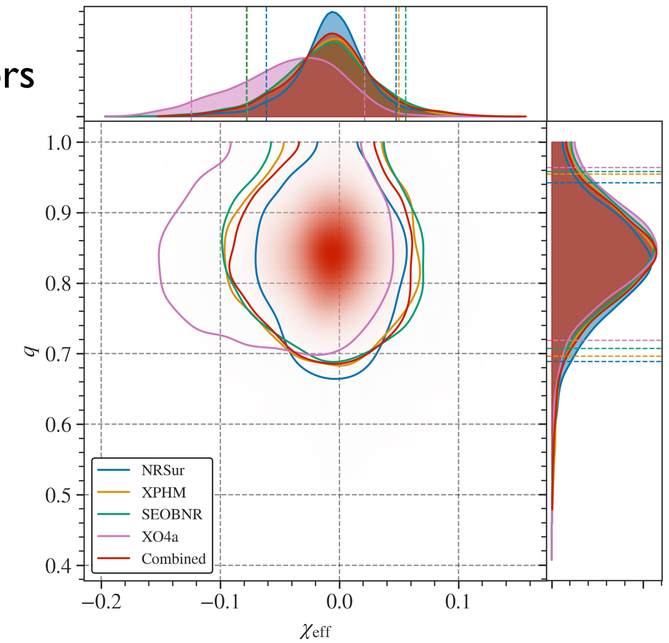
LVK status: waveform systematics



[LVK GWTC-5 2026]

LVK waveform systematics example: GW230814

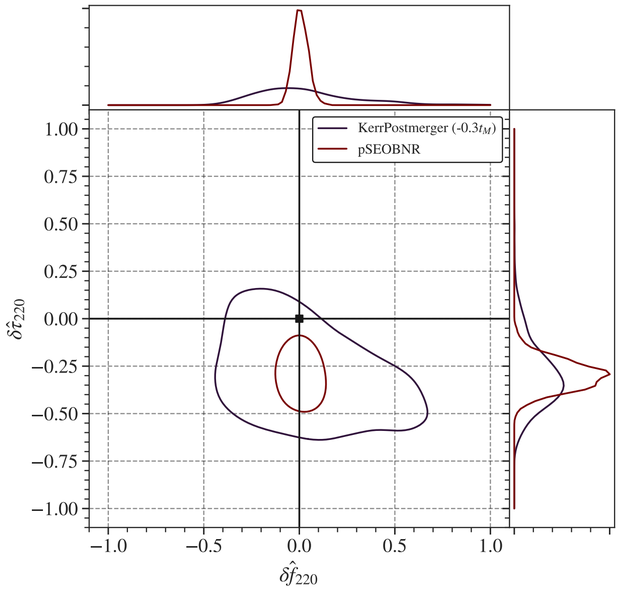
GR posteriors



[LVK GW230814 2025]

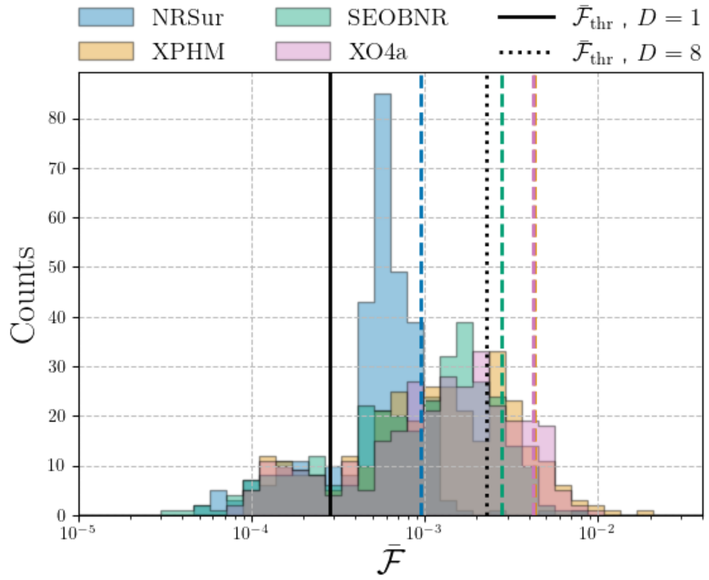
- GW230814: high-SNR (40), single-detector
- pSEOBNR: ringdown deviations in IMR waveform
- Mild apparent deviation from GR ?

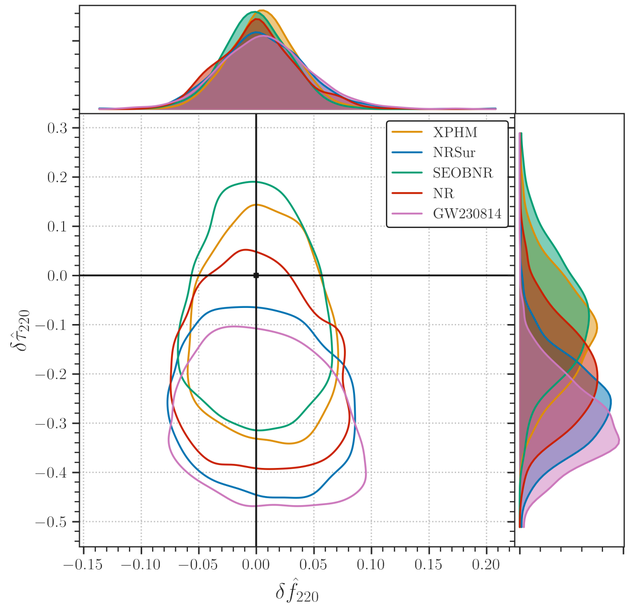
- Possibility of noise artefact also open



pSEOBNR IMR results show deviation

Faithfulness, compatible NR sims

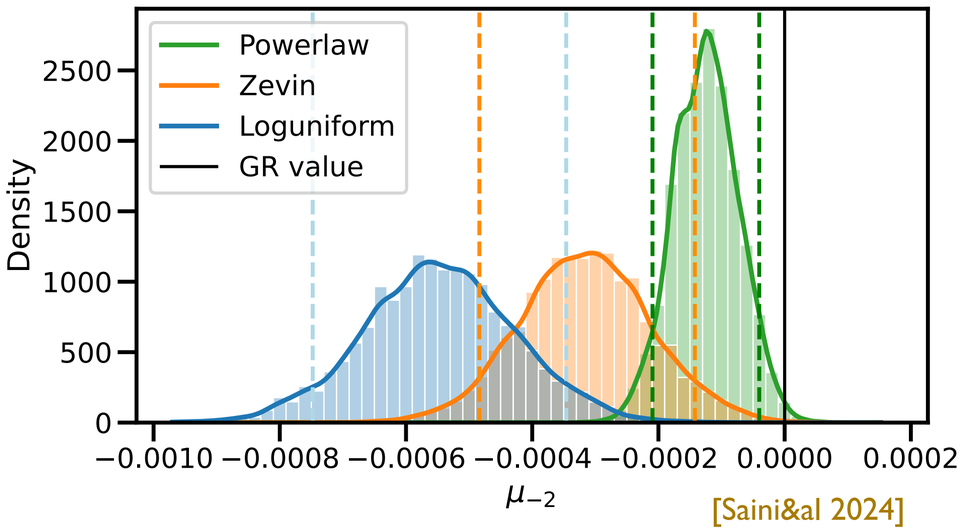




pSEOBNR results on injections with other waveforms

LVK: effect of missing physics on TGR -- Eccentricity

Hierarchical mean of deviation for -1PN parameter, varying eccentricity distr.

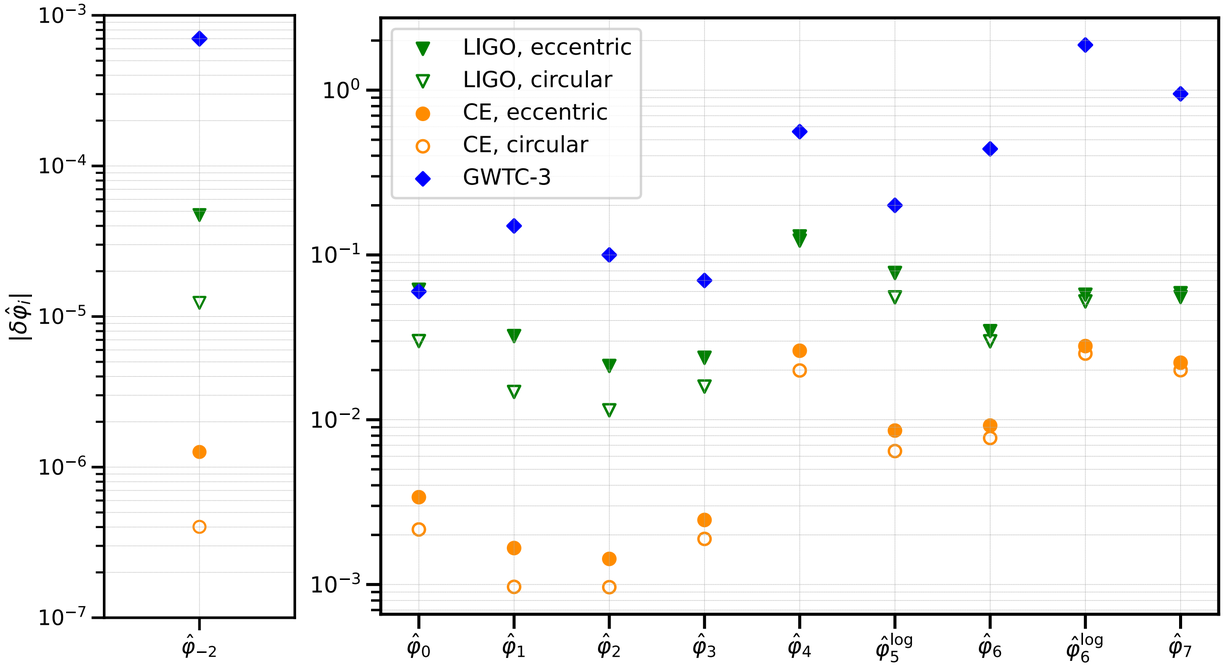


[Saini&al 2024]

[Saini&al 2024] Simplified simulation setup: eccentricity represented only in the inspiral phasing, parametrized inspiral test

- Ignoring physics (like ecc.) can lead to false apparent deviations from GR
- Marginalizing over the extra degrees of freedom will degrade constraints

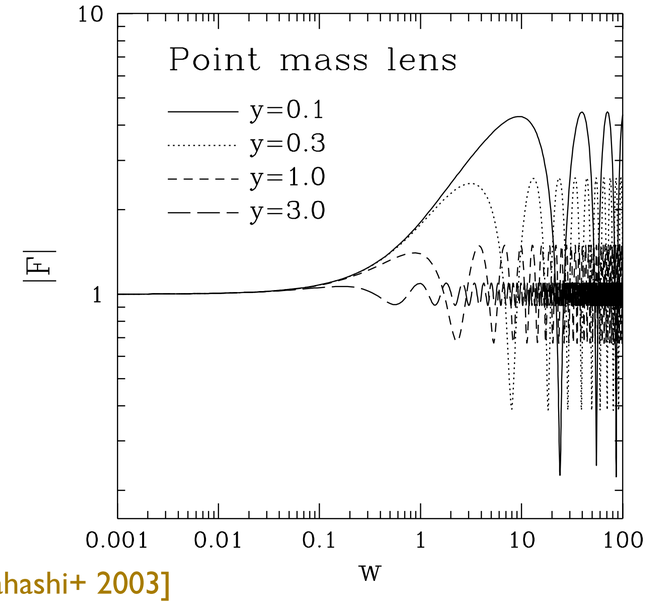
Simulated hierarchical contraints for 300 BBH with and without eccentricity let free



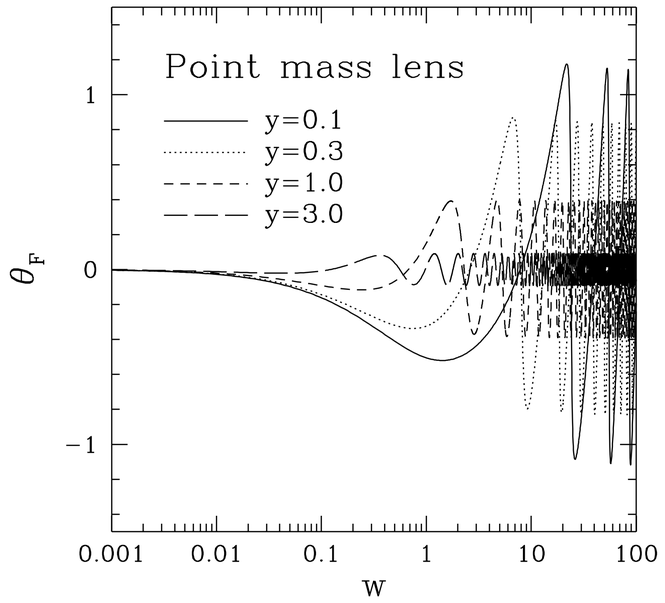
[Saini&al 2024]

$$\delta\psi_{\ell m}(f) = \frac{3}{128\eta v^5}\frac{-m}{2}\Bigg[\sum_{\substack{n=0\\ n\neq 1}}^{7}\delta\hat{\varphi}_n\psi_n^{\rm PN}v^n + \delta\kappa_s(\psi_{4,\kappa_s}^{\rm PN}v^4 + \psi_{6,\kappa_s}^{\rm PN}v^6) + \sum_{n=5}^{6}\delta\hat{\varphi}_{n(\ell)}\psi_{n(\ell)}^{\rm PN}v^n\log v + \sum_{n=-2,1}\delta\hat{\varphi}_n v^n\Bigg]$$

LVK: effect of missing physics on TGR -- Lensing



[Takahashi+ 2003]

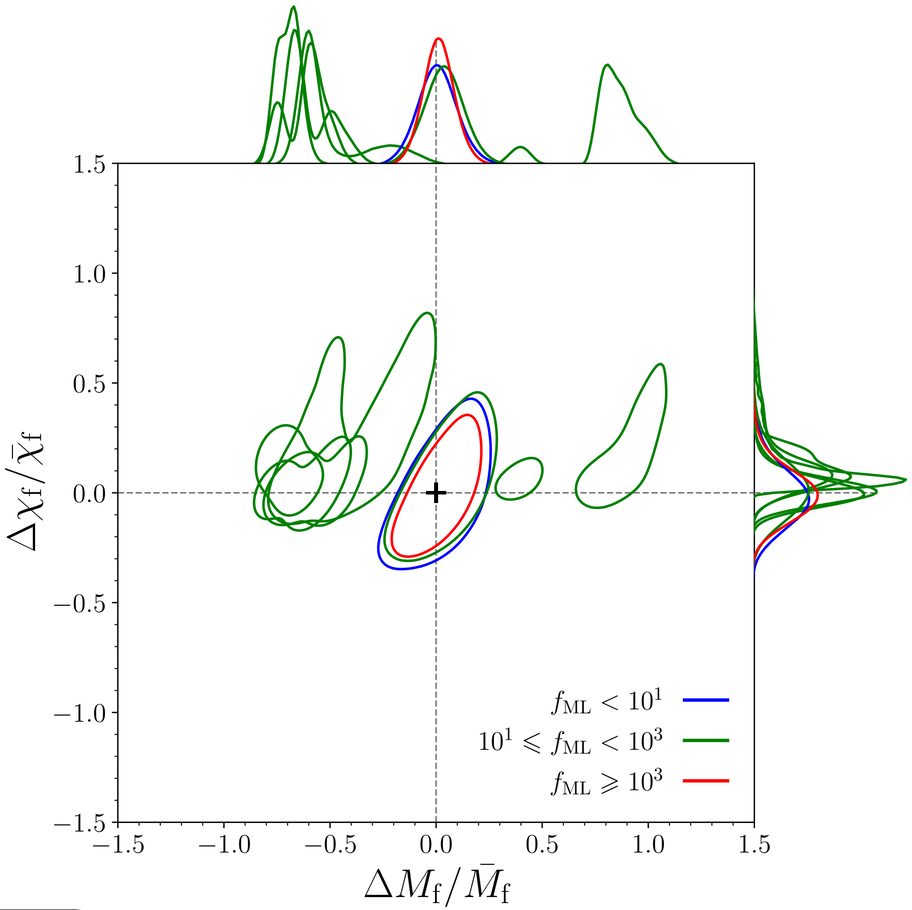


Wave-optics (diffractive) lensing: lensing transfer function $F(w, y)$ with $w \sim M_L f$ dimensionless frequency and y impact parameter

Lensing is expected to be rare astrophysically: $< 1\,\%$ of sources

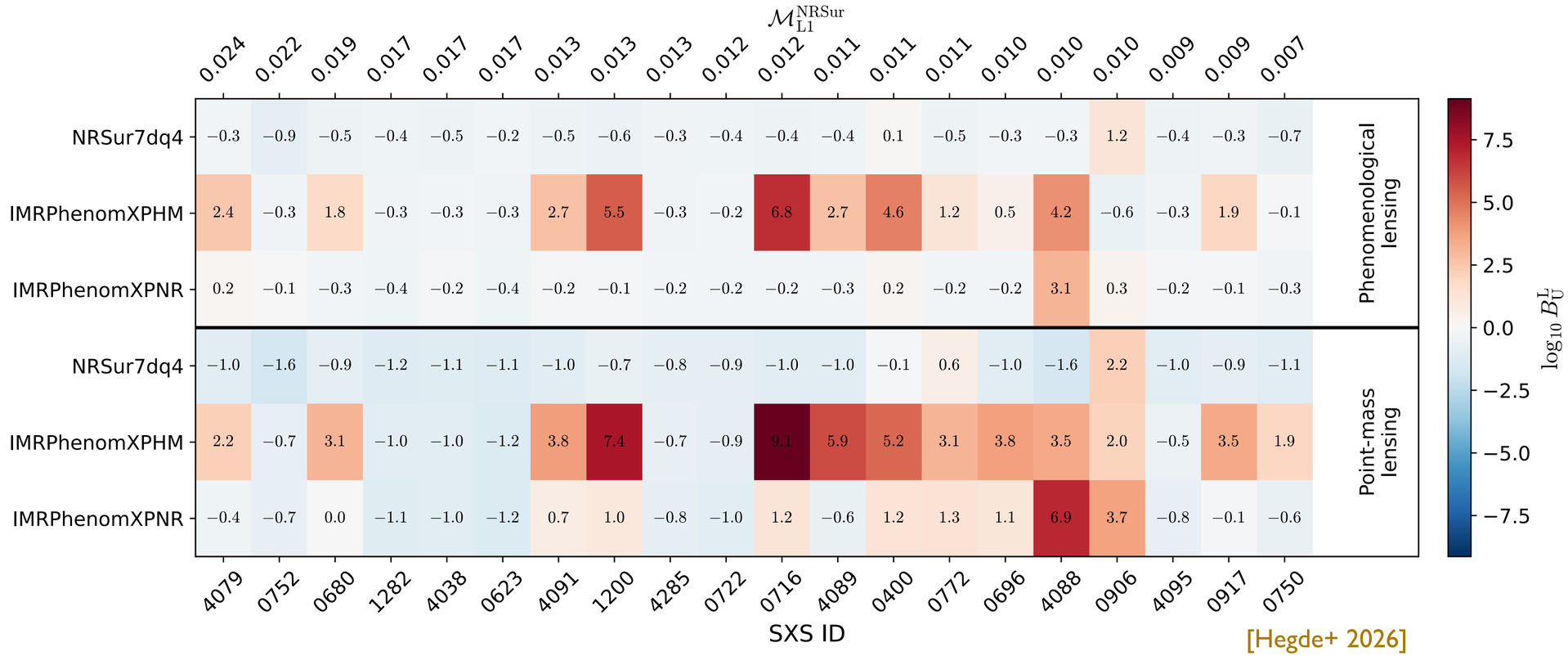
[Kopty+ 2026] Simulation of various tests of GR (here IMR consistency) in presence of diffractive lensing of the injected signal

- Wave optics lensing can lead to false detection of deviations from GR



[Kopty+ 2026]

Lensing: interplay with waveform systematics



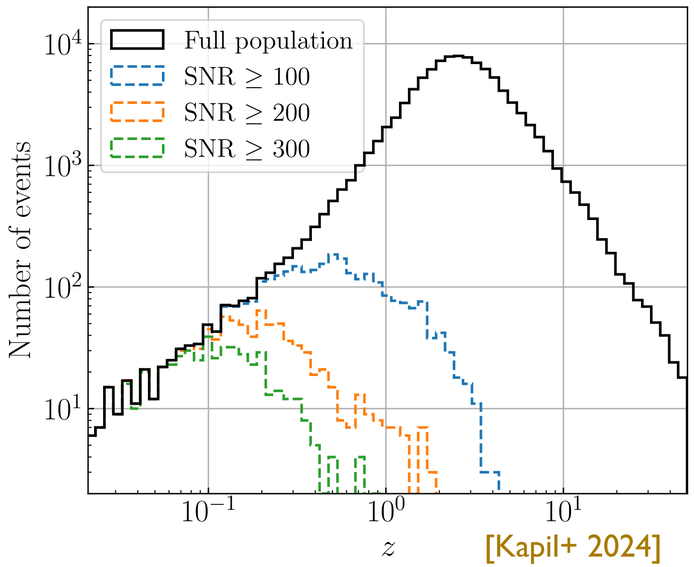
[Hegde+ 2026]

[Hegde+ 2026] Bayes factors for the presence of a lens (both geometric optics and wave optics) on NR injections with different waveforms as templates

- Waveform systematics can lead to false detection of gravitational lensing

3G detectors: waveform systematics

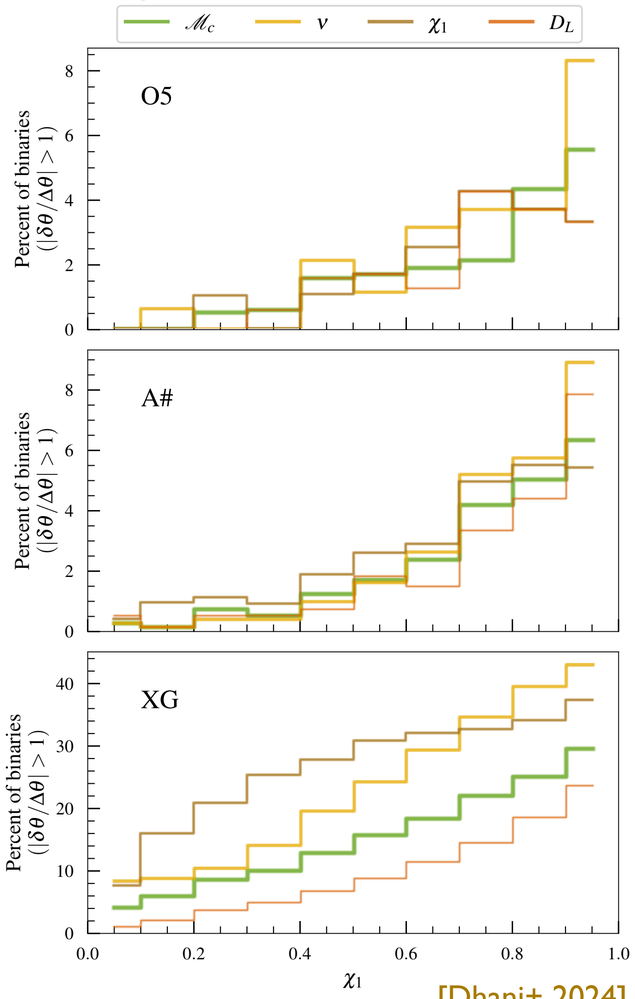
Source histograms per SNR level



[Kapil+ 2024]

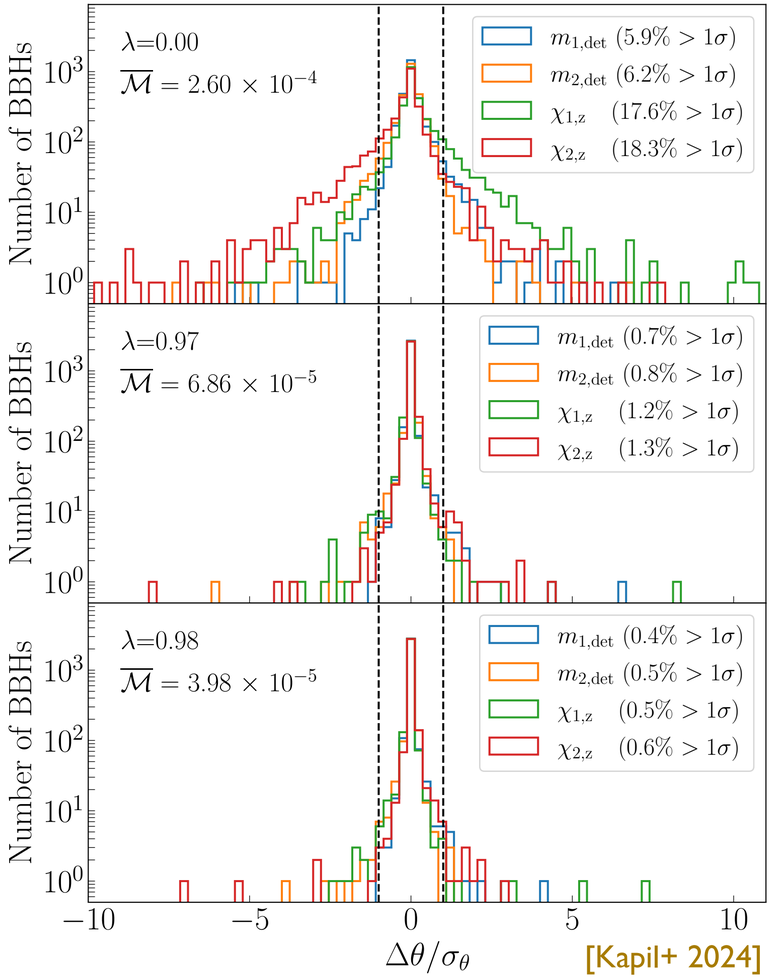
- 3G detectors will have a tail of golden events with high SNRs
- Significant fraction of analyses expected to be biased with current waveform accuracy

Proportion of biased sources



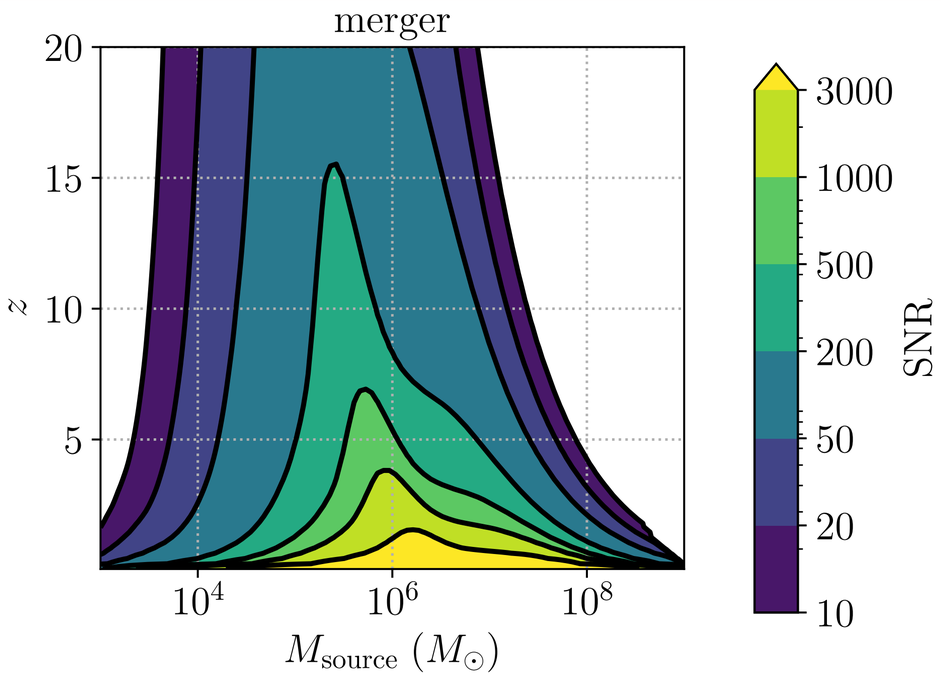
[Dhani+ 2024]

Requirement on mismatch to eliminate biases



[Kapil+ 2024]

MBHB signals in LISA



Science case for MBHB systematics

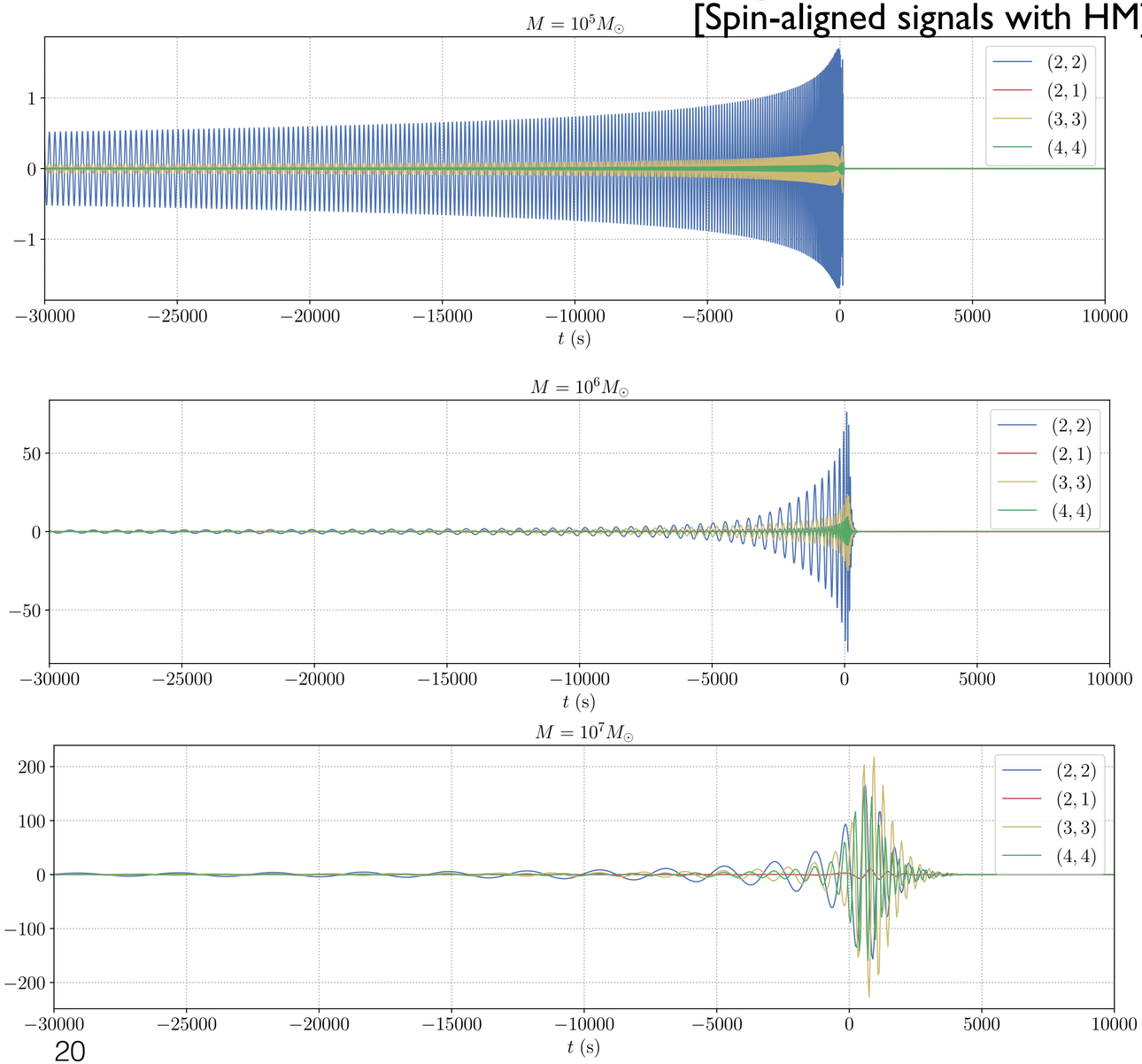
MBHBs high-SNR sources for LISA, waveform systematics crucial for:

- golden events for EM counterparts
- golden events for TGR
- population inference and cosmology
- global fit and residuals

Broader parameter space than LVK: high-q, high-spins, precession, eccentricity

Whitened time-domain signals

[Spin-aligned signals with HM]



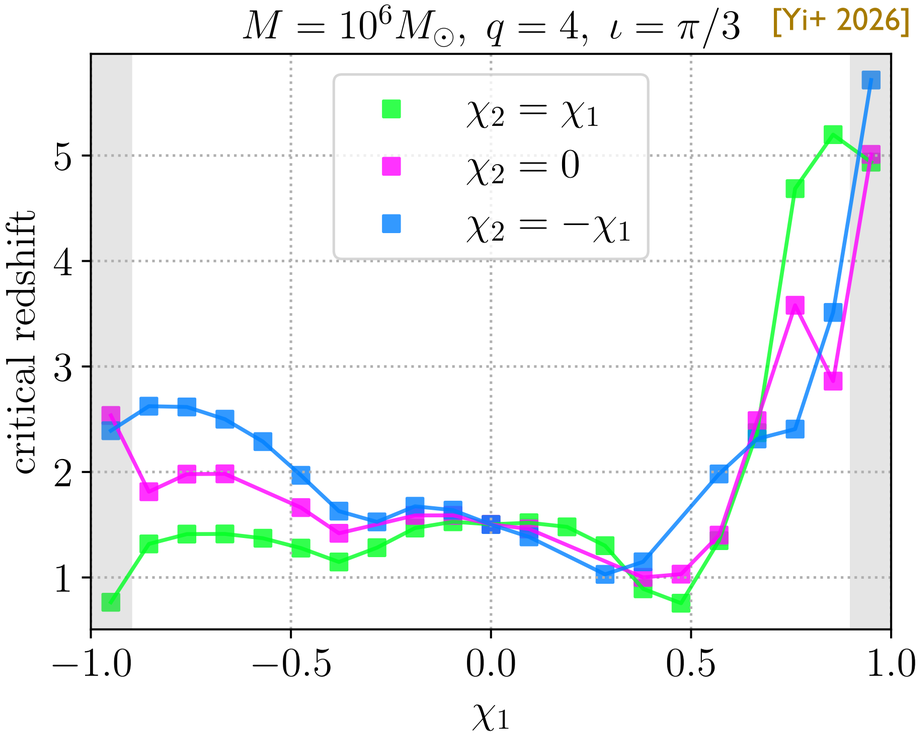
MBHB biases caused by missing higher harmonics

Question: how many $h_{\ell m}$ should be included in waveform templates ? Current models go up to h_{55} at most.

[Pitte+ 2023]
[Yi+ 2025]
[Yi+ 2026]

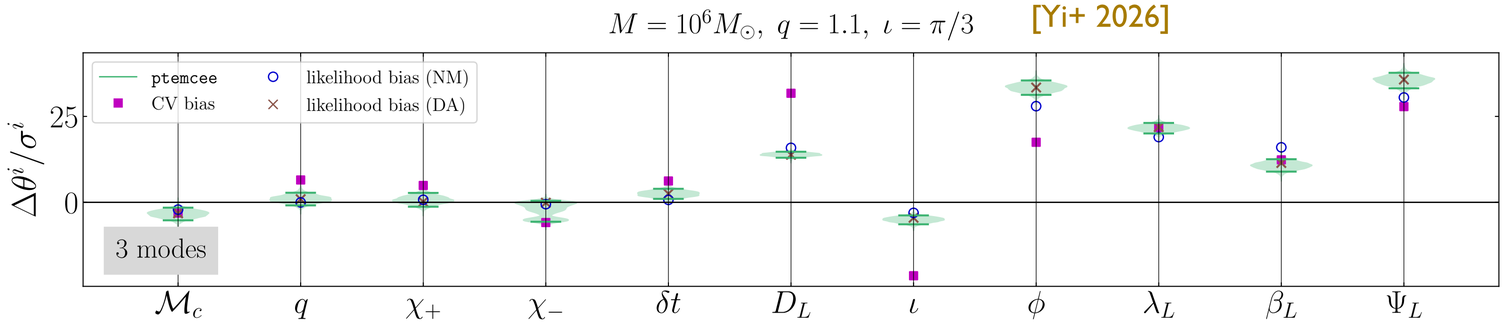
LISA consortium project [Waveform Working Group] coordinator: **Sophia Yi** (Johns Hopkins U.)

Min. redshift unbiased when ignoring (3,2) mode:



[Yi+ 2026]

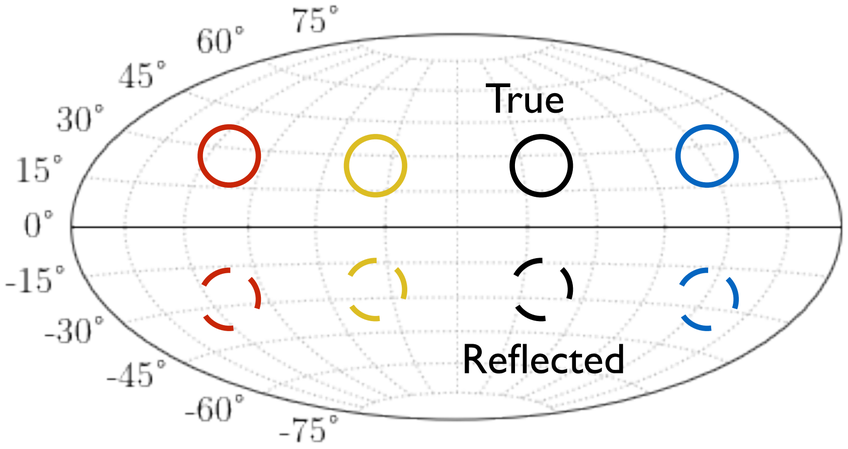
Estimating biases: Cutler-Vallisneri (can fail), direct Nelder-Mead

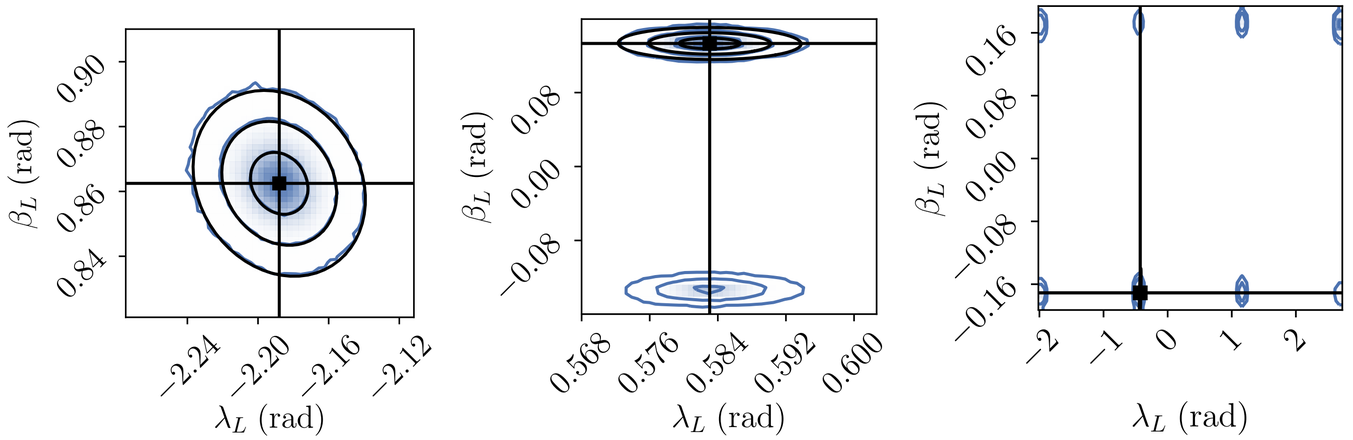


[Yi+ 2026]

LISA response and multimodality in the sky

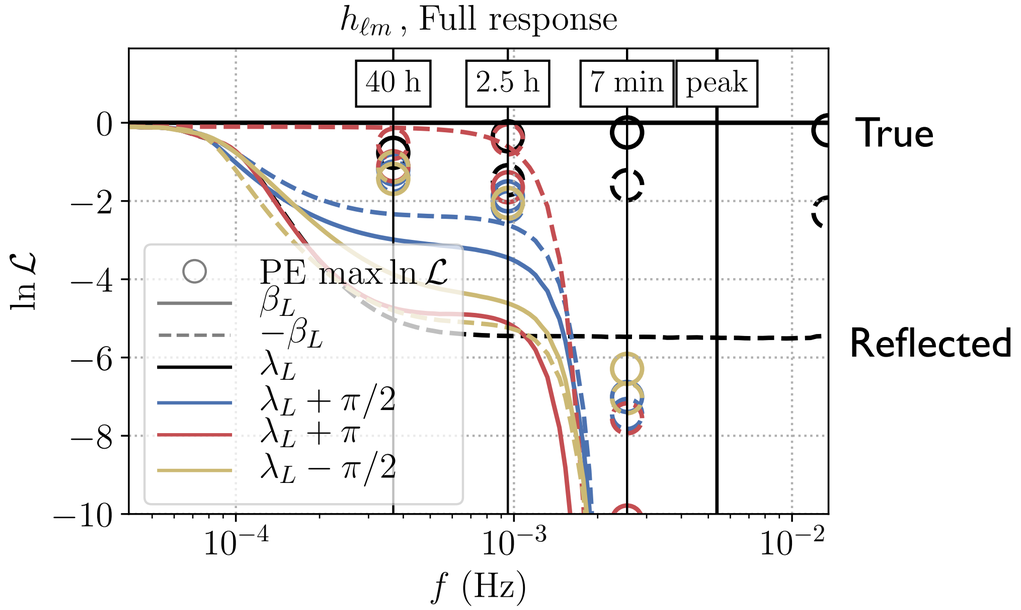
Multimodality pattern:





Degeneracy breaking:

- motion of LISA: eliminates all modes but the antipodal, weak for short high-mass signals
- high-frequency effects in the response: eliminates all modes but the reflected, only at high frequencies

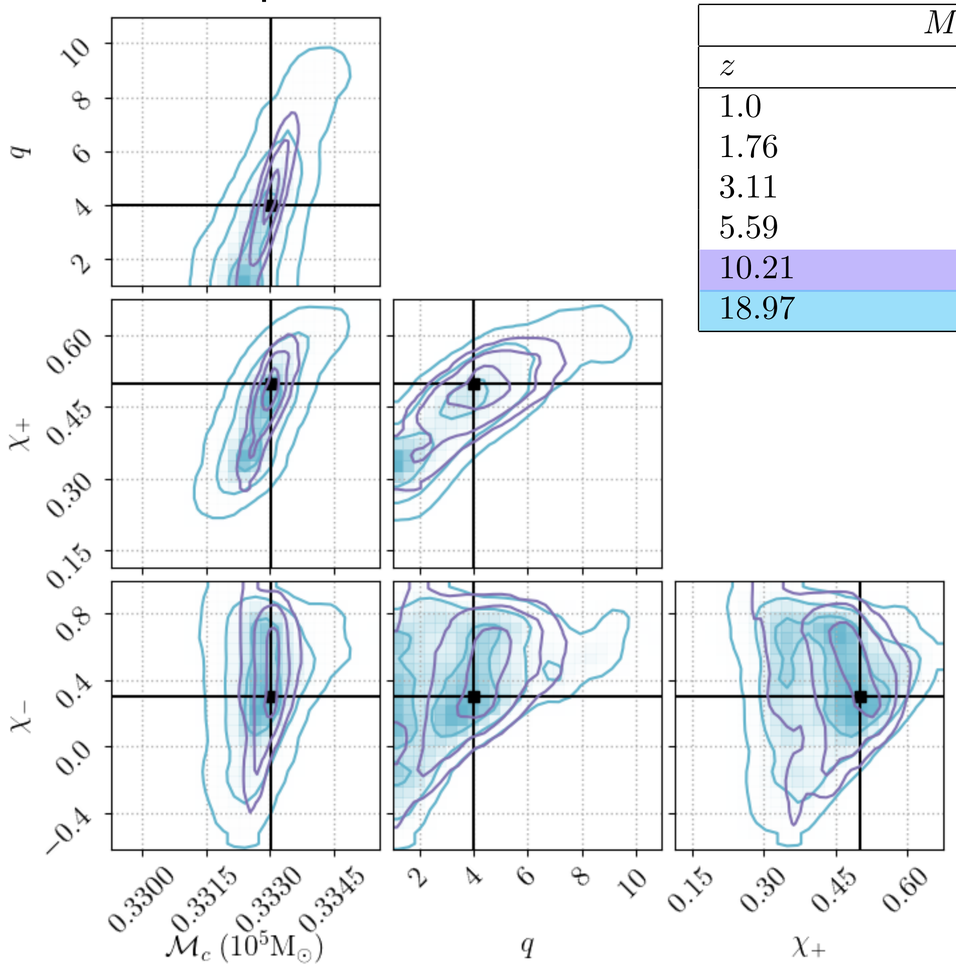


- Multimodality broken by subdominant effects in response (motion, high-f)

Example Parameter estimation with systematics I

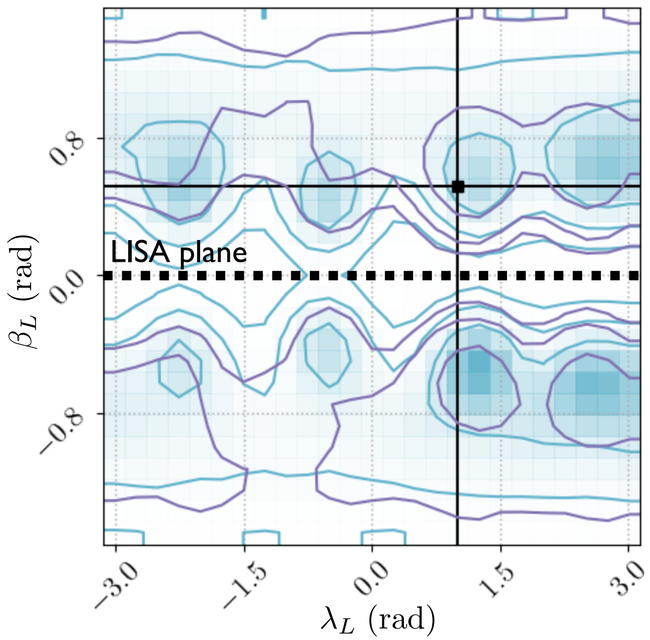
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^5\ M_\odot$	
z	SNR
1.0	317
1.76	158
3.11	79
5.59	40
10.21	20
18.97	10

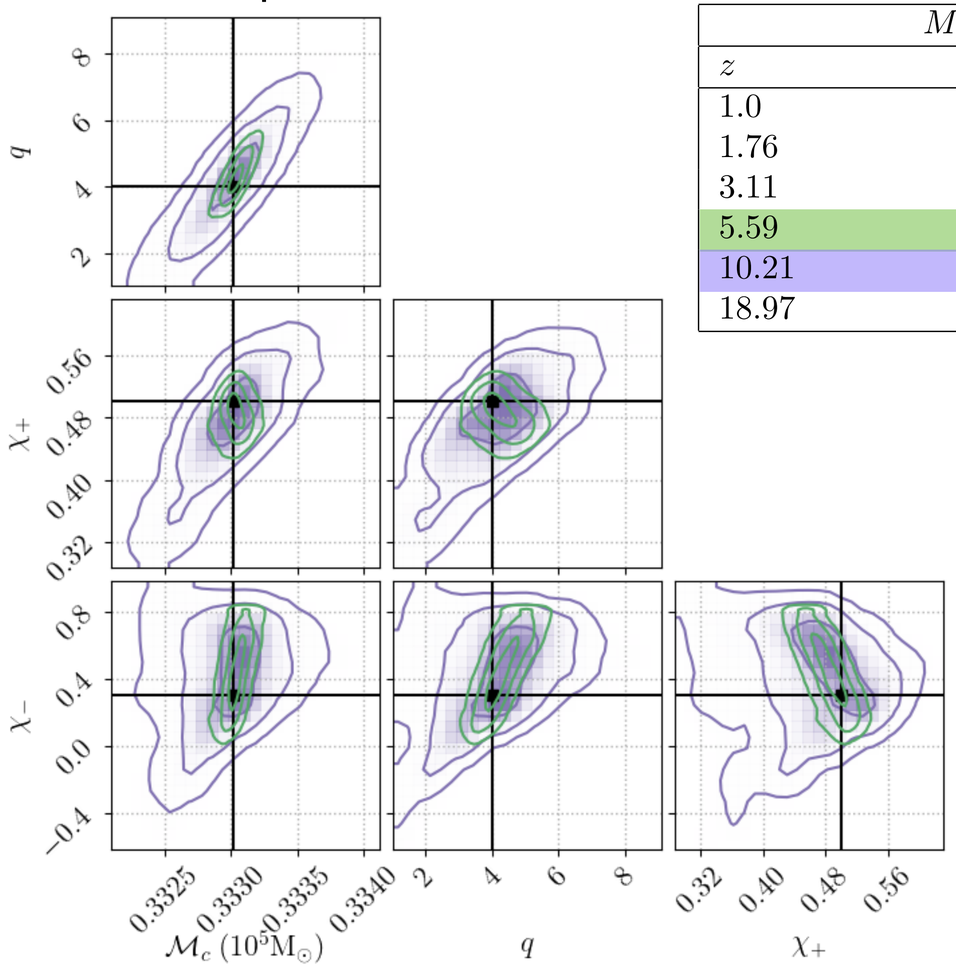
Sky localisation



Example Parameter estimation with systematics I

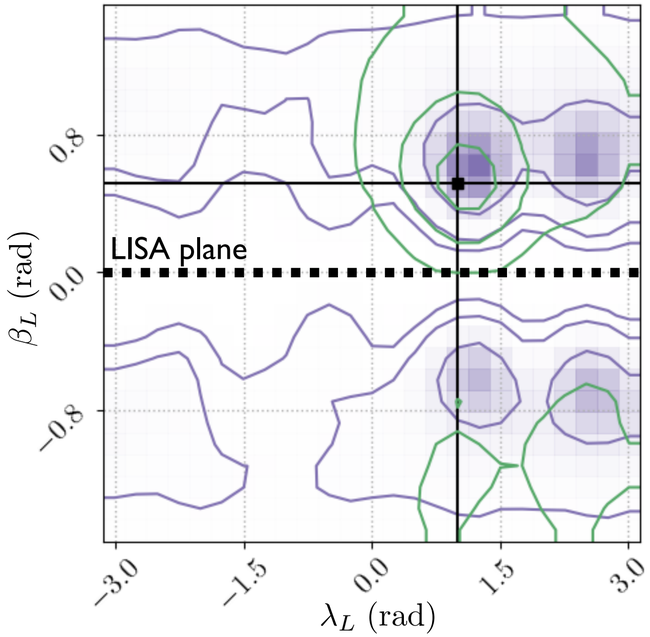
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^5\ M_\odot$	
z	SNR
1.0	317
1.76	158
3.11	79
5.59	40
10.21	20
18.97	10

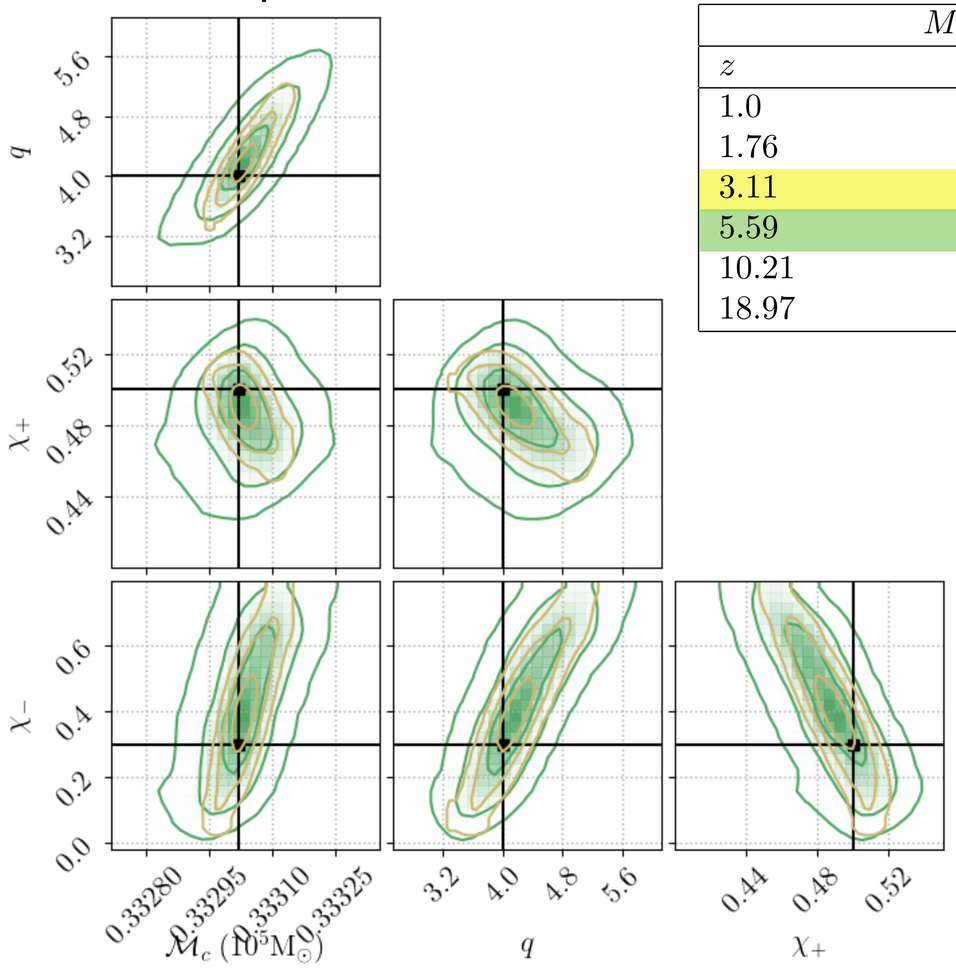
Sky localisation



Example Parameter estimation with systematics I

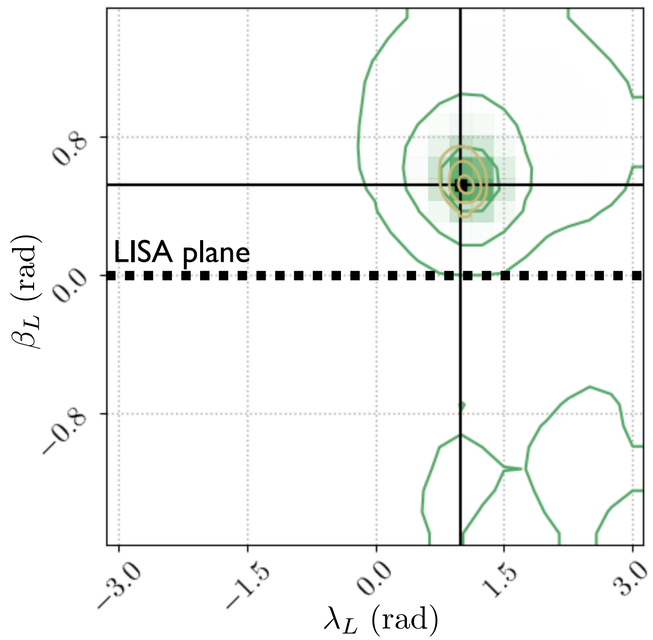
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^5\ M_\odot$	
z	SNR
1.0	317
1.76	158
3.11	79
5.59	40
10.21	20
18.97	10

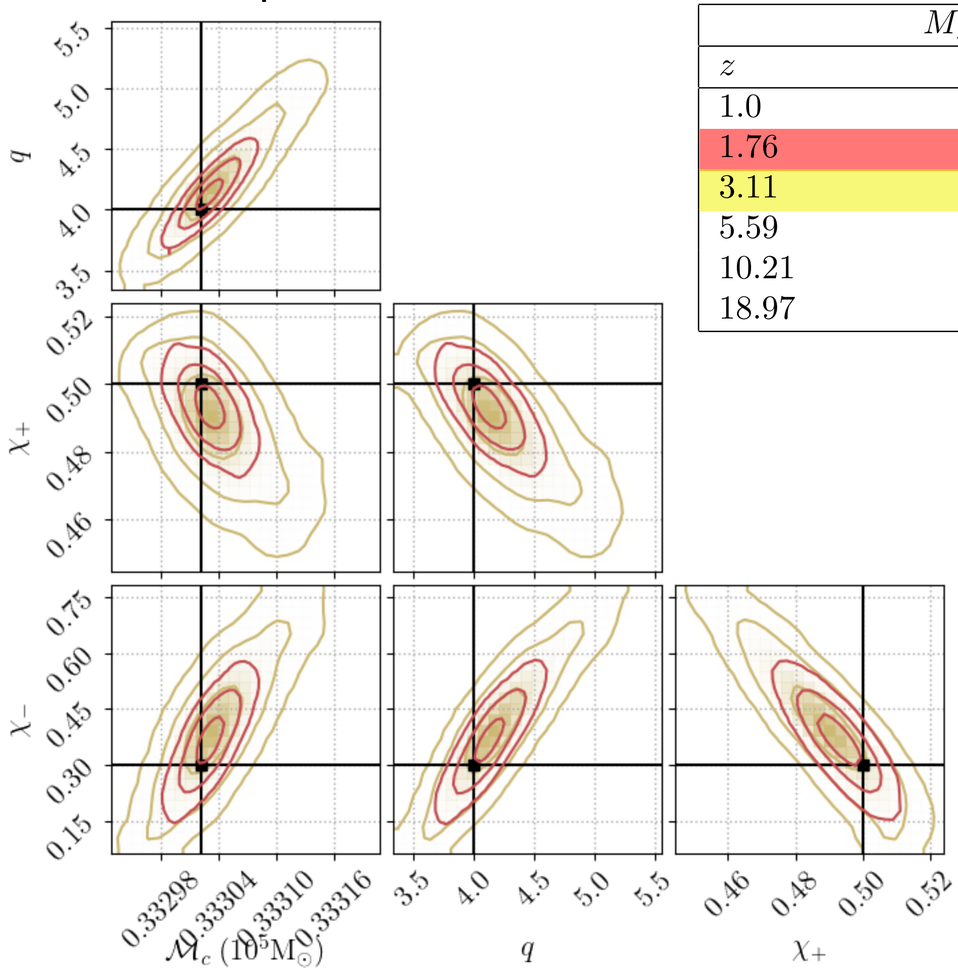
Sky localisation



Example Parameter estimation with systematics I

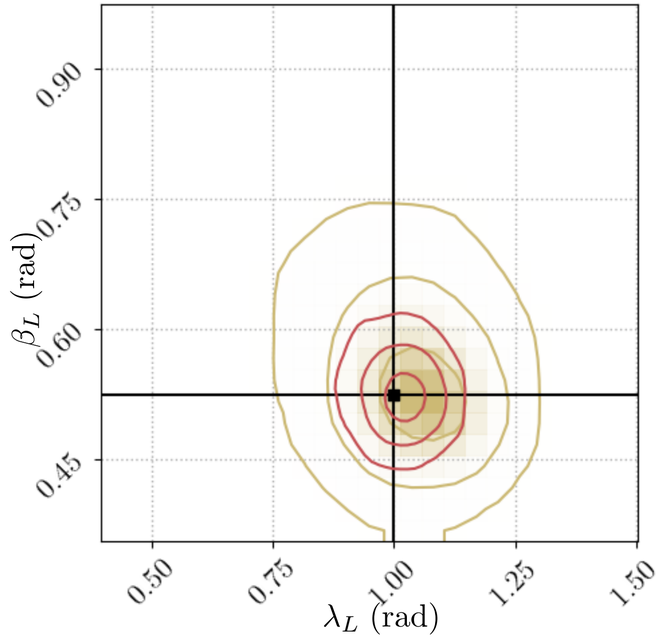
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^5\ M_\odot$	
z	SNR
1.0	317
1.76	158
3.11	79
5.59	40
10.21	20
18.97	10

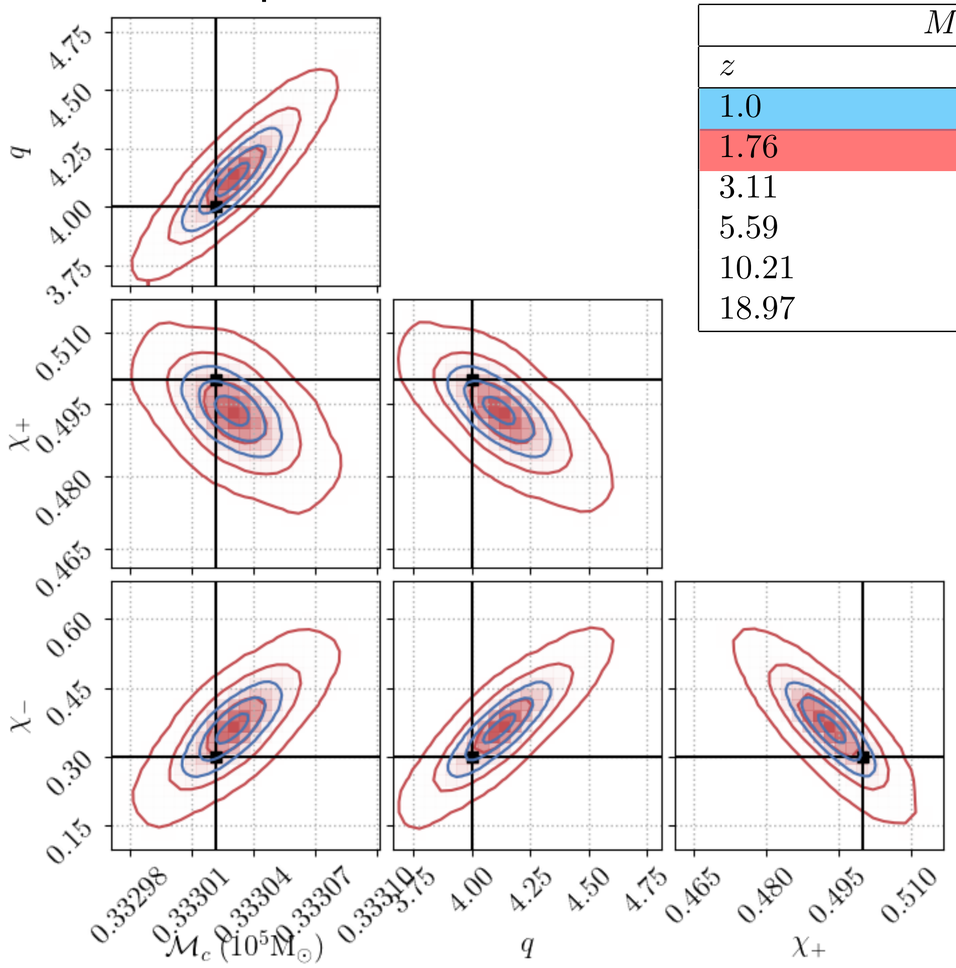
Sky localisation



Example Parameter estimation with systematics I

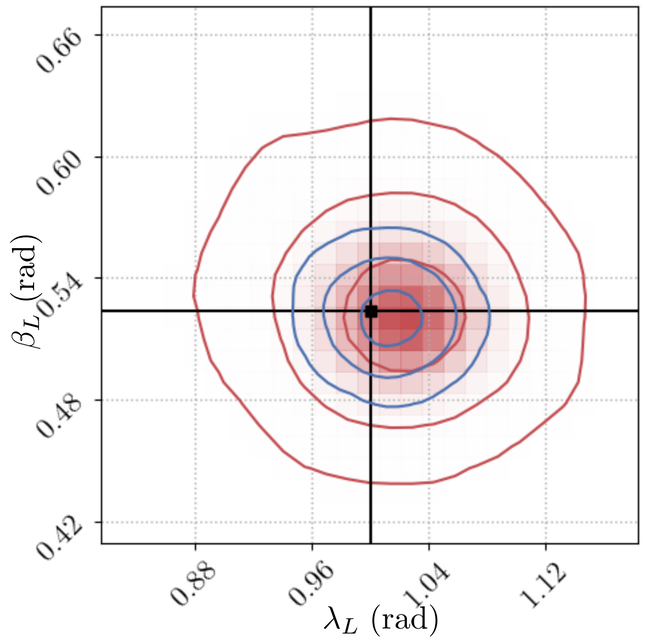
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^5\ M_\odot$	
z	SNR
1.0	317
1.76	158
3.11	79
5.59	40
10.21	20
18.97	10

Sky localisation

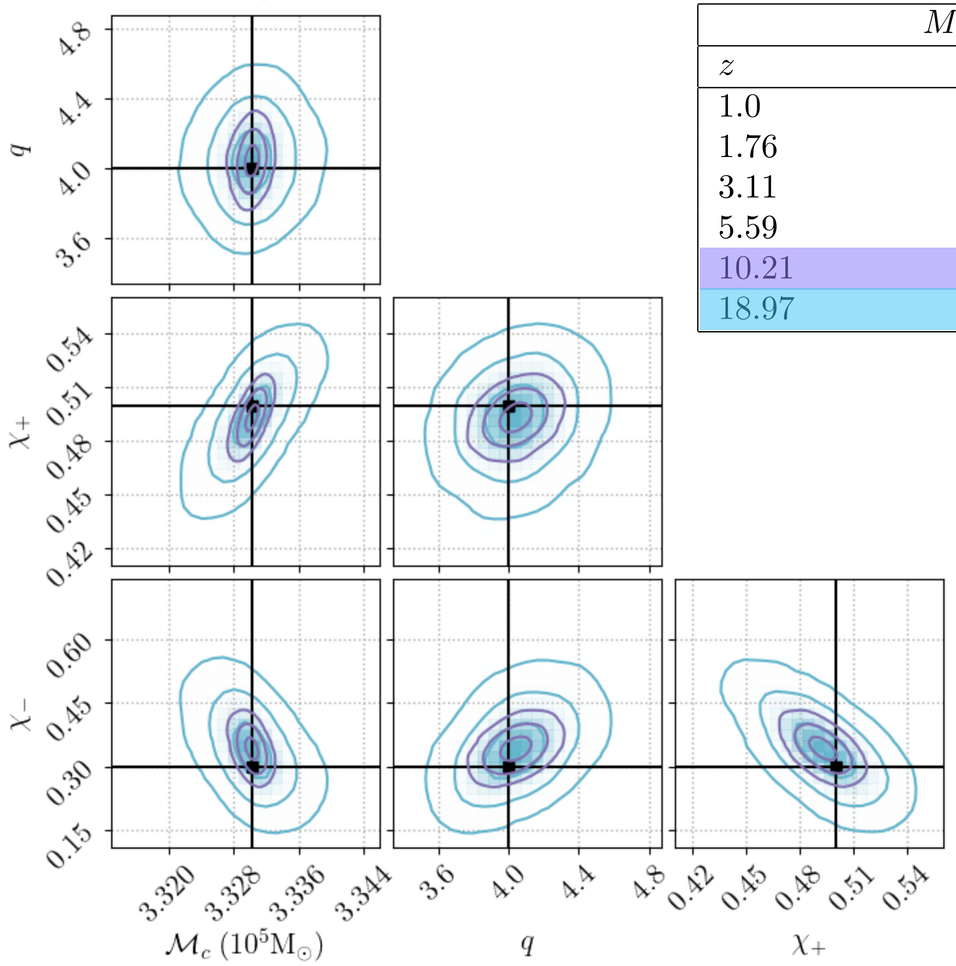


The good:
- converges on the true parameters
- mild bias at $z = 1, \mathrm{SNR} = 317$

Example Parameter estimation with systematics II

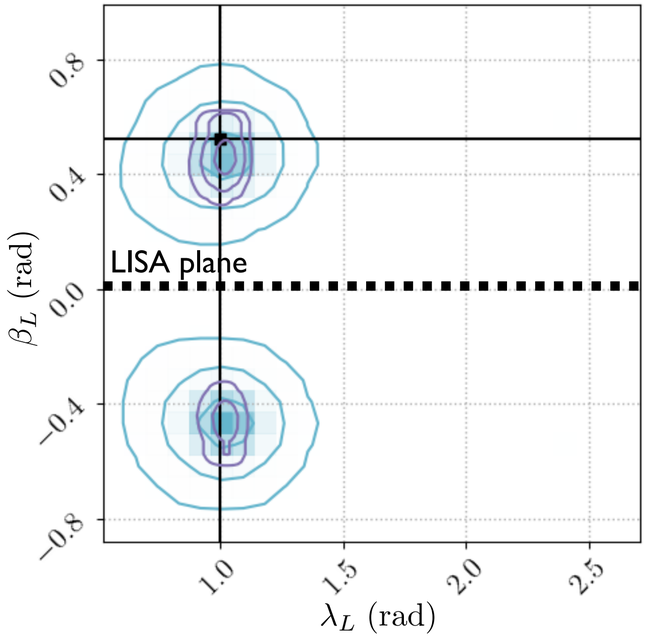
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

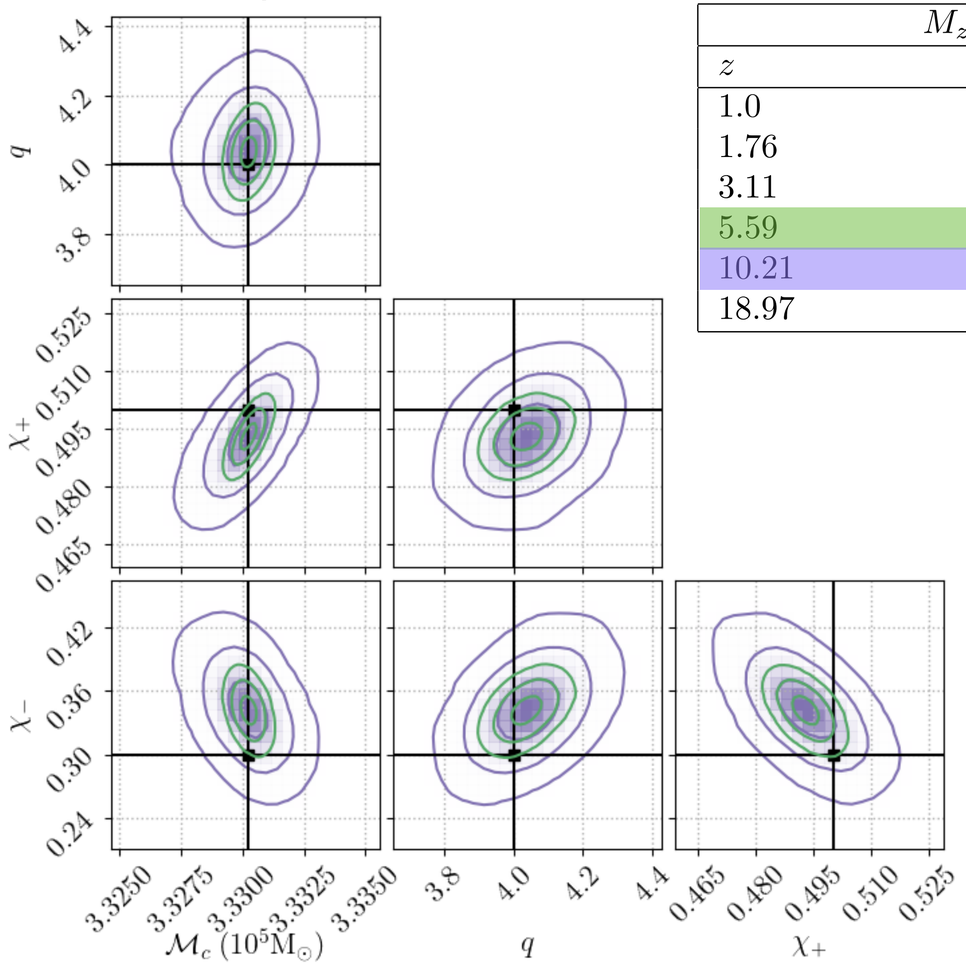
Sky localisation



Example Parameter estimation with systematics II

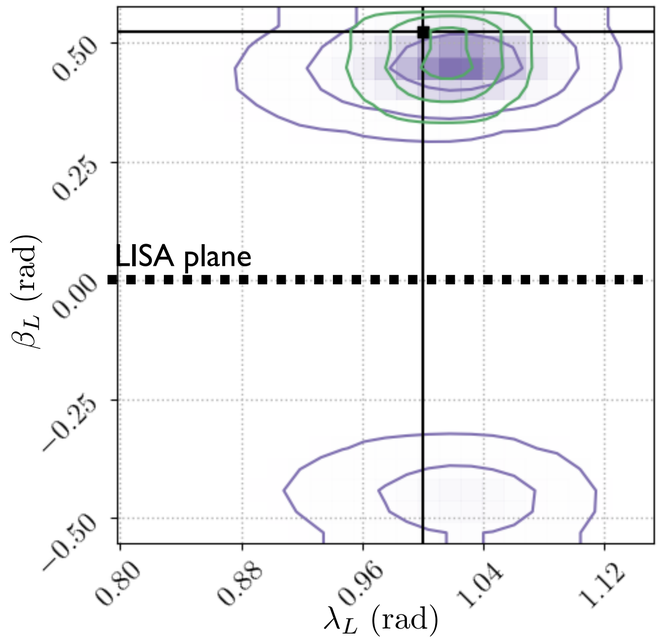
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

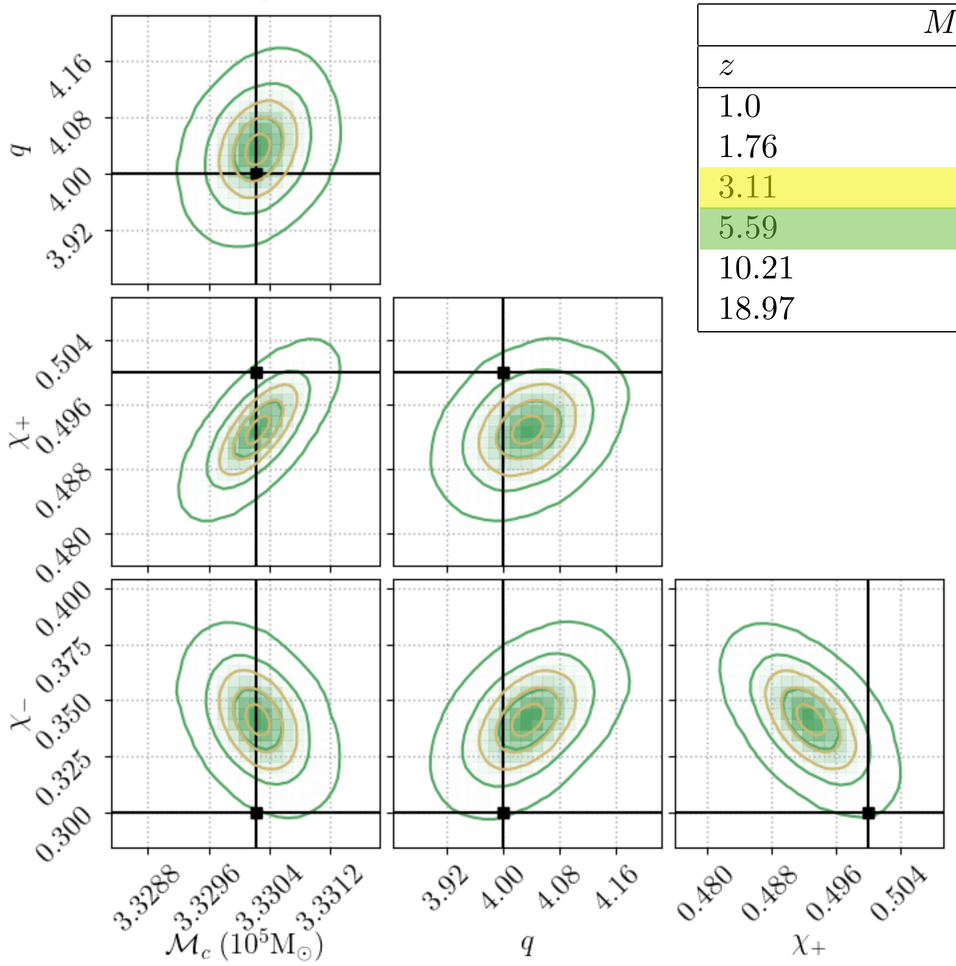
Sky localisation



Example Parameter estimation with systematics II

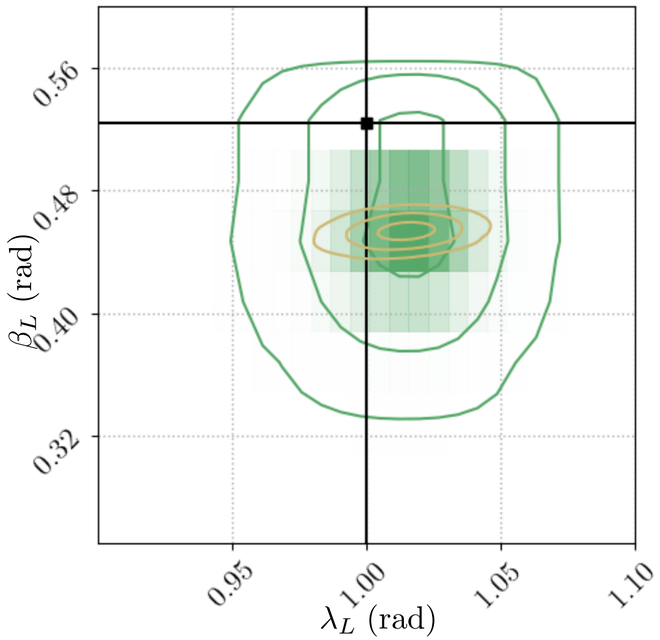
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

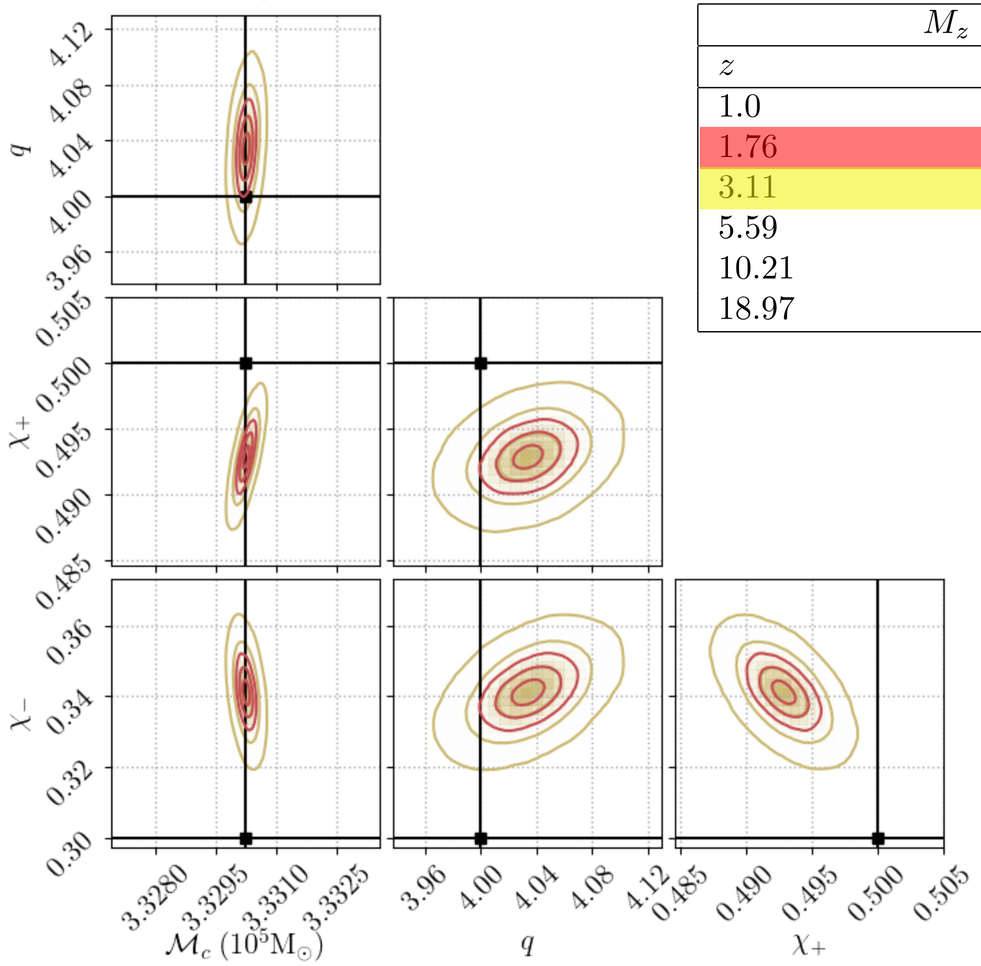
Sky localisation



Example Parameter estimation with systematics II

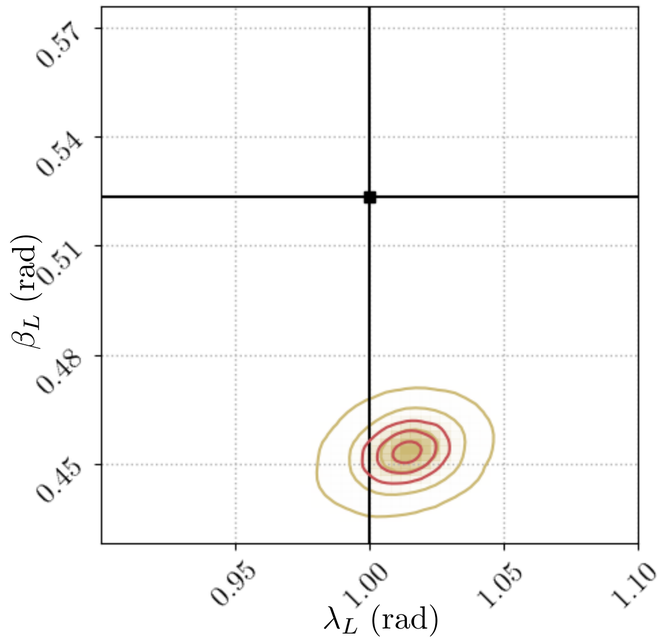
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

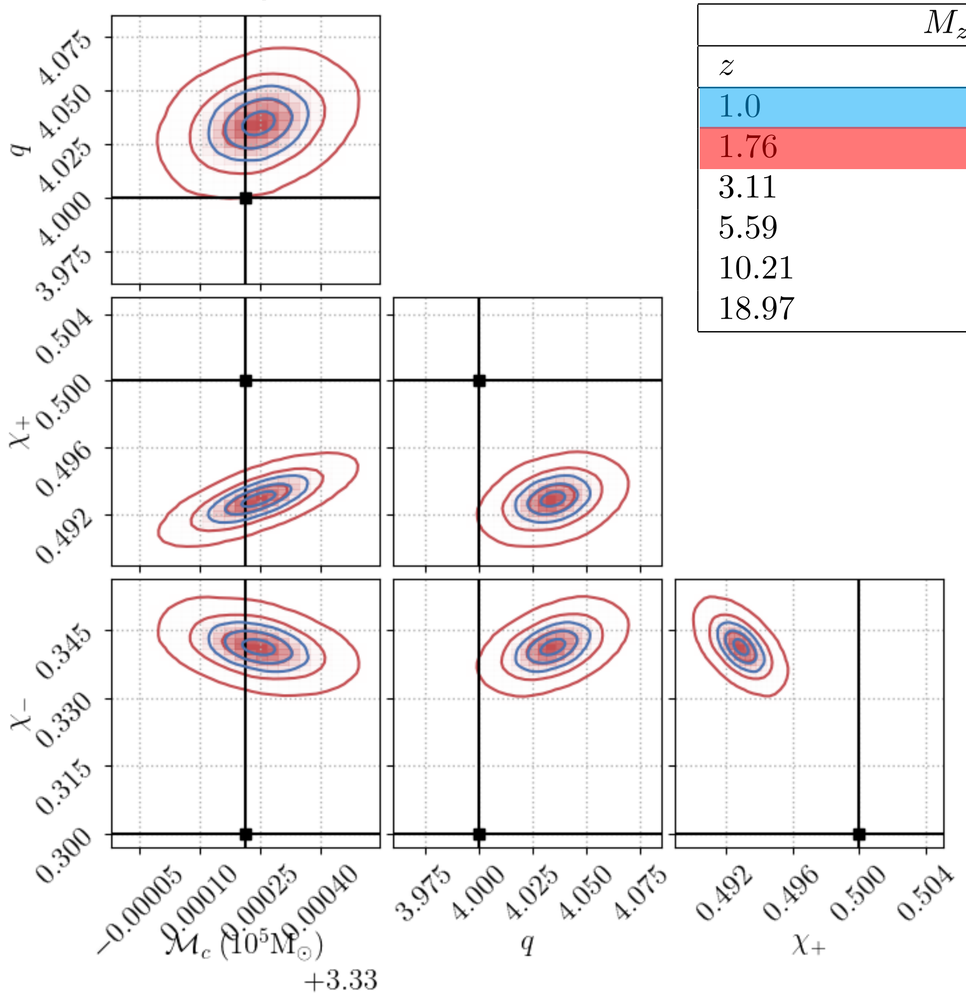
Sky localisation



Example Parameter estimation with systematics II

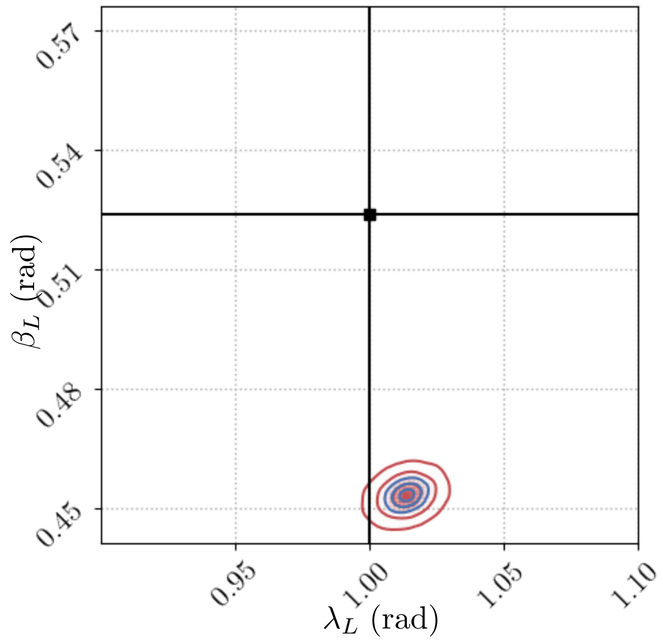
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

Sky localisation



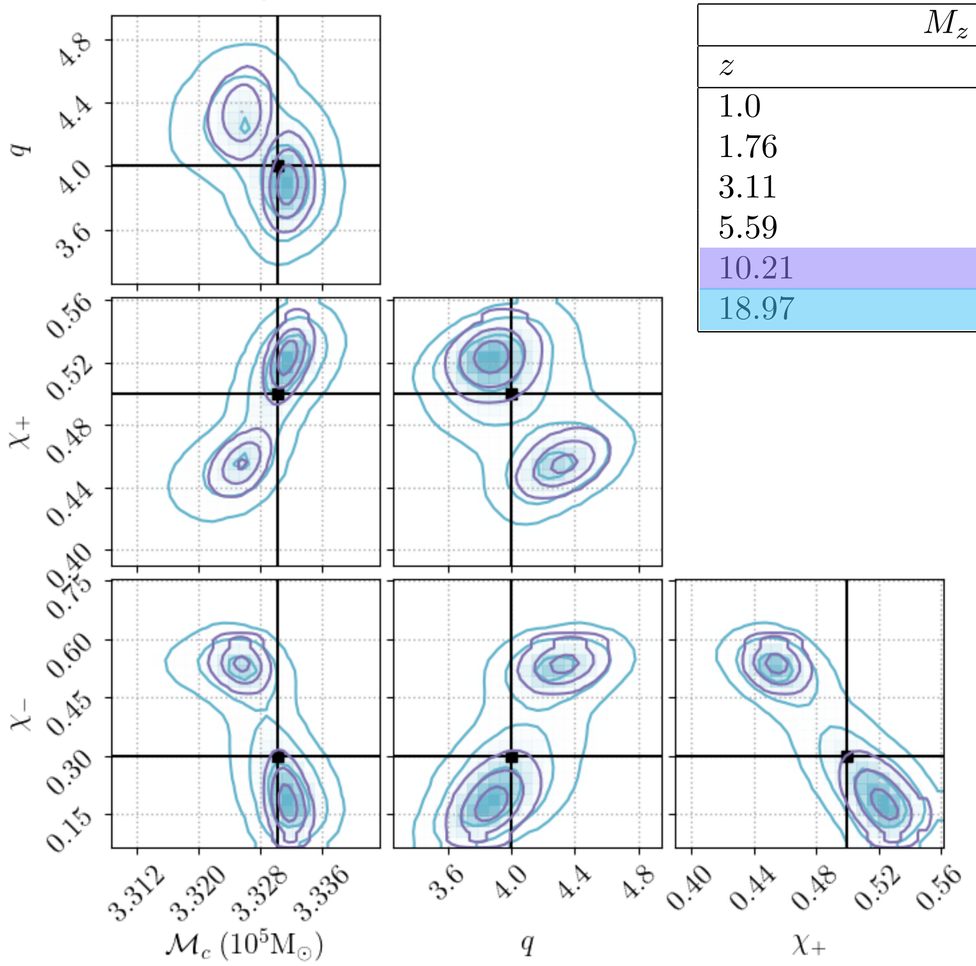
The bad:

- converges 'near' the true parameters
- significant bias at $z = 1, \mathrm{SNR} = 1907$

Example Parameter estimation with systematics III

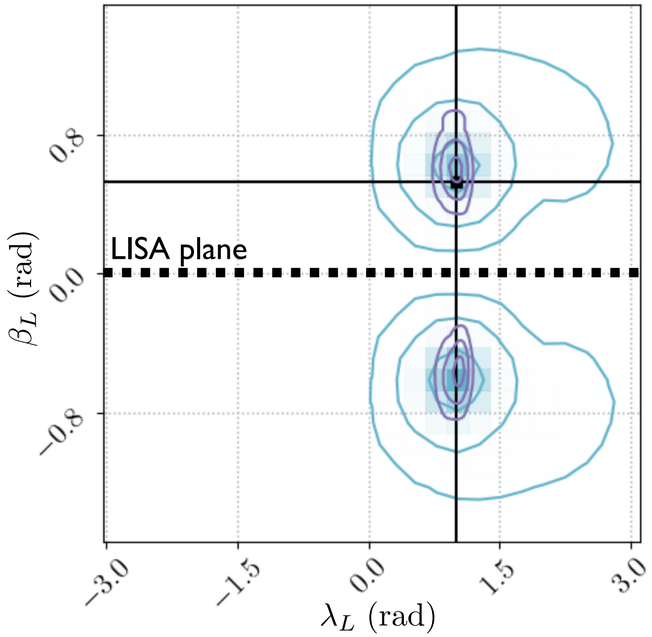
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

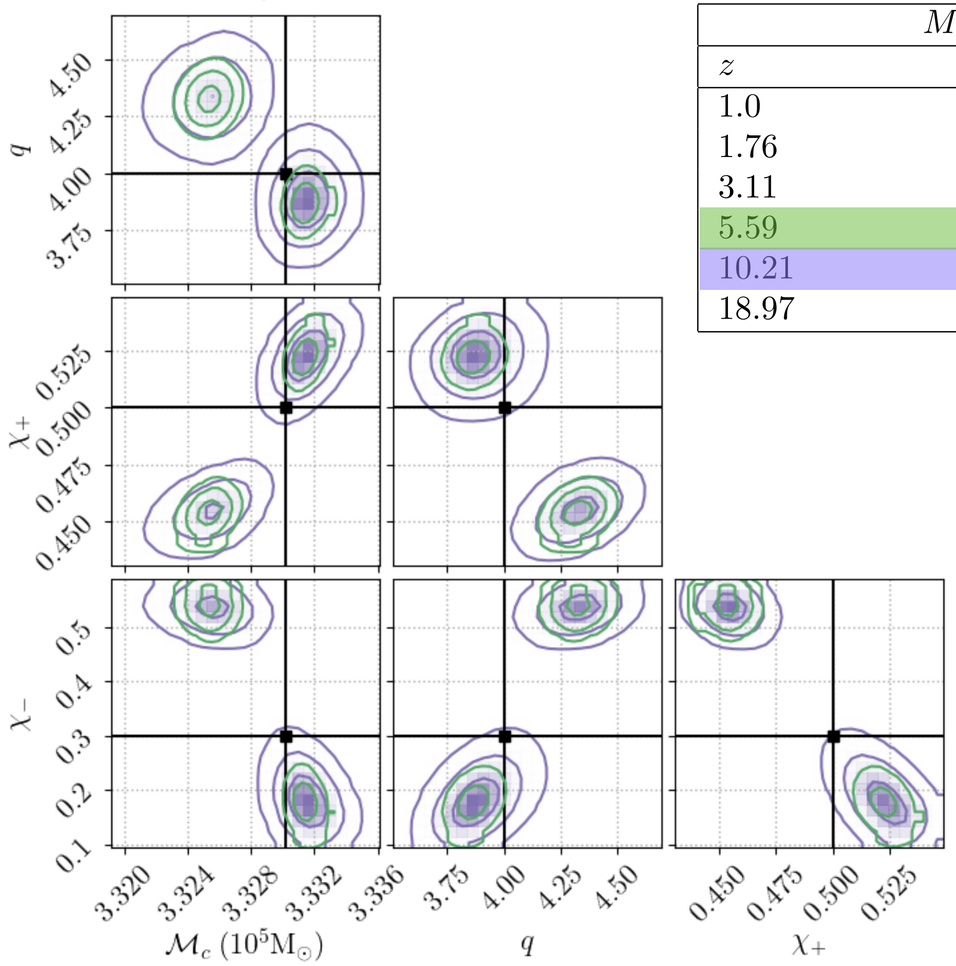
Sky localisation



Example Parameter estimation with systematics III

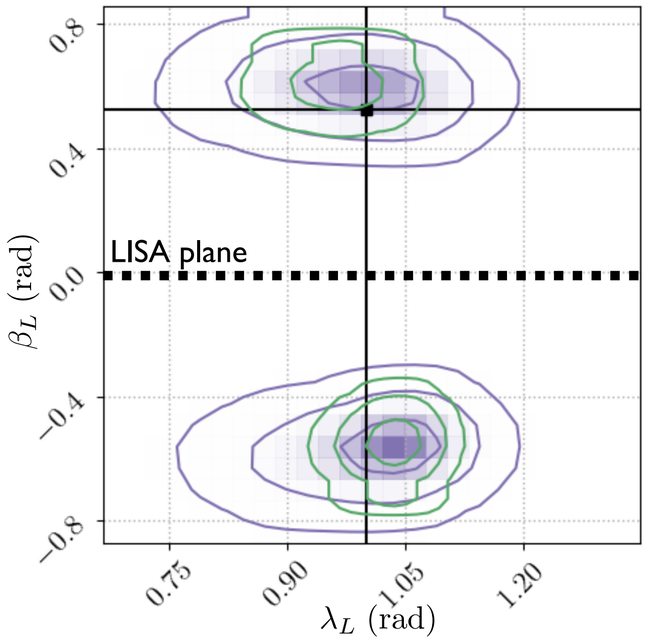
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

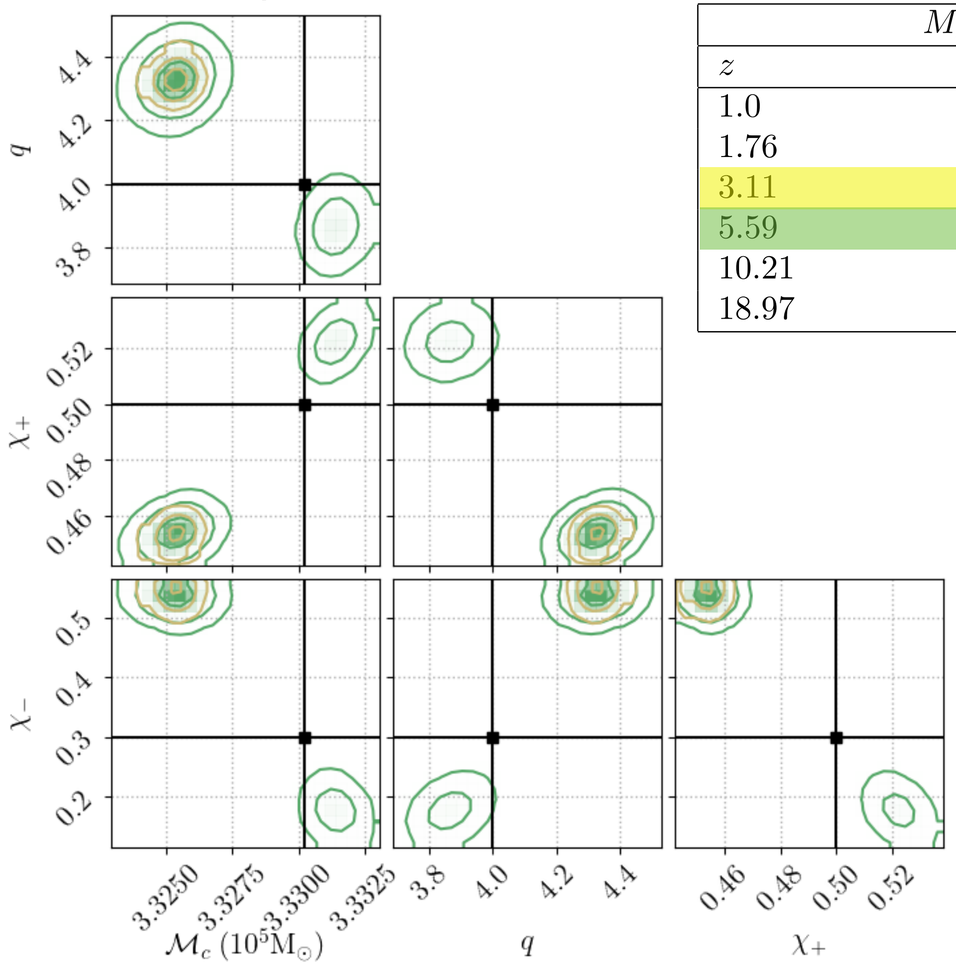
Sky localisation



Example Parameter estimation with systematics III

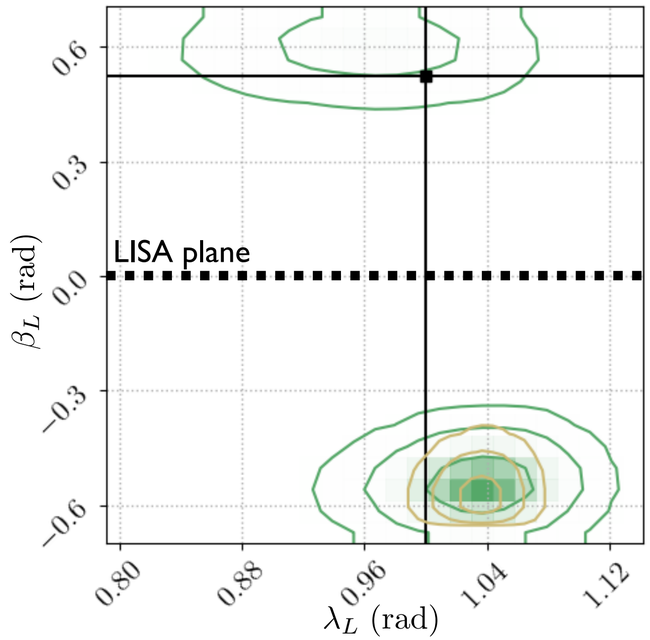
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

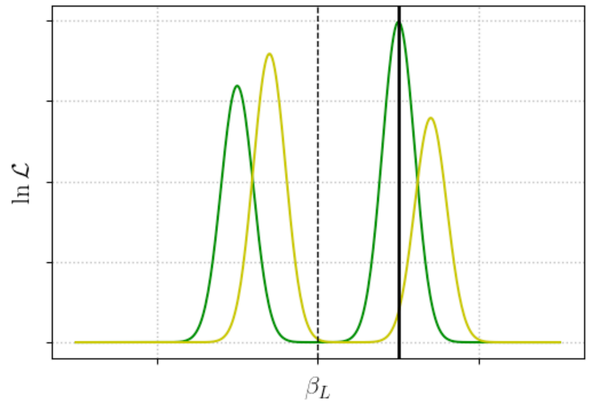
Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

Sky localisation

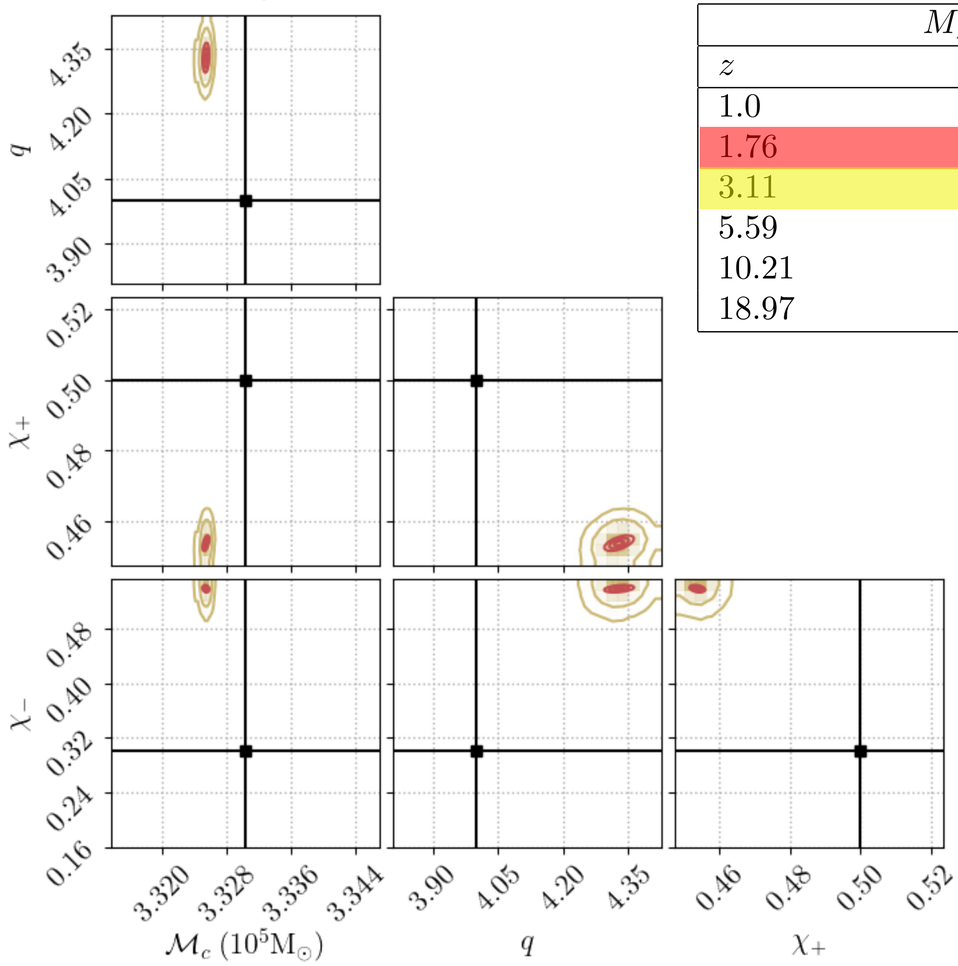




Example Parameter estimation with systematics III

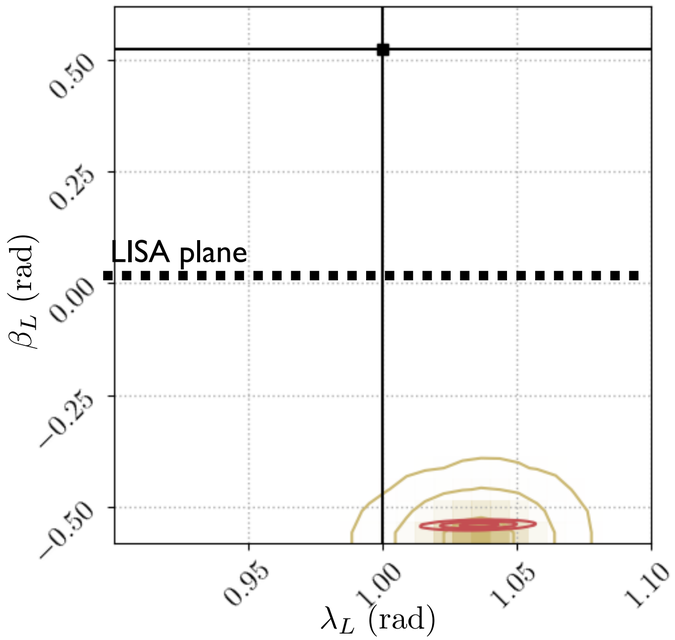
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

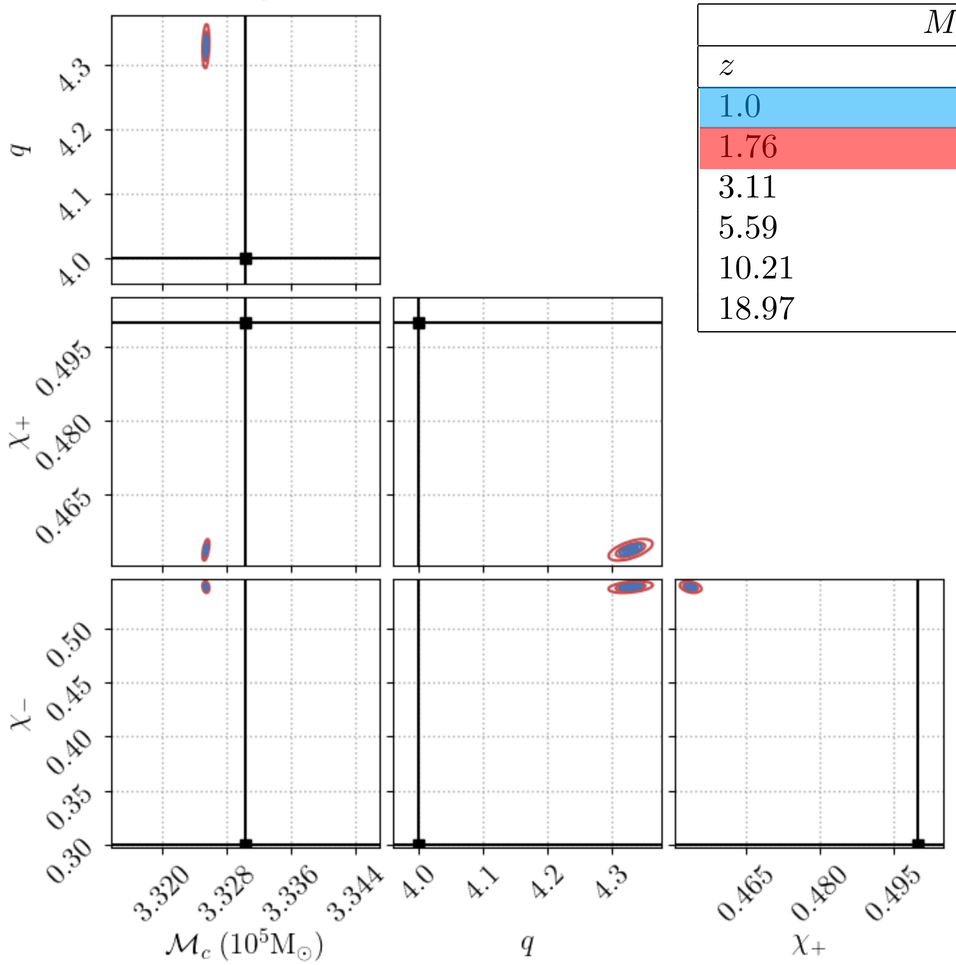
Sky localisation



Example Parameter estimation with systematics III

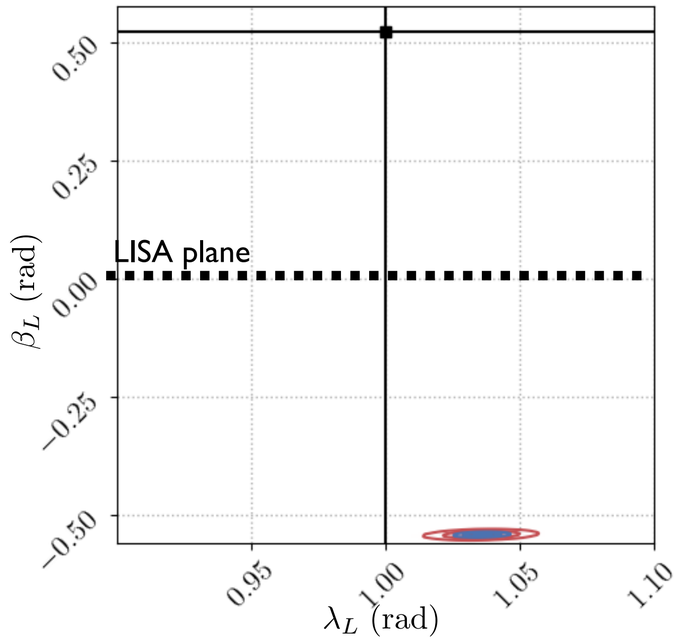
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

Intrinsic params.



$M_z = 10^6\ M_\odot$	
z	SNR
1.0	1907
1.76	954
3.11	477
5.59	238
10.21	119
18.97	59

Sky localisation

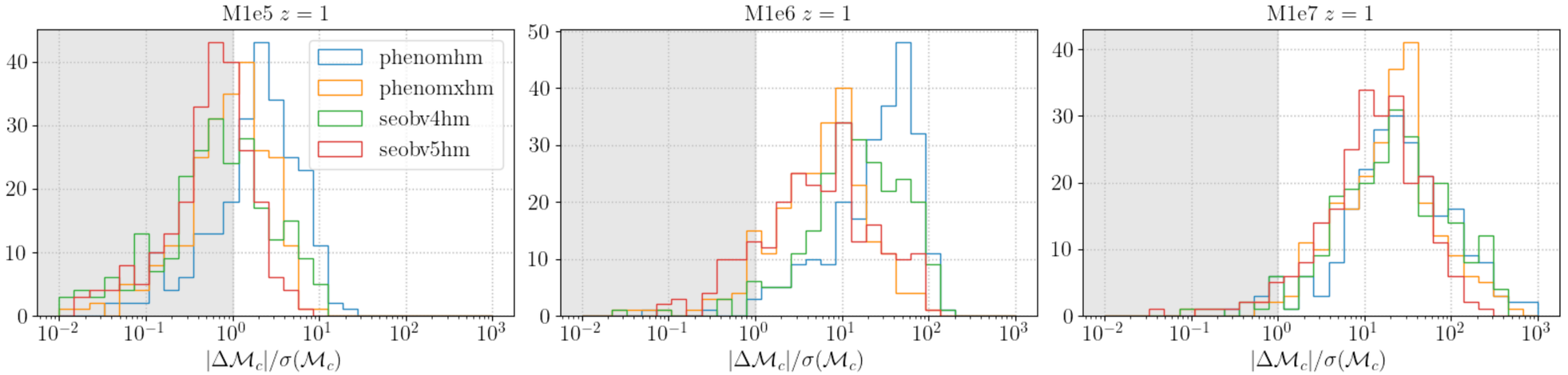


The ugly:

- converges to the **wrong sky mode**
- important bias at $z = 1, \mathrm{SNR} = 1907$

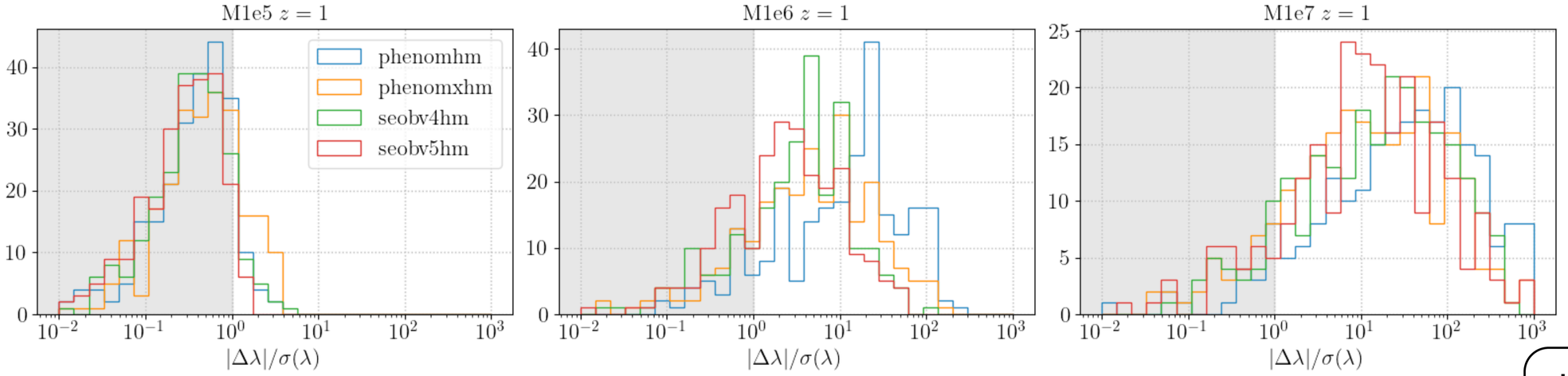
Statistical significance of biases: intrinsic parameters

Bias in chirp mass:



Injections:
NRHybSur3dq8

Bias in longitude (on corrected skymode):

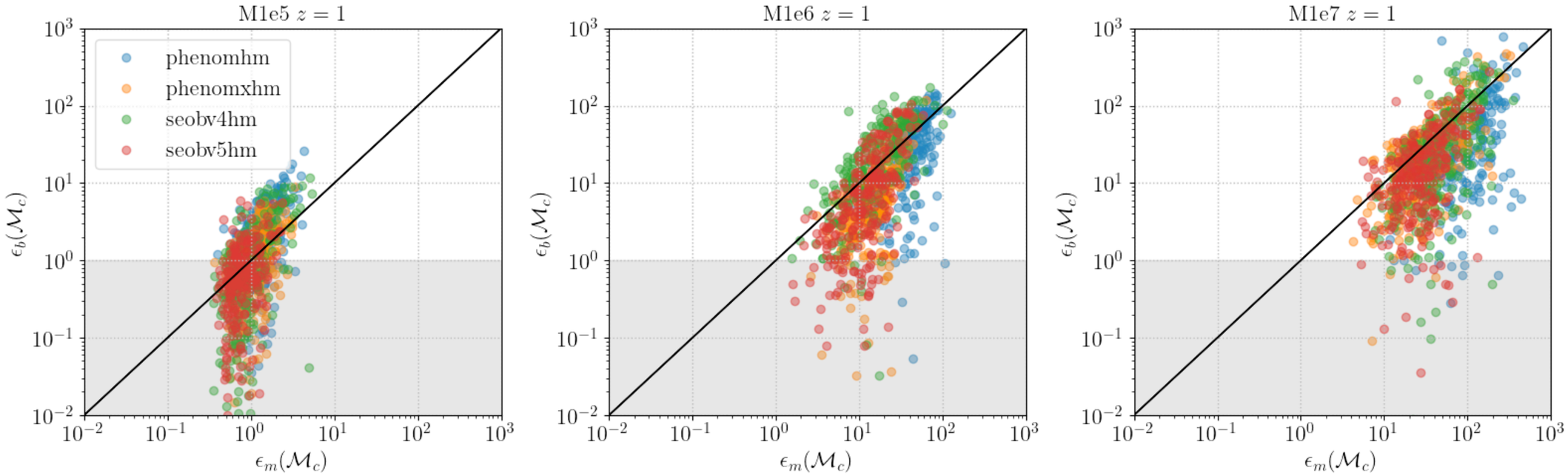


Wrong skymode recovered:

phenomxhm	
$M_z(M_\odot)$	% wrong
10^5	0 %
10^6	10 %
10^7	54 %

Large biases at high mass
Wrong skymodes common

LISA systematics: linking mismatches and biases



From indistinguishability criterion:

$$\mathrm{MM} < \frac{D}{2}\frac{1}{\mathrm{SNR}^2}$$

$$\epsilon_m = \sqrt{\frac{2}{D}\mathrm{SNR}^2\mathrm{MM}}$$

$\epsilon_m > 1$ means that the mismatch is large enough to indicate a significant bias

From bias measured in PE:

$$\epsilon_b = \frac{\Delta\theta}{\sigma(\theta)}$$

$\epsilon_b > 1$ indicates means that PE measures a significant bias

Both $\epsilon_b,\ \epsilon_m \propto \mathrm{SNR}$

Mismatch criterion usually pessimistic, can also be saturated

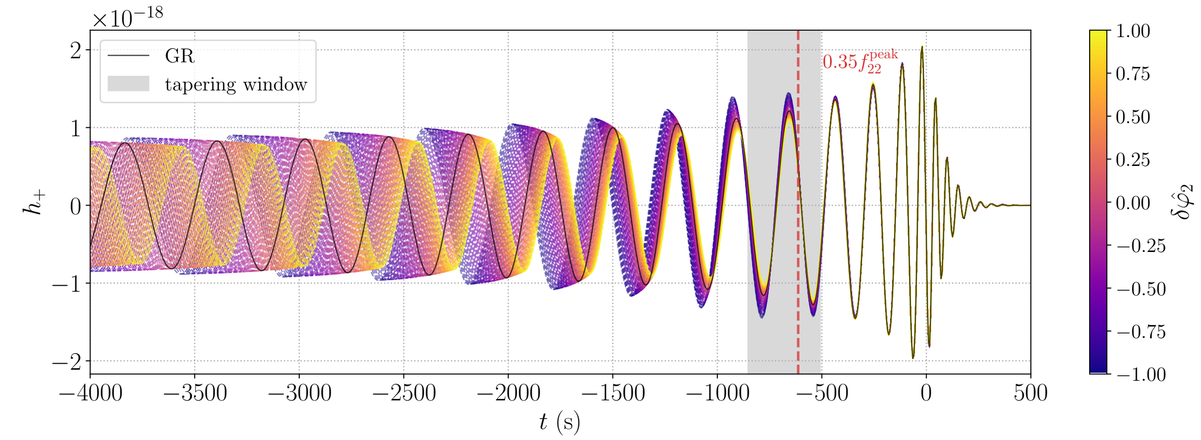
MBHB inspiral tests of GR

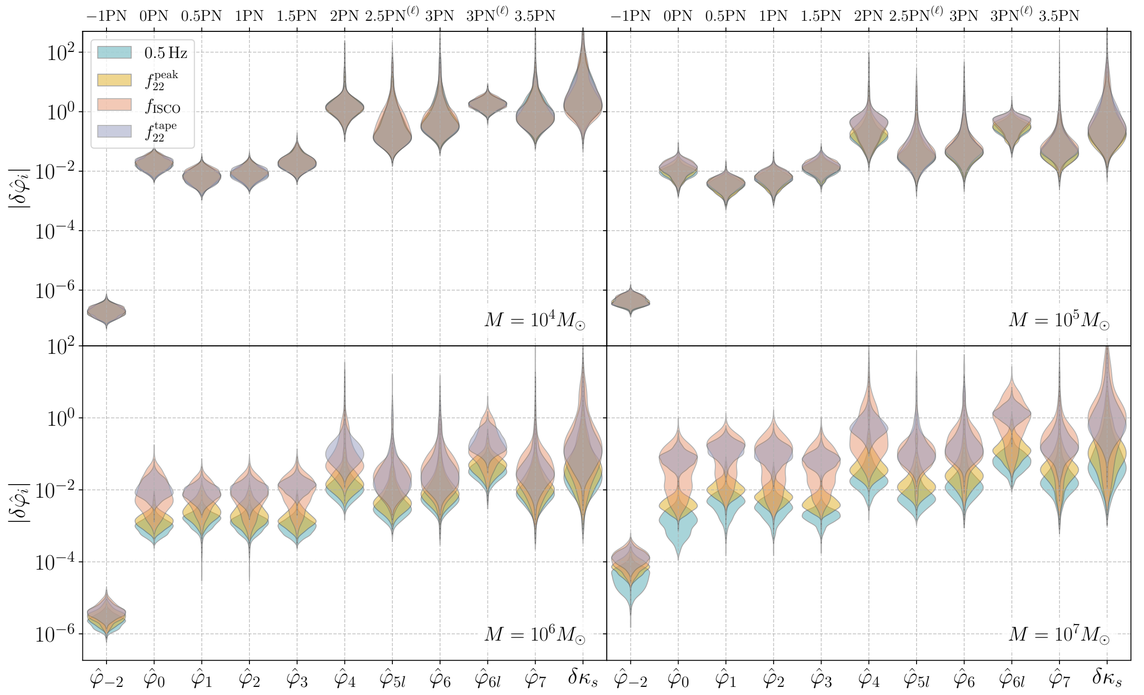
Theory-agnostic inspiral tests of GR, LISA MBHBs:

$$\delta\psi_{lm}(f) = \frac{3}{128\eta v^5}\frac{m}{2}\Bigg[\sum_{\substack{n=0\\ n\neq 1}}^{7}\delta\hat{\varphi}_n\psi_n^{\rm PN}v^n + \delta\kappa_s(\psi_{4,\kappa_s}^{\rm PN}v^4 + \psi_{6,\kappa_s}^{\rm PN}v^6) + \sum_{n=5}^{6}\delta\hat{\varphi}_{n(\ell)}\psi_{n(\ell)}^{\rm PN}v^n\log v + \sum_{n=-2,-1,1}\delta\hat{\varphi}_n v^n\Bigg].$$

- Interplay between length and SNR for best measurements
- Interplay between merger and inspiral information: deviation in inspiral but mass/spin information from merger

[Piarulli+ 2025]





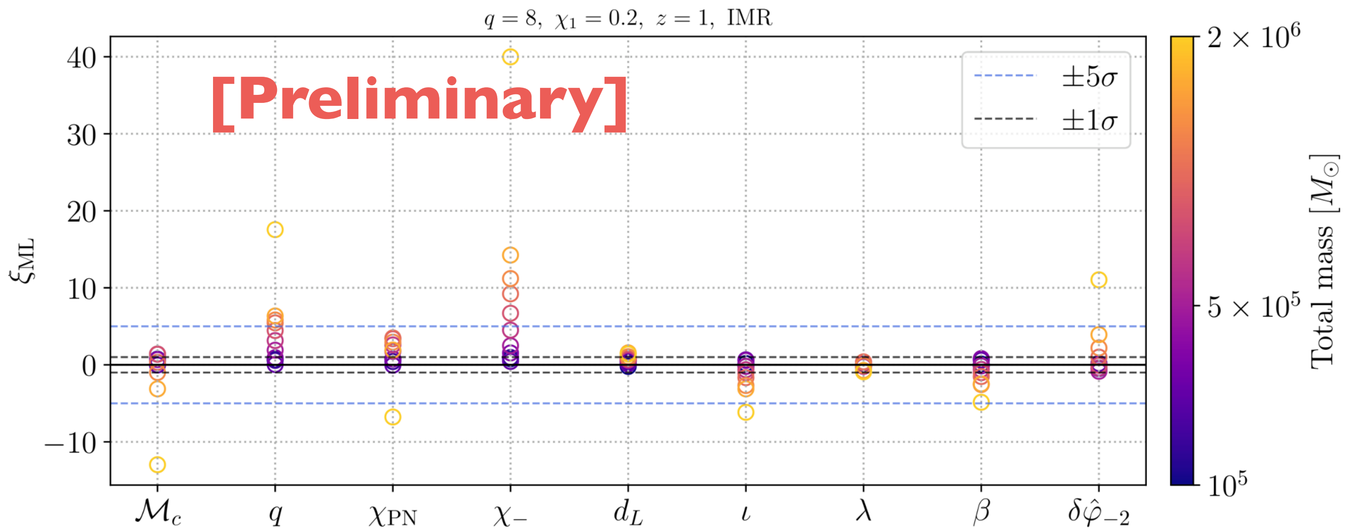
[Piarulli+ 2025]

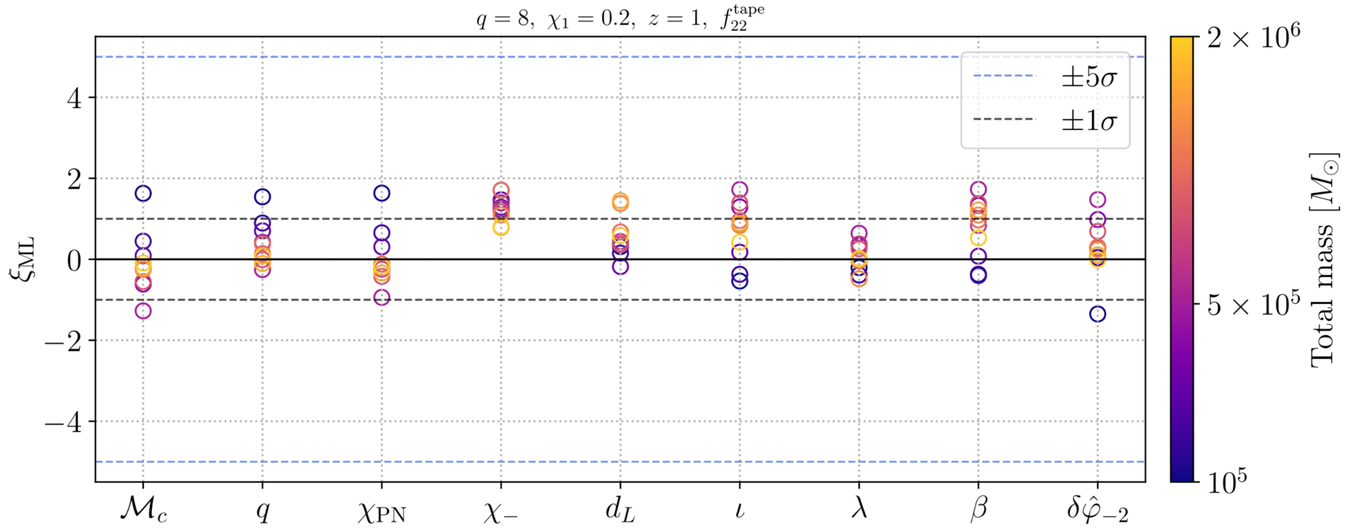
MBHB biases: impact on tests of GR

Question: how does waveform inaccuracy affect tests of GR, and can waveform errors be mistaken for false deviations from GR ?

LISA consortium project [Fundamental Physics Working Group] coordinators: **Manuel Piarulli** (L2IT), **Elisa Maggio** (La Sapienza, Roma)

- Injection: NRsurrogate
- Recovery: SEOB-ROM
- Exploration of role of total mass
- Exploration of the role of inspiral vs merger-ringdown
- Recovery with free inspiral deviation parameters

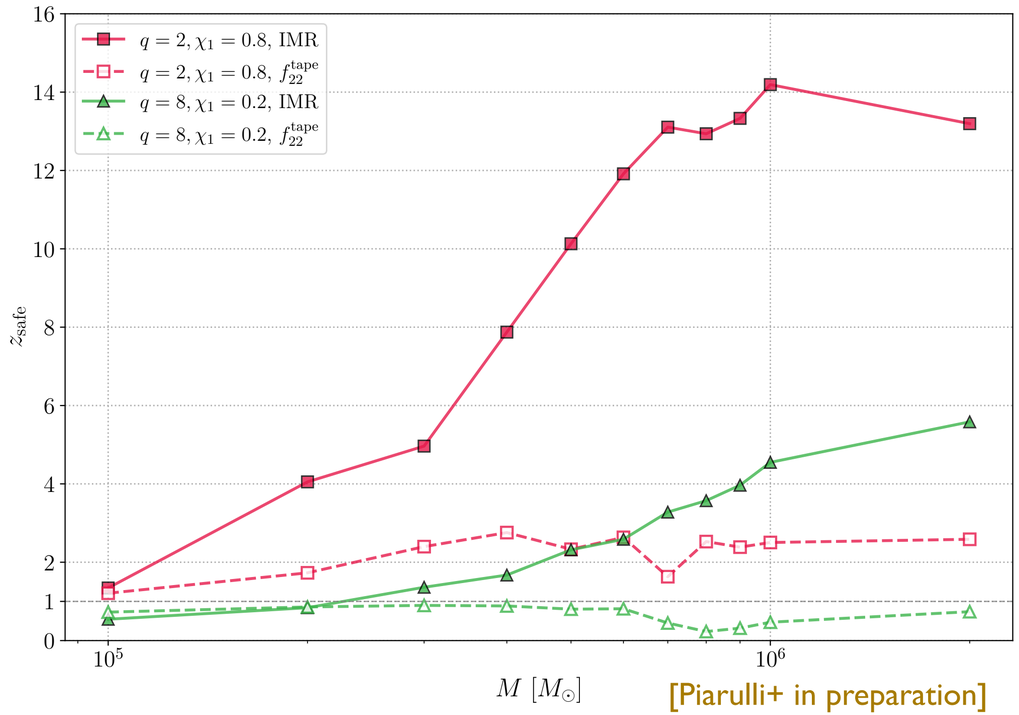




[Piarulli+ in preparation]

MBHB biases: impact on tests of GR

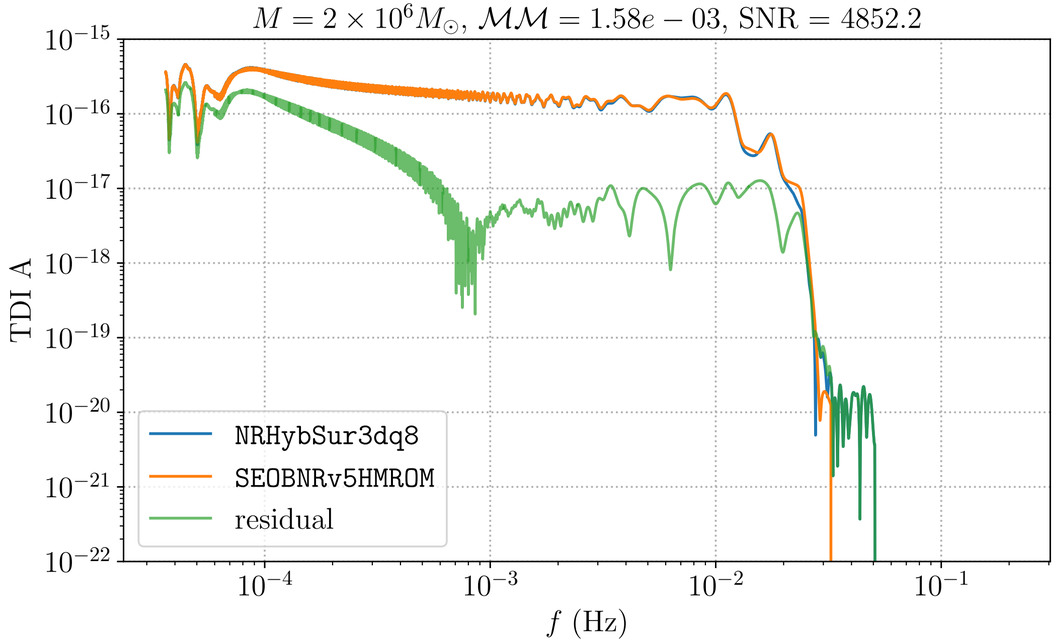
Safe redshift: where the bias would become smaller than rescaled statistical uncertainty



[Piarulli+ in preparation]

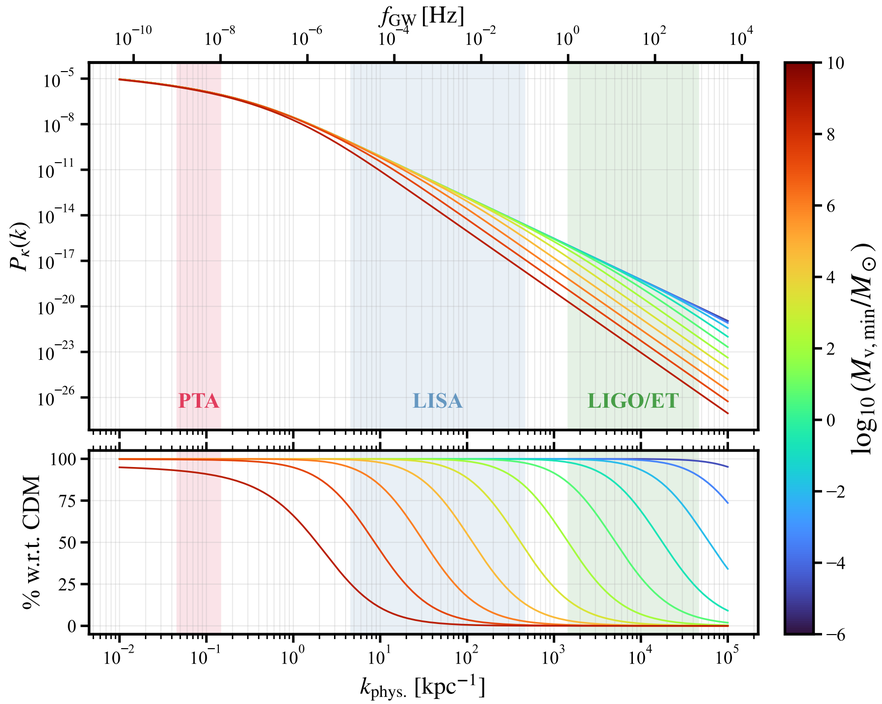
[Preliminary]

Signal and residuals

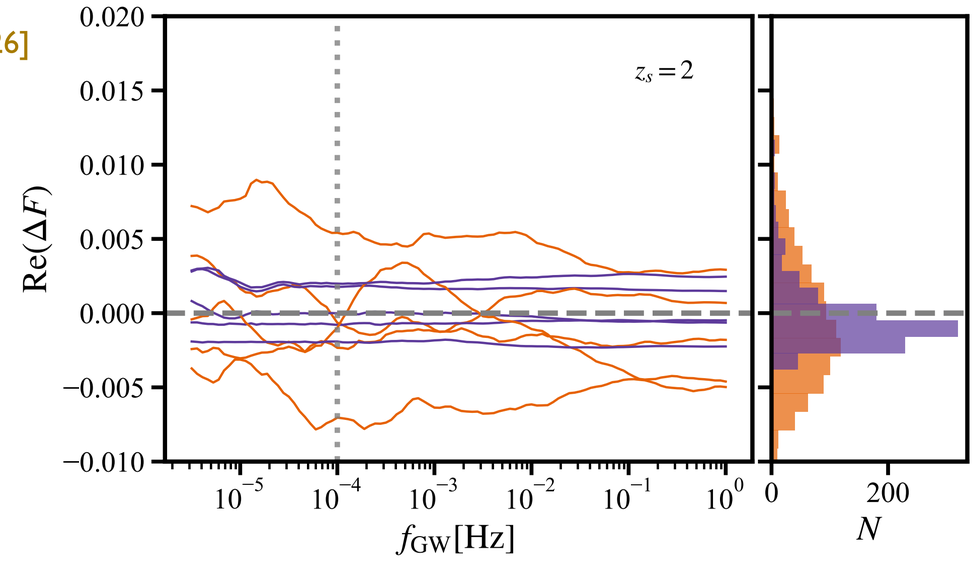


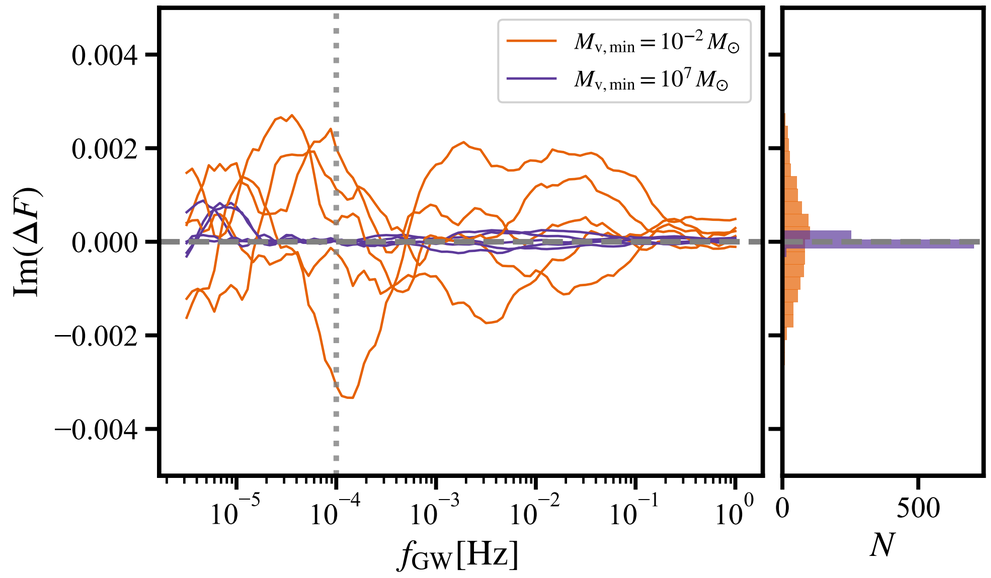
- Significant false deviations from GR in inspiral tests
- Strongly dependent on total mass: higher masses are merger-dominated, also higher SNR
- Ignoring the merger-ringdown reduces biases

Lensing and Environmental effects



[Choi+ 2026]

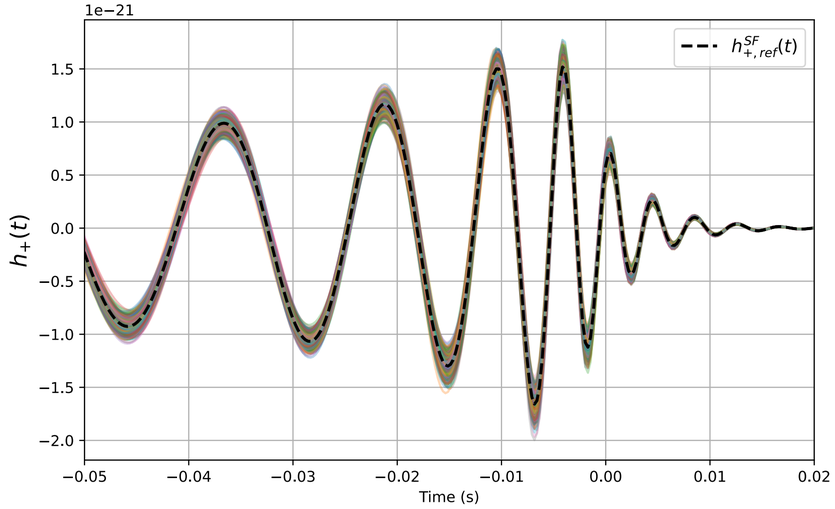




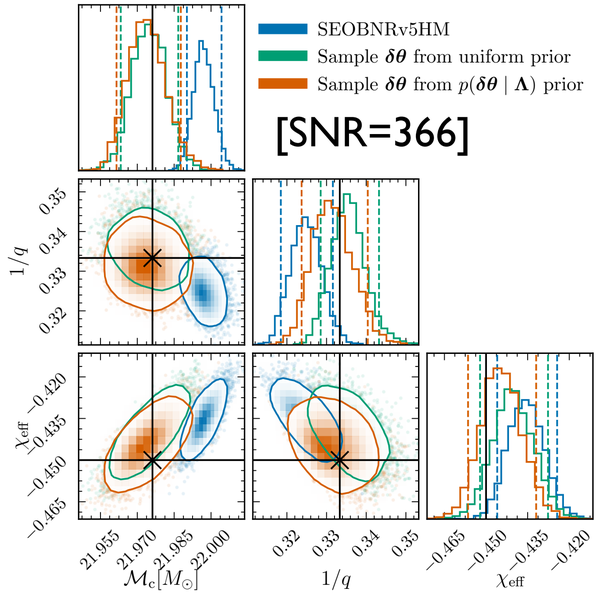
- Small dark matter halos along line-of-sight could combine to produce stochastic lensing signature
- Could act as a 'background noise' on signals, limit in precision (similar to instrument calibration)

How to mitigate waveform systematics ?

Marginalization over waveform errors:



Marginalization over SEOB NR calibration:



Calibration params
$\theta \sim p(\theta \mid \mathrm{NR})$

[Pompili+ 2024]

Probabilistic waveforms:

- amplitude/phase envelopes, as in detector calibration
$h(f) = (A + \delta A)e^{i(\phi+\delta\phi)}$ [Kumar+ 2025]

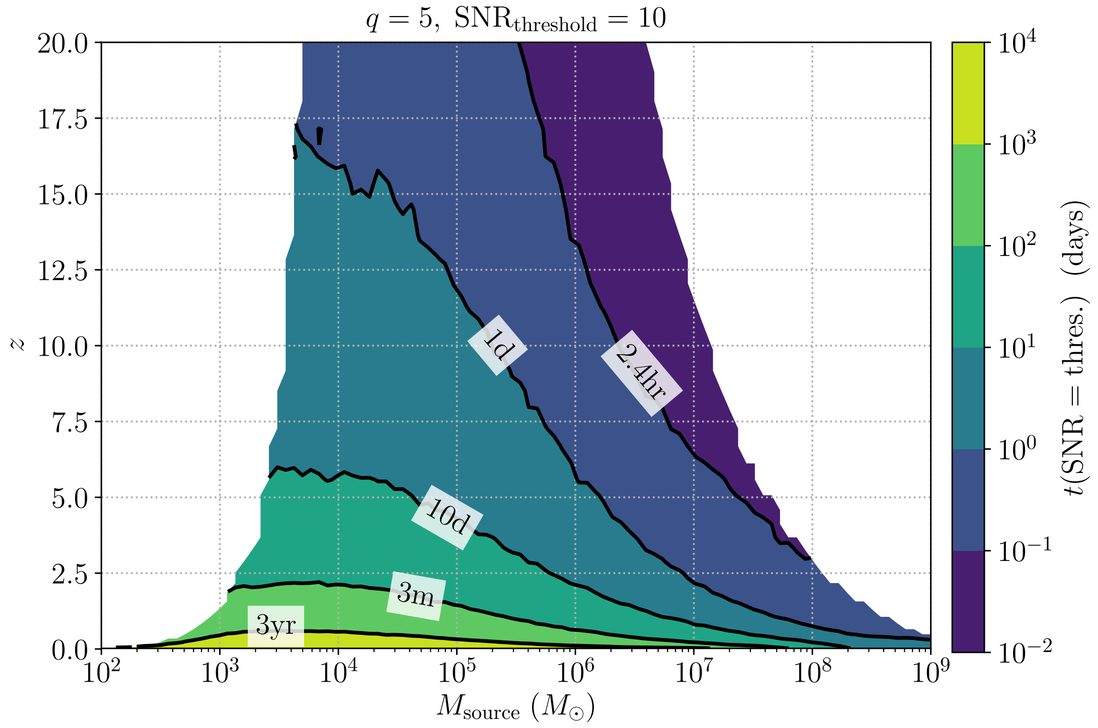
- constructed as Gaussian Process Regression (also guides next simulations)

[Moore-Gair 2014]
[Williams+2019]
[Bachhar+ 2024]

- Waveform should come with uncertainties !
- How to build a stochastic model for this deterministic error ? Calibration envelopes ?
- The science case of future instruments should be reevaluated with this waveform uncertainties

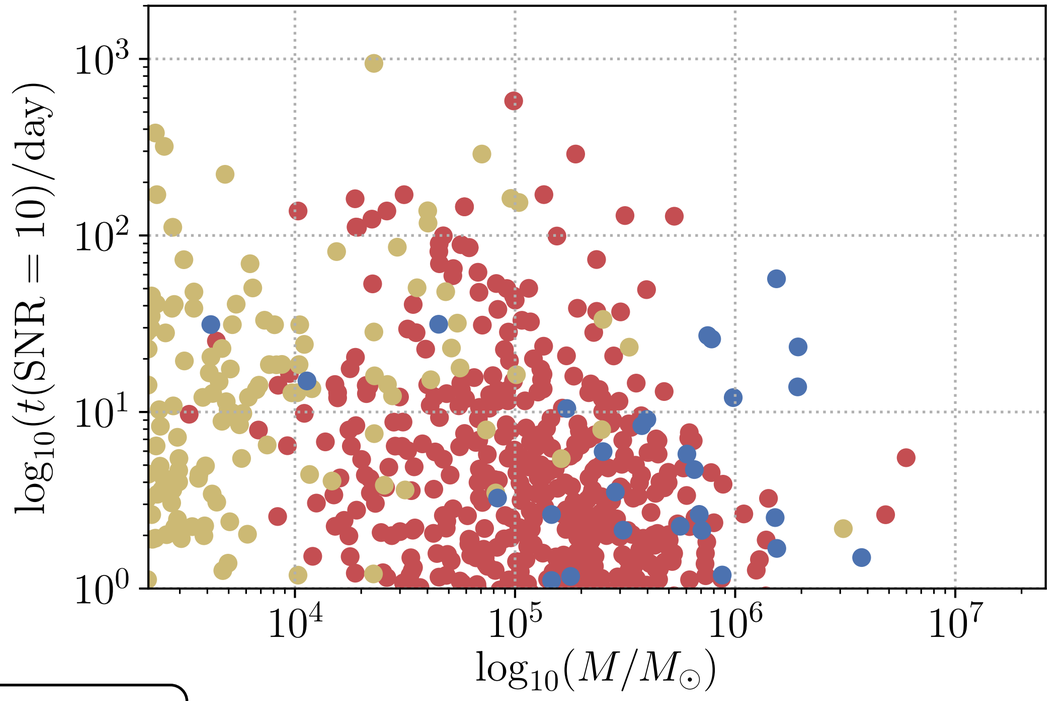
The length of MBHB signals

- How long before merger can we detect the signal ?
- SNR=10 to claim detection



Astrophysical models [Barausse 2012]:

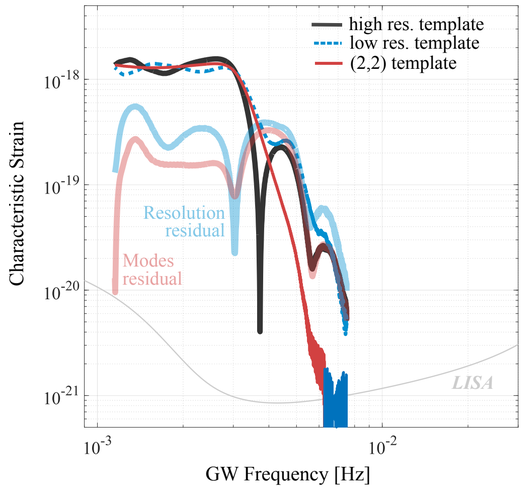
- Heavy seeds - delay
- Heavy seeds - no delay
- PopIII seeds - delay



MBHB detected signals:
Bulk shorter than ~10days
Tail extending to ~3months

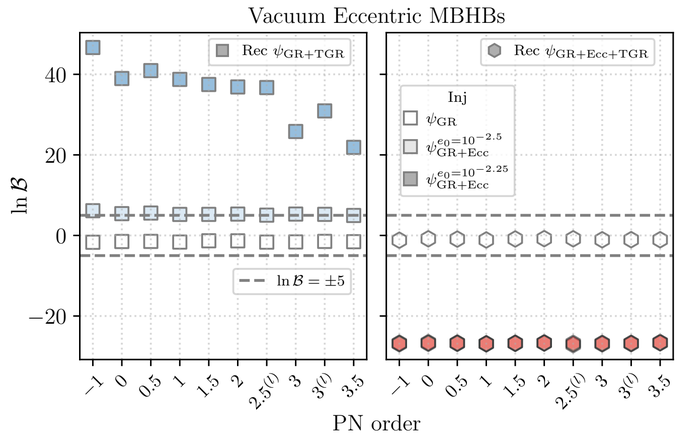
Waveform systematics studies for LISA MBHBs

NR residuals (Maya code): [Ferguson+ 2020]

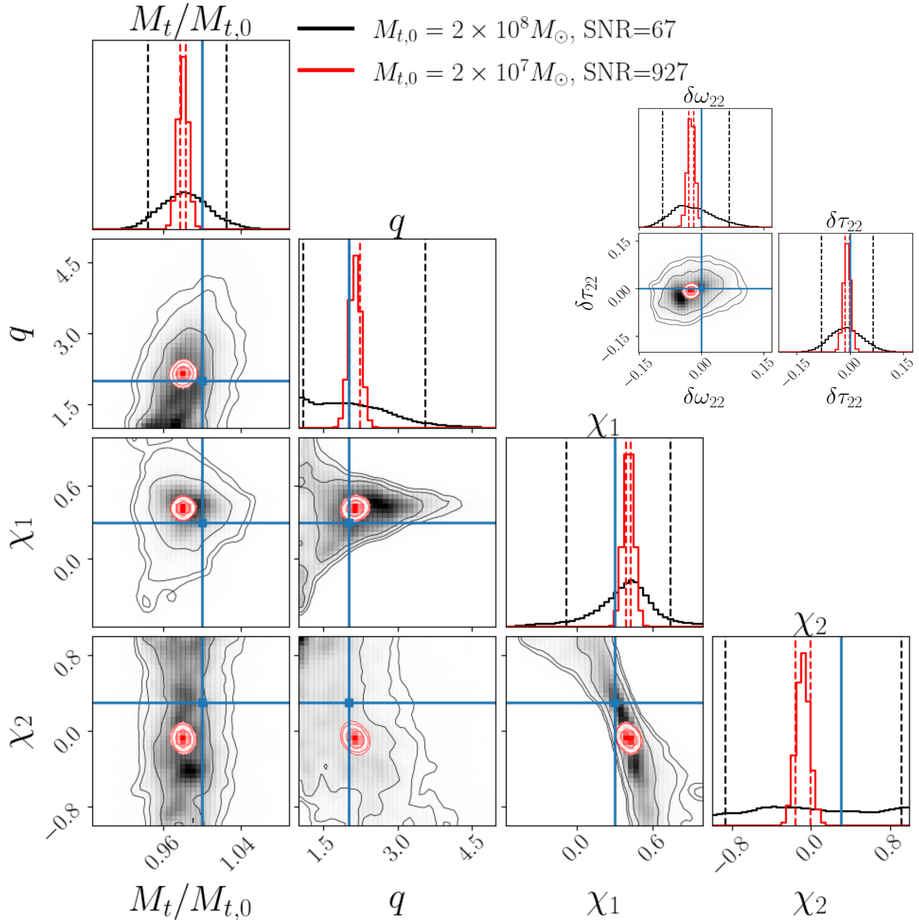


TGR systematics when ignoring ecc/env.:

[Garg+ 2024]

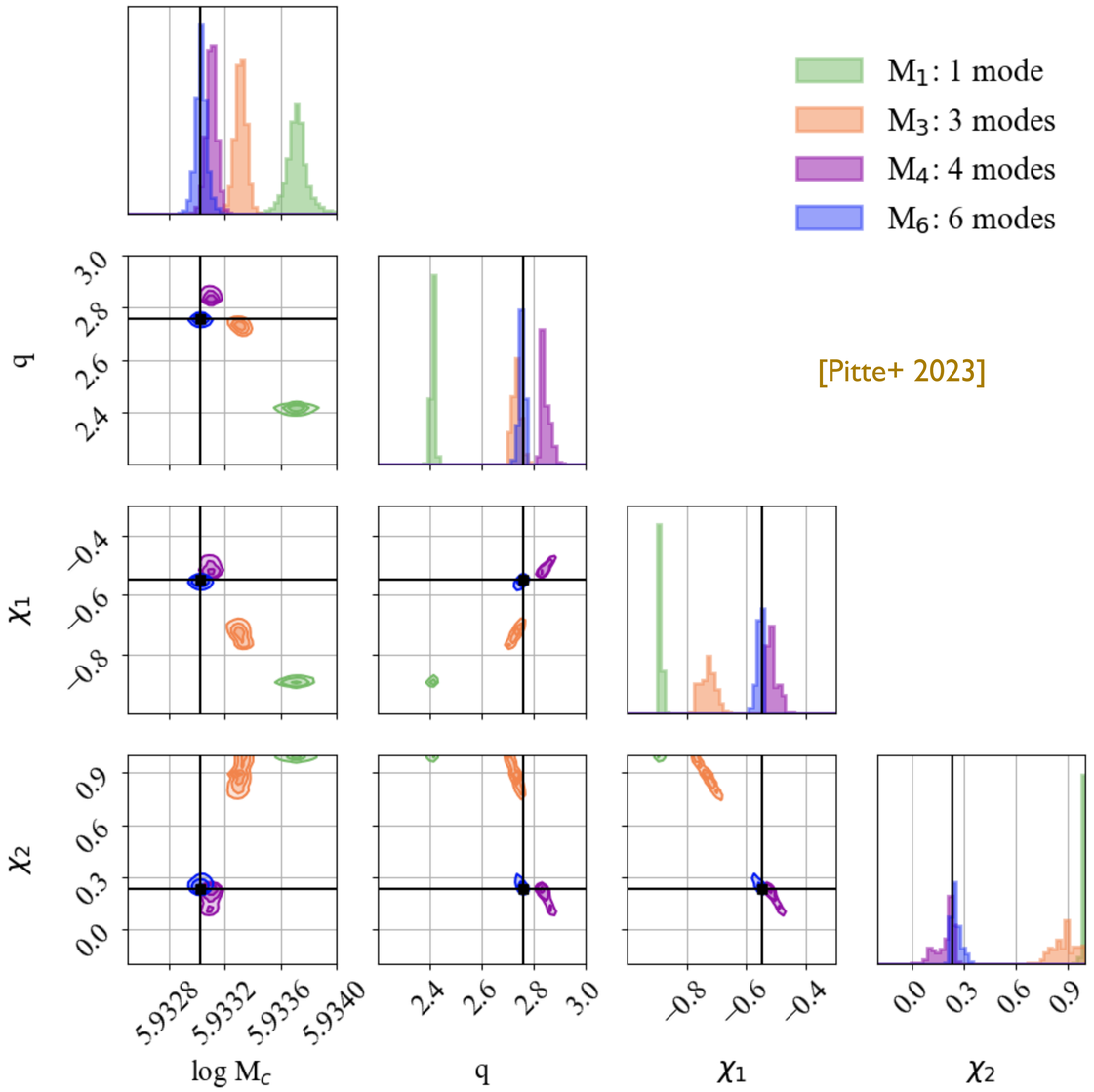


pSEOBNR vs NR:



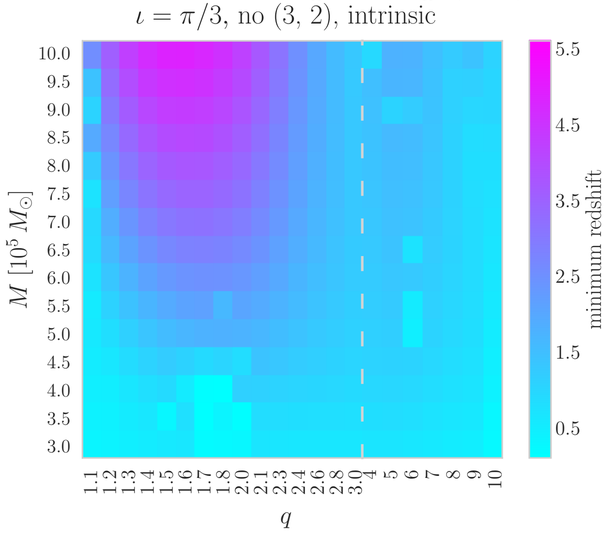
[Toubiana+ 2023]

MBHB biases caused by missing higher harmonics



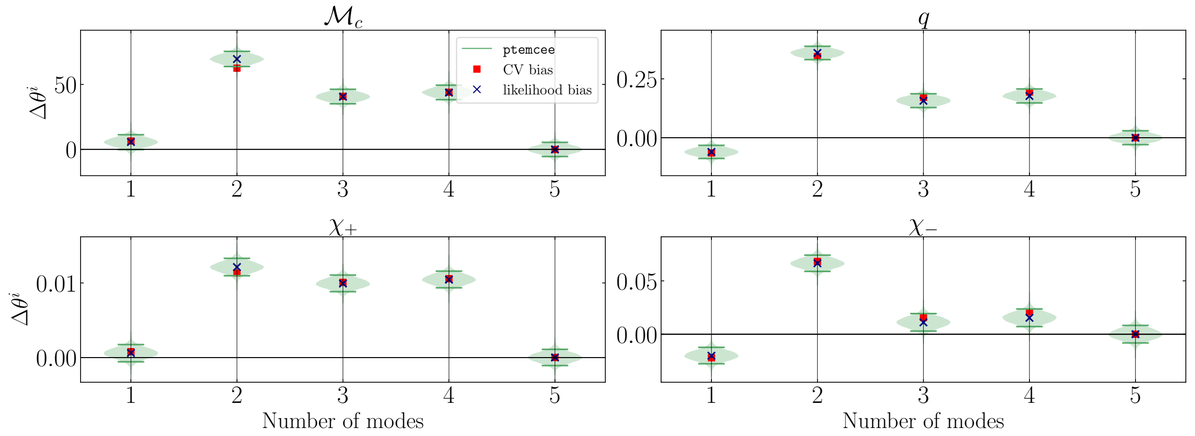
[Pitte+ 2023]

Min. redshift unbiased when ignoring (3,2) mode:



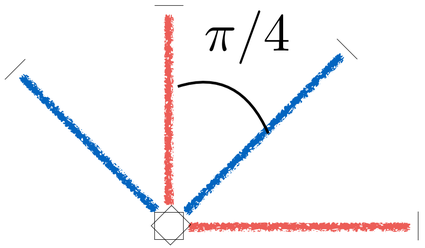
[Yi+ 2025]

Estimating biases: Cutler-Vallisneri (can fail), direct Nelder-Mead



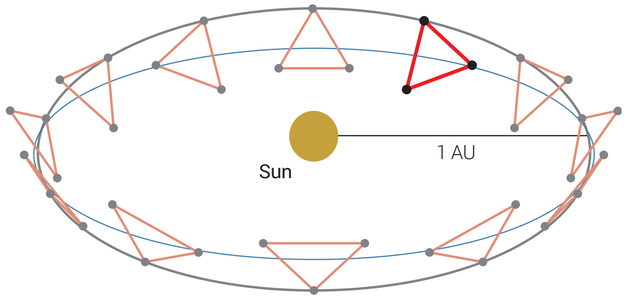
The LISA instrumental response

The low-frequency response



- At low frequencies TDIA, TDIE become equivalent to 2 LIGO-like channels
- For short transients the LISA motion has only a weak effect

The full response



Single-link response:

$$\mathcal{T}_{slr} = \frac{i\pi f L}{2}\mathrm{sinc}\left[\pi f L\left(1 - k\cdot n_l\right)\right]\exp\left[i\pi f\left(L + k\cdot\left(p_r + p_s\right)\right)\right] n_l \cdot P \cdot n_l(t_f)$$

+ Doppler phase: $\exp\left[2i\pi f k\cdot p_0(t_f)\right]$ + TDI combinations

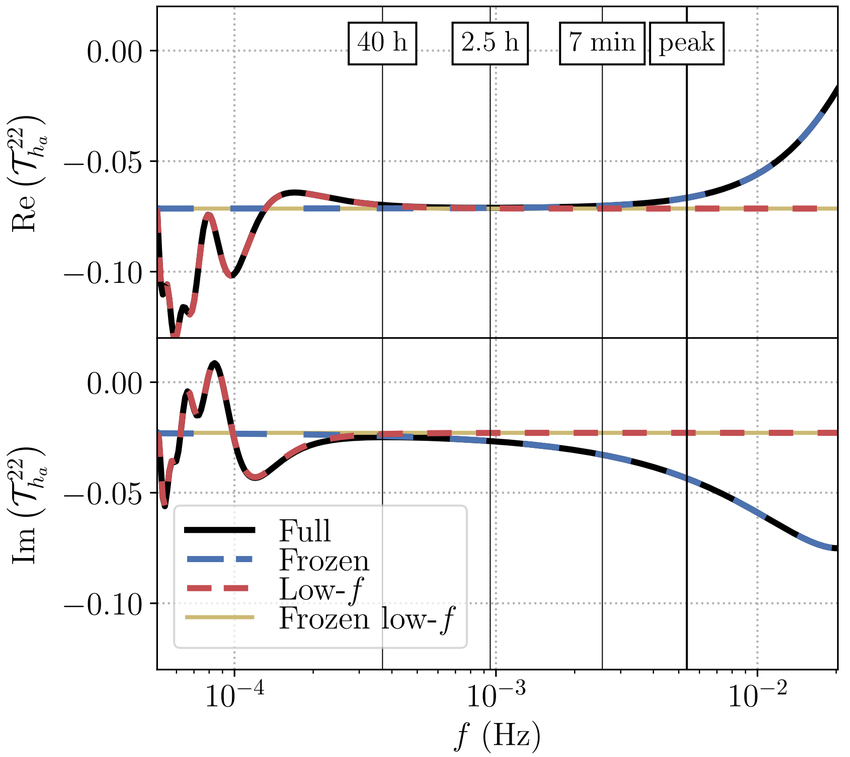
Time and frequency-dependency in transfer functions
Time: motion of LISA on its orbit
Frequency: departure from long-wavelength approx.

Pattern function response with HM:

$$h = \sum_{a,e}\sum_{\ell m}\frac{1}{d}F^{\ell m}_{a,e}h_{\ell m}$$

$$F^{\ell m}_{a,e} = \frac{1}{2}{}_{-2}Y_{\ell m}e^{-2i\psi_L}\left(F^{+}_{a,e} + iF^{\times}_{a,e}\right) + \frac{1}{2}(-1)^{\ell}{}_{-2}Y^{*}_{\ell,-m}e^{+2i\psi_L}\left(F^{+}_{a,e} - iF^{\times}_{a,e}\right)$$

Example FD transfer function



LISA response accuracy

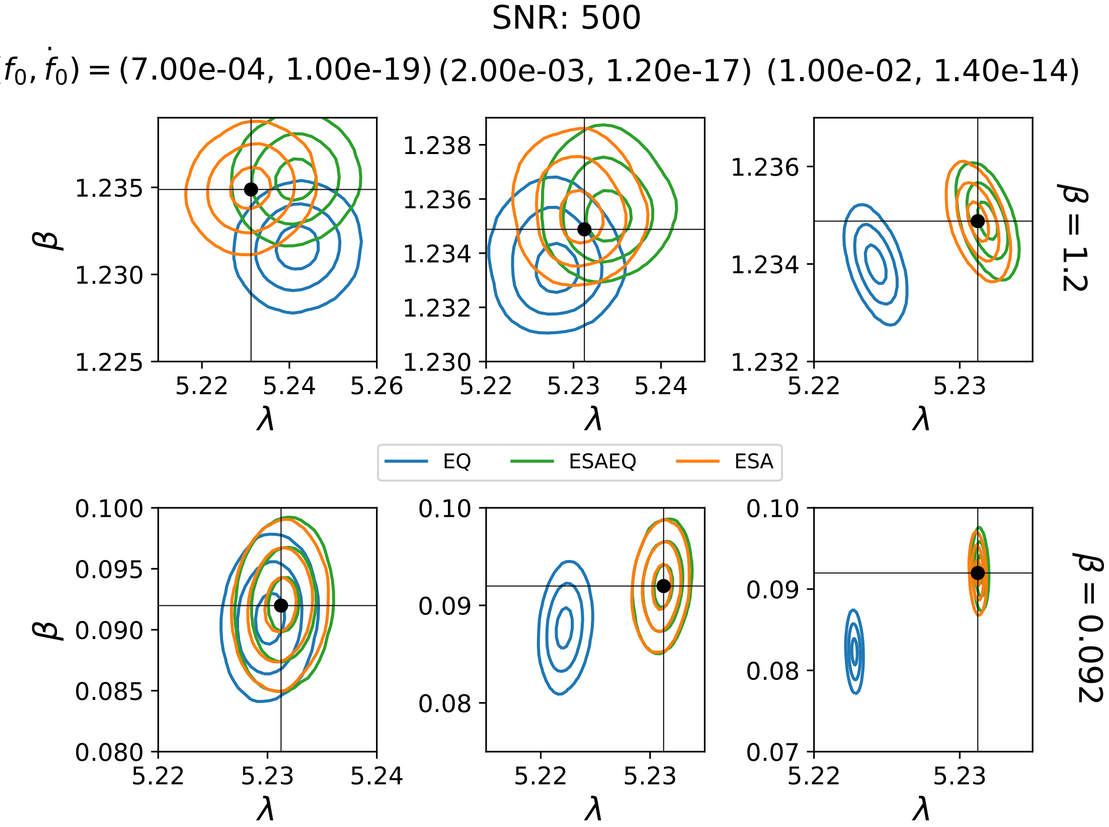
TDI combinations:

$$
\begin{aligned}
X_1 &= y_{13} + \mathbf{D}_{13}y_{31} + \mathbf{D}_{131}y_{12} + \mathbf{D}_{1312}y_{21} \\
&\quad - \left[y_{12} + \mathbf{D}_{12}y_{21} + \mathbf{D}_{121}y_{13} + \mathbf{D}_{1213}y_{31}\right], \\
X_2 &= X_1 + \mathbf{D}_{13121}y_{12} + \mathbf{D}_{131212}y_{21} + \mathbf{D}_{1312121}y_{13} \\
&\quad + \mathbf{D}_{13121213}y_{31} - \left[\mathbf{D}_{12131}y_{13} + \mathbf{D}_{121313}y_{31}\right. \\
&\quad \left. + \mathbf{D}_{1213131}y_{12} + \mathbf{D}_{12131312}y_{21}\right].
\end{aligned}
$$

Levels of approximations:

- the armlength is not constant: $L_1(t) \neq L_2(t) \neq L_3(t)$ [10^{-2} level]
- Pointing-ahead, delays do not commute: $D_{12} \neq D_{21}$ [10^{-4} level]
- non-stationarities, non-independence of A, E, T

High-SNR GB bias with approx. orbits: [Katz+ 2022]



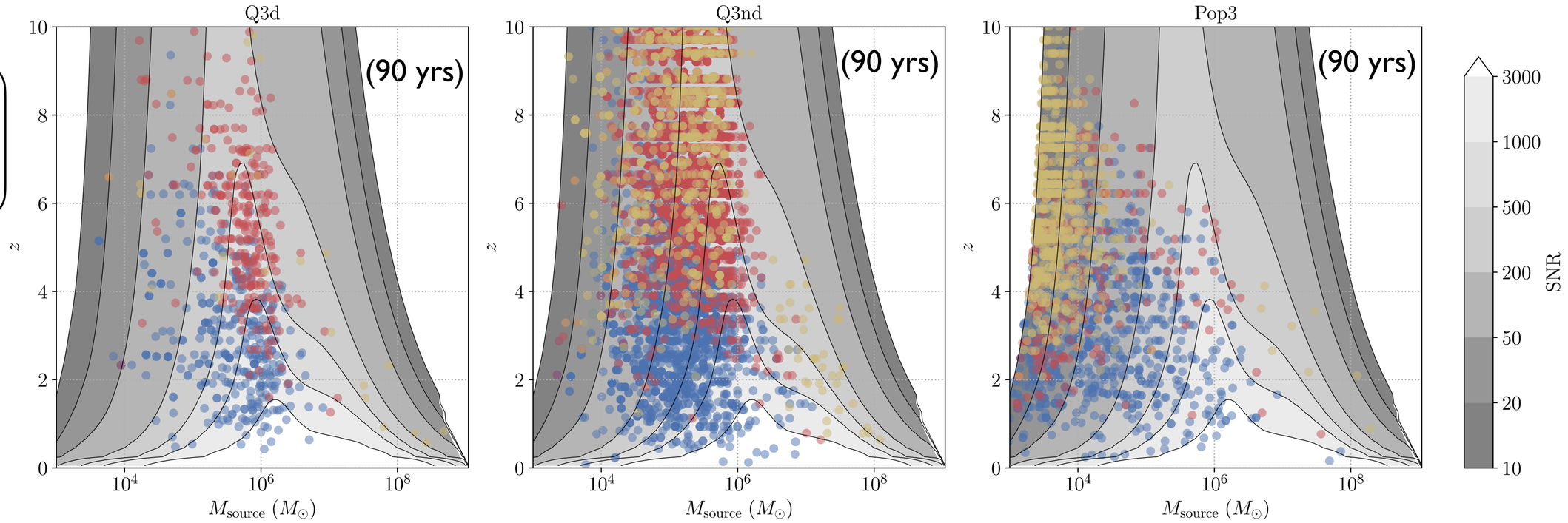
Sky multimodalities for LISA MBHBs

- **Bayesian PE** required to explore multimodal posteriors
- Simulation of 90yrs catalogs
- Custom proposals for degeneracies

Astrophysical models [Barausse 2012]:
- Heavy seeds - delay (Q3d)
- Heavy seeds - no delay (Q3nd)
- PopIII seeds - delay (Pop3)

- 1 mode
- 2 modes
- >2 modes

Threshold:
5% prob.



Applications: EM counterparts and cosmological inference
[Mangiagli&al 2022, Mangiagli&al 2023]

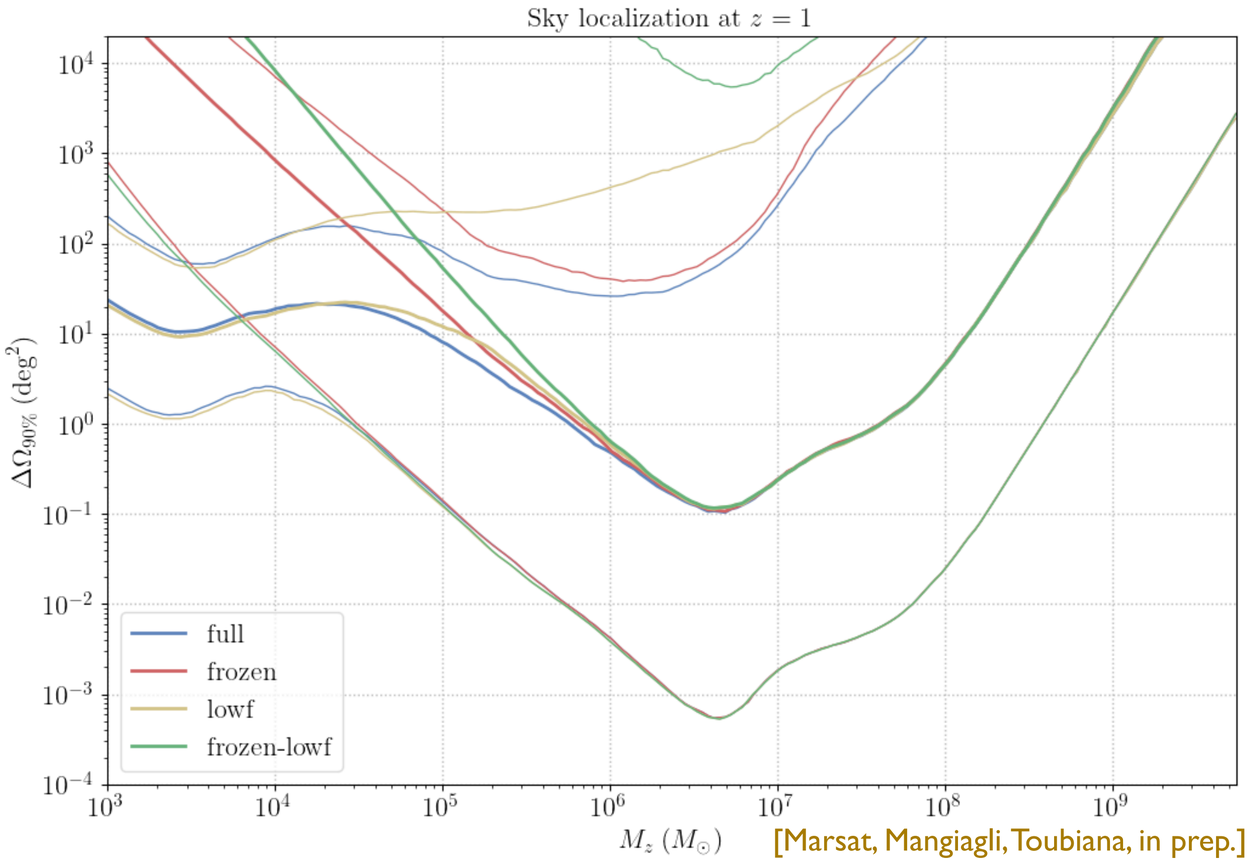
Multimodality in the sky present, but rare for counterpart candidates post-merger

Fisher localization: which part of the response is important ?

Analysis settings:

- Fisher matrix localization: sky area of the main mode of the posterior
- Randomization over 1000 orientations, mass ratios, spins
- Change the response model: keep or ignore the motion and high-f effects

- ‘Pattern function’ response is the main source of main-mode localization at high mass, from subdominant HM
- Multimodality broken in turn by subdominant effects in response (motion, high-f)



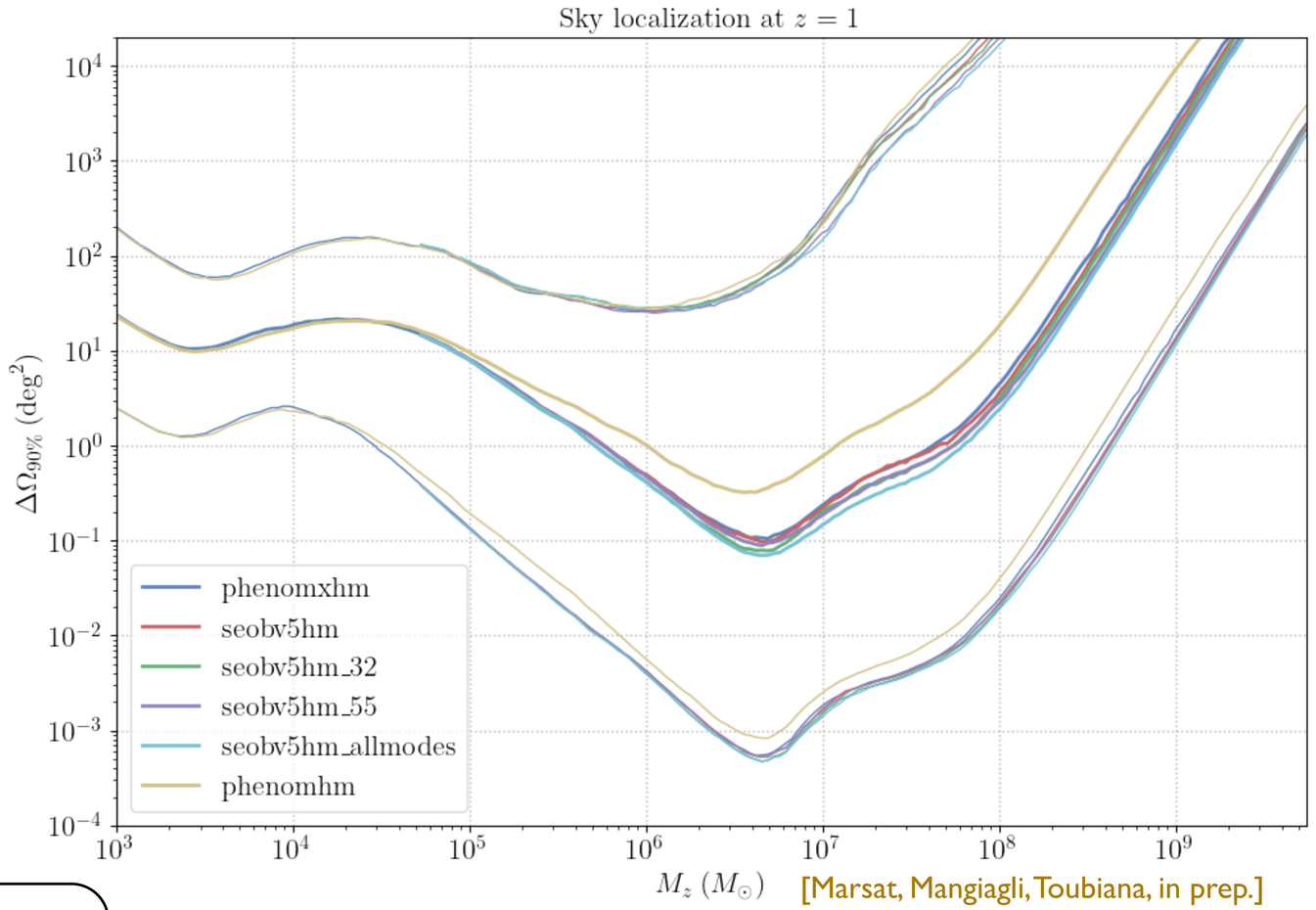
[Marsat, Mangiagli, Toubiana, in prep.]

- Sky localization at high mass: weak effects, high SNR
- Unlike LVK localization from triangulation, LISA localization potentially vulnerable to systematics

Fisher localization: do waveform models agree on prospective ?

Analysis settings:

- Fisher matrix localization: sky area of the main mode of the posterior
- Randomization over 1000 orientations, mass ratios, spins
- Change the waveform model: PhenomHM, PhenomXHM, SEOBNRv5HM_ROM



[Marsat, Mangiagli, Toubiana, in prep.]

- In the high-mass range (HM important), older waveform models can be inaccurate also for prospective PE
- Modern waveform models agree well for prospective

Analysis settings: waveform models

Injections:

`NRHybSur3dq8` [Varma&al 2018]

- SXS NR simulations hybridized with long EOB inspirals (covers ~6 months for $M = 10^5 M_\odot$)
- Surrogate interpolant, time-domain

Templates:

Efficient Fourier-domain models from 2 families:

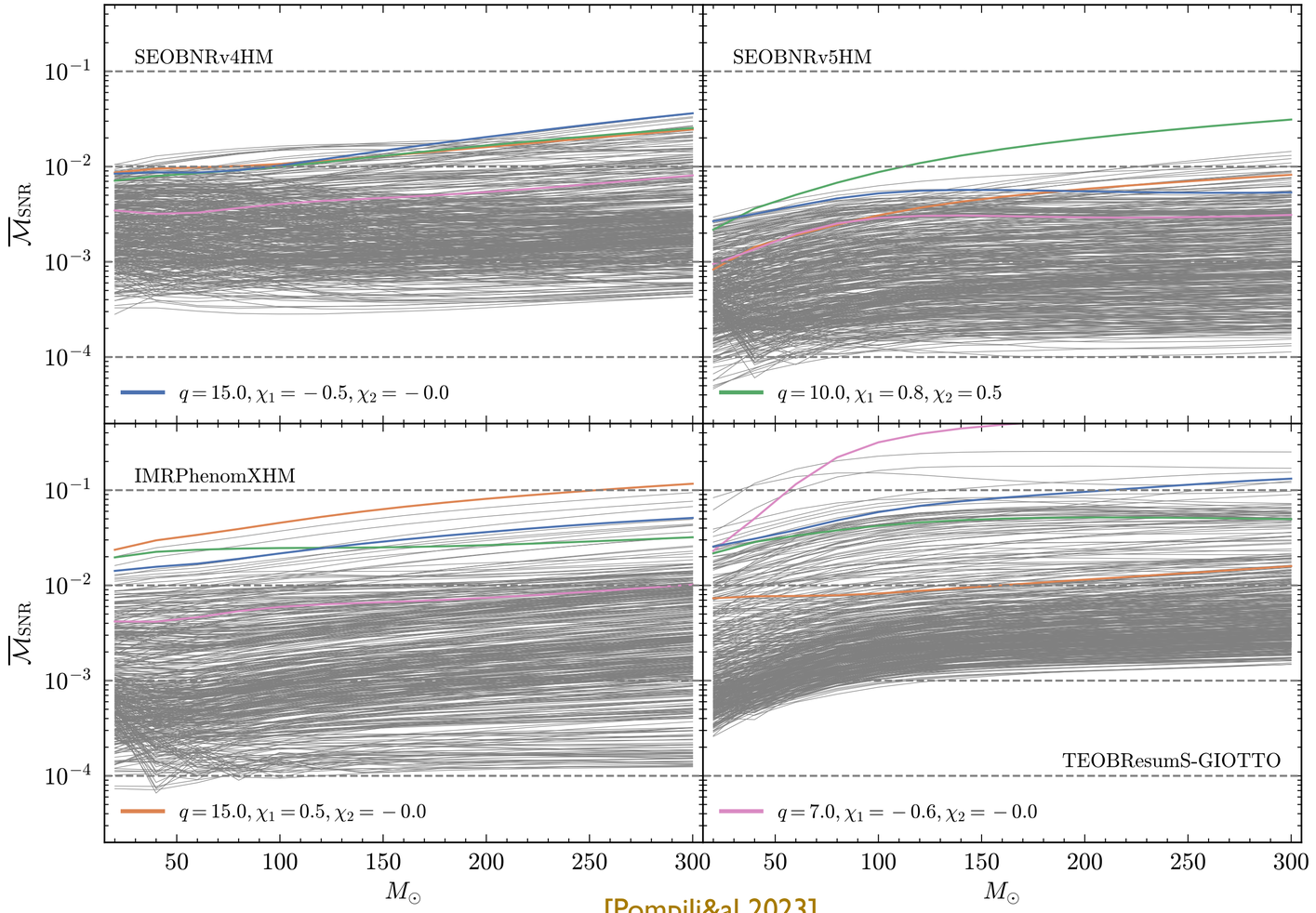
- `PhenomHM` [London&al 2017]
- `PhenomXHM` [García-Quirós&al 2020]
- `SEOBNRv4HM_ROM` [Cotesta&al 2018]
- `SEOBNRv5HM_ROM` [Pompili&al 2023]

Mode content for all: 22, 21, 33, 44

Limitations:

- **aligned spins only**
- in the inspiral, all based on PN/EOB

Aligned spin case: mismatch with NR ~ $10^{-4} - 10^{-2}$



[Pompili&al 2023]

Analysis settings: waveform models

Parameter estimation:

We use PE sampling tools: `lisabeta` [Marsat&al 2020]

PE provides posteriors, $\max \ln \mathscr{L}$ provides the bias

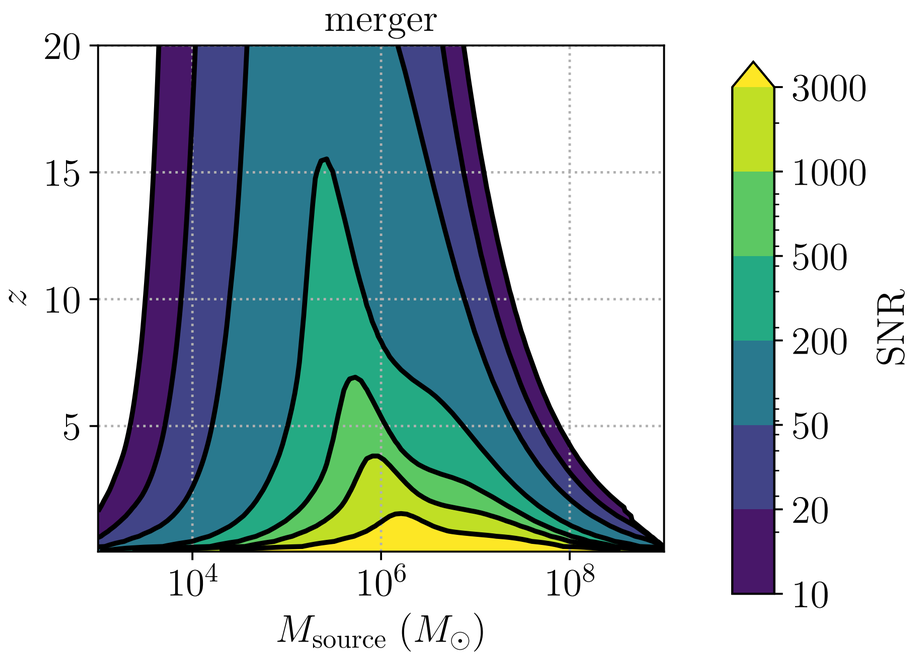
Fourier-domain full response (same for injection and templates)

Direct FD Whittle likelihood (no approx.)

Parameter space exploration:

- $M_z = [10^5, 10^6, 10^7] M_\odot$
- $z_{\min} = 1$

- $N = 240$ simulations
- uniform $q \in [1,8]$
- uniform $\chi_1 \in [-0.8, 0.8]$
- uniform $\chi_2 \in [-0.8, 0.8]$
- randomize orientations

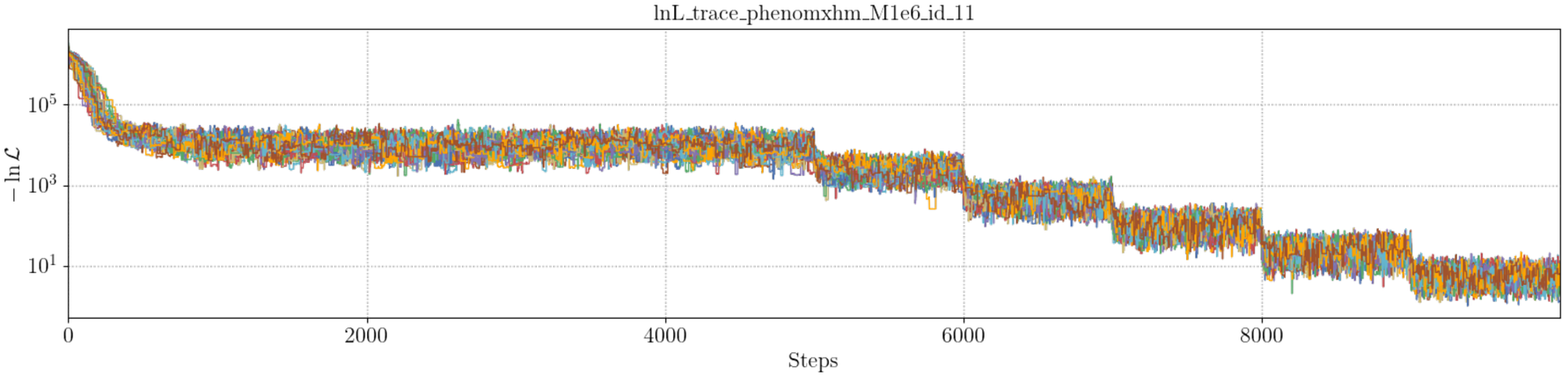


Tempering the analysis:

Start at large z / small SNR, then increase SNR gradually

Equivalent to a tempering scheme with non-0 base temperature

Additional parallel tempering with 4 temps.

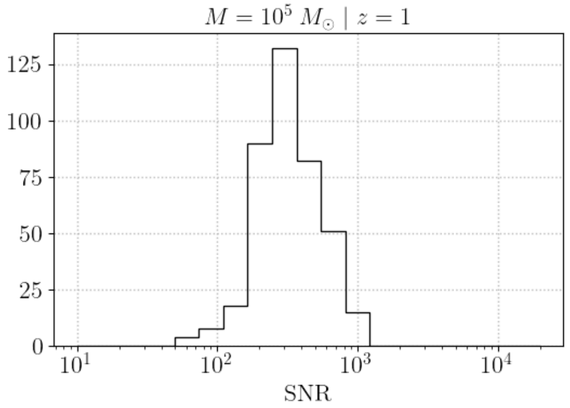


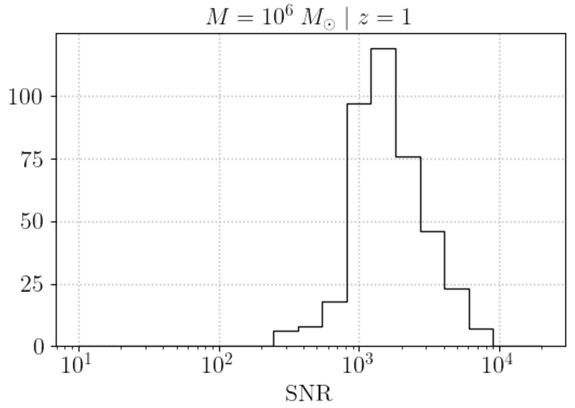
SNRs and mismatches

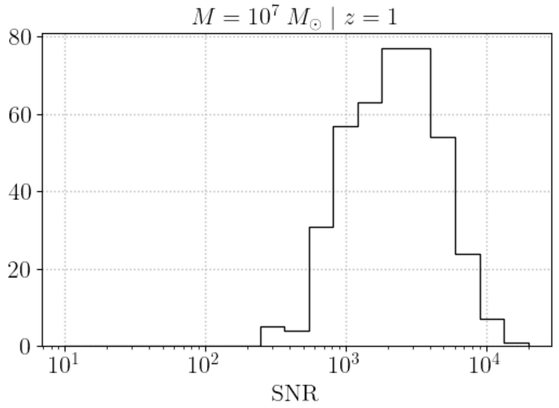
Signal-to-noise ratios:

$$\mathrm{SNR}^2 = (h_0|h_0)$$

SNRs up to thousands at $z = 1$





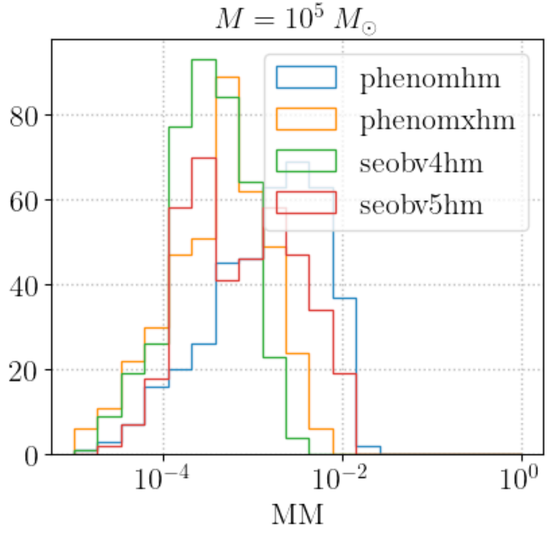


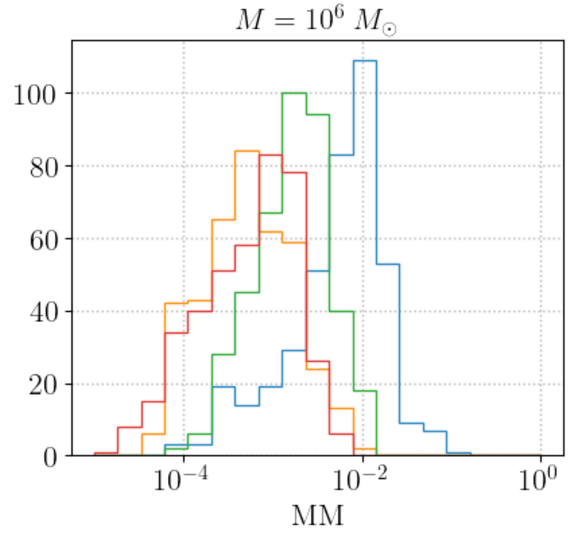
Mismatch:

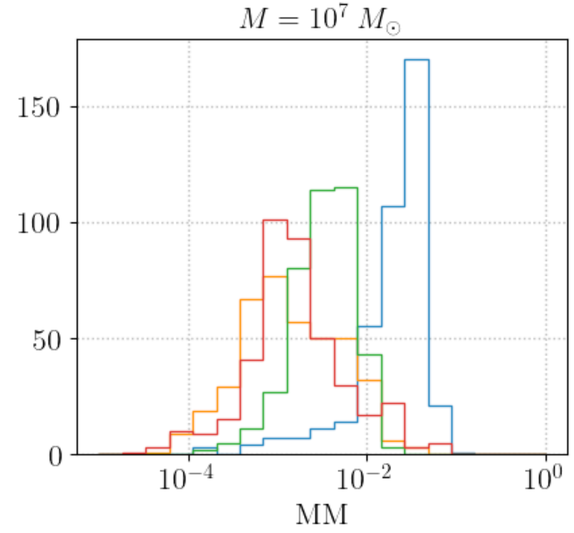
- averaged response
- single-detector
- optimized: time, phase, polarization, sky

$$\mathrm{MM} = 1 - \max_{t,\varphi,\psi,\ldots} \frac{(h_m|h_0)}{\sqrt{(h_m|h_m)}\sqrt{h_0|h_0}}$$

- Mismatches order-of-magnitude similar to LVK
- Improvement for more recent models







Example Parameter estimation with systematics III: Cutler-Vallisneri bias

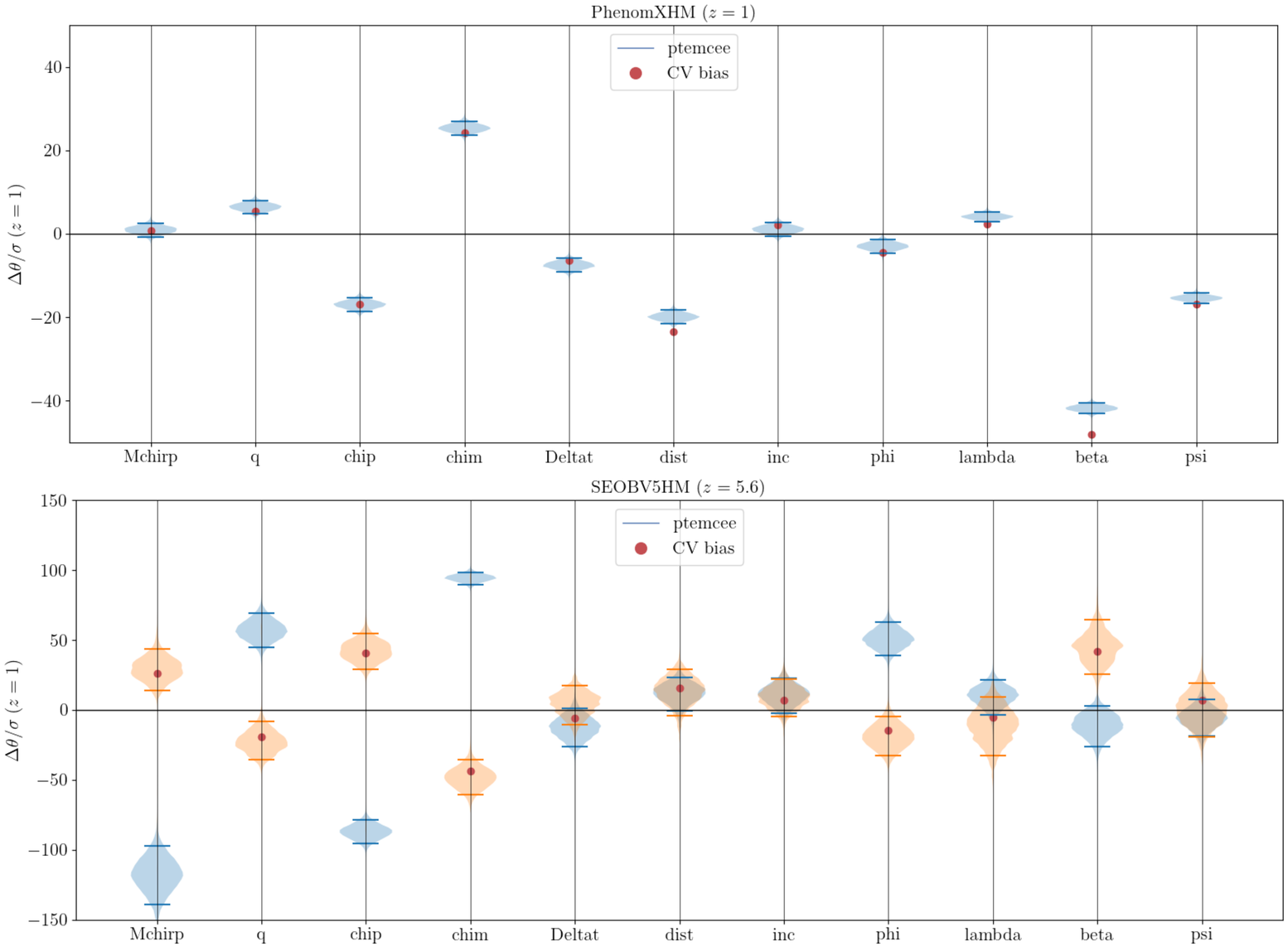
Linearized biases (Cutler-Vallisneri):

[Flanagan-Hughes 1997]
[Cutler-Vallisneri 2007]

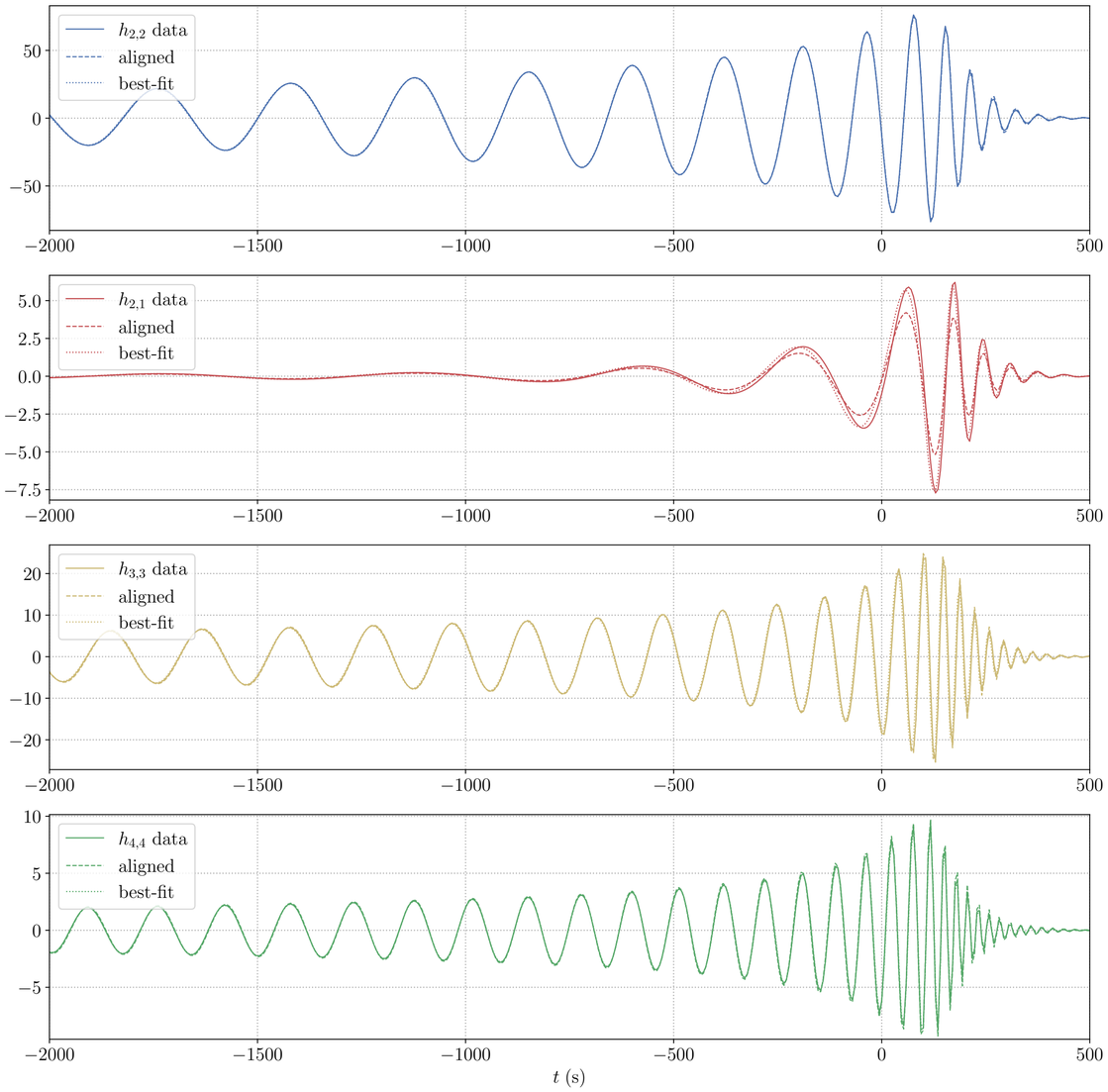
In the linear signal approximation, estimation of bias:

$$F_{ij} = (\partial_i h | \partial_j h)$$
$$\Delta\theta_i = F_{ij}^{-1}(\partial_j h | \delta h)$$

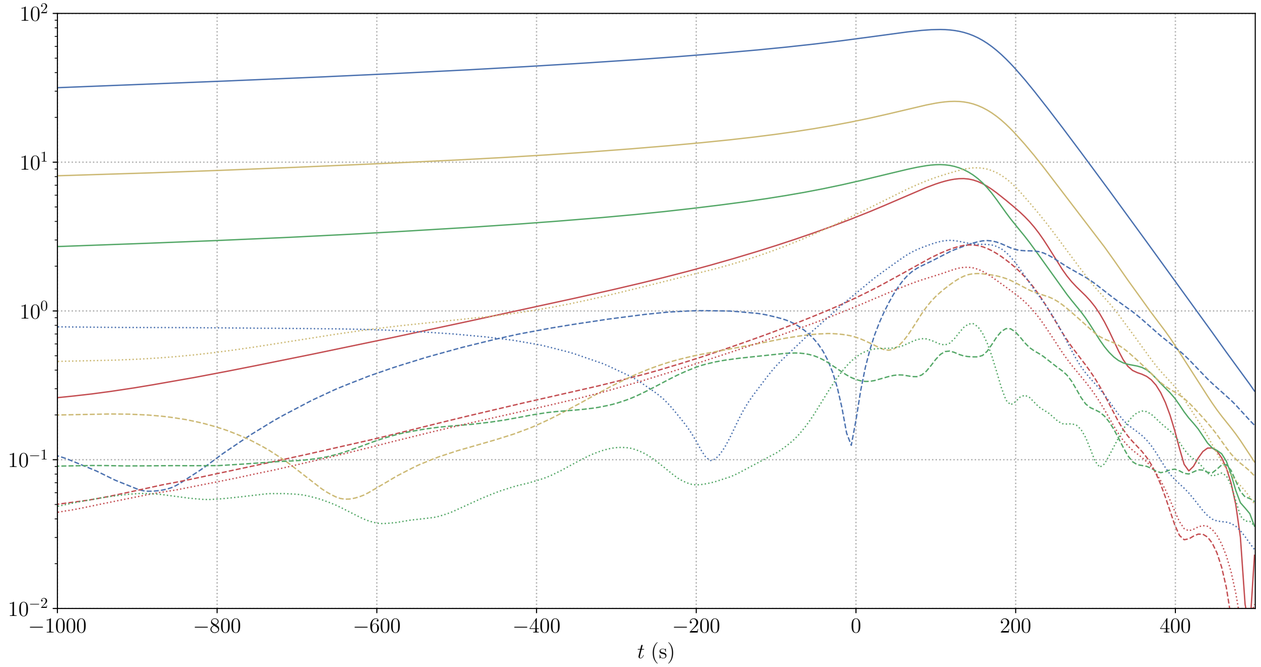
- Cutler-Vallisneri biases give reasonable estimates in mild cases
- Fail to capture distant secondary mode — needs to be adapted
- Can fail for severe biases



Example Parameter estimation with systematics III: TD signals and residuals



Residuals for the ‘ugly’ case



- Residuals are ‘visually’ small…
- Details of HM at merger are important

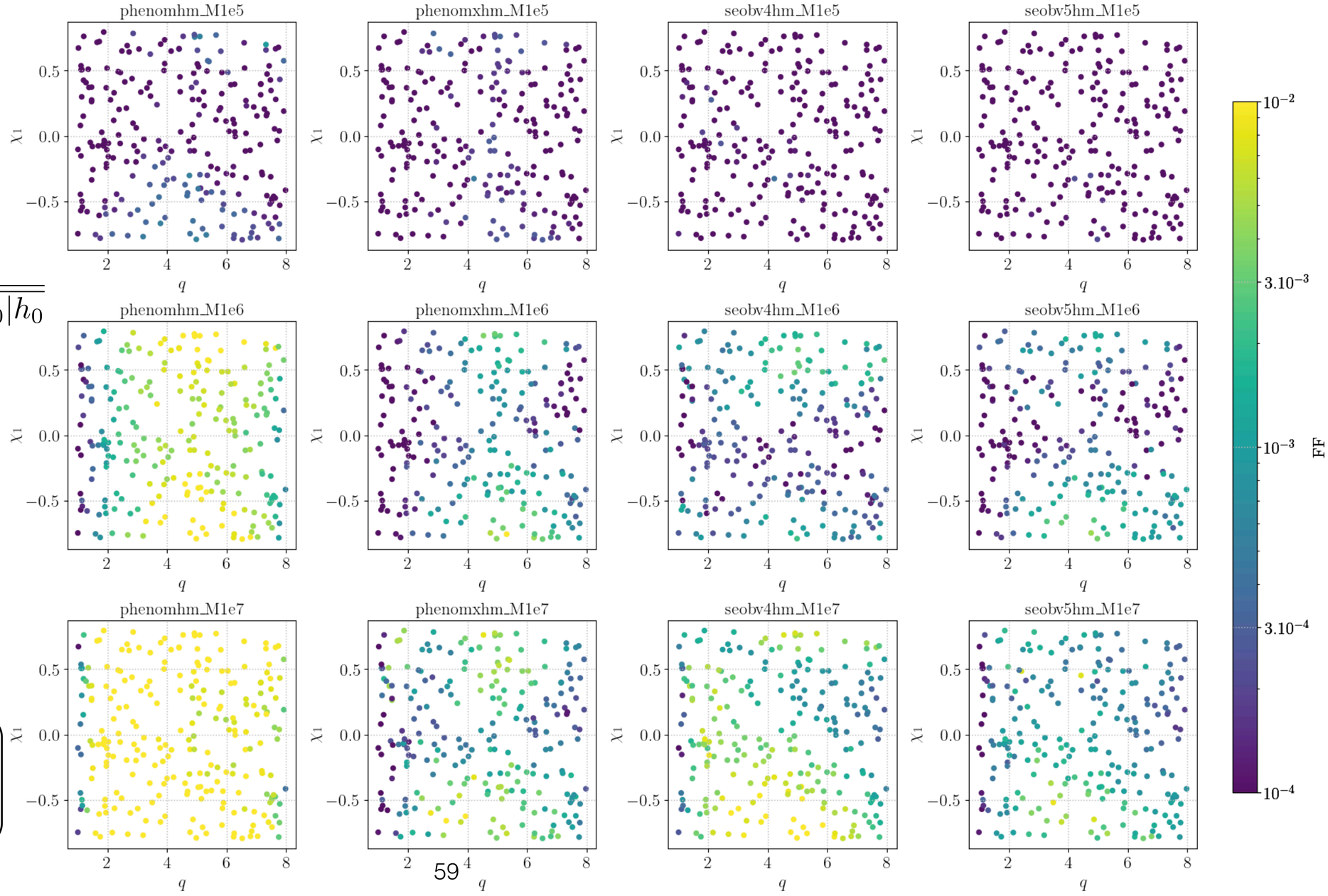
Fitting factors in parameter space

Fitting factor, computed at best-fit params (optimization over all parameters):

$$\mathrm{FF} = 1 - \max_{\boldsymbol{\theta}} \frac{(h_m|h_0)}{\sqrt{(h_m|h_m)}\sqrt{h_0|h_0}}$$

$$\ln \mathcal{L}_{\max} = -\mathrm{SNR}^2 \times \mathrm{FF}$$

Trend: strong dependence on M
Trend: larger errors at large spins
Trend in q ?



Complement: NR vs NR surrogate

Injections:

NR from SXS catalog [SXS 2025]

- SXS NR simulations hybridized with long EOB inspirals (covers ~6 months for $M = 10^5 M_\odot$)
- Surrogate interpolant, time-domain

Templates:

NRHybSur3dq8 [Varma&al 2018]

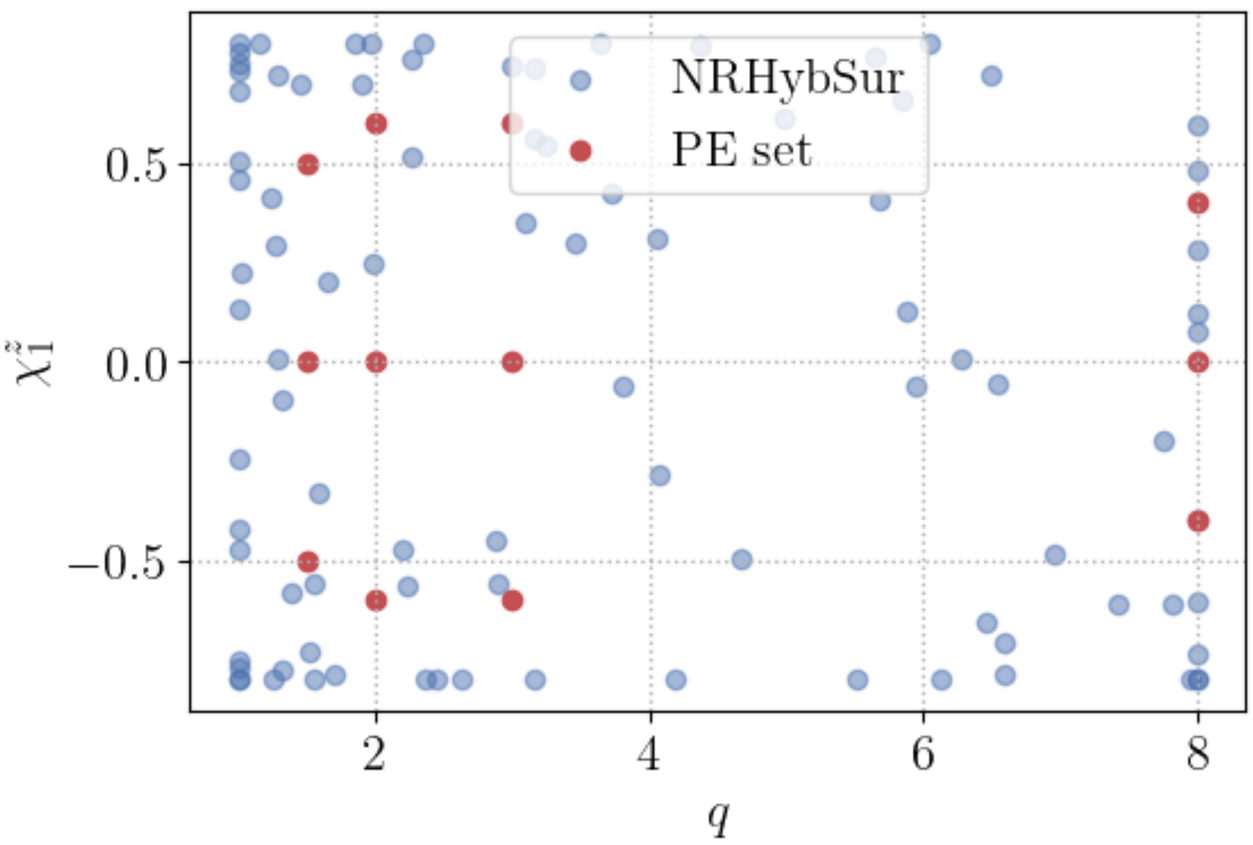
- SXS NR simulations hybridized with long EOB inspirals (covers ~6 months for $M = 10^5 M_\odot$)
- Surrogate interpolant, time-domain

Mode content for all: 22, 21, 33, 44

Time-domain LISA response

Limitations:
- aligned spins only
- M=1e7 only, 2 days of signal

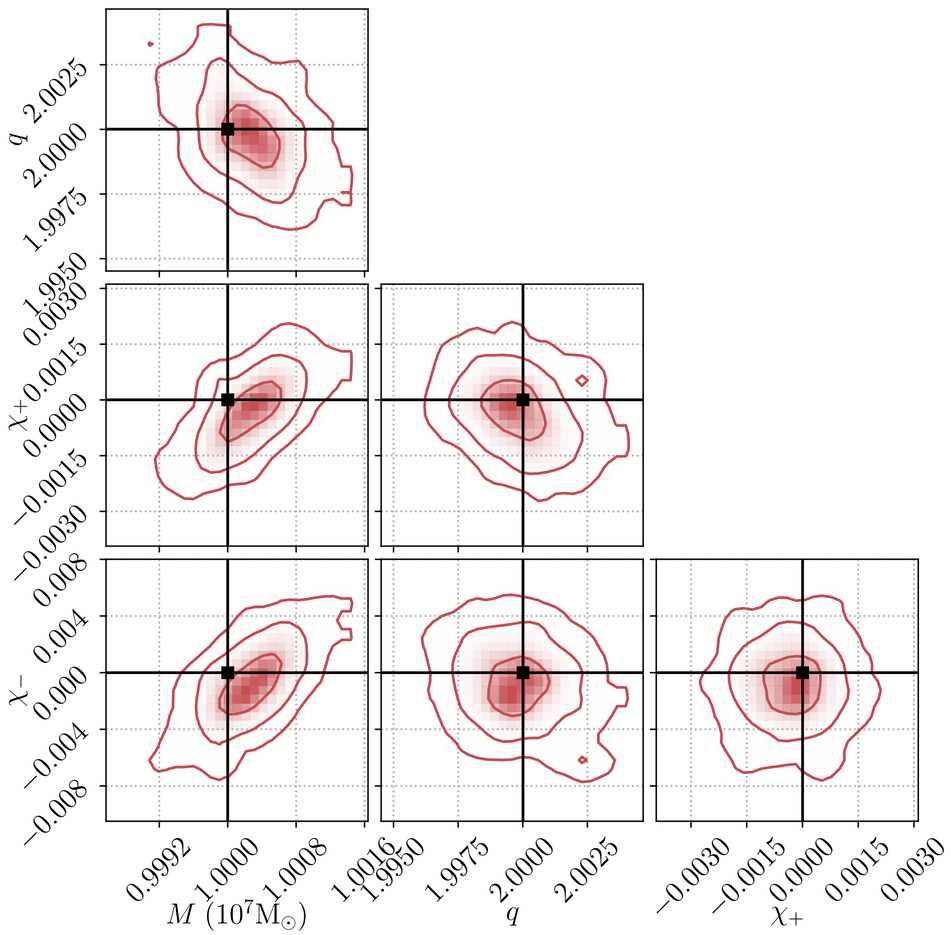
[Preliminary]

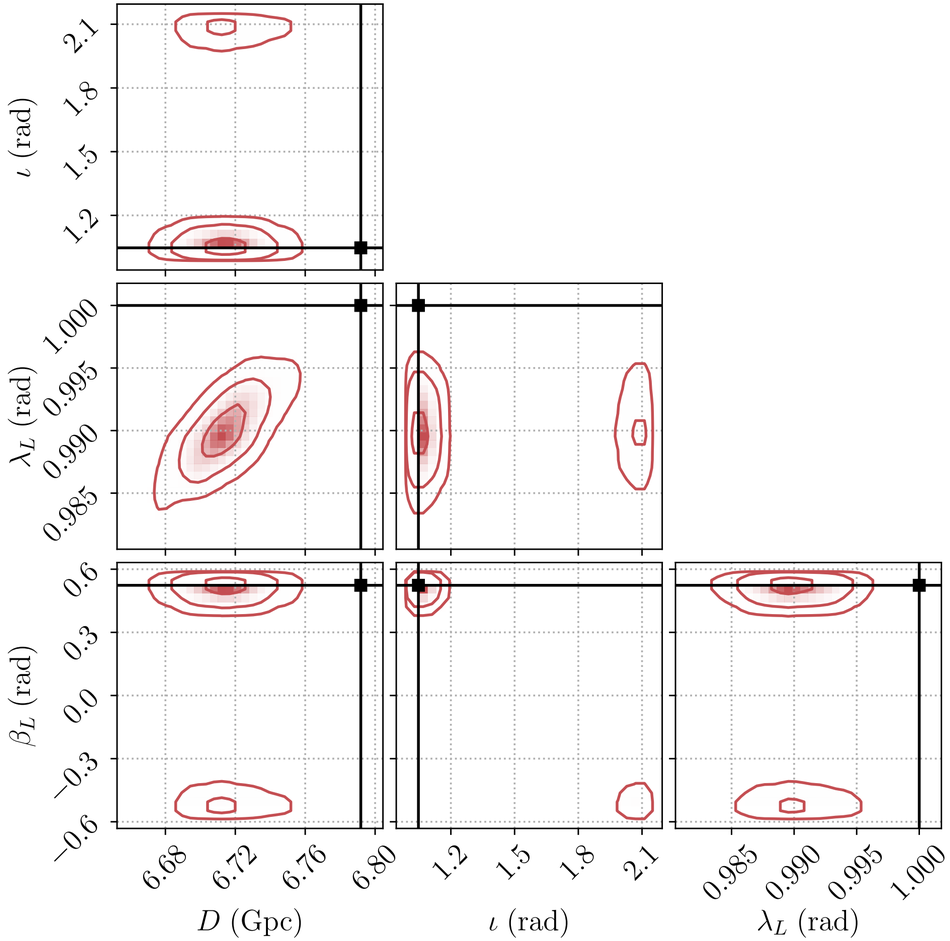


Complement: NR vs NR surrogate

SXS:BBH:2497

$q = 2, \chi_1 = 0, \chi_2 = 0$

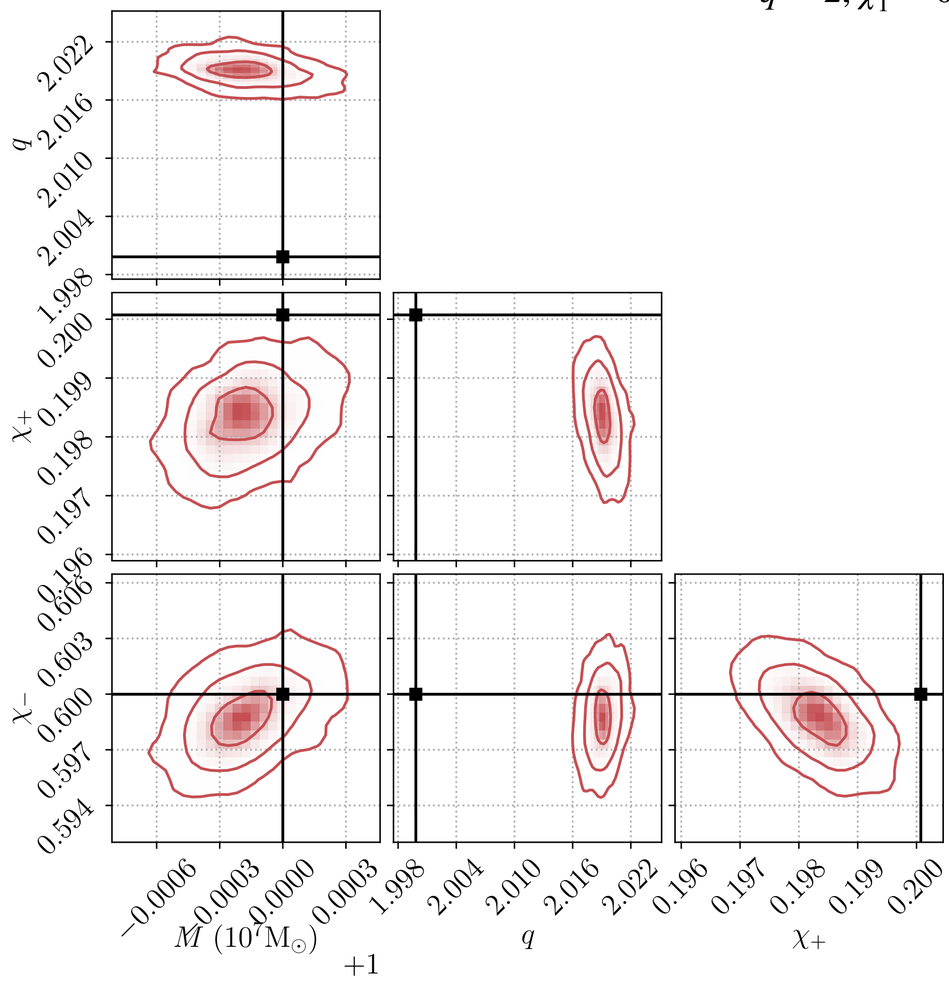


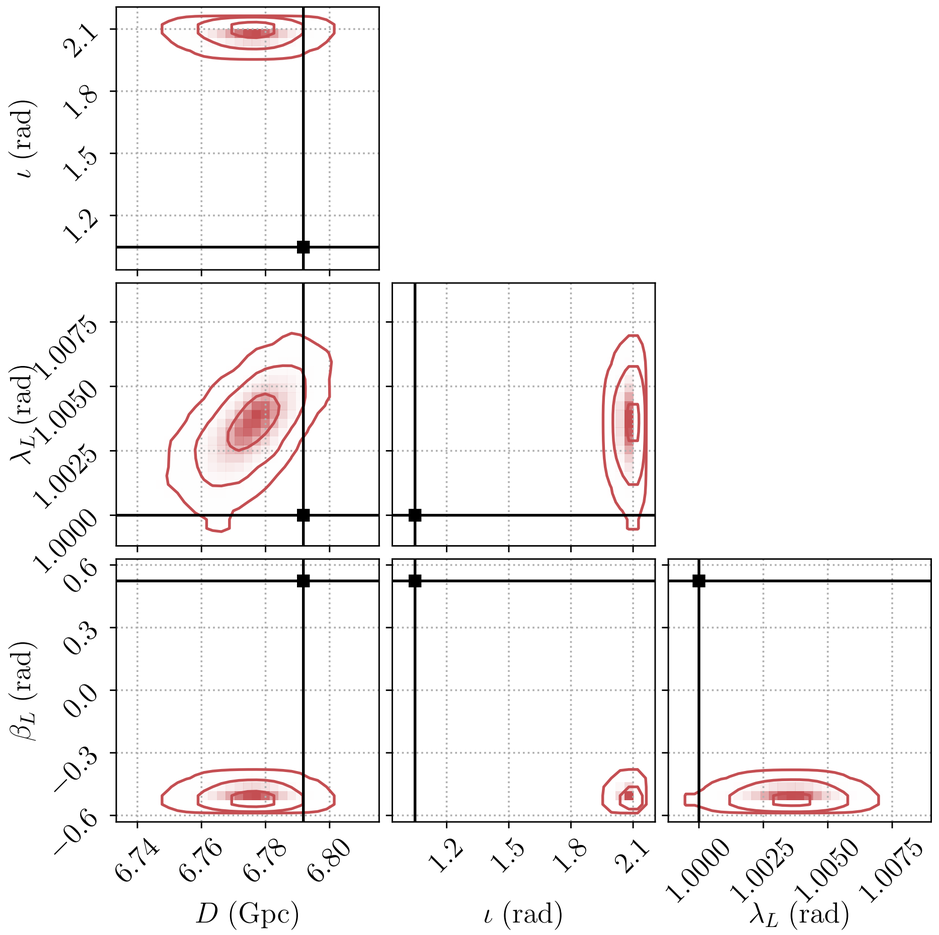


Complement: NR vs NR surrogate

SXS:BBH:2128

$q = 2, \chi_1 = 0.6, \chi_2 = -0.6$

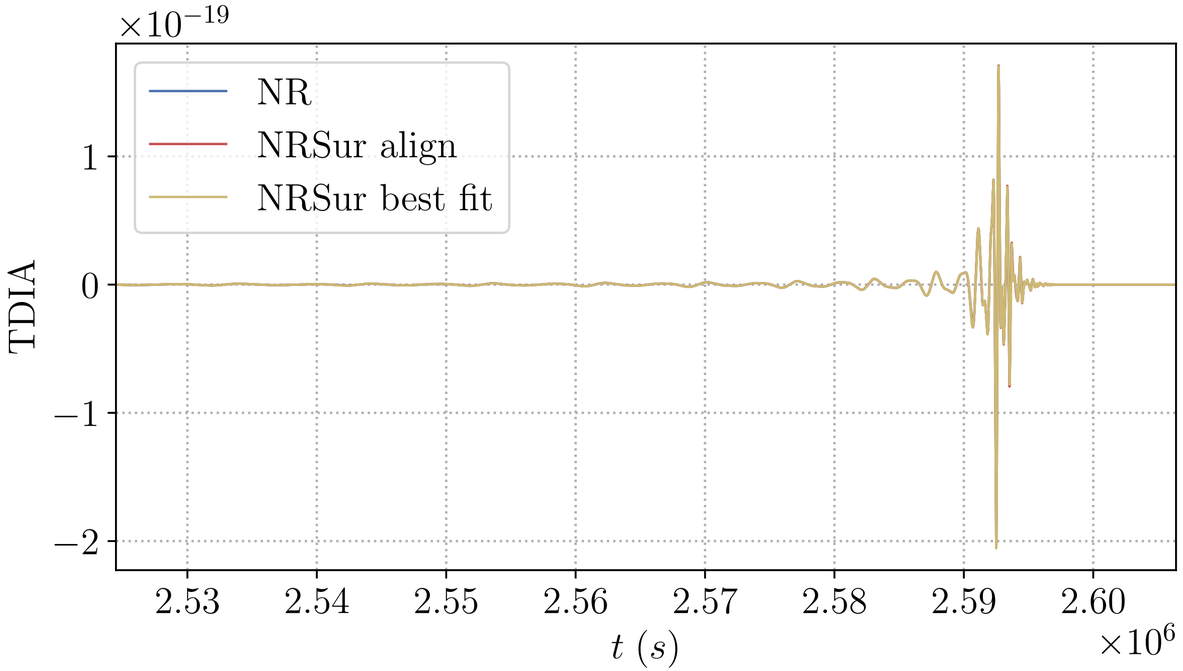


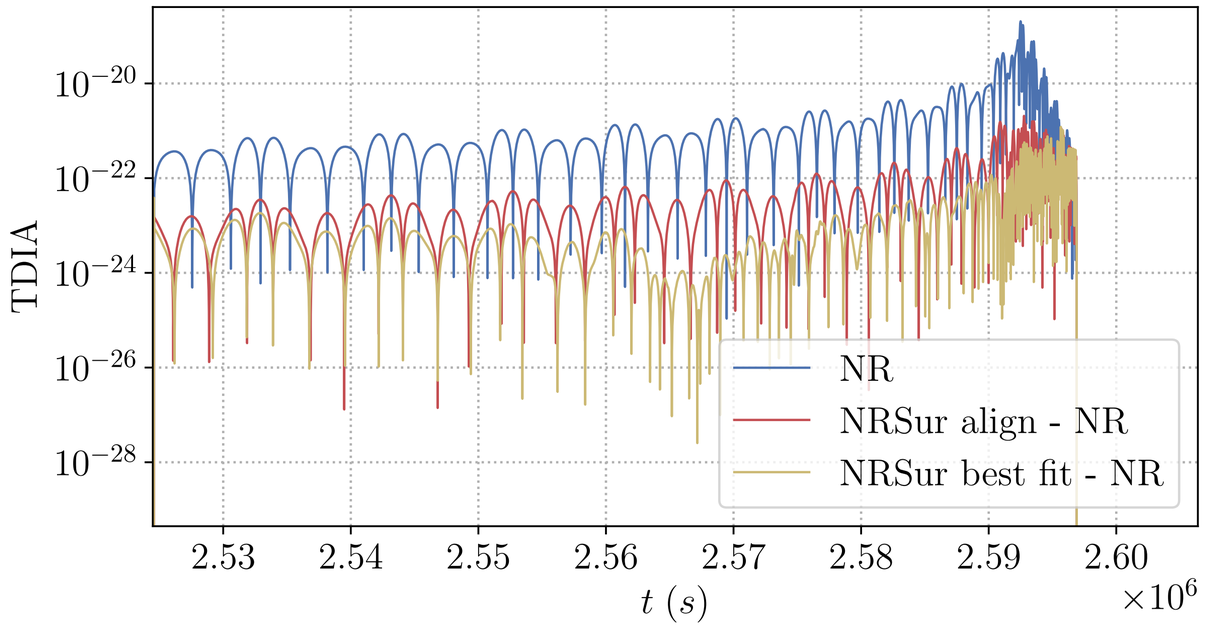


Complement: NR vs NR surrogate

SXS:BBH:2160

$q = 3, \chi_1 = 0.6, \chi_2 = -0.4$





Biases caused by missing higher harmonics

Estimating the bias is an optimization problem

[Preliminary]

- Cutler Vallisneri is one step of Newton's gradient descent, approximating the Hessian with the Fisher matrix

$$\Delta\theta = H^{-1} \cdot \nabla \ln \mathcal{L}$$
$$H_{ij} = \partial_i \partial_j \ln \mathcal{L} \sim (\partial_i h | \partial_j h)$$

- Are there better optimization algorithms ? (e.g. simplex method)
- One simple idea is to repeat CV as iterated gradient descent

There should be better bias estimators than CV - robustness ?

[Sophia Yi &al, in prep.]

