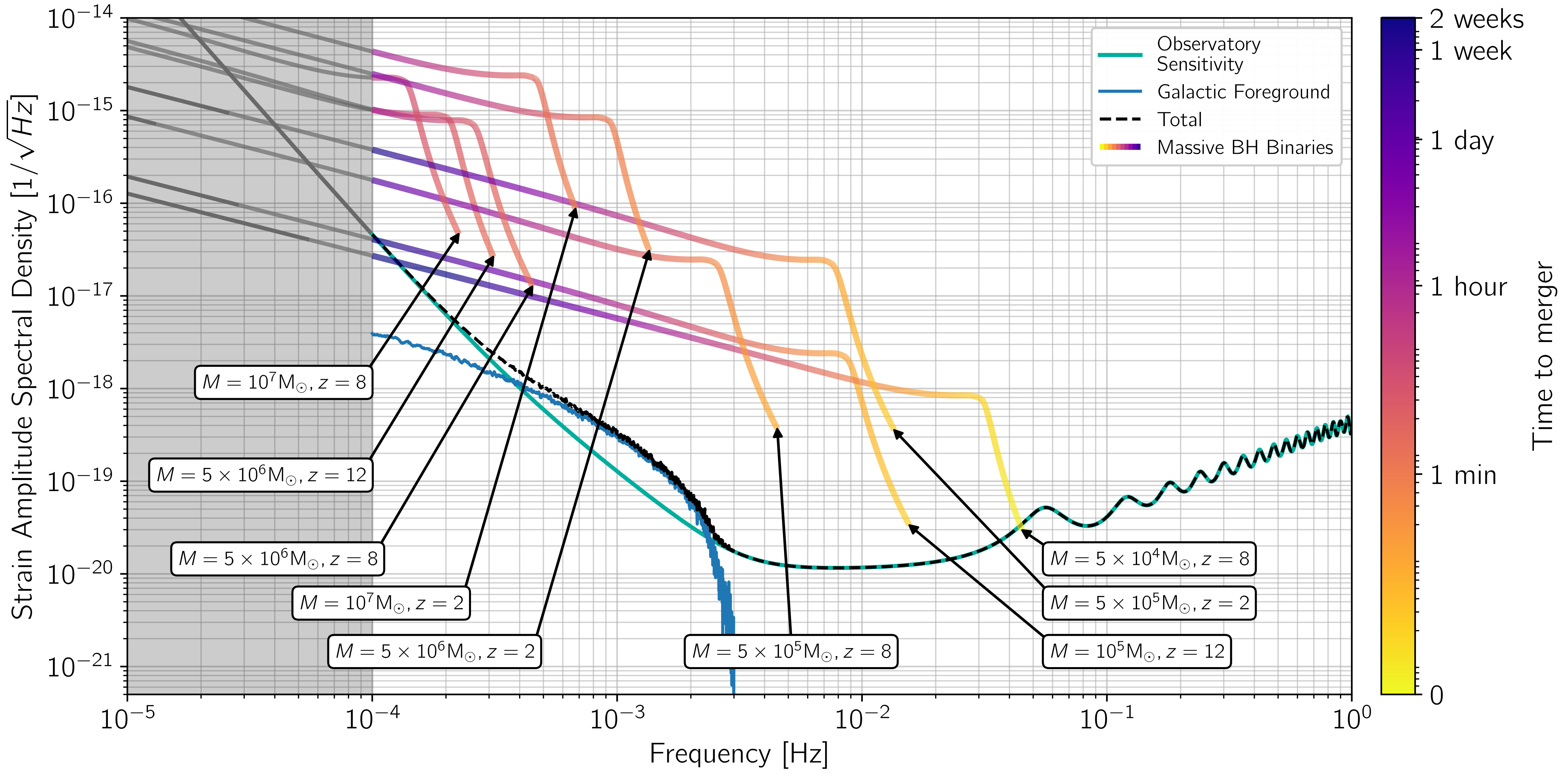
An exploration of the impact of waveform systematics for LISA massive black hole binaries

Sylvain Marsat (L2IT, Toulouse)

Massive black hole binaries for LISA



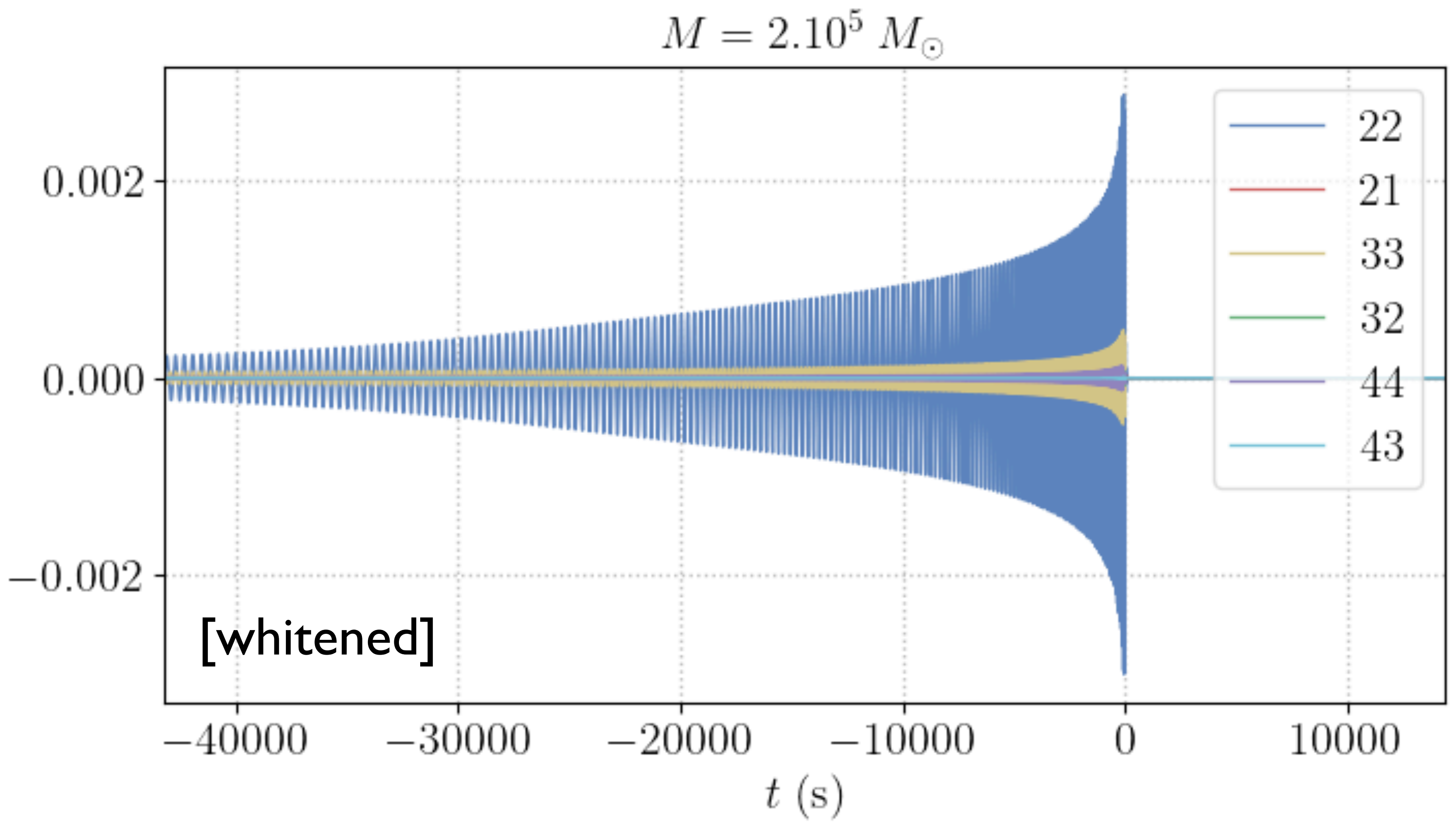
Science case

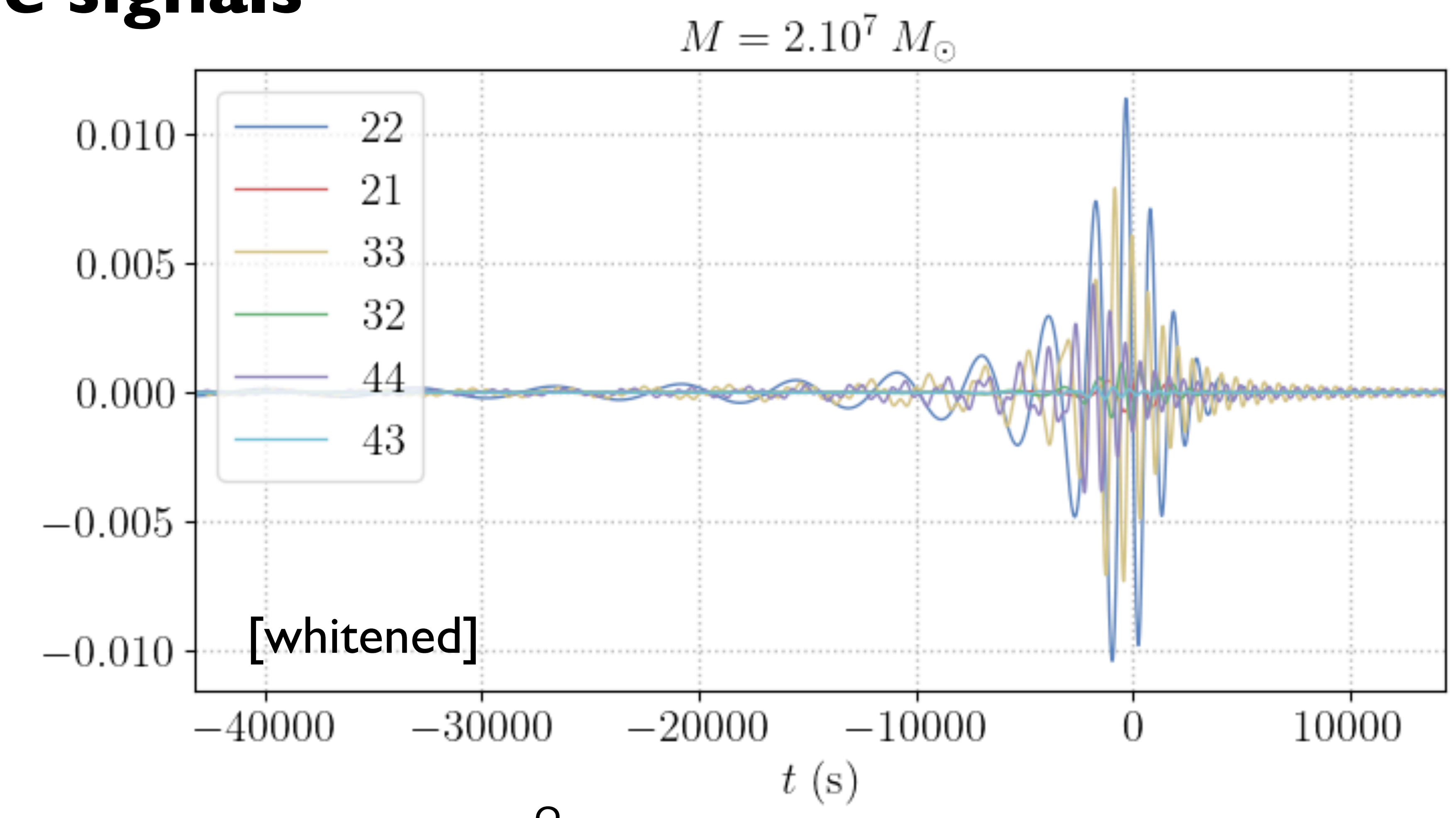
MBHBs cardinal sources for LISA, SNRs of thousands !

Waveform errors (systematics) crucial for:

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

Example signals





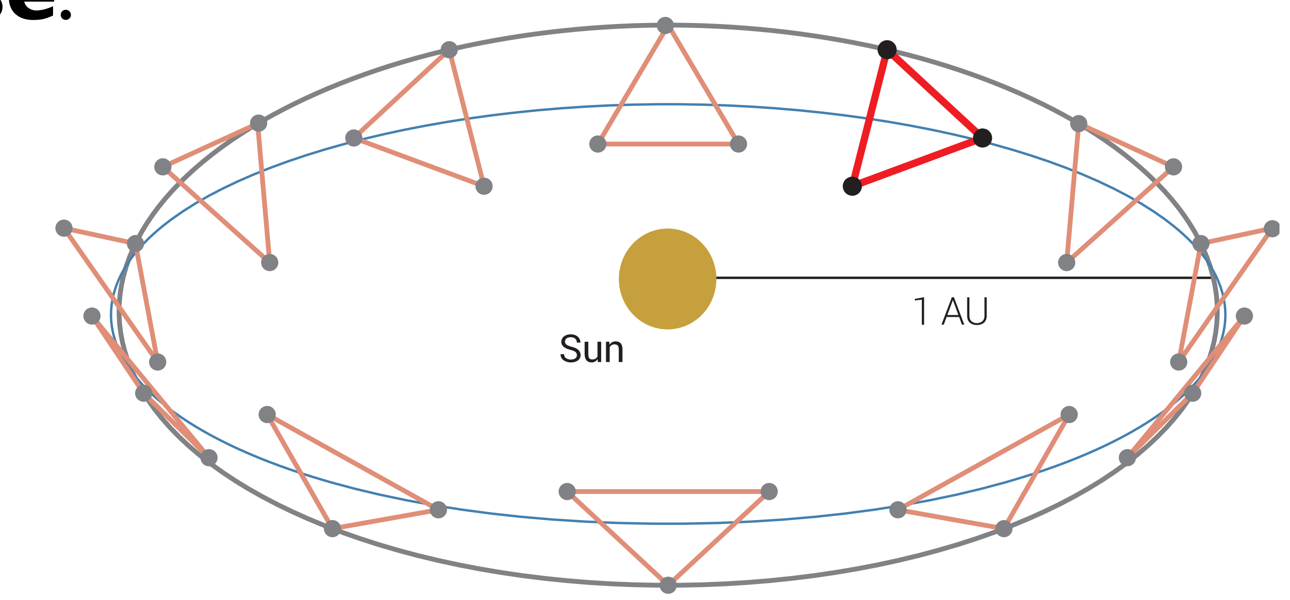
In the literature:

- a lot of recent activity in LVK/LISA
- similarity 3G/LISA

[Toubiana&al 2023]
[Pitte&al 2023]
[Jan&al 2024]
[Dhani&al 2024]
[Kapil&al 2024]

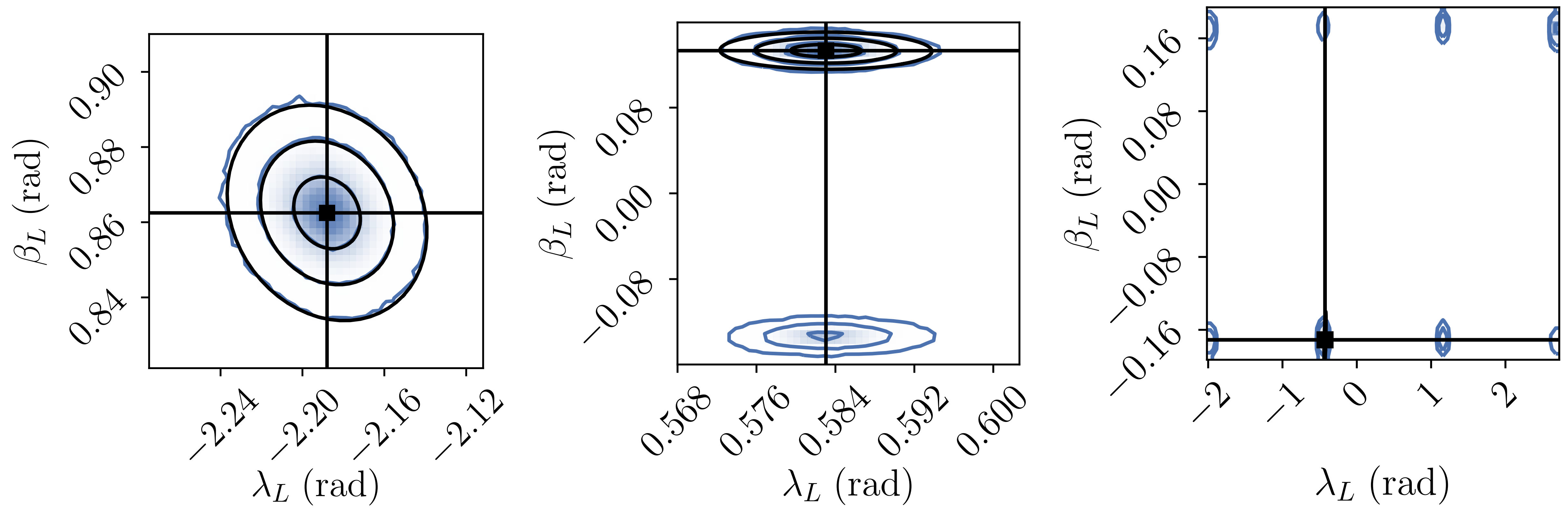
LISA response and sky localization of MBHBs

Response:

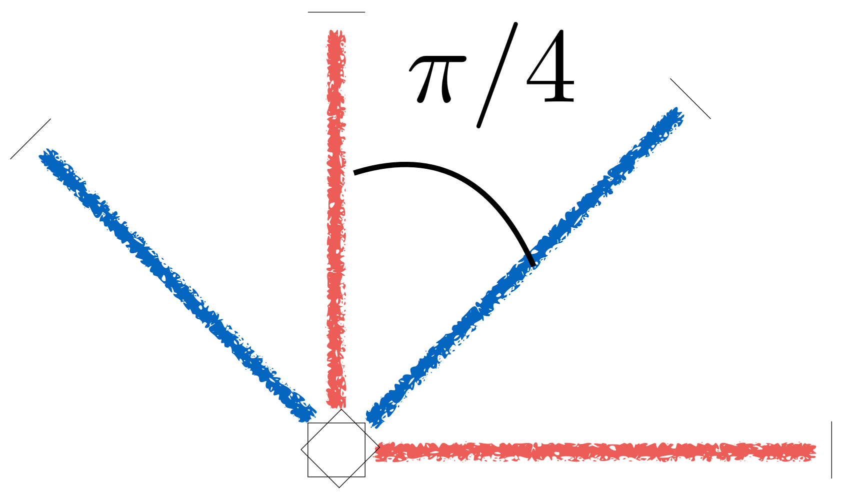


Time and frequency-dependency
Time: motion of LISA on its orbit
Frequency: inter-spacecraft delays

Multimodality patterns:



'Pattern function' localization:



Without these effects TDIA, TDIE become 2 LIGO-like channels

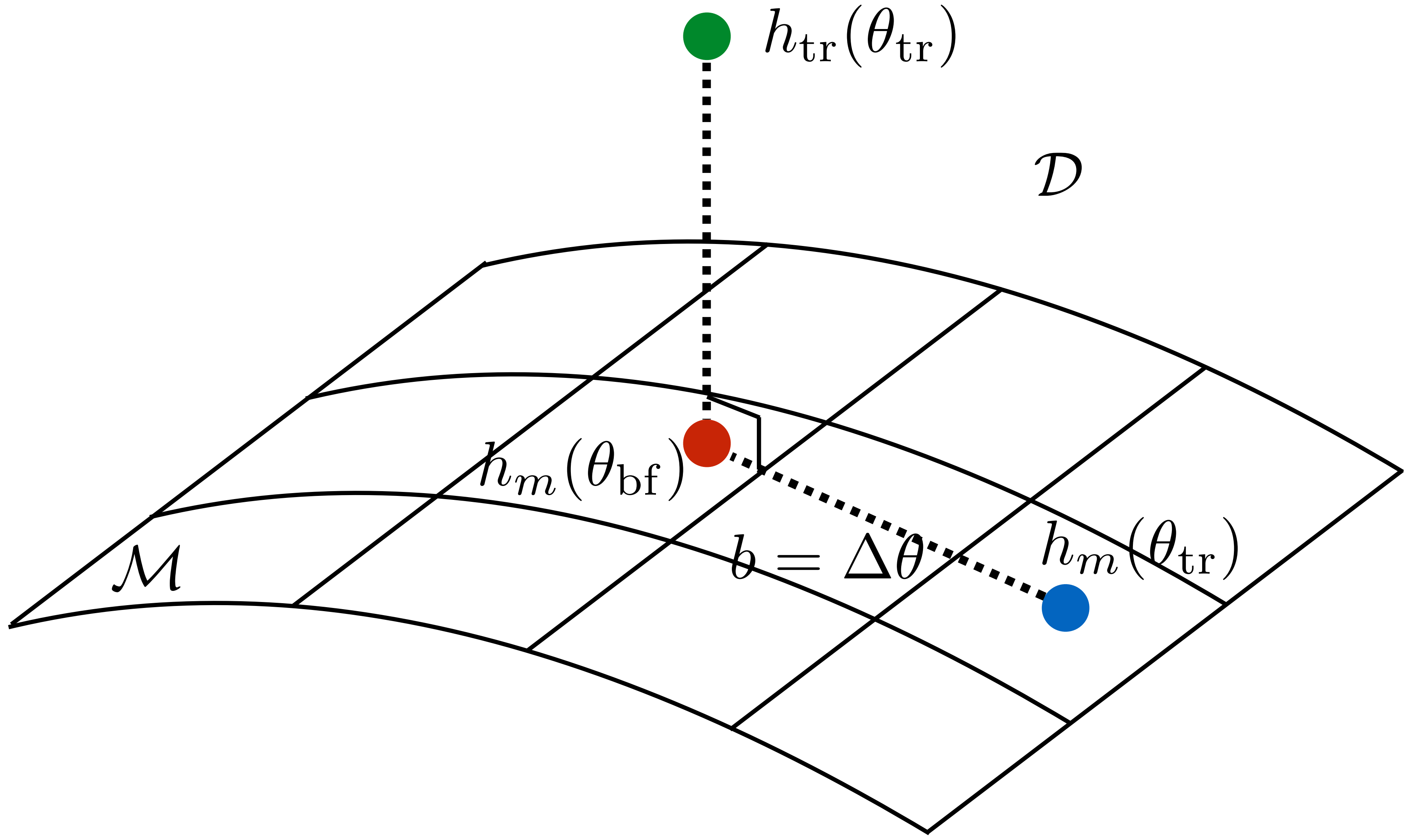
The size of the main mode in the sky is driven by the pattern function response $F_+, F_\times$ for 2 channels, with higher harmonics: amplitude/phase of $h_{\ell m}$important

Degeneracy breaking:

- motion of LISA: weak for short high-mass signals
- high-frequency effects in the response: only at high frequencies

- 'Pattern function' response is a source of information at high mass, depends on $h_{\ell m}$ amplitudes
- Multimodality broken by subdominant effects in response (motion, high-f)
- LISA sky localization at high mass: weak effects, high SNR

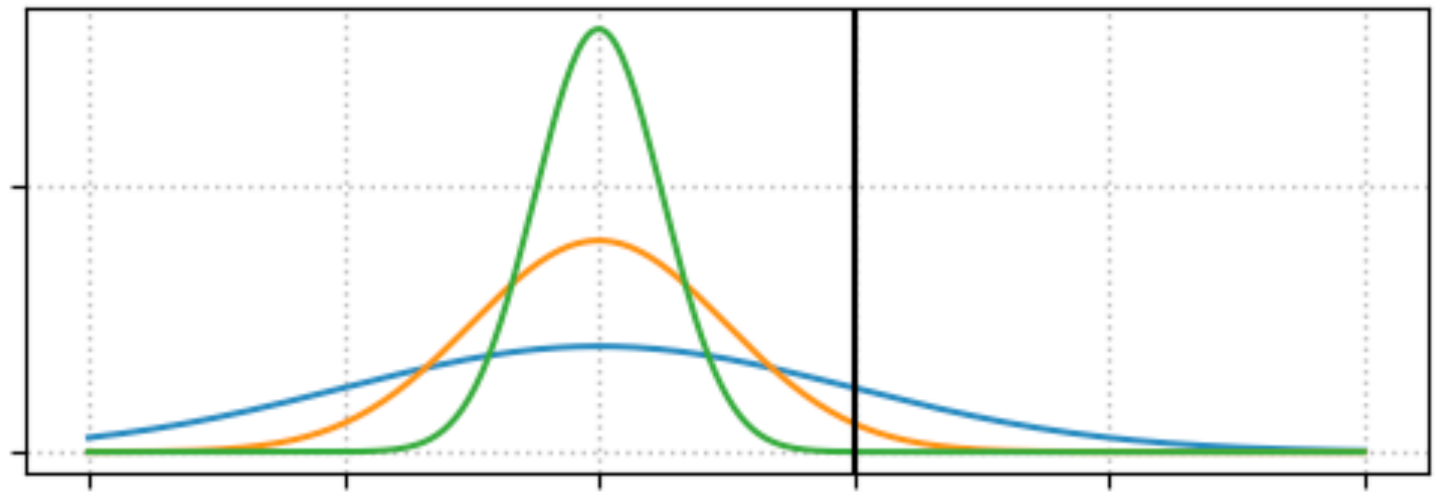
Waveform systematics and parameter estimation



Systematic biases:

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

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



Mismatch (unfaithfulness):

Mismatch, optimization over time/phase/polarization:

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

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

Indistinguishability criterion:

[Lindblom&al 2008]
[Chatziioannou&al 2019]

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

$$\ln \mathcal{L}(\theta_{bf}) \sim \ln \mathcal{L}(\theta_{1-\sigma}) \qquad \mathrm{MM} < \frac{D}{2}\frac{1}{\mathrm{SNR}^2}$$

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

Goals:
- Assess waveform models systematics for LISA with full PE
- Compare with simpler methods

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]

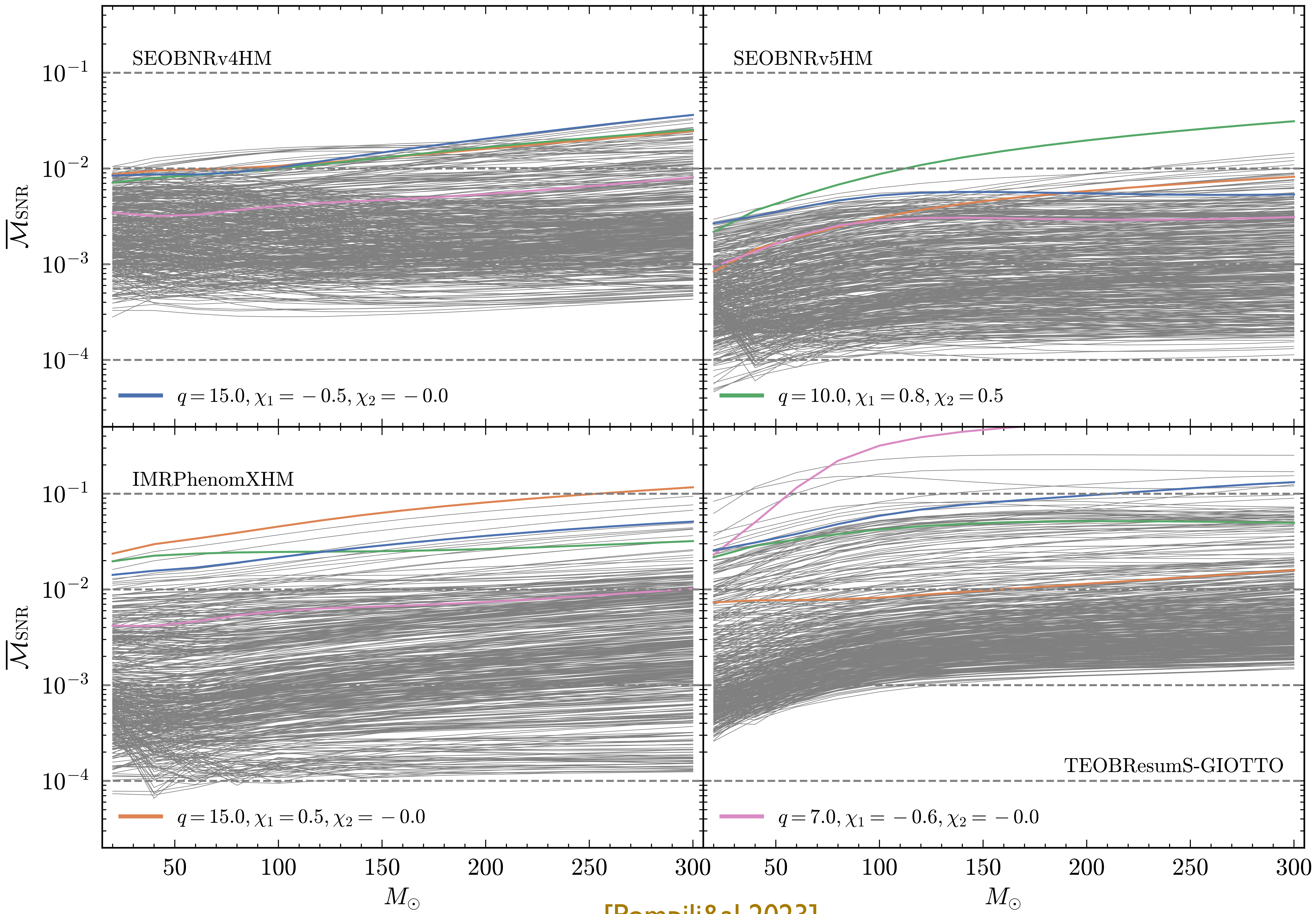
Mode content for all: 22, 21, 33, 44

Limitations:

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

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

Precessing spin case: mismatch with NR ~ $10^{-3} - 10^{-1}$



[Pompili&al 2023]

Analysis settings: waveform models

Parameter estimation:

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

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

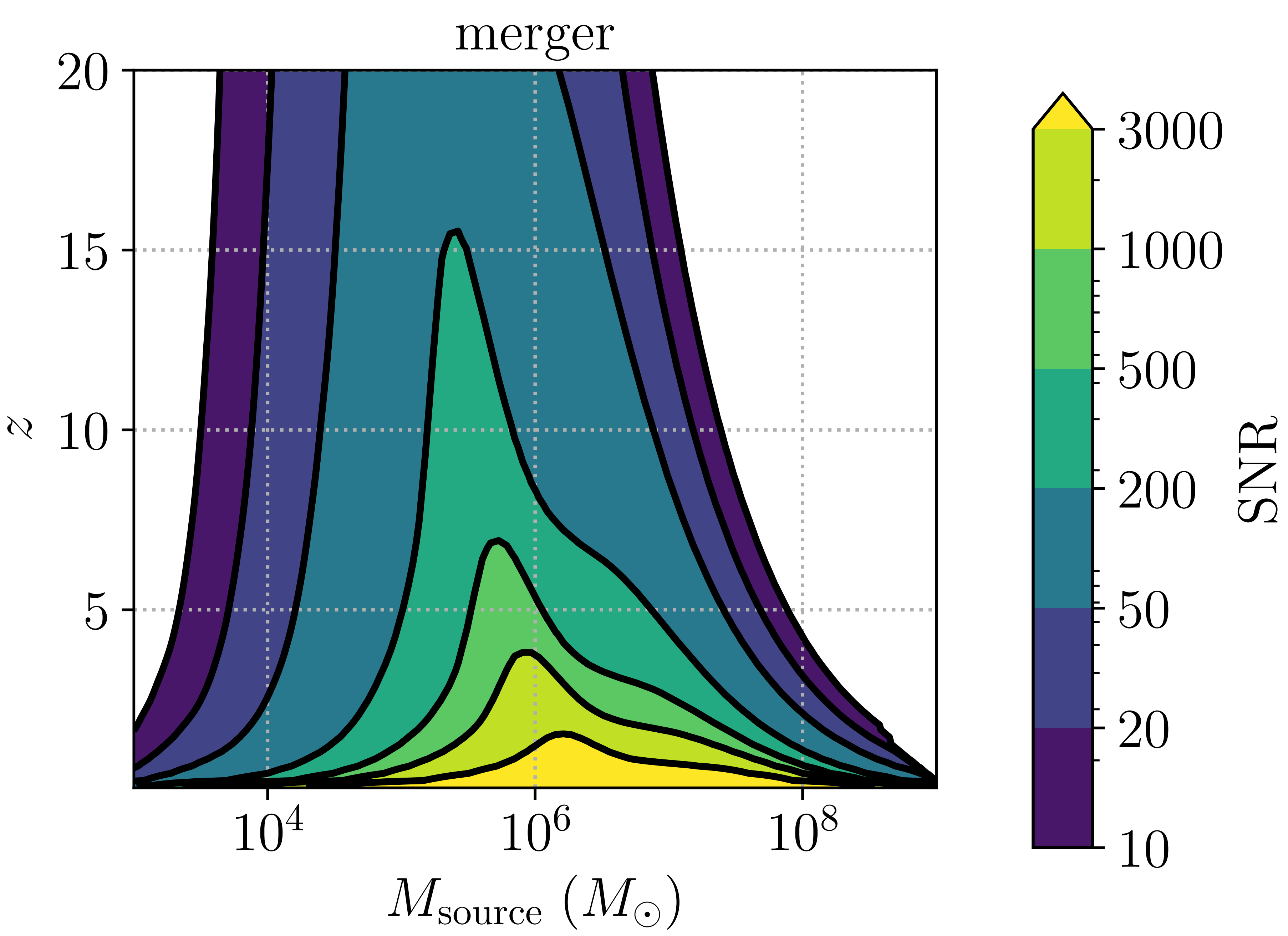
Fourier-domain full response (same for injection and templates)

Direct FD Whittle likelihood (no approx.)

Parameter space exploration:

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

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



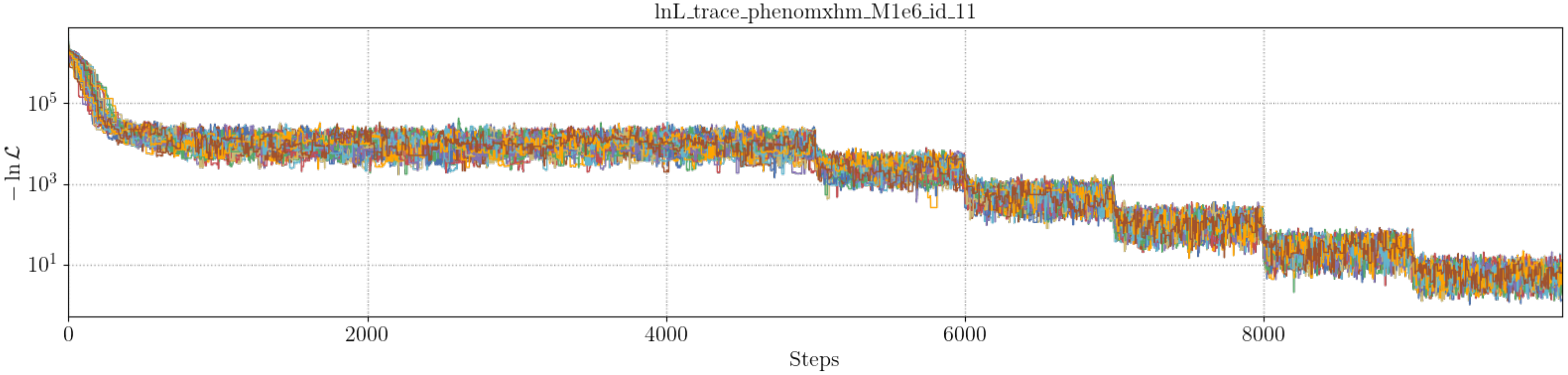
[Preliminary]

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.

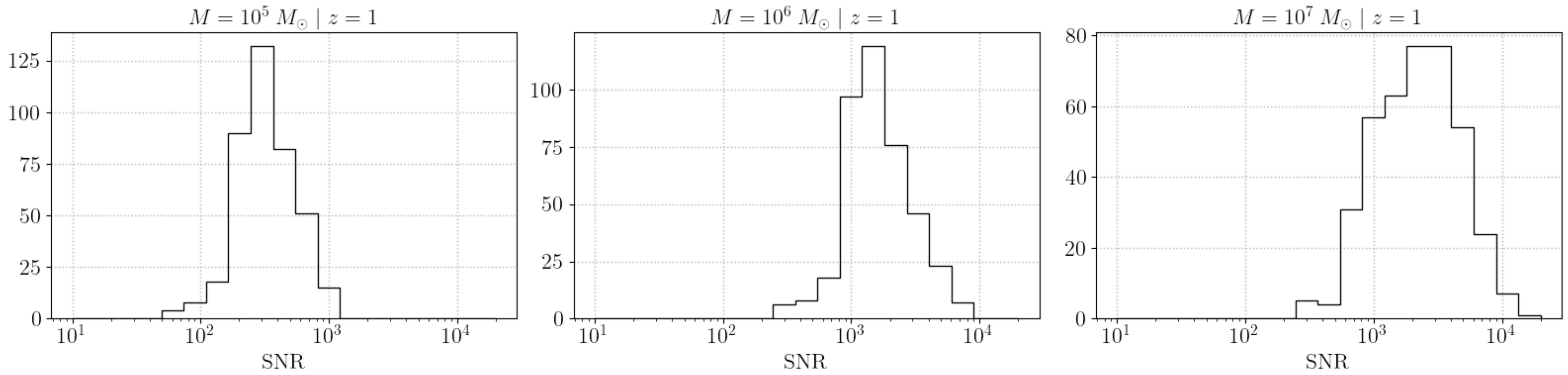


SNRs and mismatches

Signal-to-noise ratios:

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

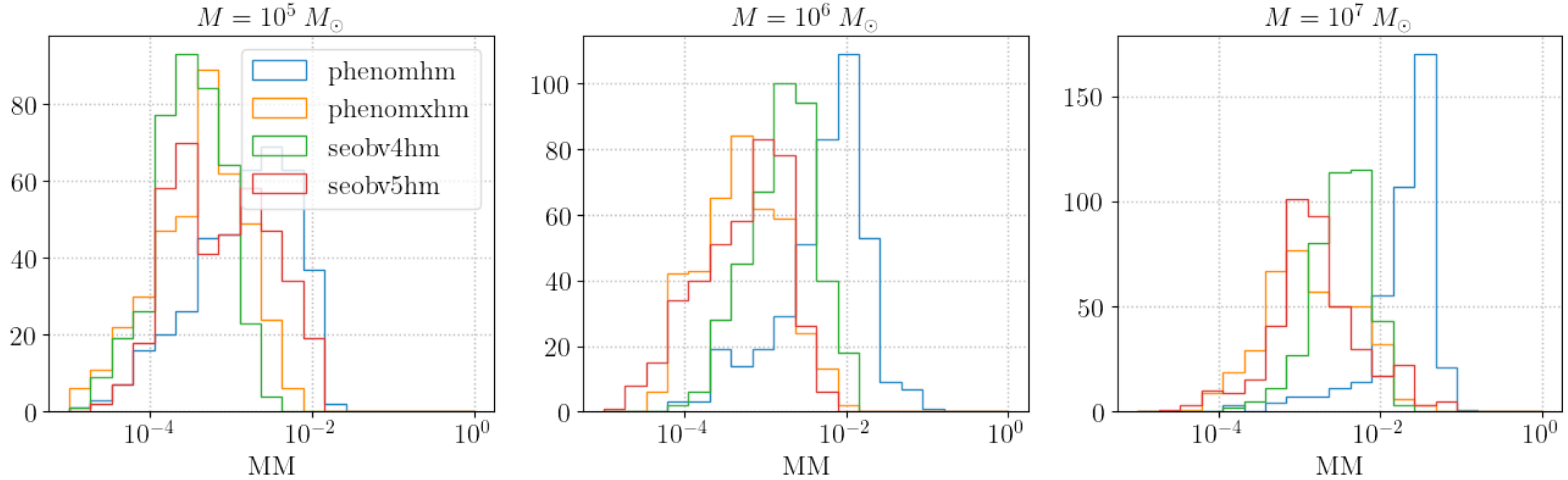
SNRs up to thousands at $z = 1$



Mismatch:

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

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

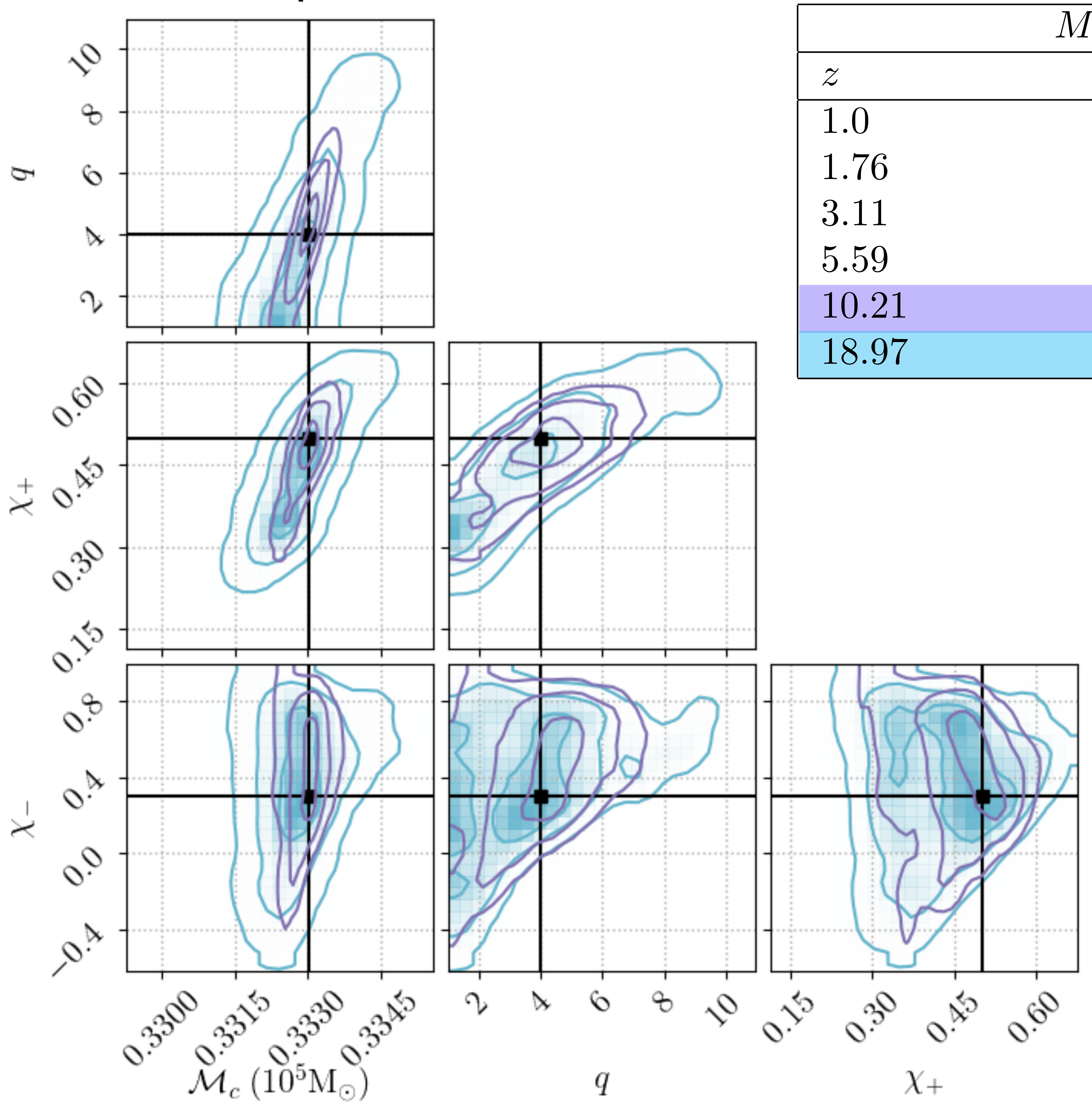


- Mismatches order-of-magnitude similar to LVK
- Improvement for more recent models

Example Parameter estimation with systematics I

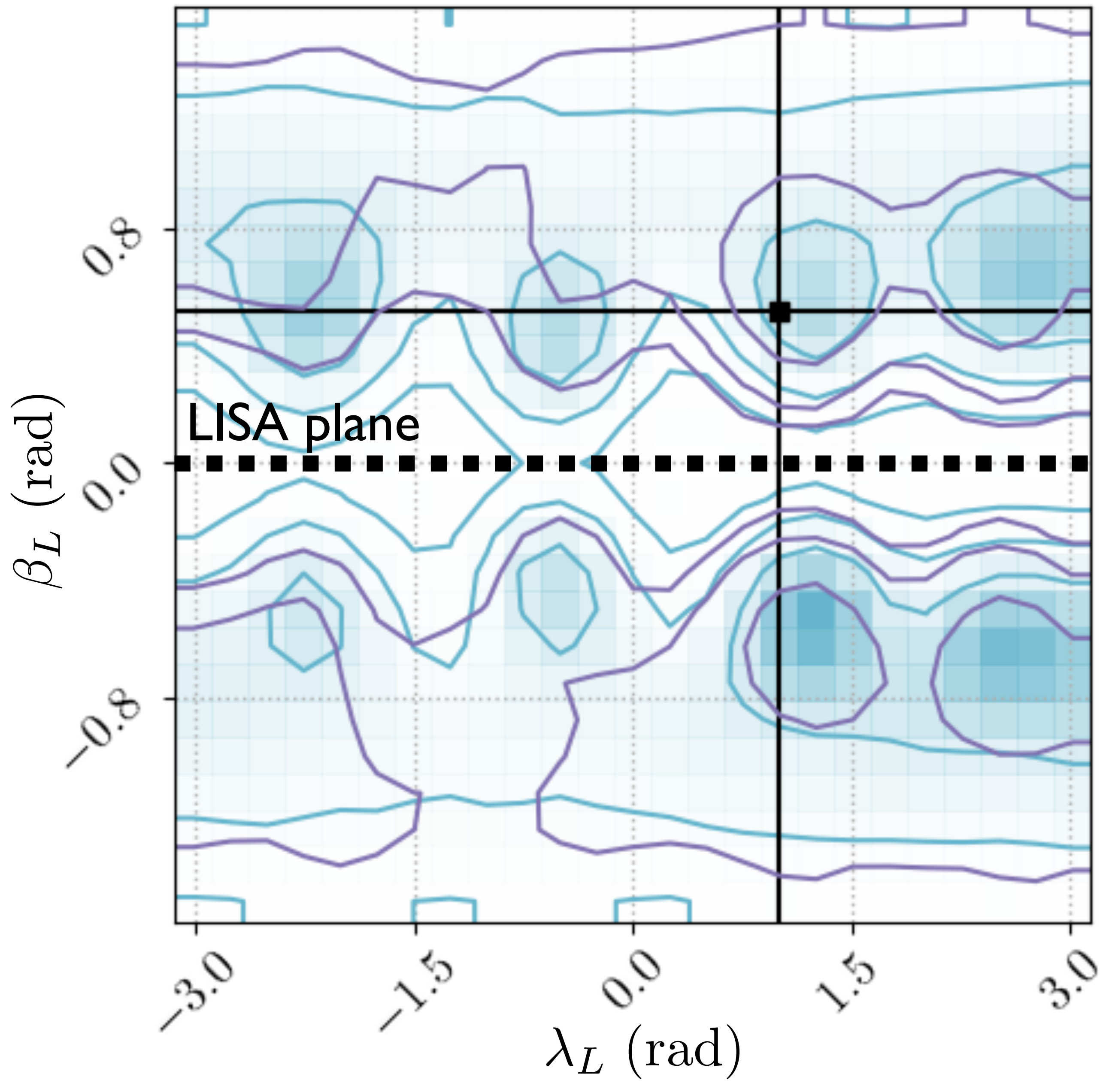
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

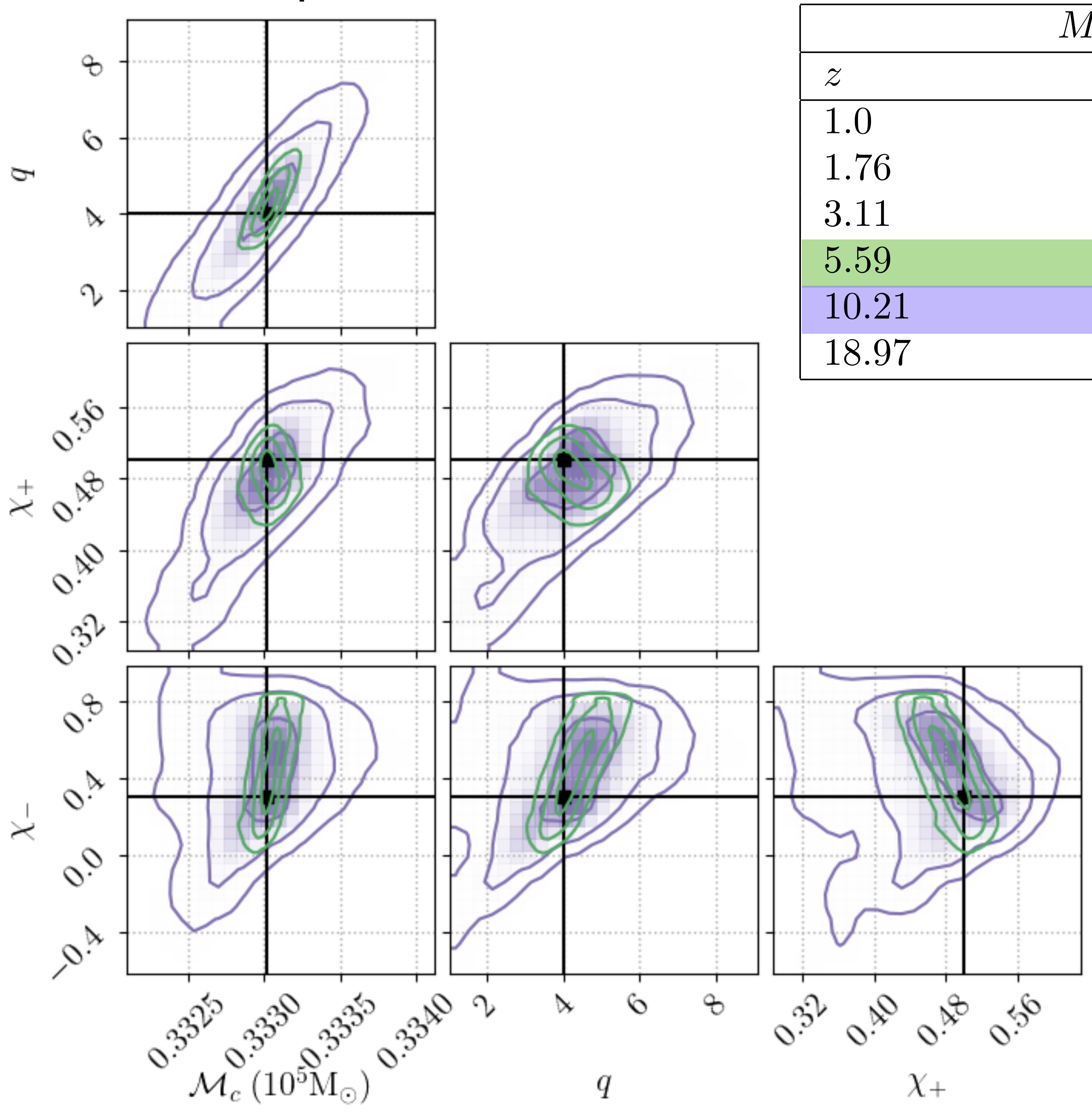
Sky localisation



Example Parameter estimation with systematics I

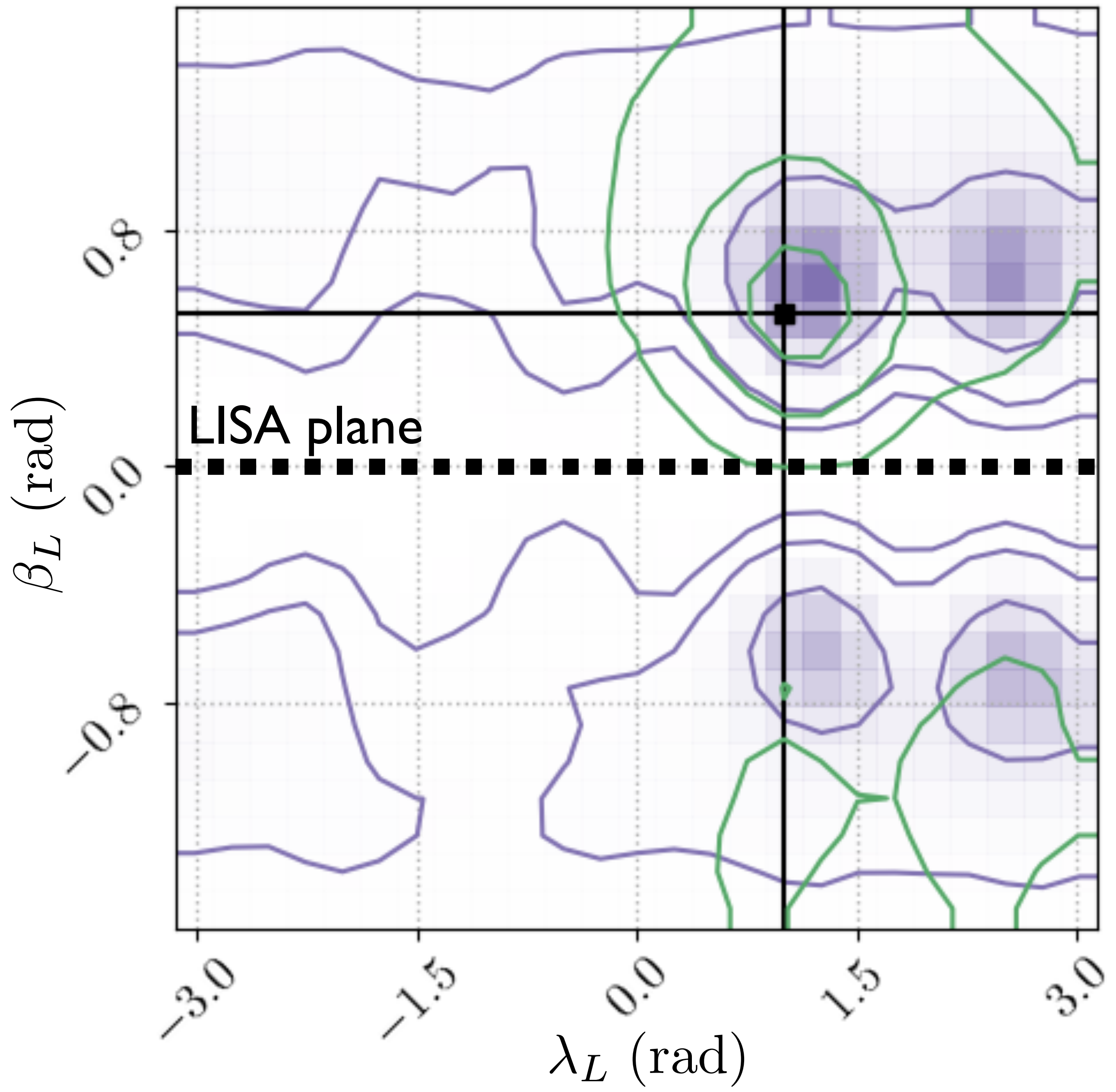
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

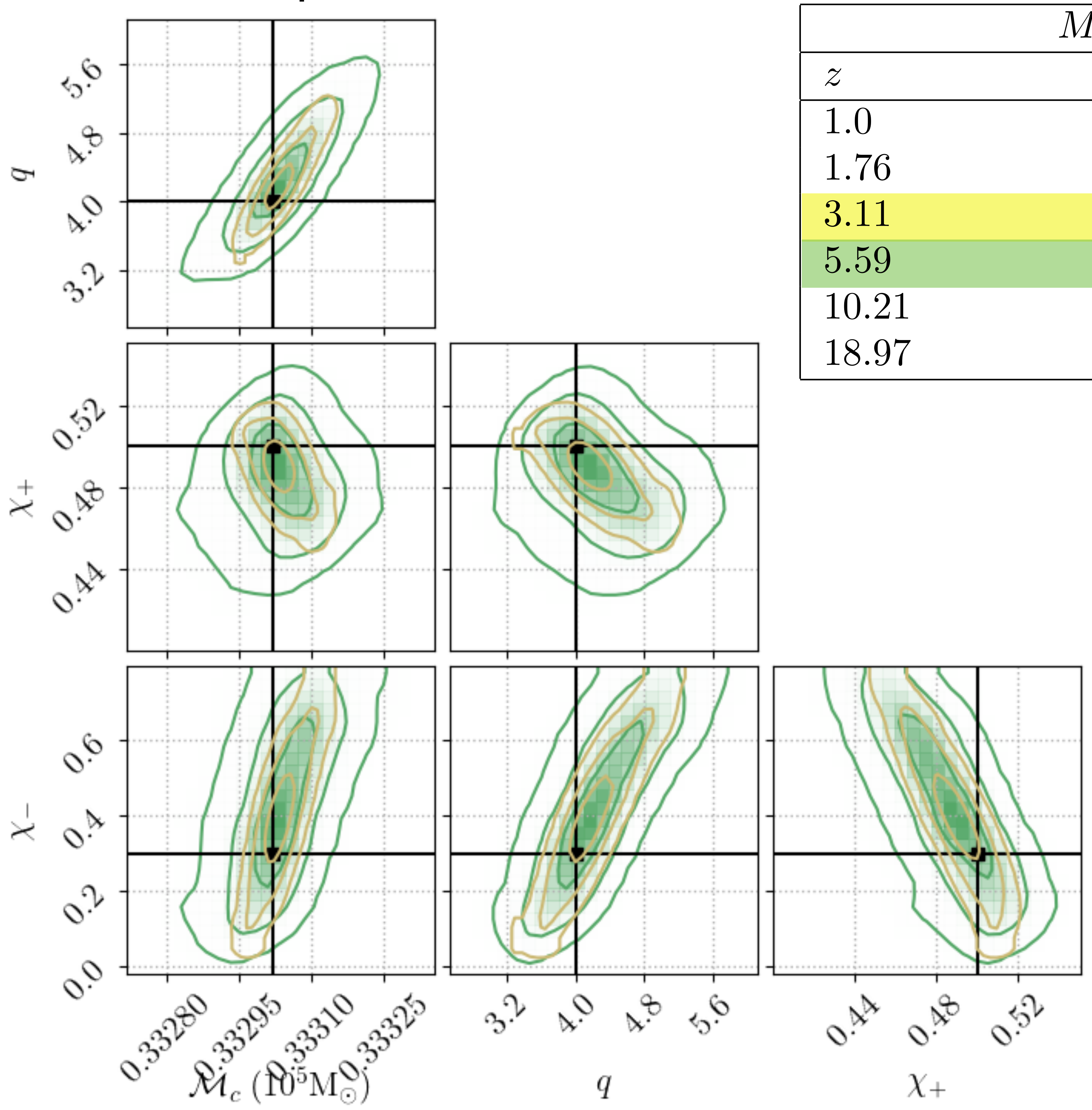
Sky localisation



Example Parameter estimation with systematics I

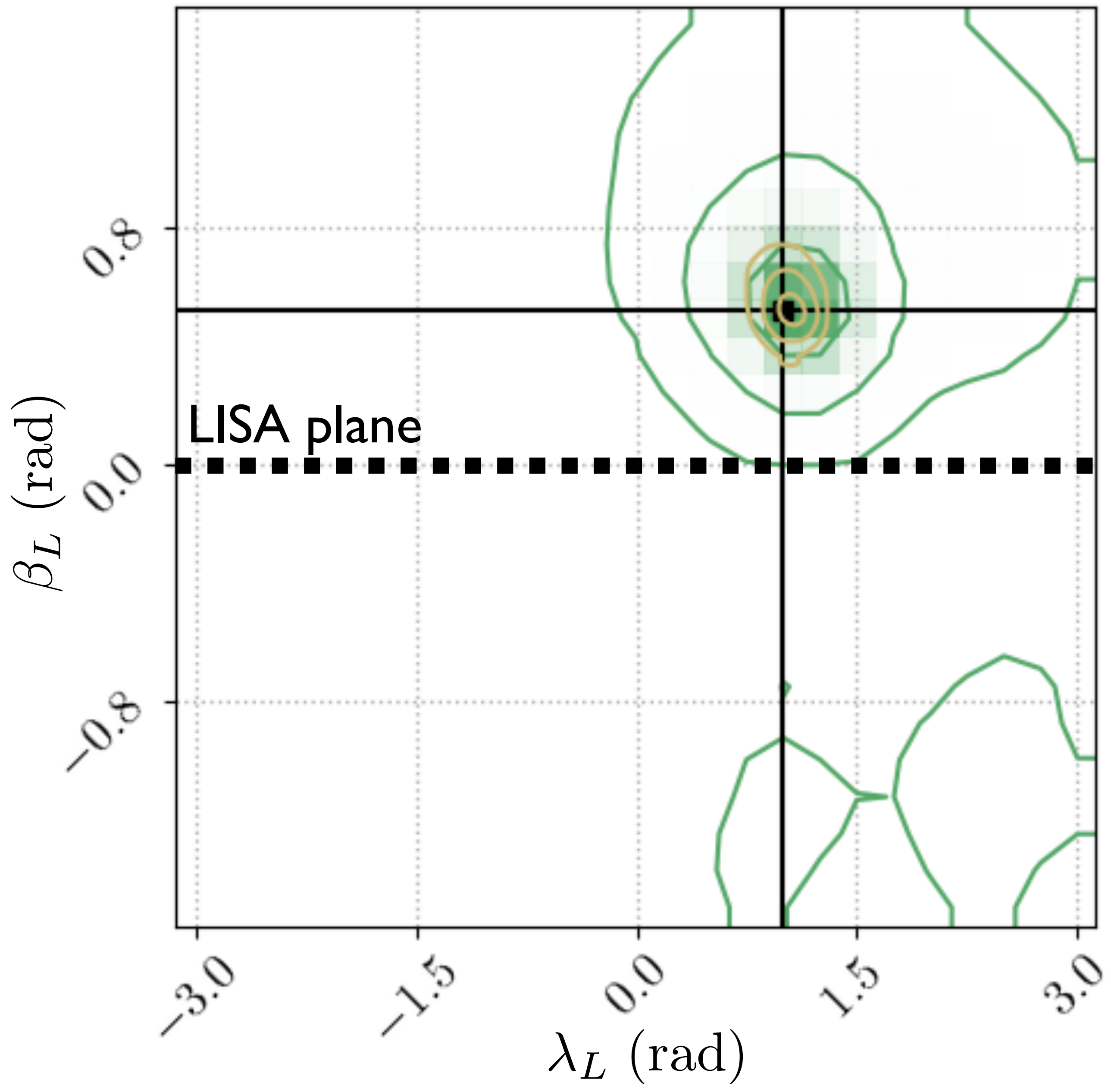
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_{\odot}, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

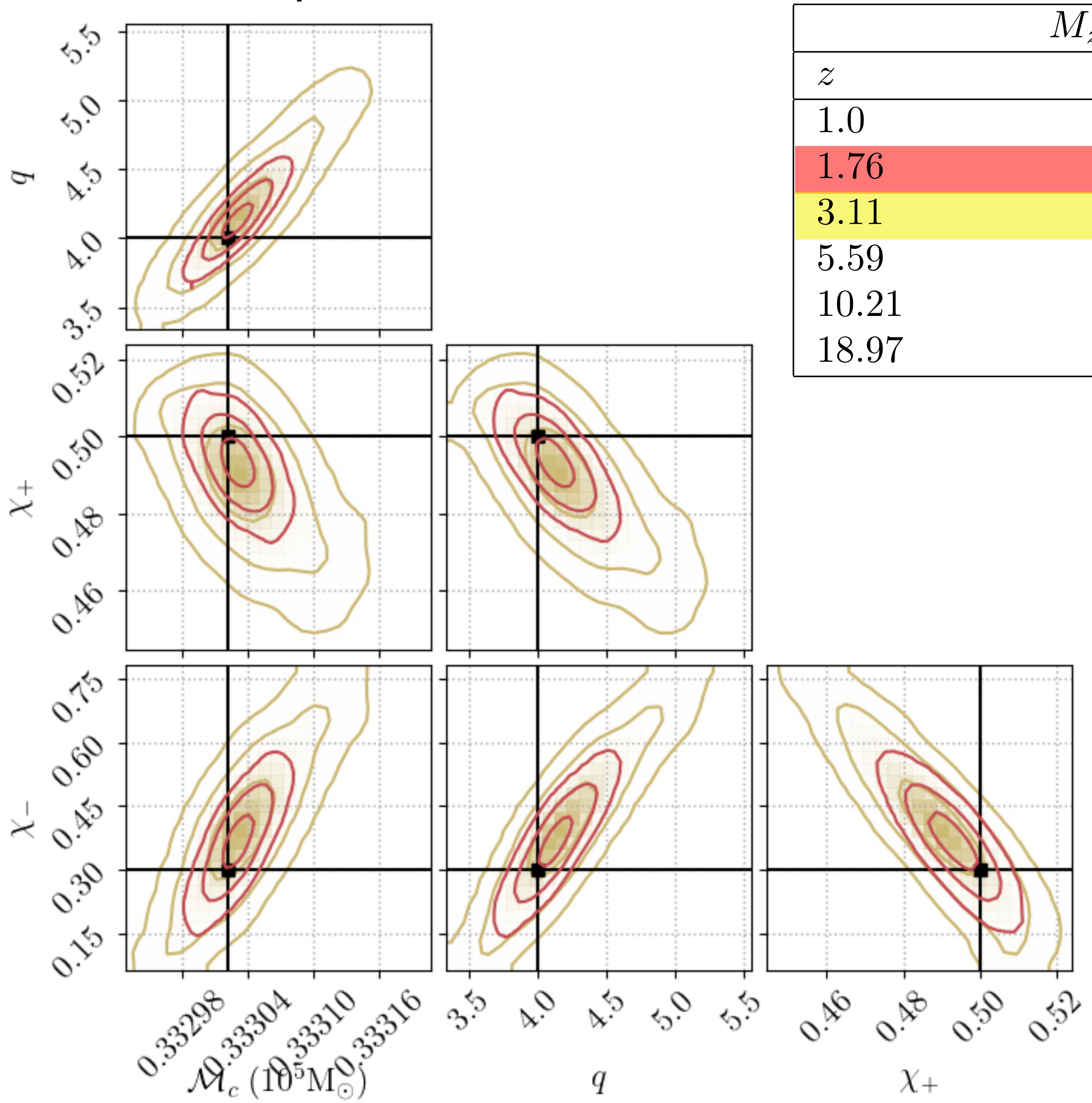
Sky localisation



Example Parameter estimation with systematics I

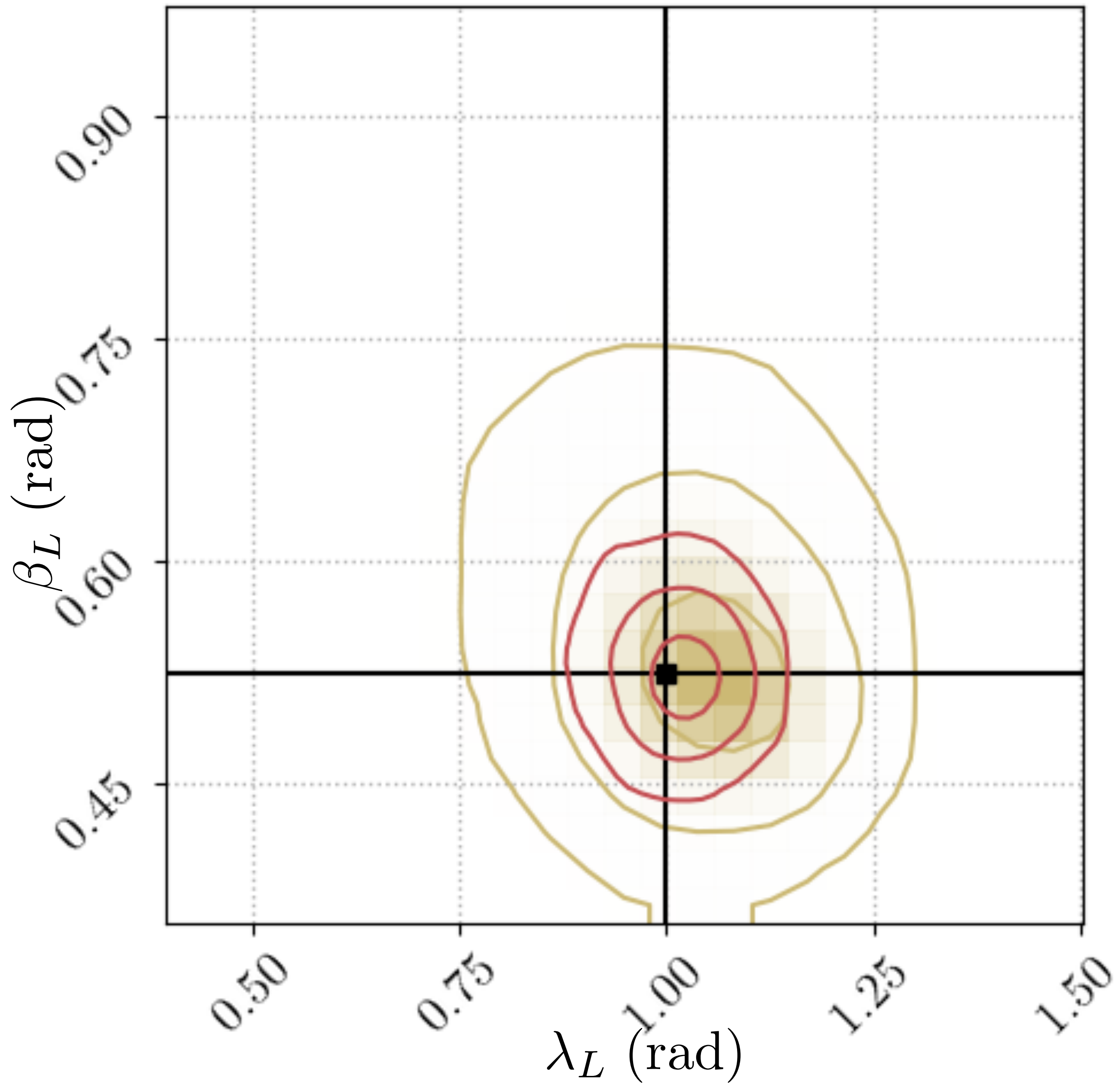
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

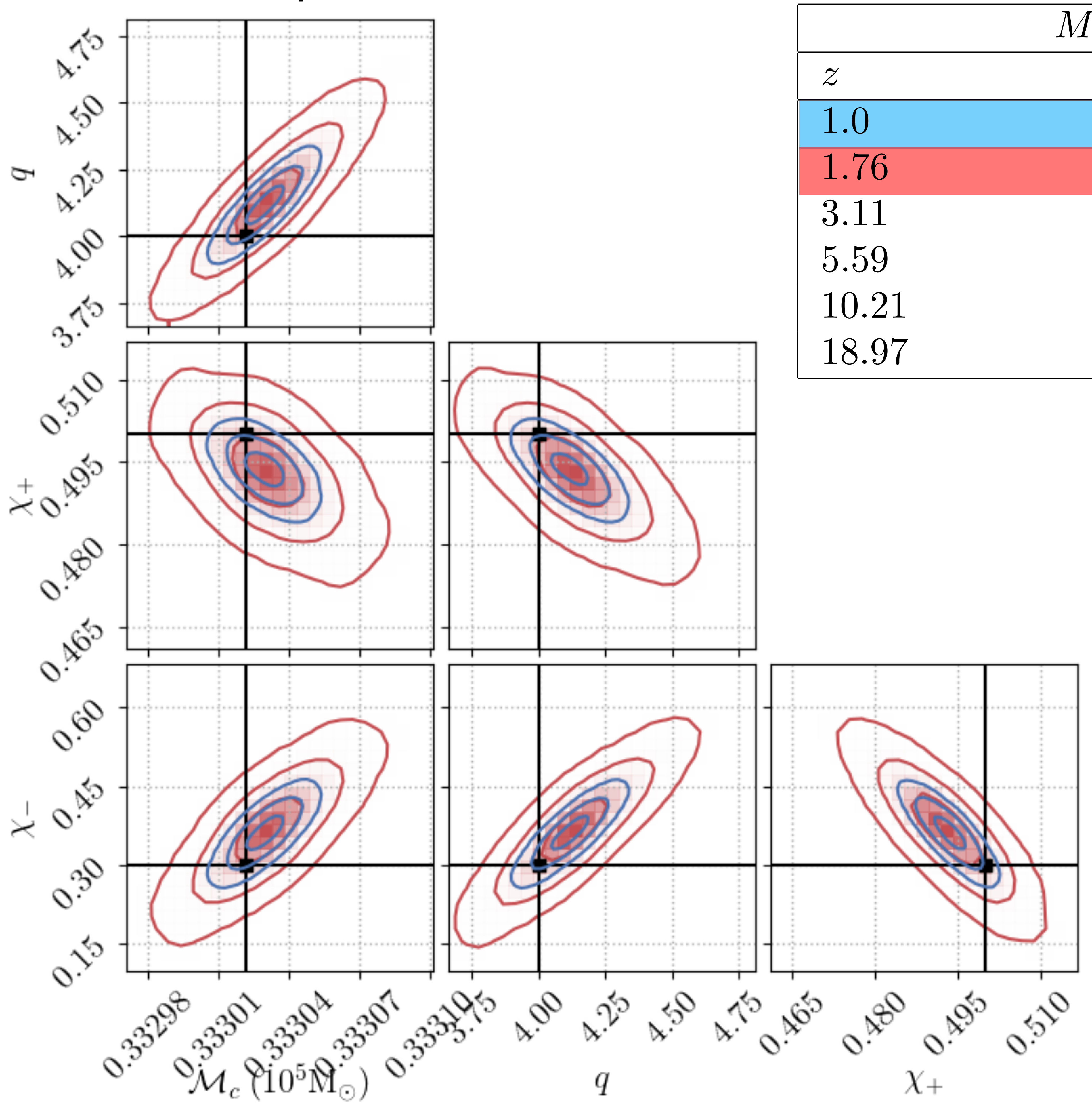
Sky localisation



Example Parameter estimation with systematics I

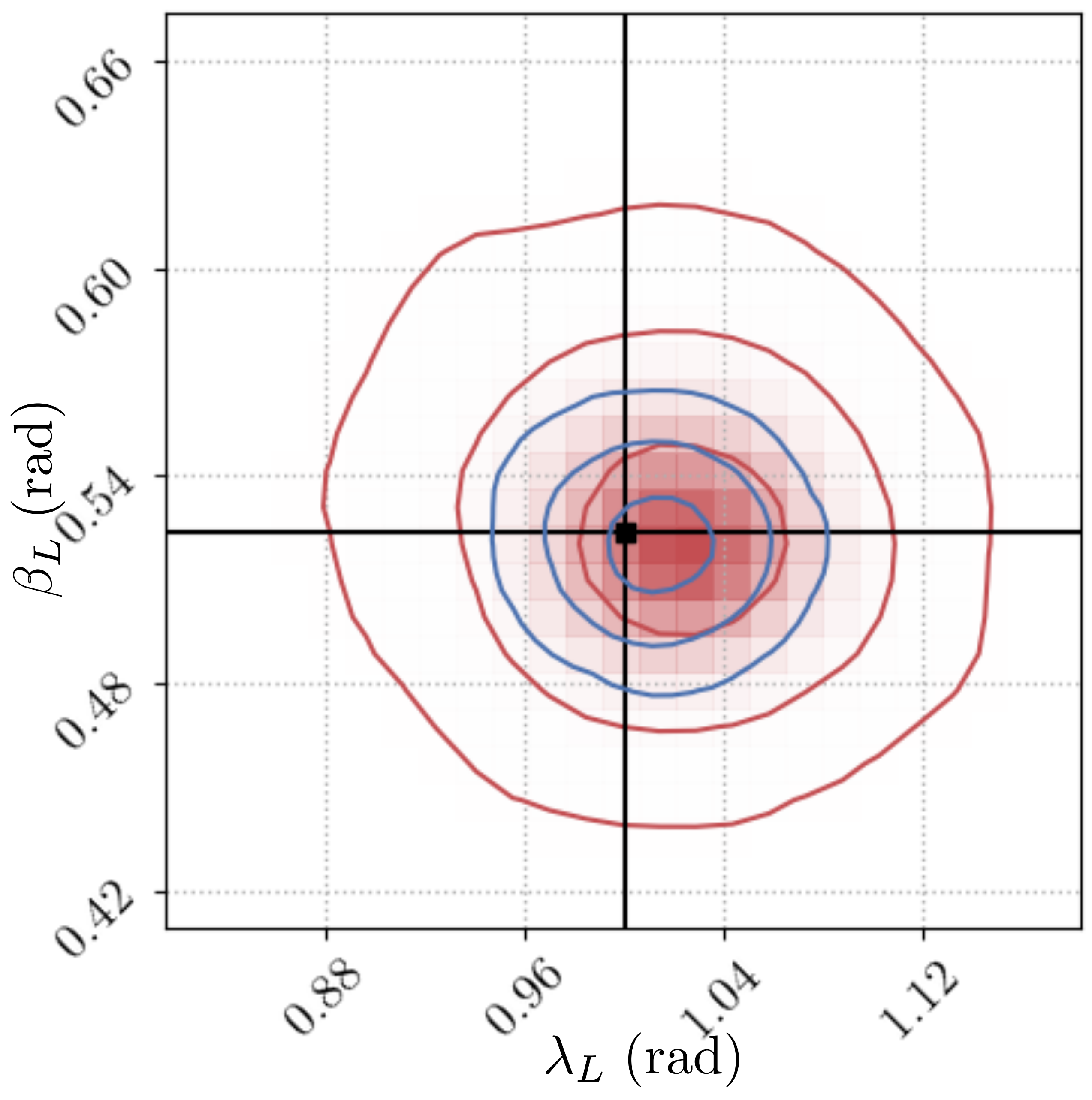
- **Injection:** NRHybSur3dq8 $\{M = 10^5 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

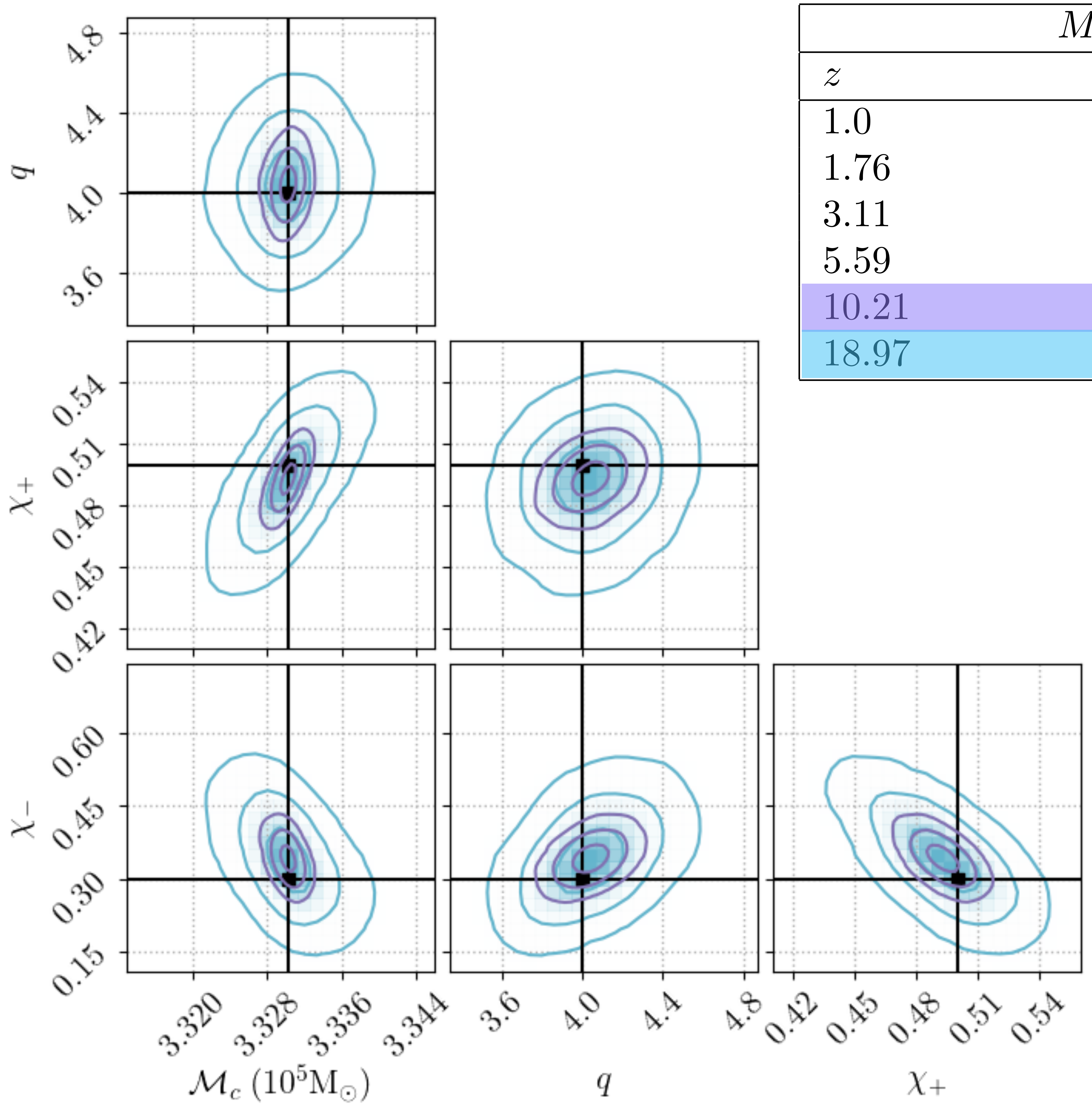


The good:
- converges on the true parameters
- mild bias at $z = 1$, SNR $= 317$

Example Parameter estimation with systematics II

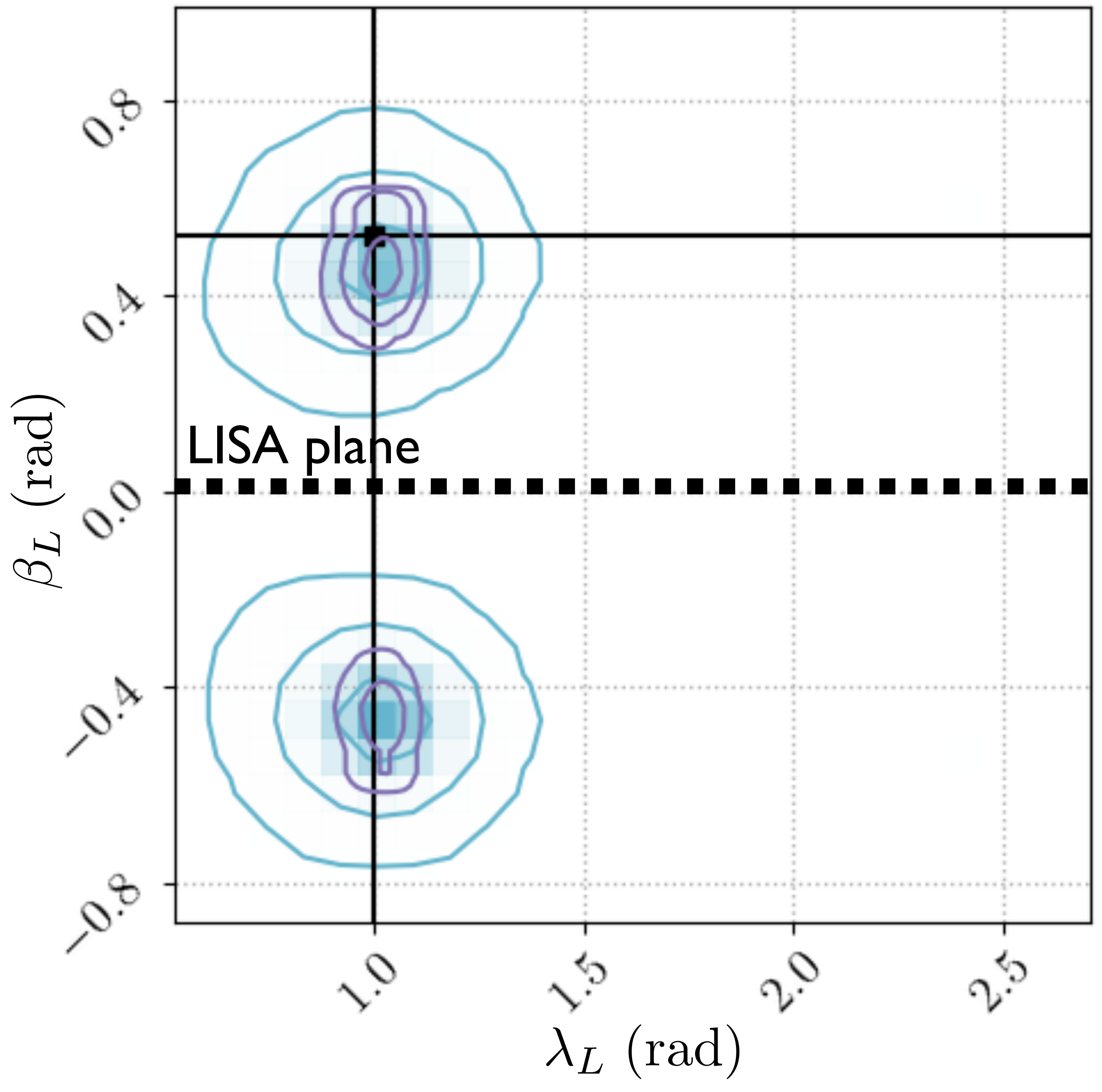
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

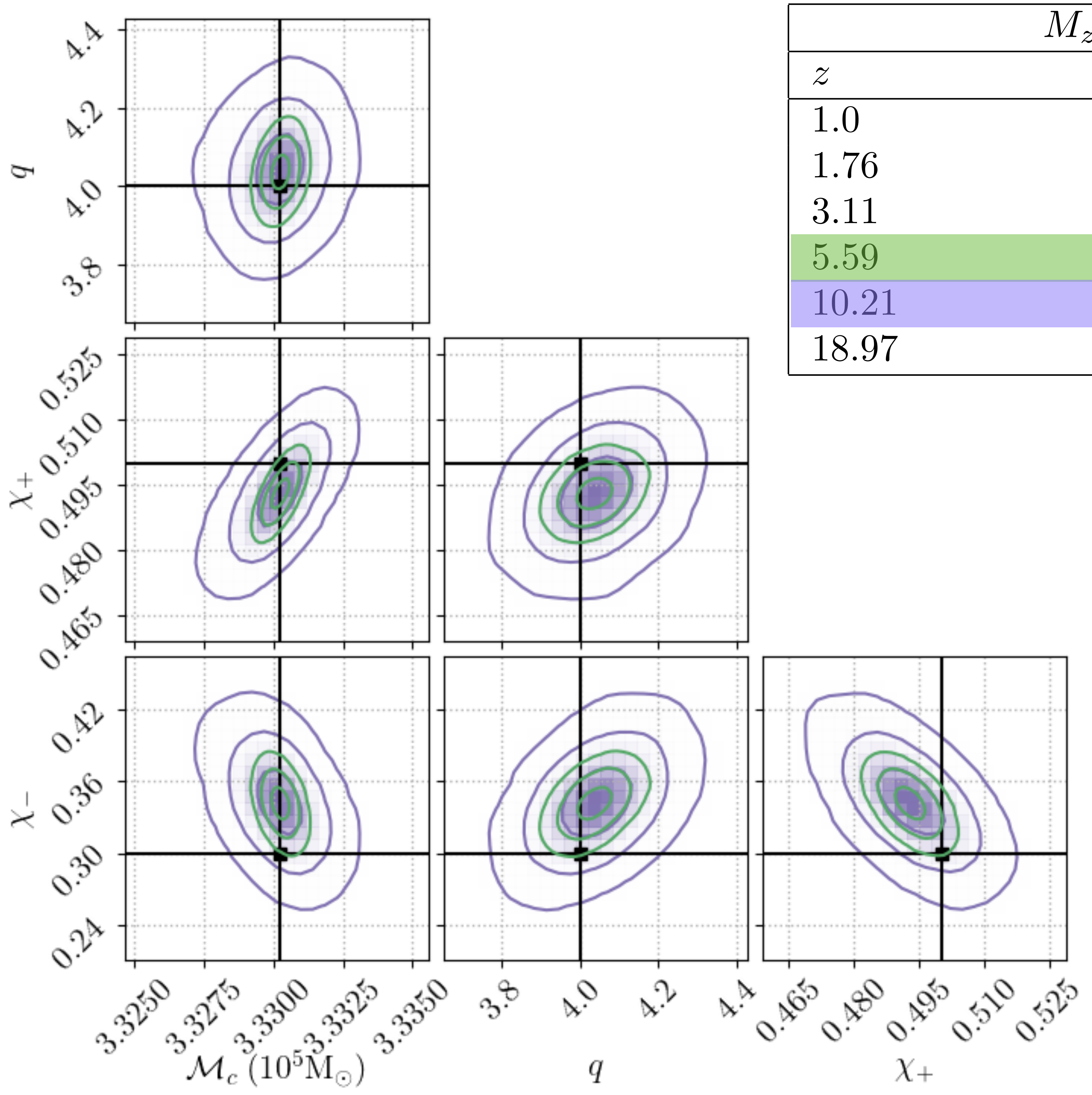
Sky localisation



Example Parameter estimation with systematics II

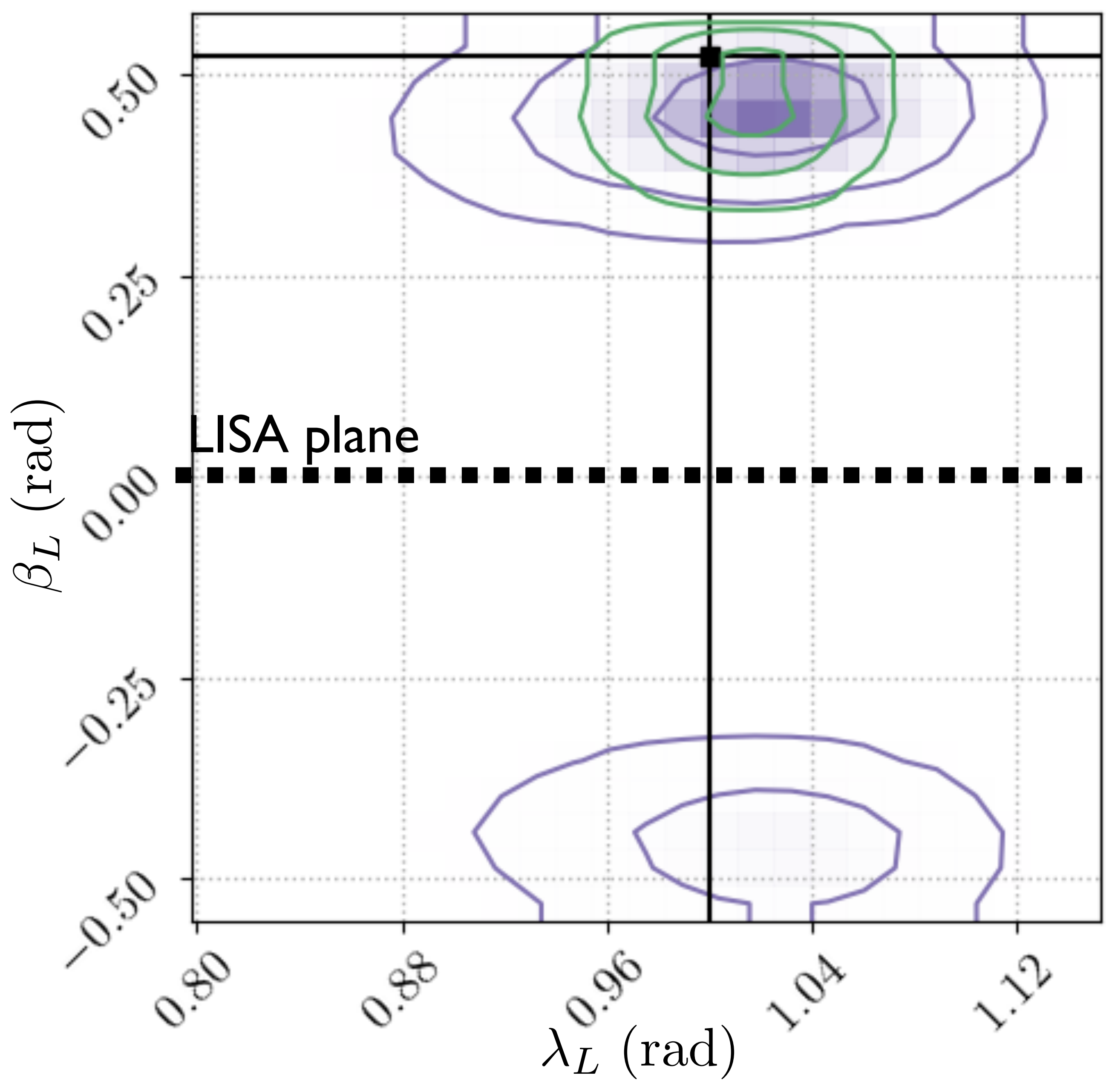
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

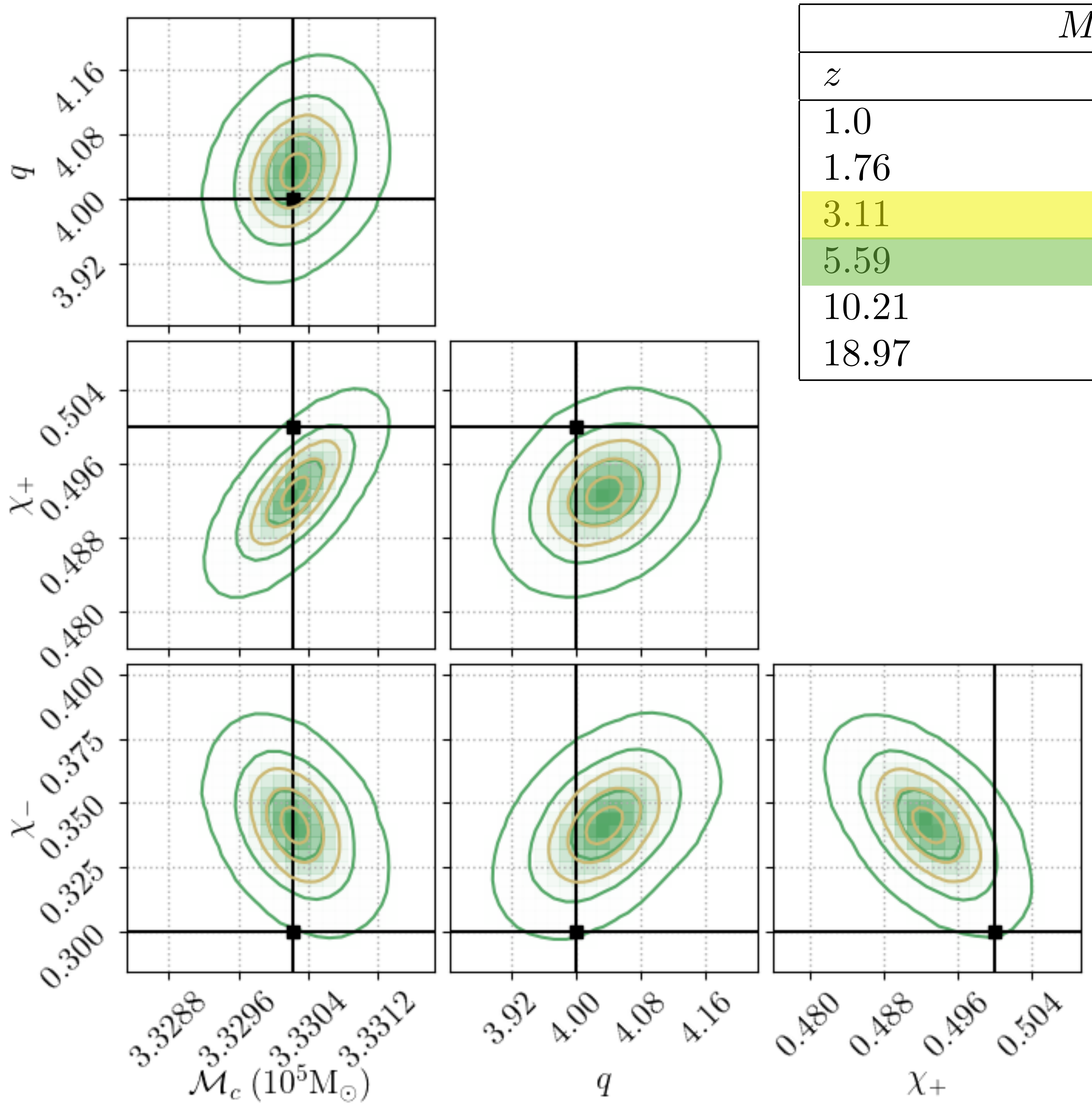
Sky localisation



Example Parameter estimation with systematics II

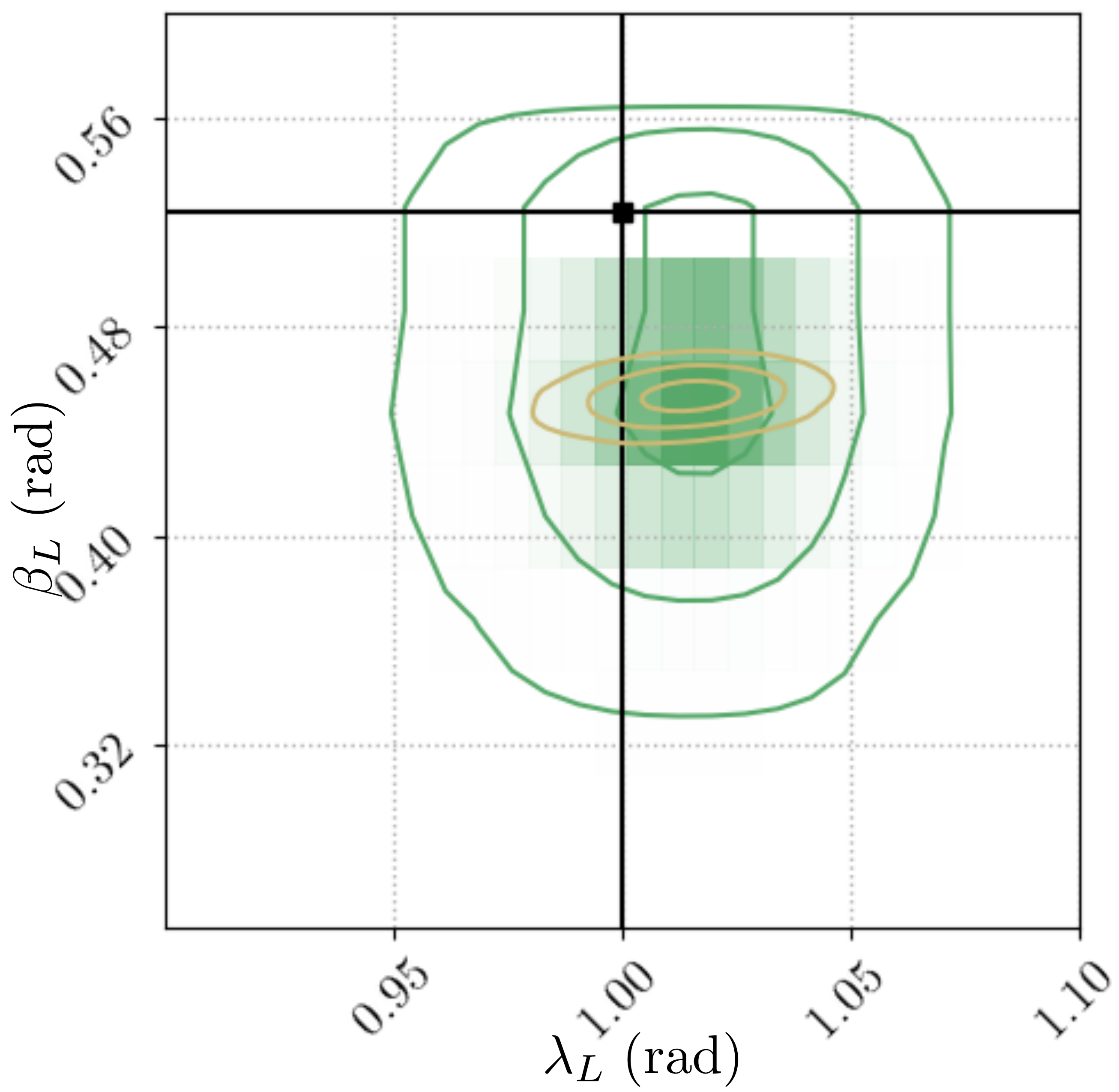
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

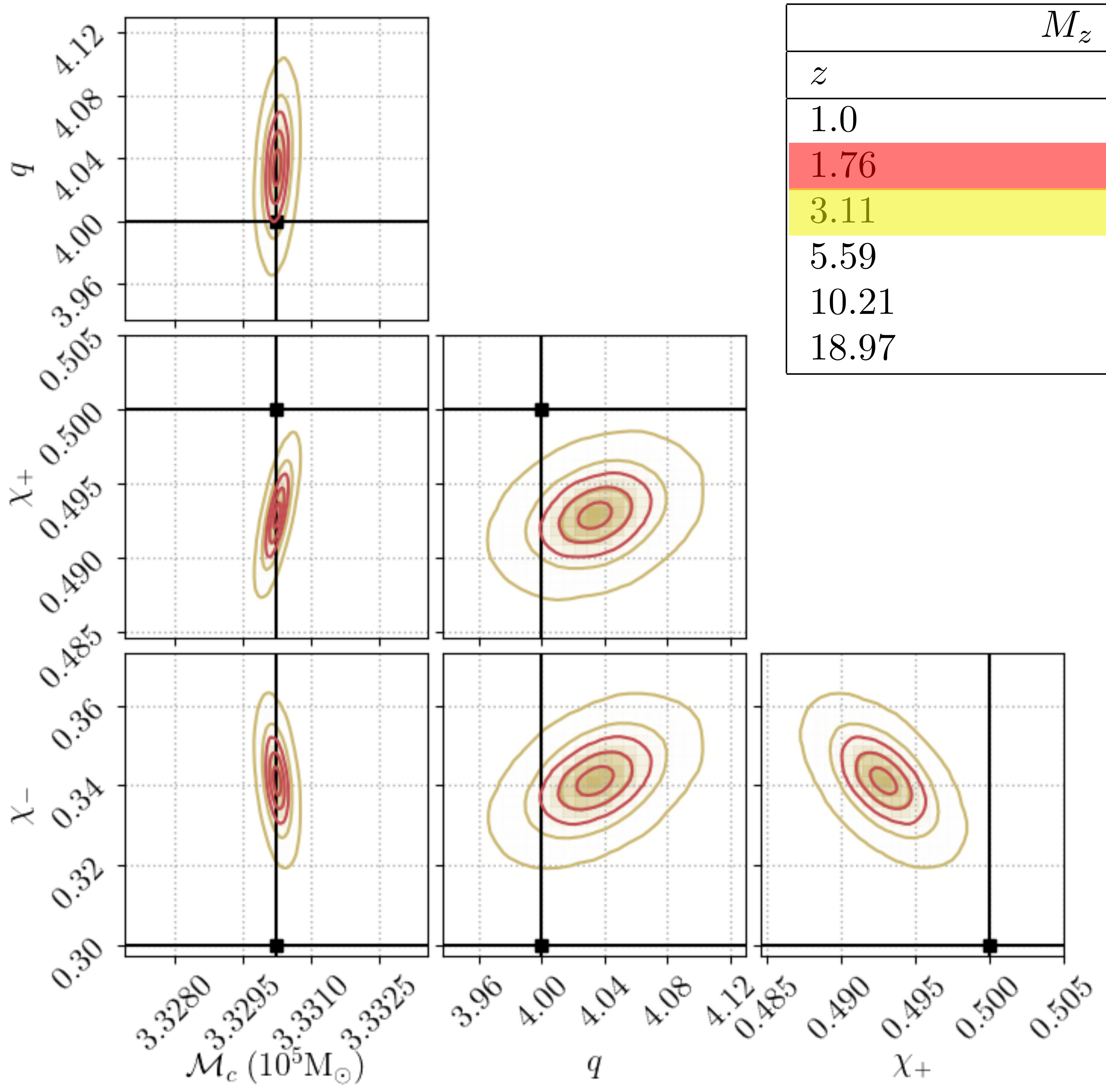
Sky localisation



Example Parameter estimation with systematics II

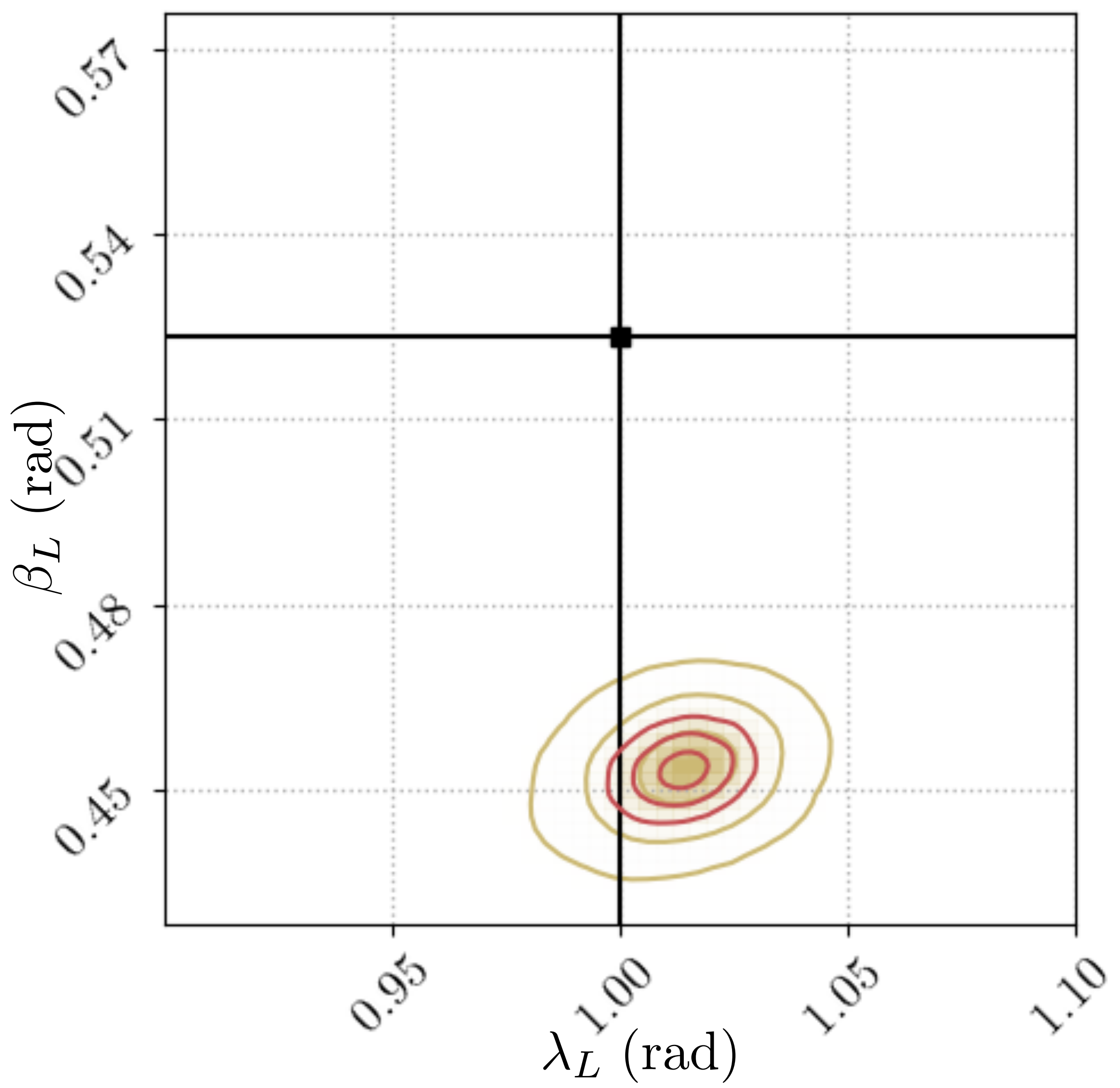
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

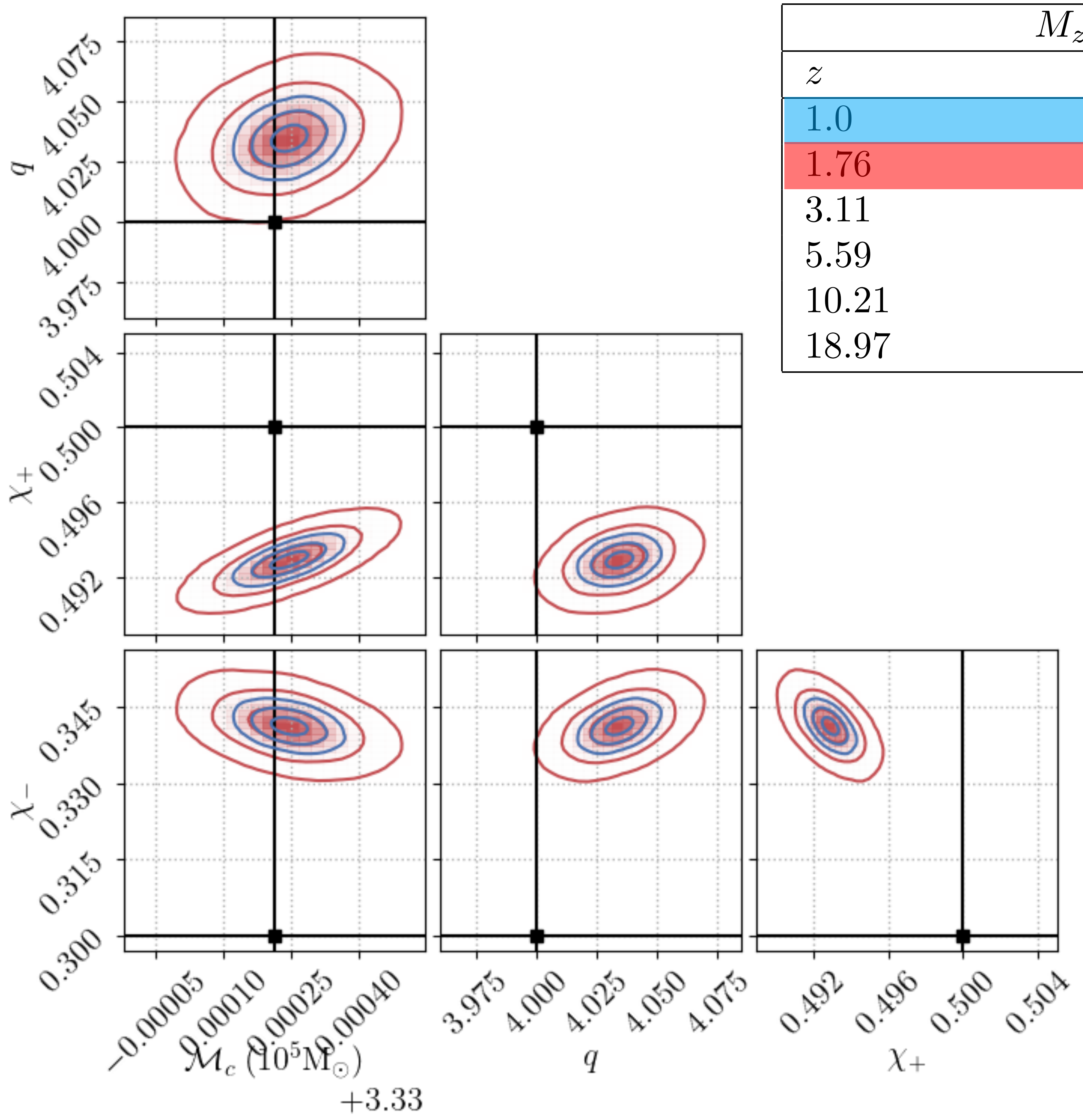
Sky localisation



Example Parameter estimation with systematics II

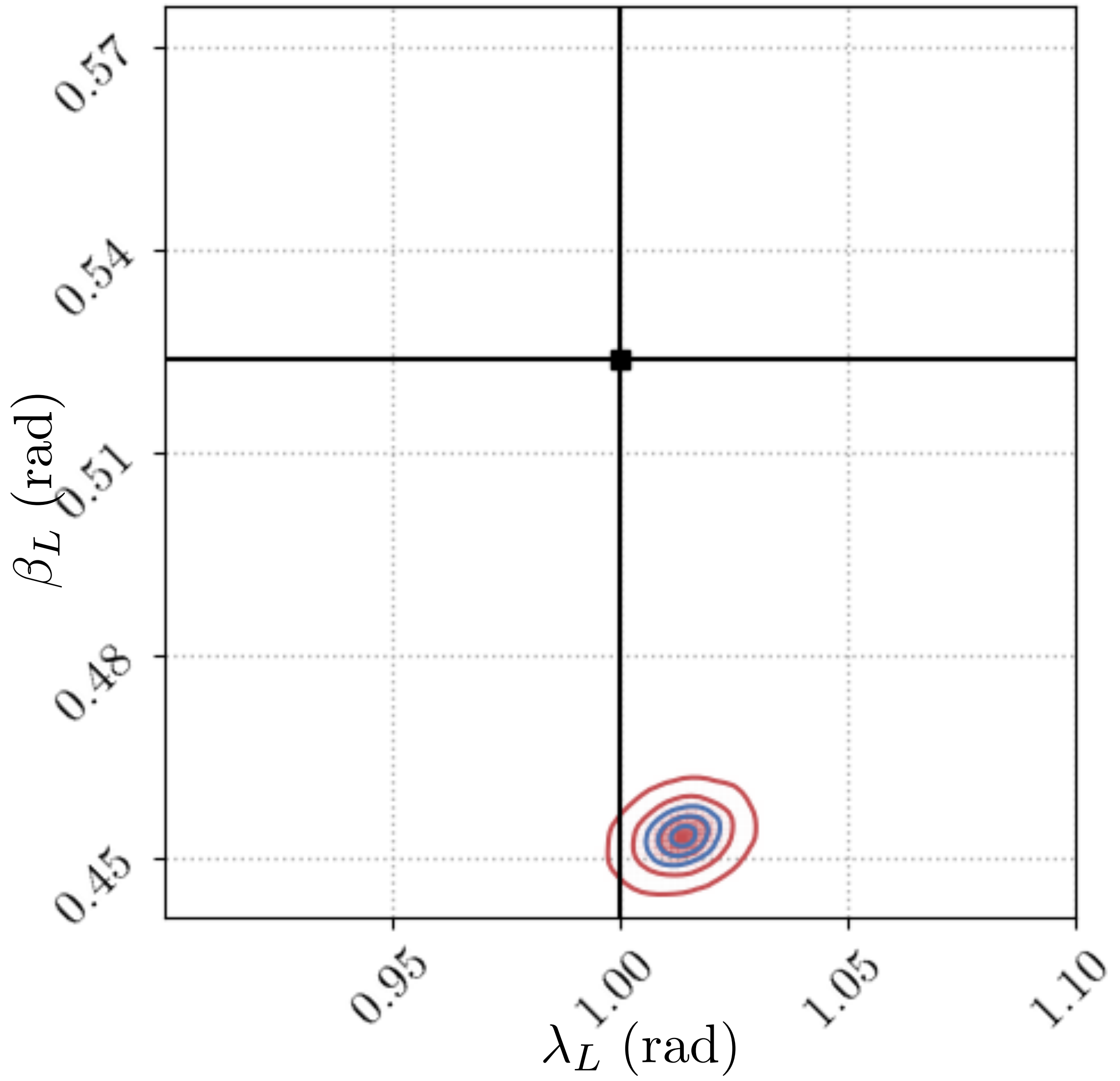
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** PhenomXHM

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

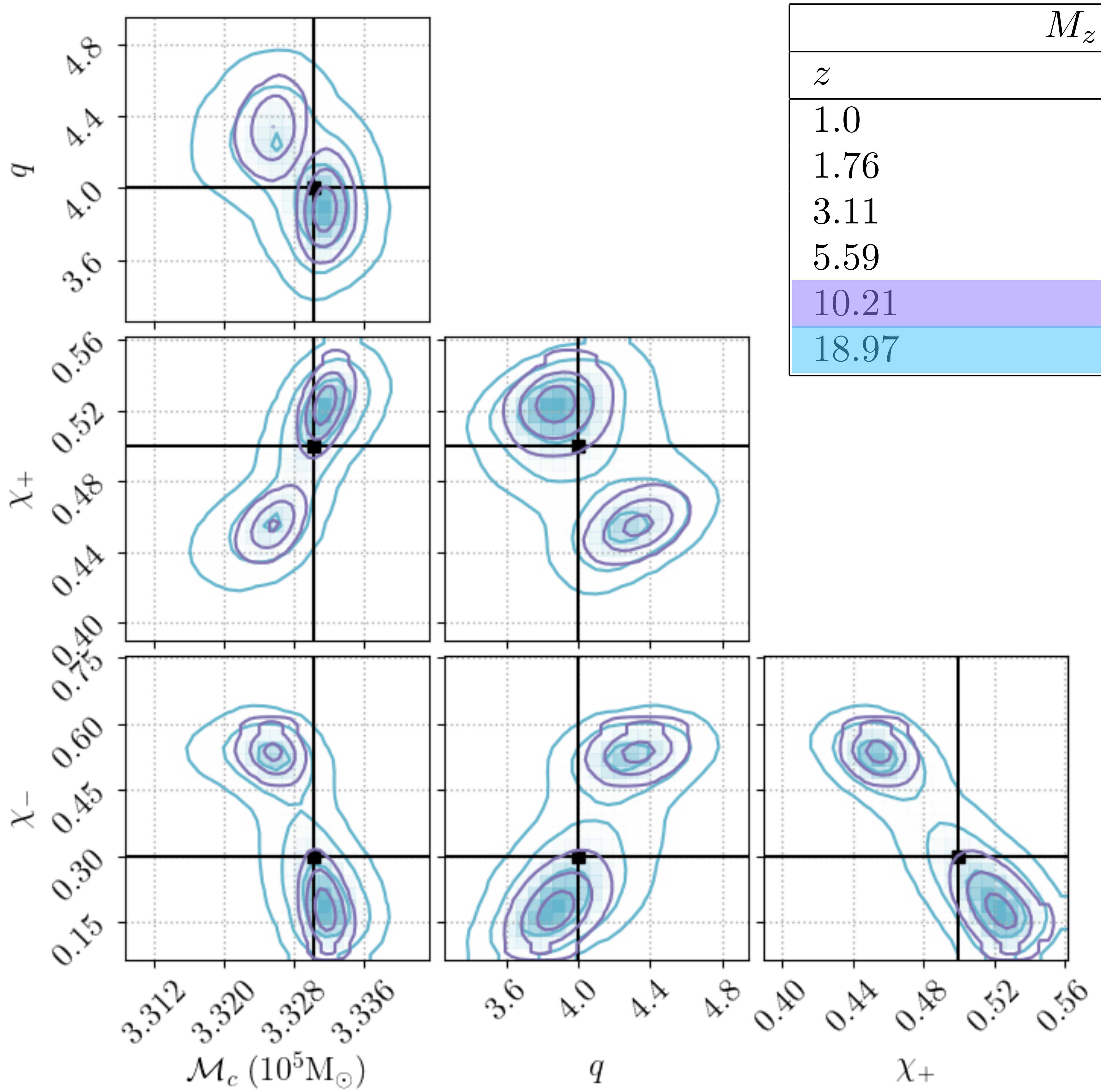


The bad:
- converges 'near' the true parameters
- significant bias at $z = 1, \text{SNR} = 1907$

Example Parameter estimation with systematics III

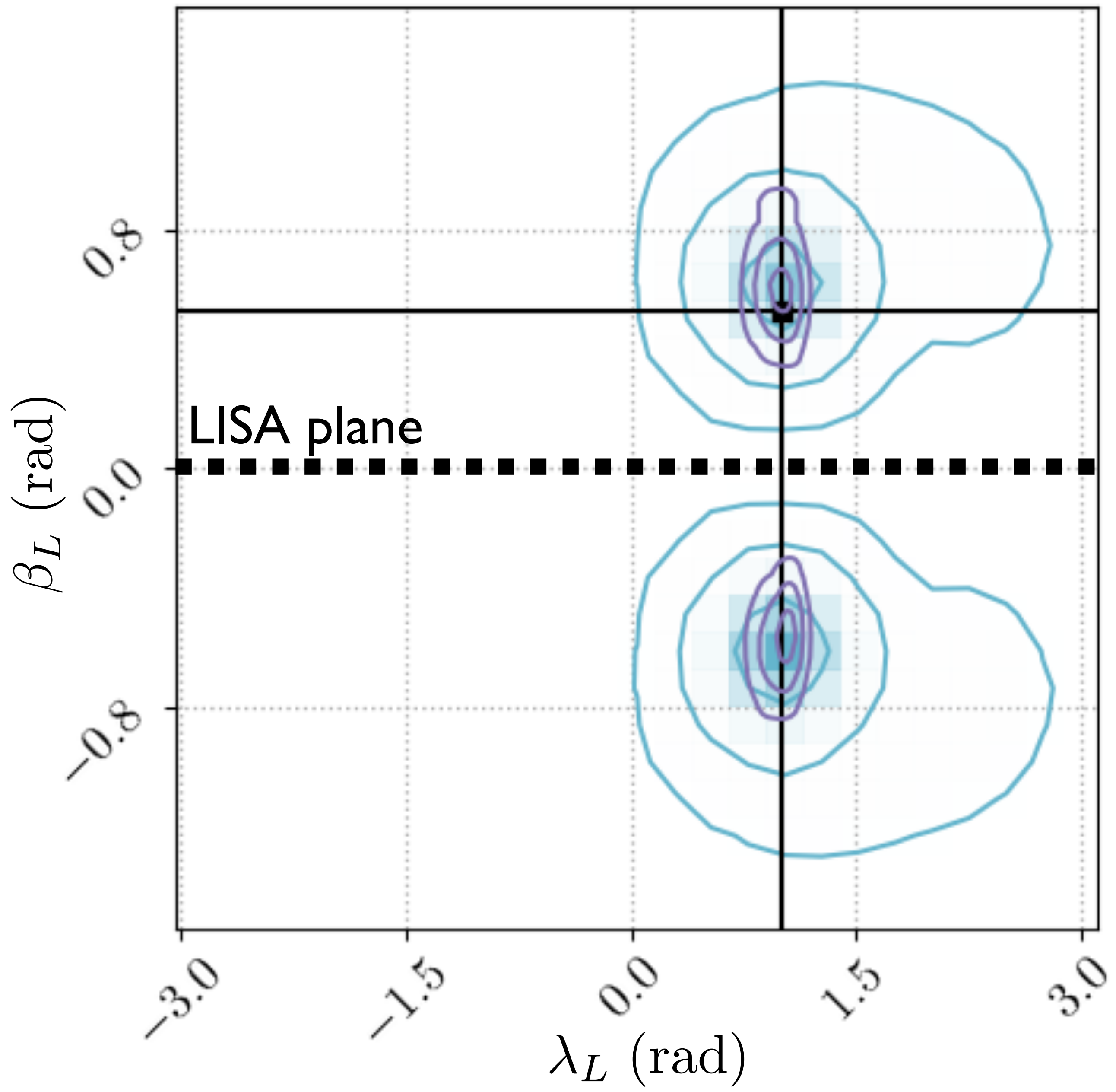
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

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

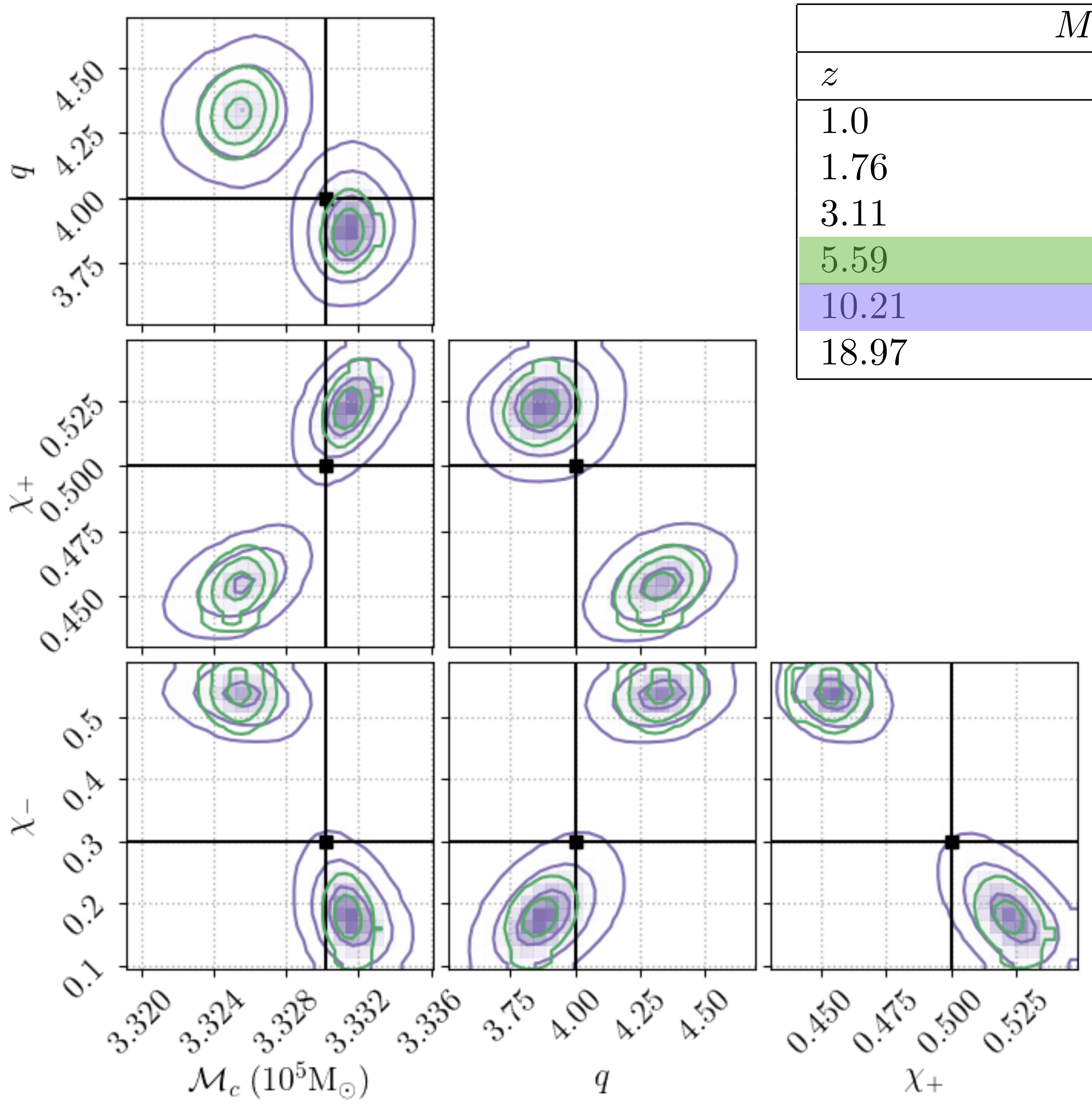
Sky localisation



Example Parameter estimation with systematics III

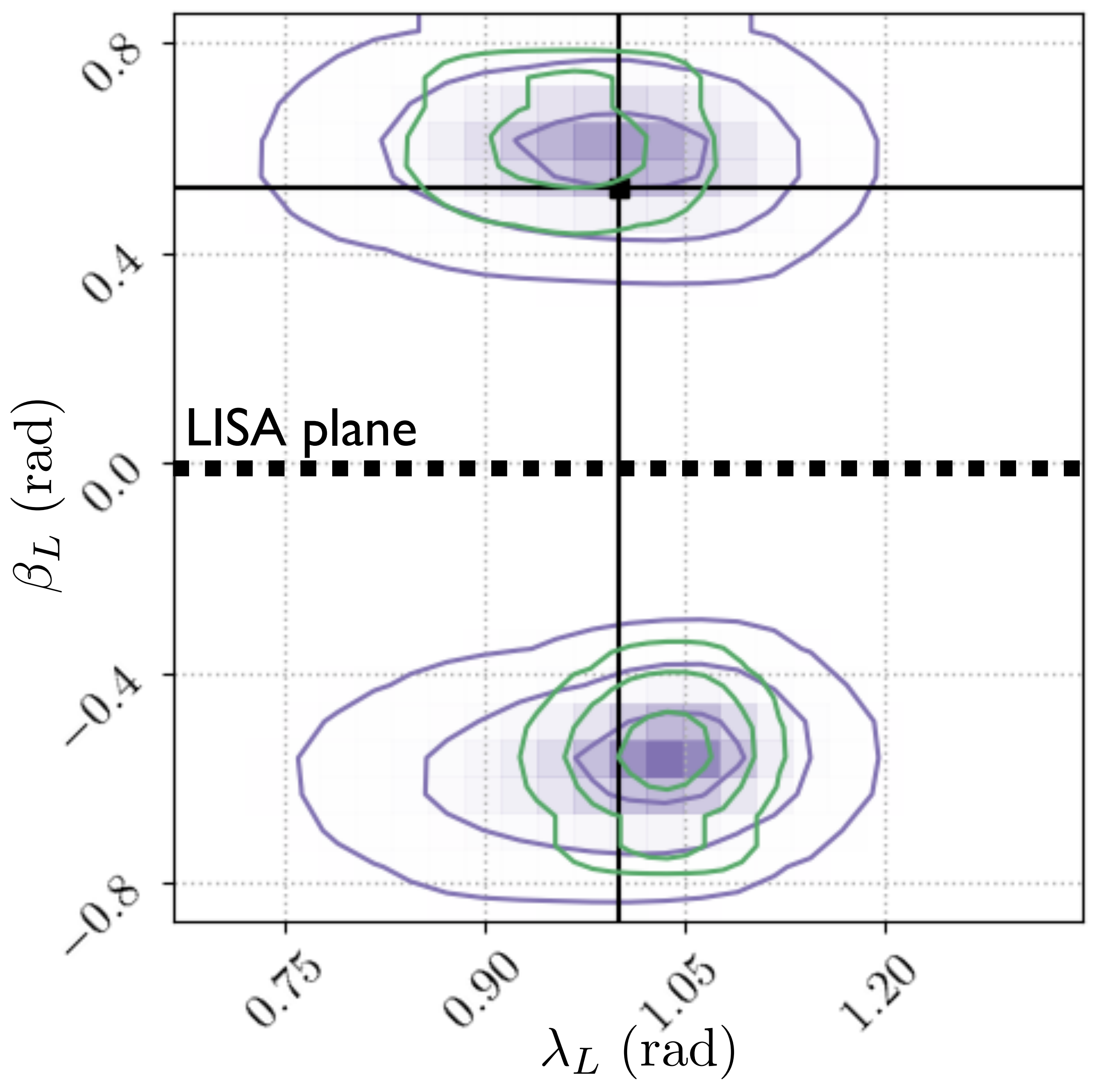
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

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

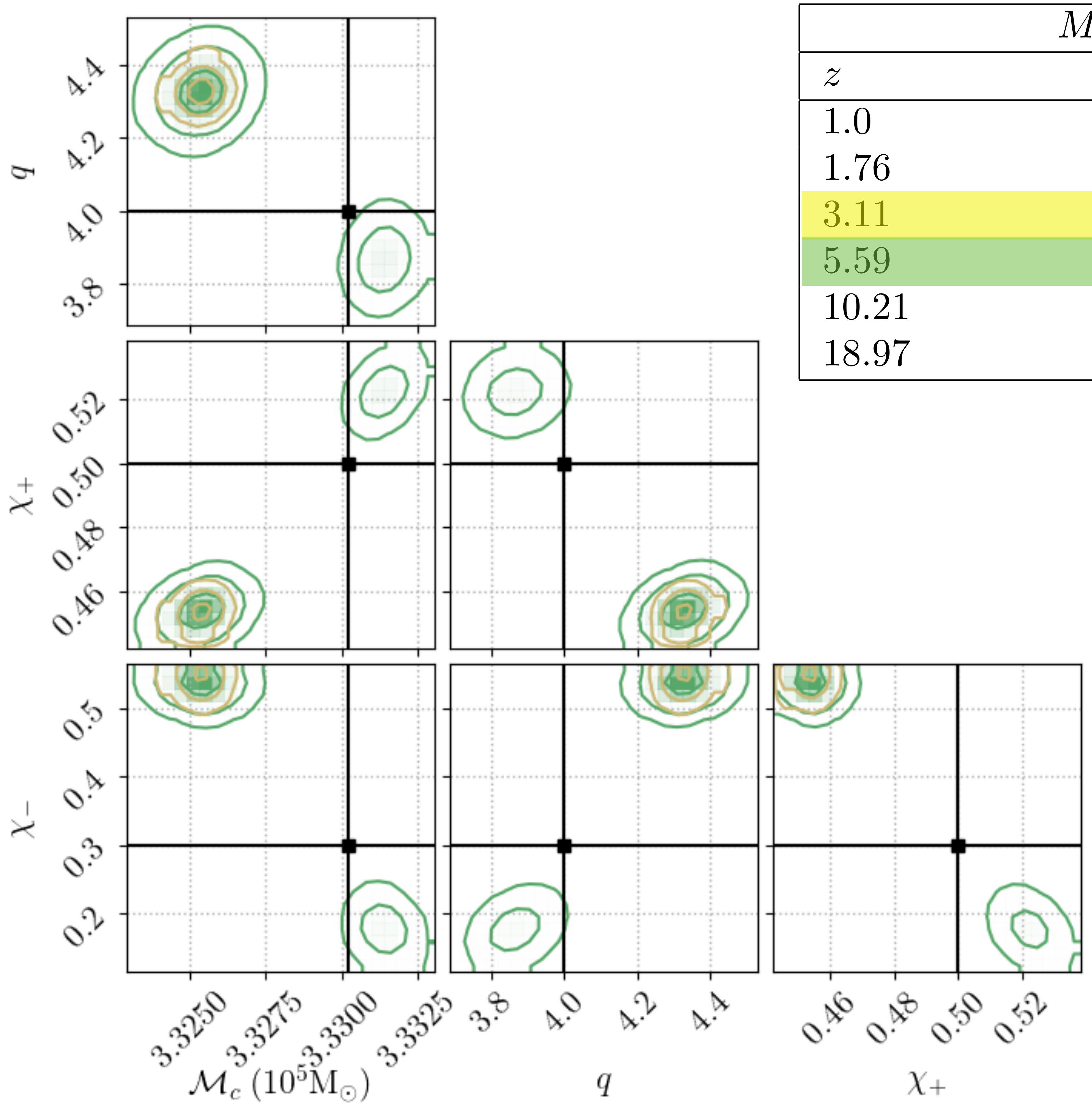
Sky localisation



Example Parameter estimation with systematics III

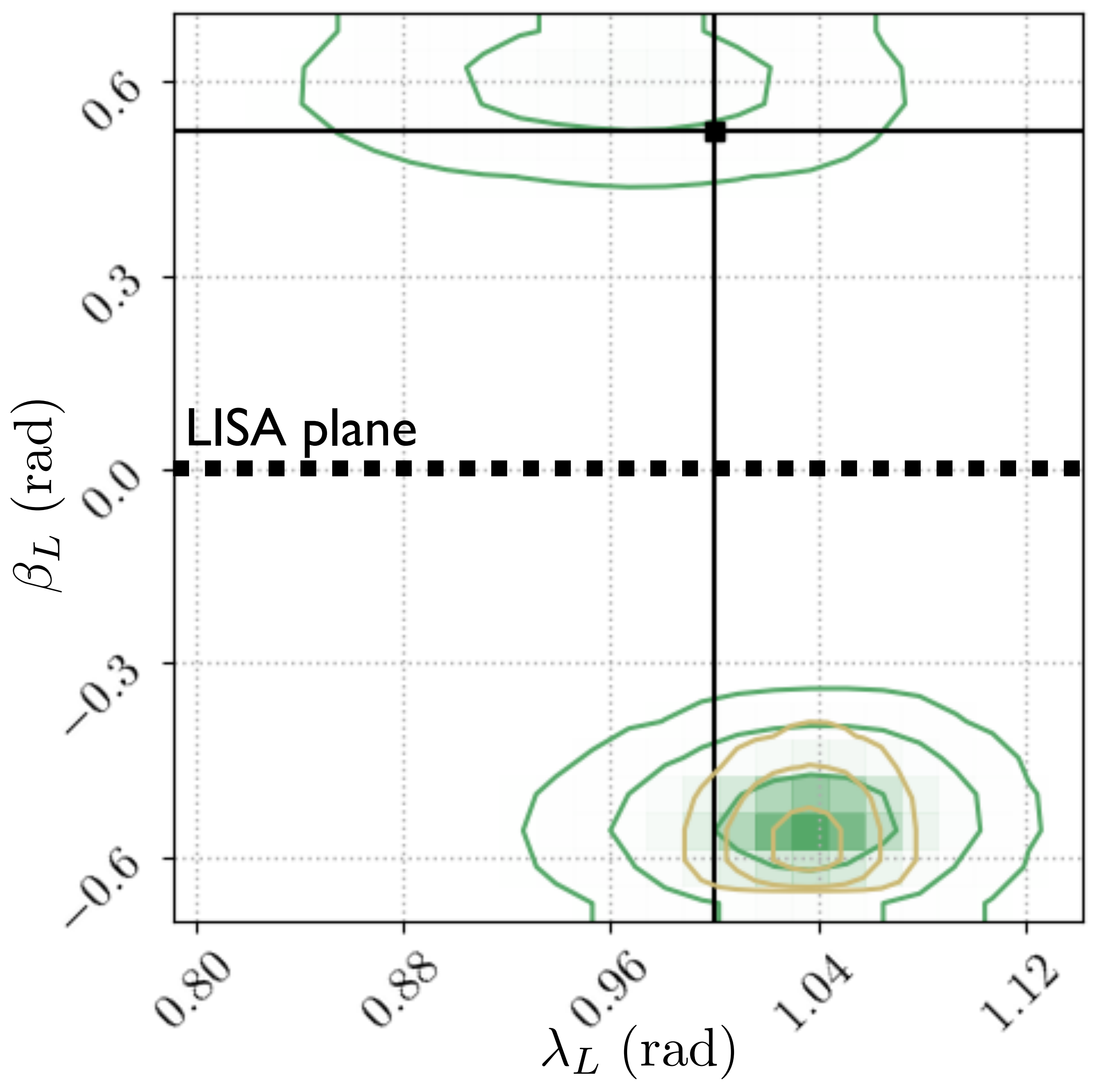
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

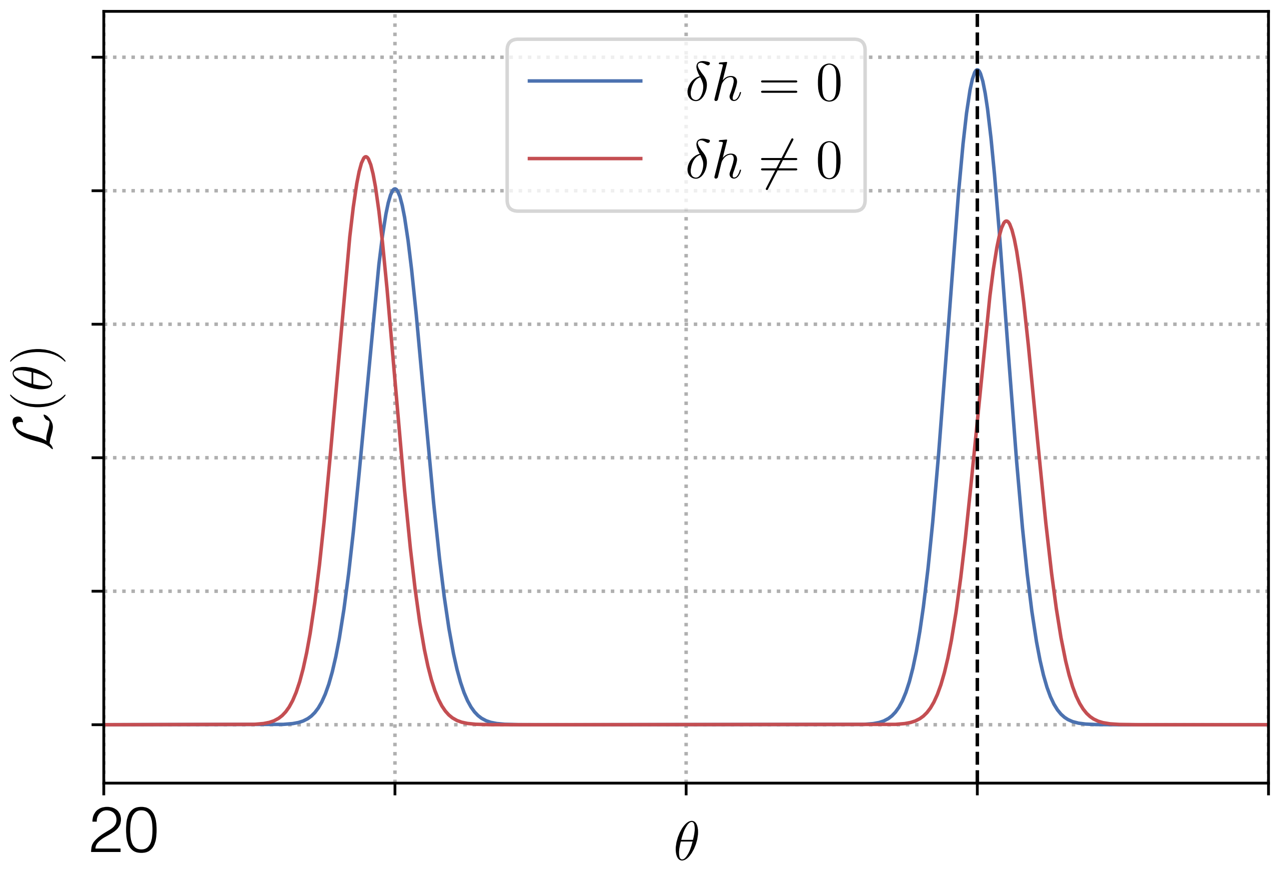
Intrinsic params.



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

Sky localisation



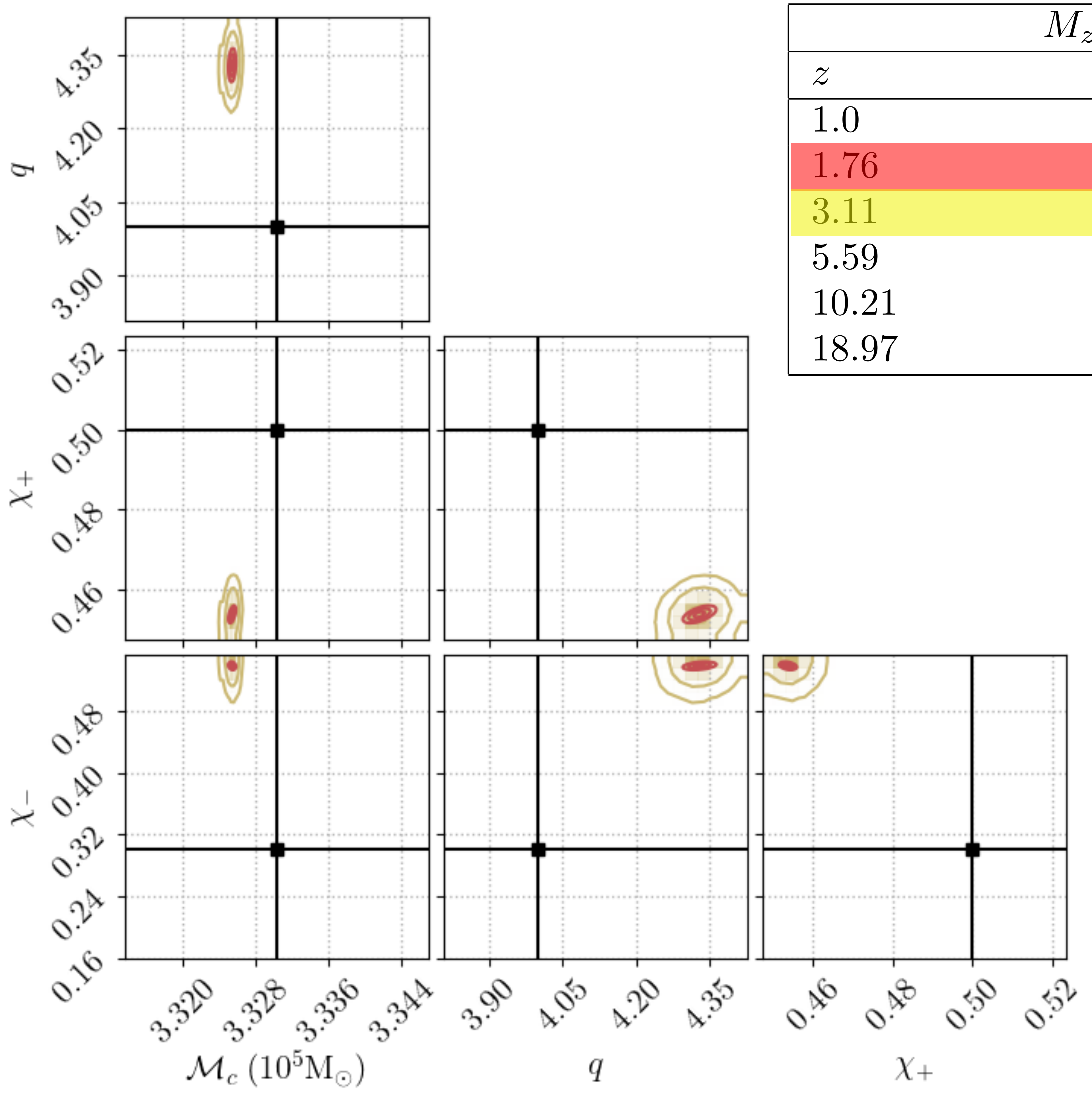


Waveform errors lead to the selection of the wrong sky mode !

Example Parameter estimation with systematics III

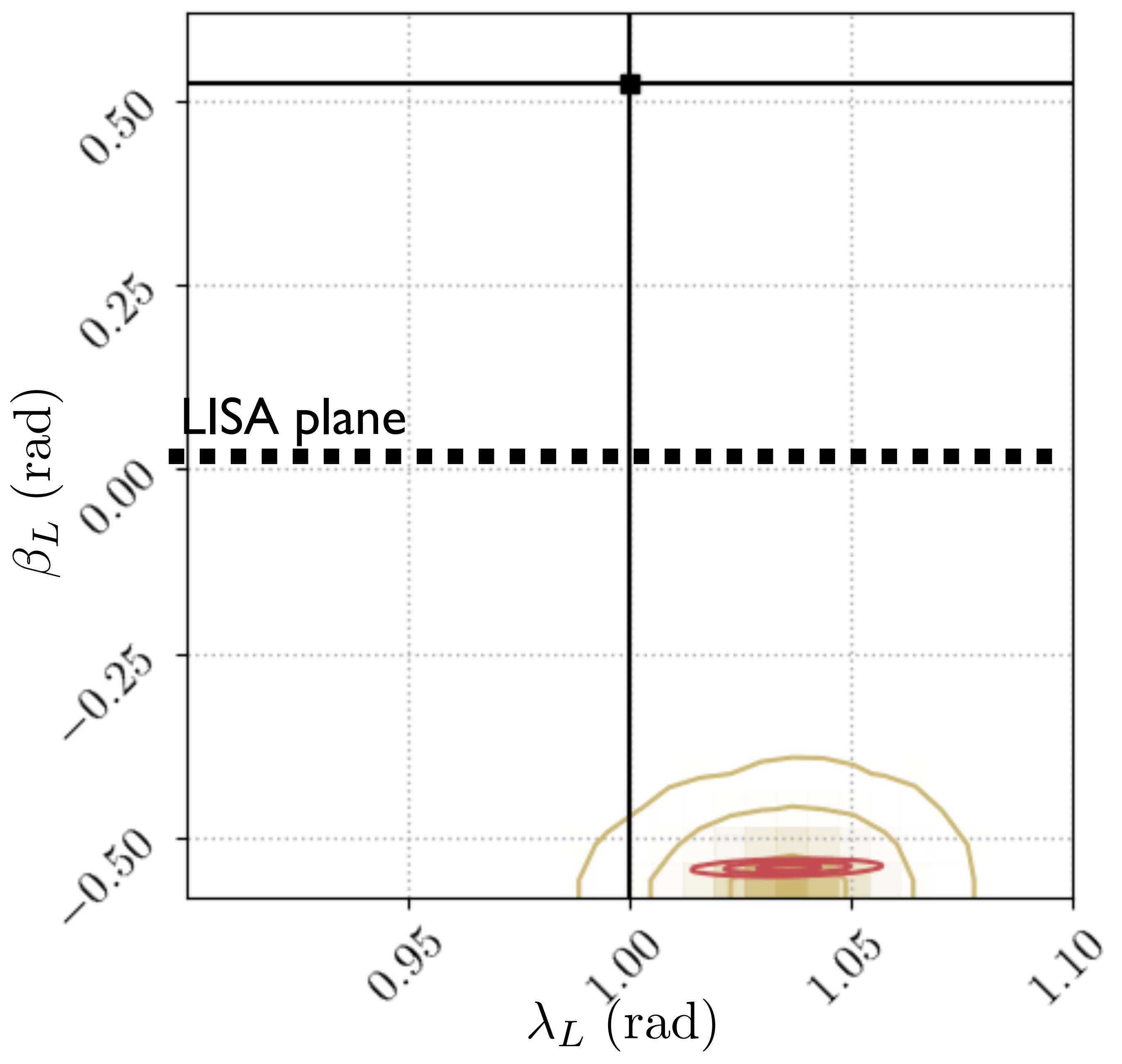
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

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

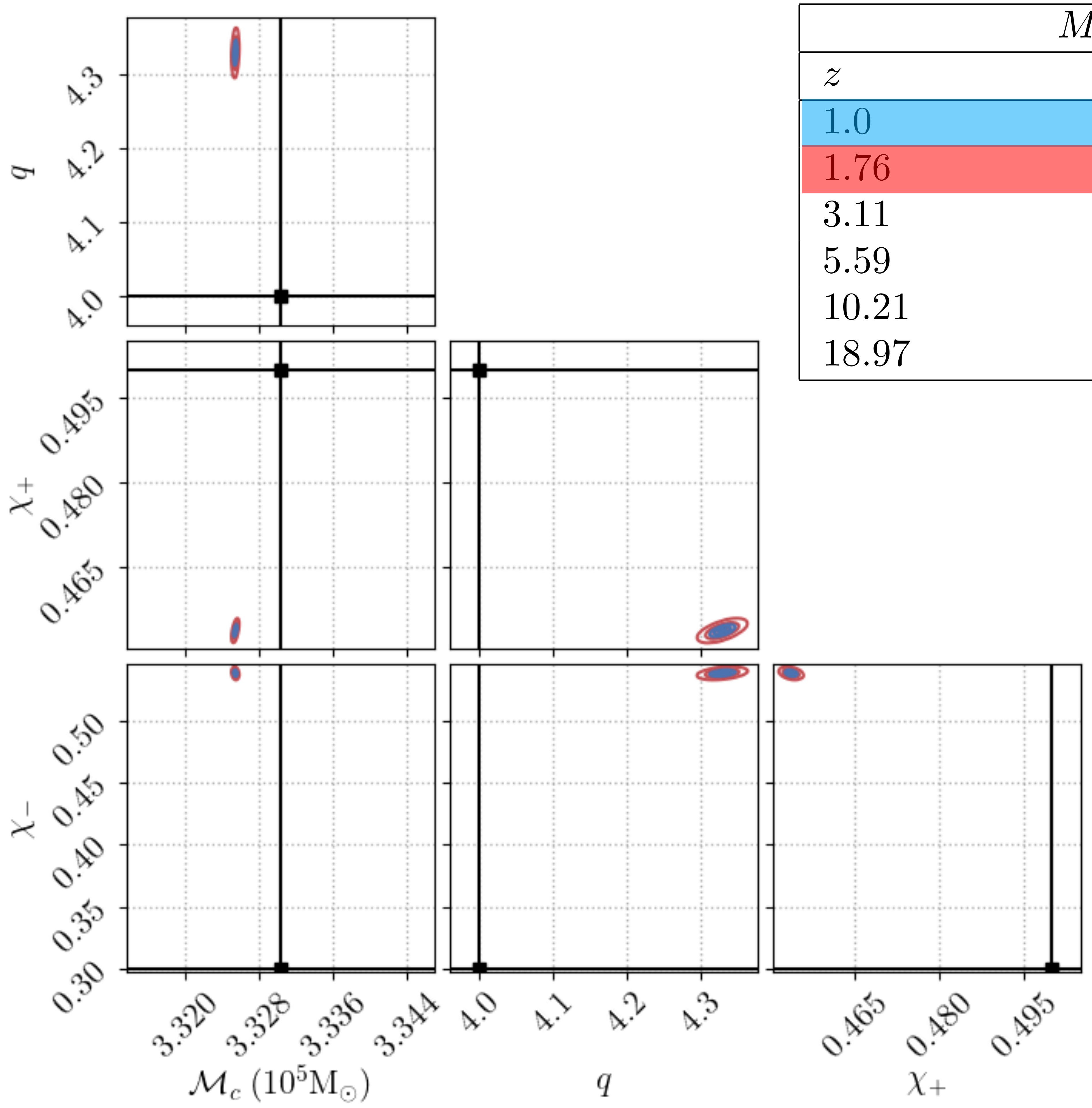
Sky localisation



Example Parameter estimation with systematics III

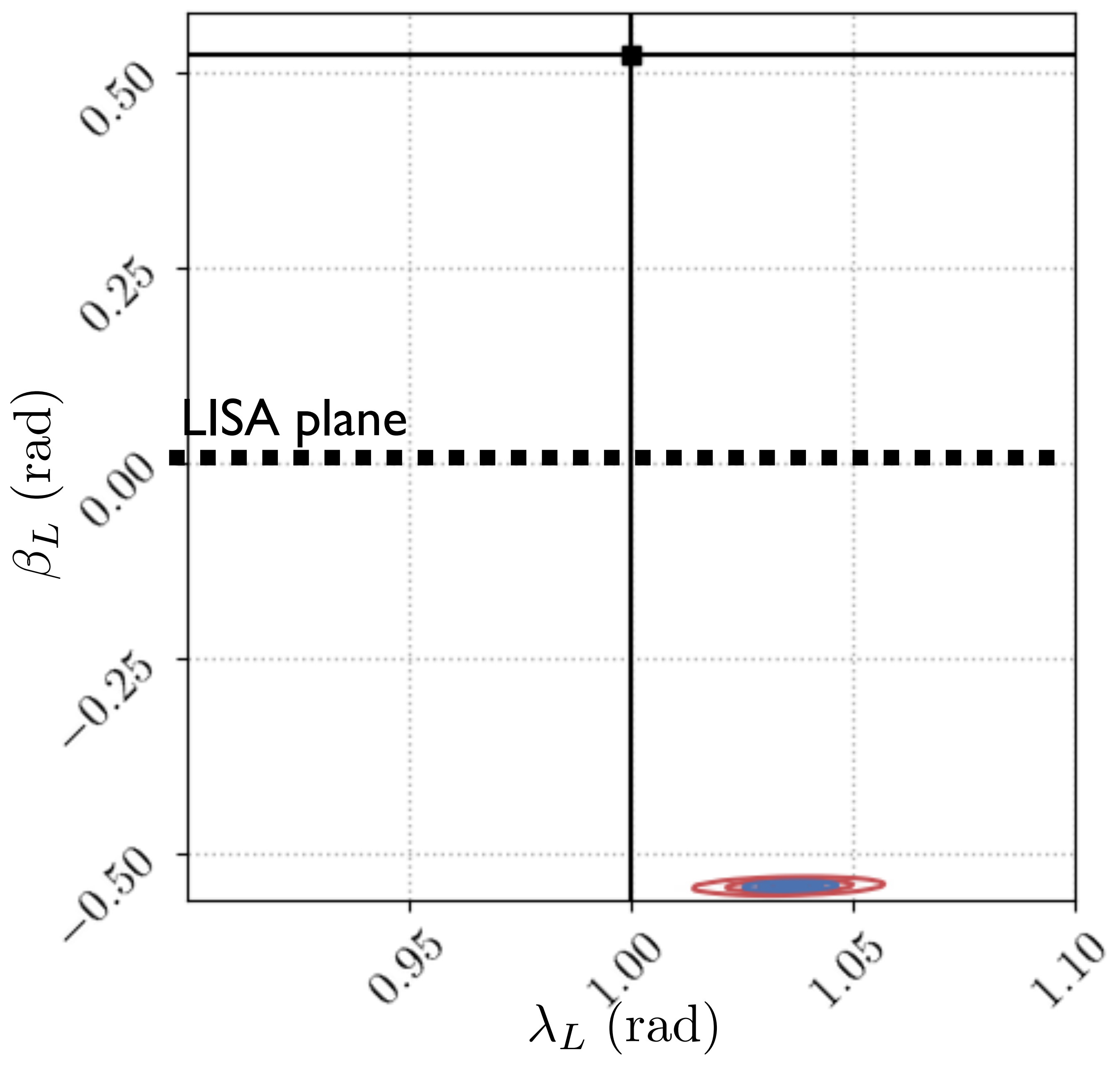
- **Injection:** NRHybSur3dq8 $\{M = 10^6 M_\odot, q = 4, \chi_1 = 0.5, \chi_2 = 0.3\}$
- **Template:** SEOBNRv5HM_ROM

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

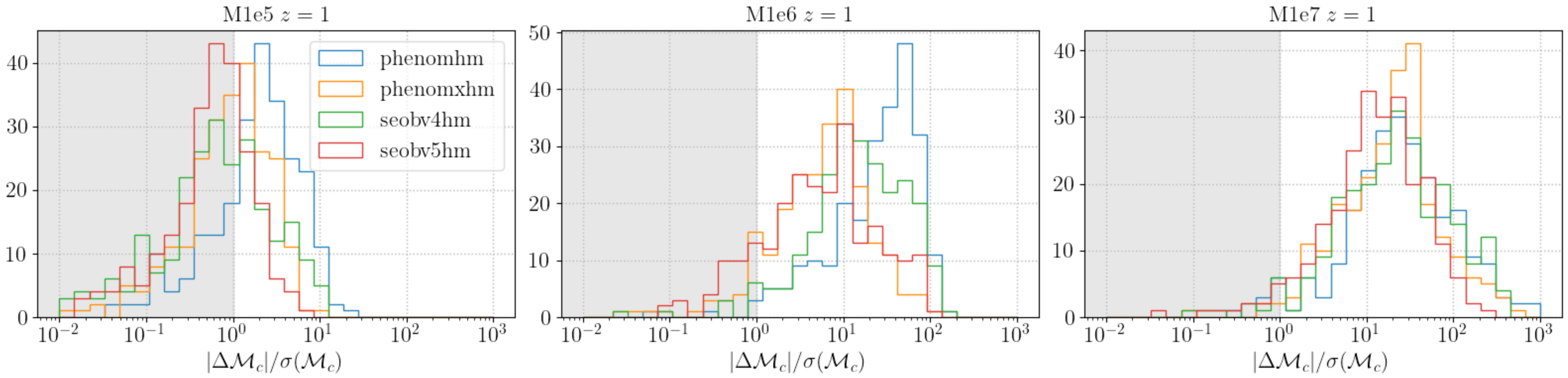
Sky localisation



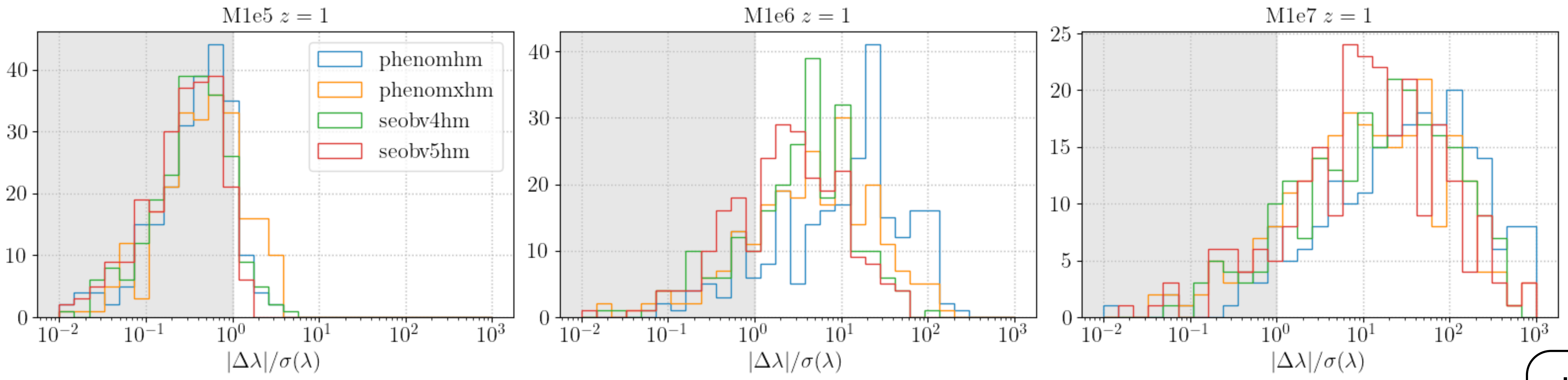
The ugly:
- converges to the **wrong sky mode**
- important bias at $z = 1, \mathrm{SNR} = 1907$

Statistical significance of biases: intrinsic parameters

Bias in chirp mass:



Bias in longitude (on corrected skymode):

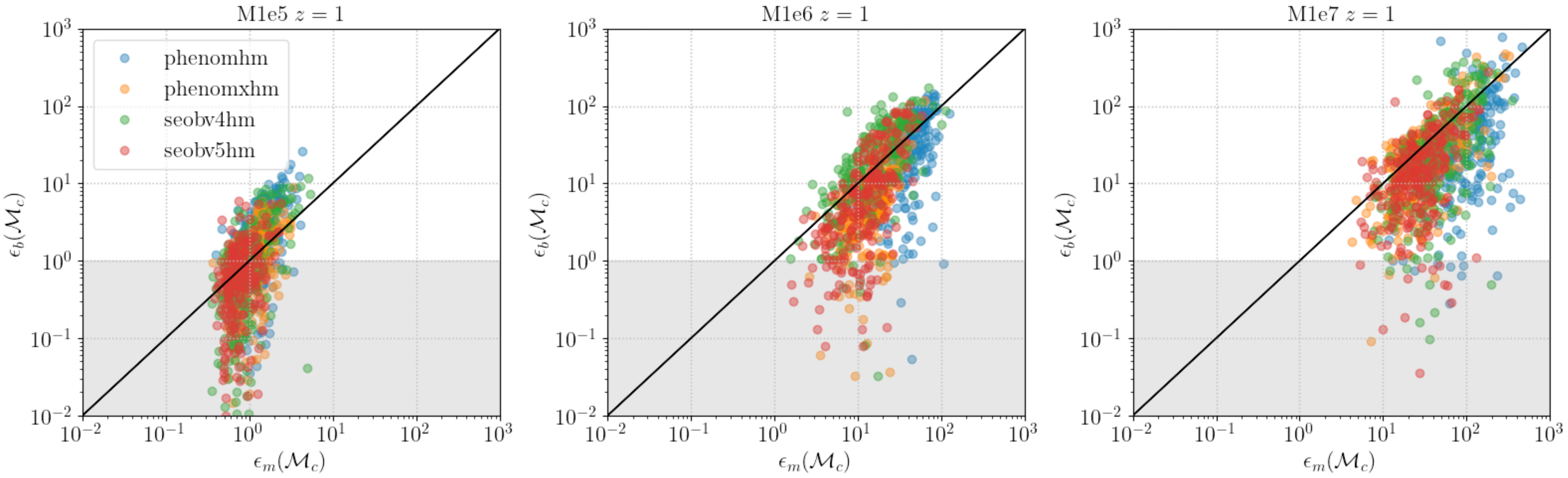


Wrong skymode recovered:

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

Large biases at high mass
Wrong skymodes common

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

Relation between mismatch and bias unclear

Example Parameter estimation with systematics III: Cutler-Vallisneri bias

Linearized biases (Cutler-Vallisneri):

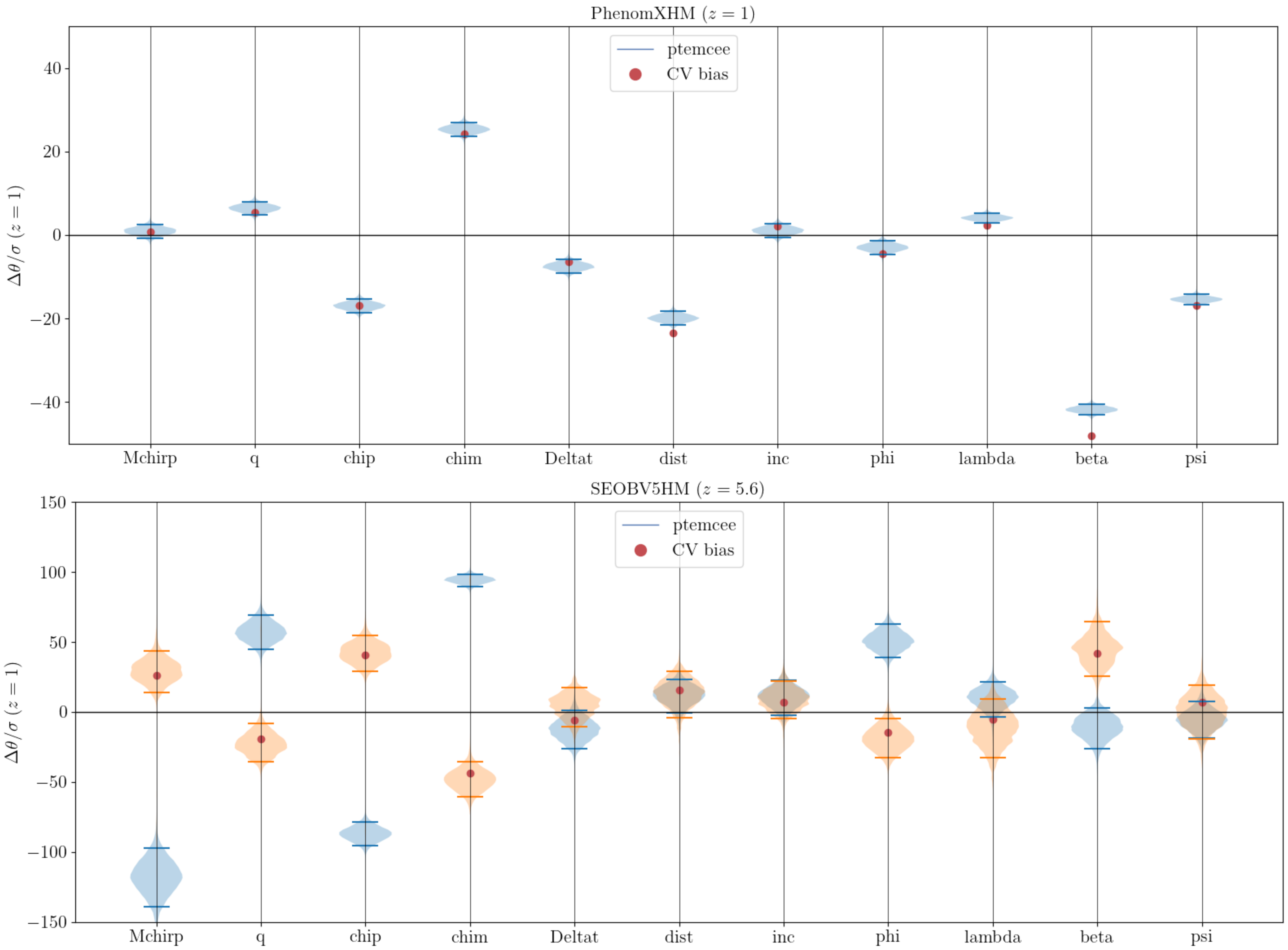
[Flanagan-Hughes 1997]
[Cutler-Vallisneri 2007]

In the linear signal approximation **[local]**, 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
- Fails to capture distant secondary mode



Results and outlook

Results

[Preliminary]

- Unsurprisingly, strong biases for loud LISA signals at $z = 1$
- Strong dependence on mass: at $z = 1$, low-mass much better than high mass
- Noticeable improvement for most recent waveform models
- Degeneracies in the sky position can lead to the wrong sky mode to be selected
- Indications that the distinguishability criterion does not apply well; Cutler-Vallisneri bias esimates can work but do not capture multimodality

Outlook

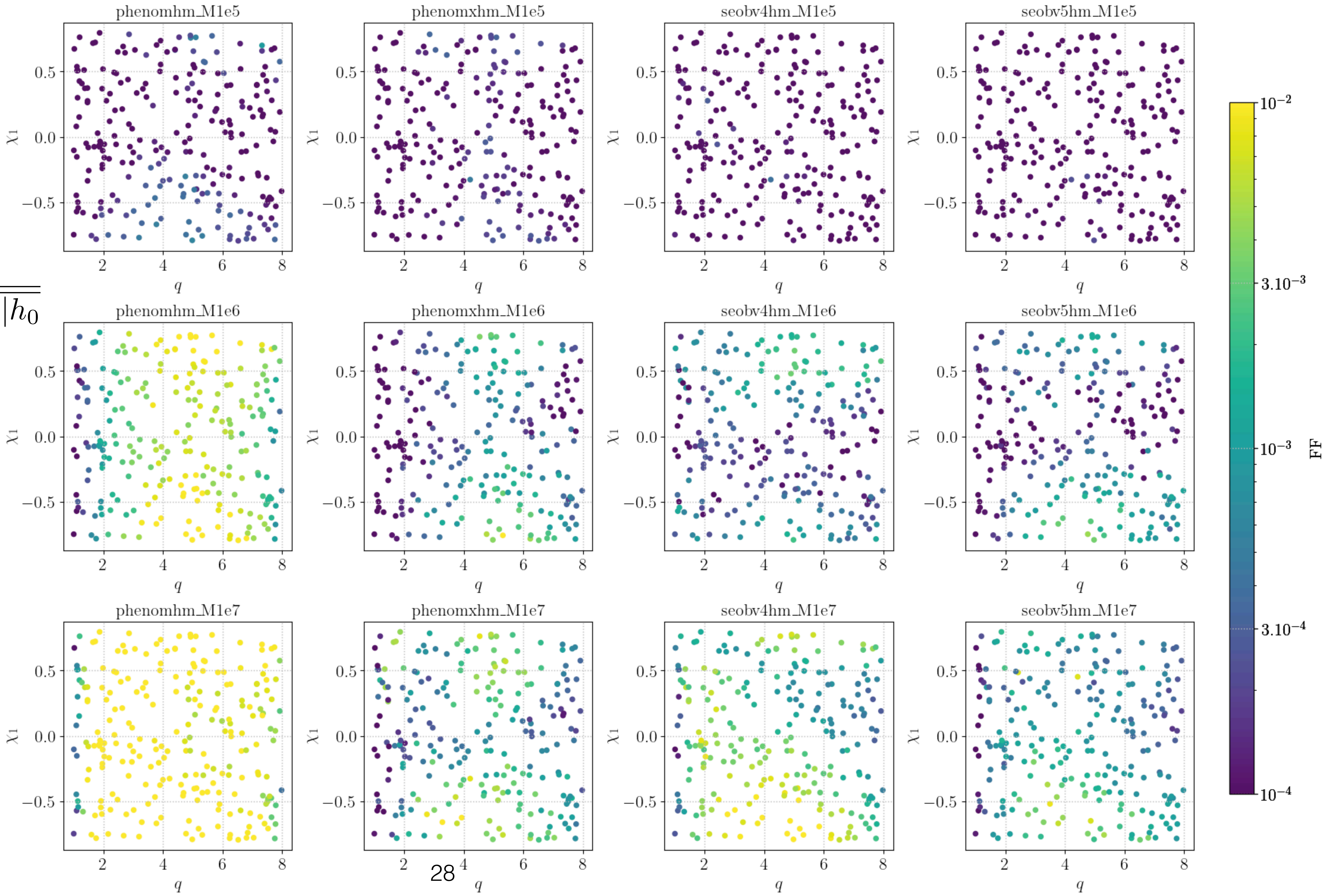
- Compare to LVK/3G studies
- Focus on high masses: NR waveforms long enough ?
- More realistic waveforms: precession (and eccentricity)

Fitting factors in parameter space

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

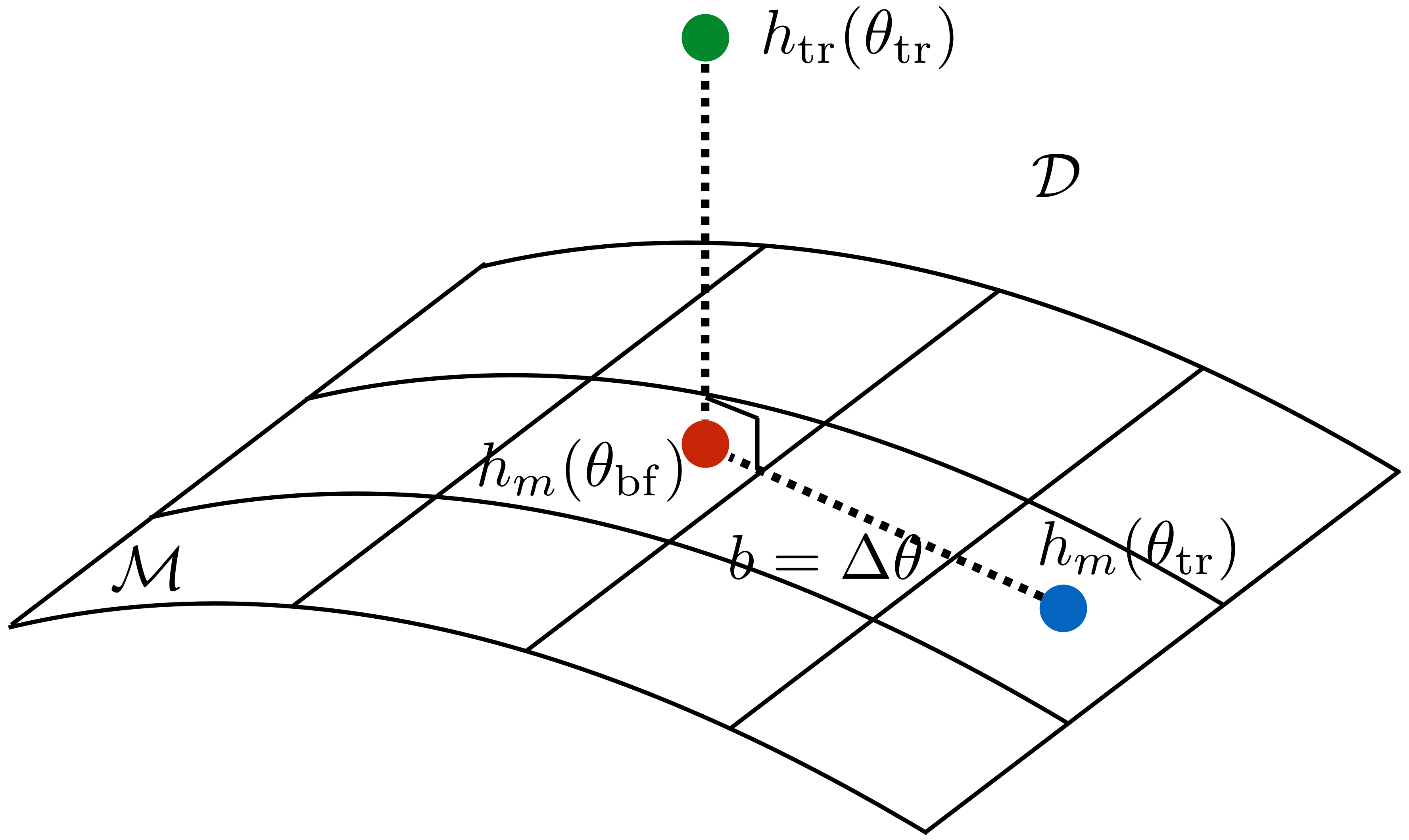
$$\mathrm{FF} = 1 - \max_{\boldsymbol{\theta}} \frac{(h_m|h_0)}{\sqrt{(h_m|h_m)}\sqrt{h_0|h_0}}$$

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



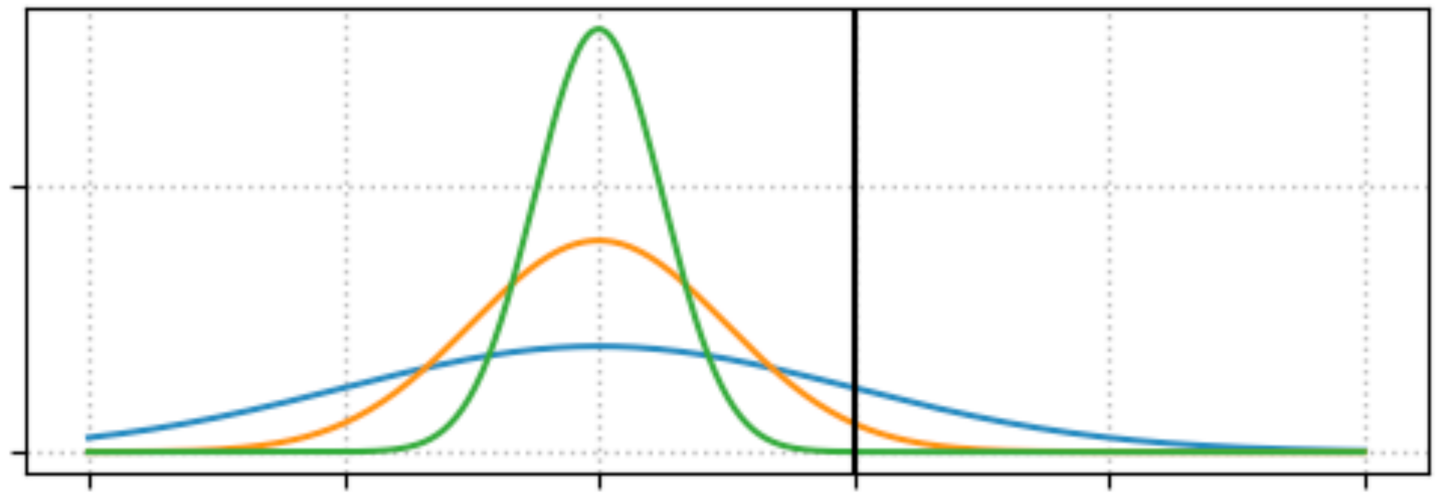
Trend: strong dependence on M
Trend: larger errors at large spins
Trend in q ?

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

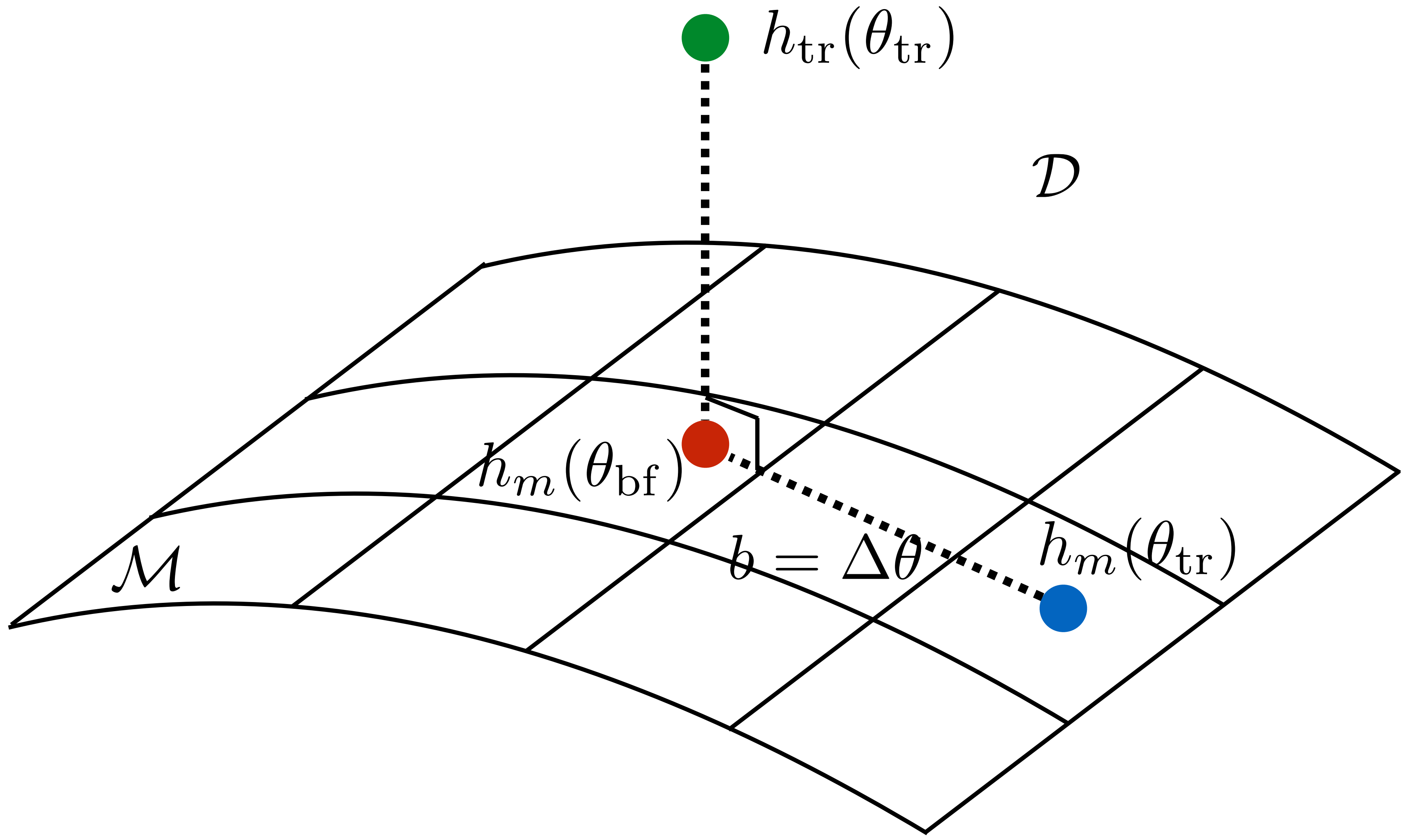


Different questions for systematics:

Are current waveform models accurate enough for LISA ?
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 global fit ?
- How will tests of GR be affected ?
- Can we mitigate the biases ? Computational aspects for low-latency ?

Waveform systematics and parameter estimation



Mismatch (unfaithfulness):

Mismatch, optimization over time/phase/polarization:

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

- 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]**:

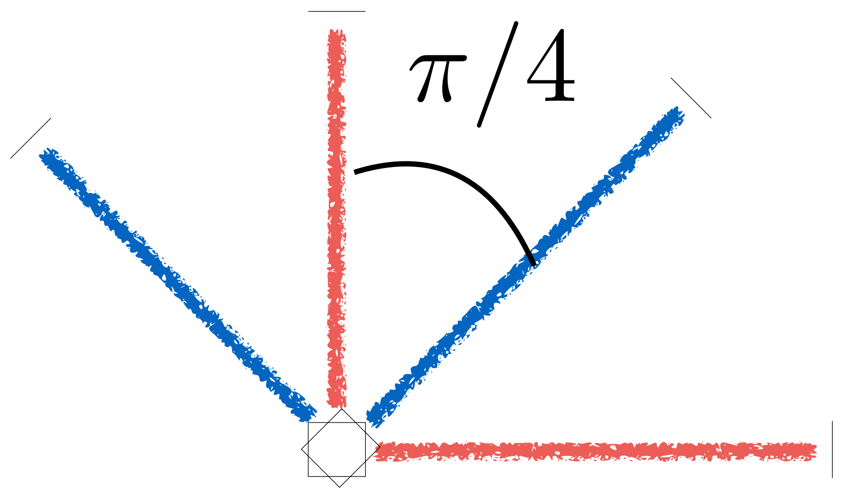
$$F_{ij} = (\partial_i h|\partial_j h)$$

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

Can we assess biases with efficient tools ?

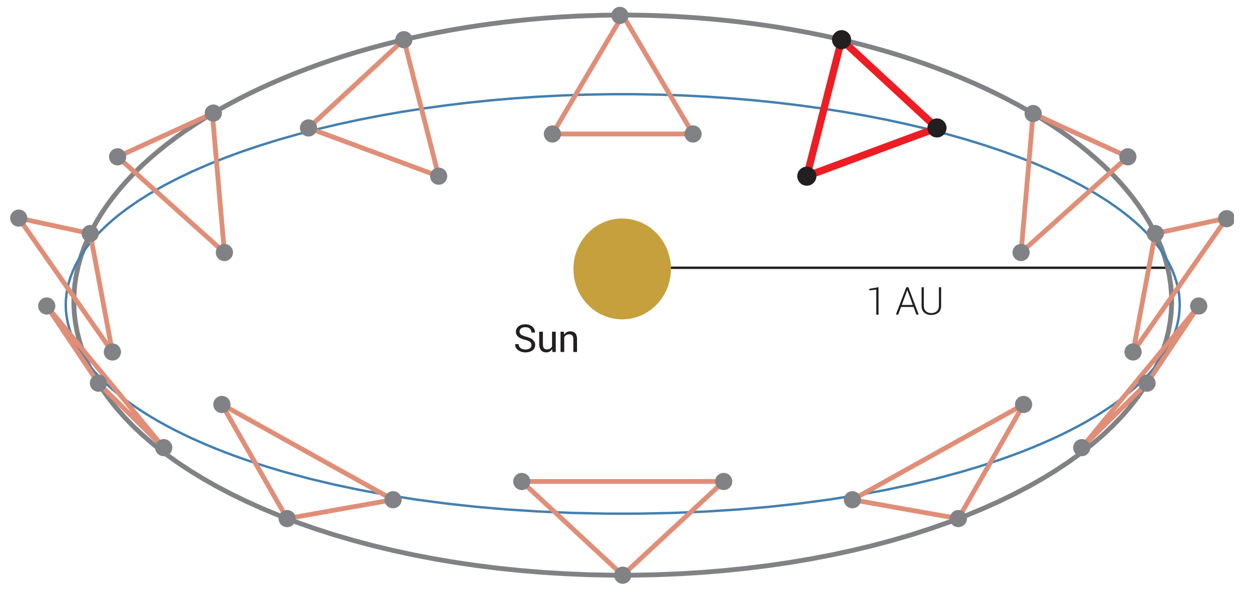
The LISA instrumental response

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:

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

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

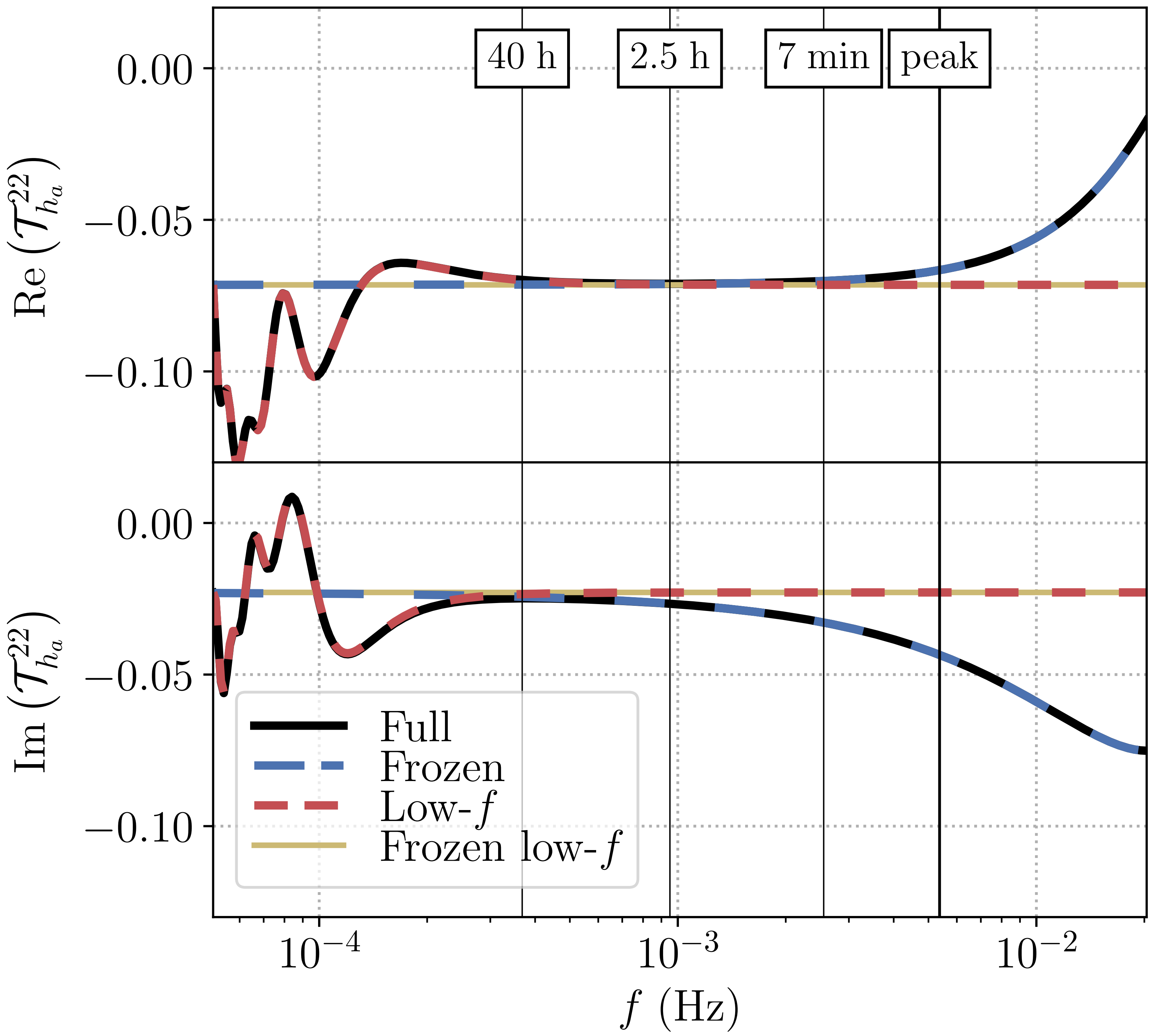
Time and frequency-dependency in transfer functions
Time: motion of LISA on its orbit
Frequency: departure from long-wavelength approx.

Pattern function response with HM:

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

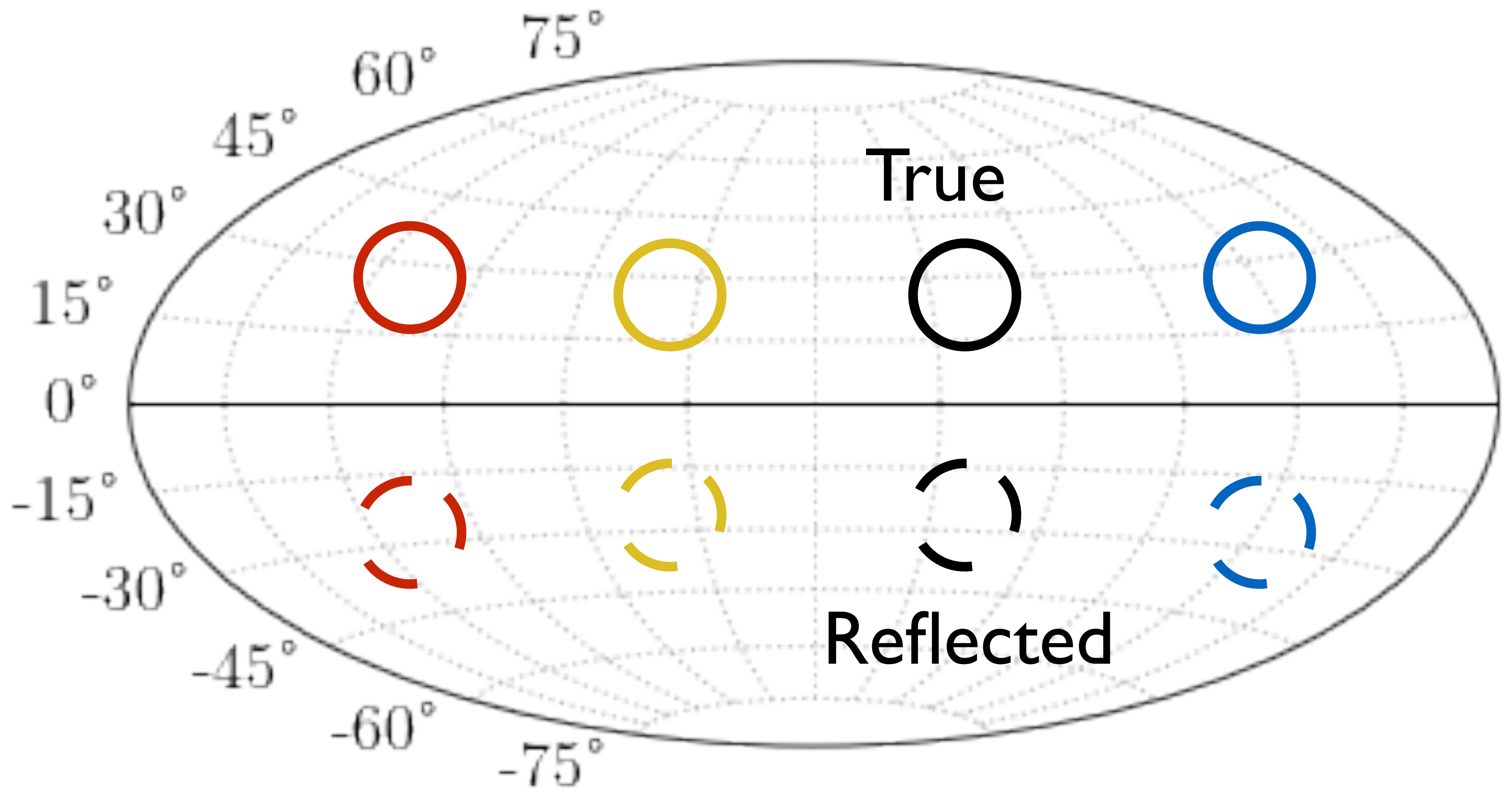
$$F^{\ell m}_{a,e} = \frac{1}{2} {}_{-2}Y_{\ell m} e^{-2i\psi_L} \left(F^{+}_{a,e} + iF^{\times}_{a,e}\right) + \frac{1}{2} (-1)^{\ell} {}_{-2}Y^{*}_{\ell,-m} e^{+2i\psi_L} \left(F^{+}_{a,e} - iF^{\times}_{a,e}\right)$$

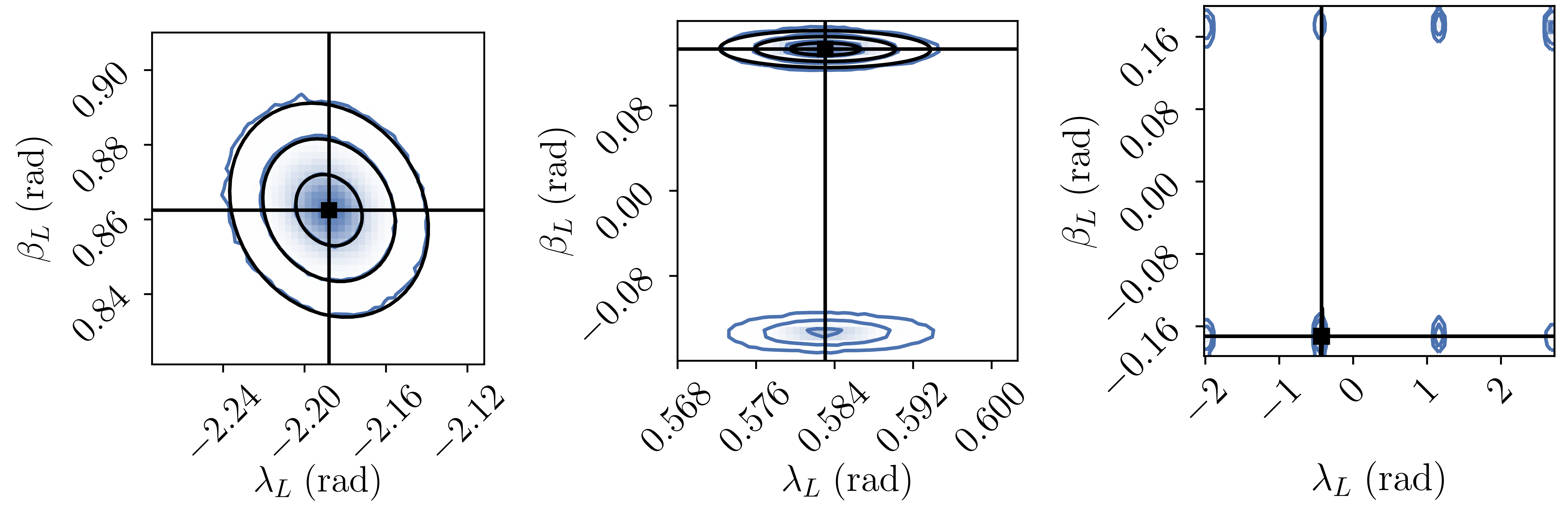
Example FD transfer function



LISA response and multimodality in the sky

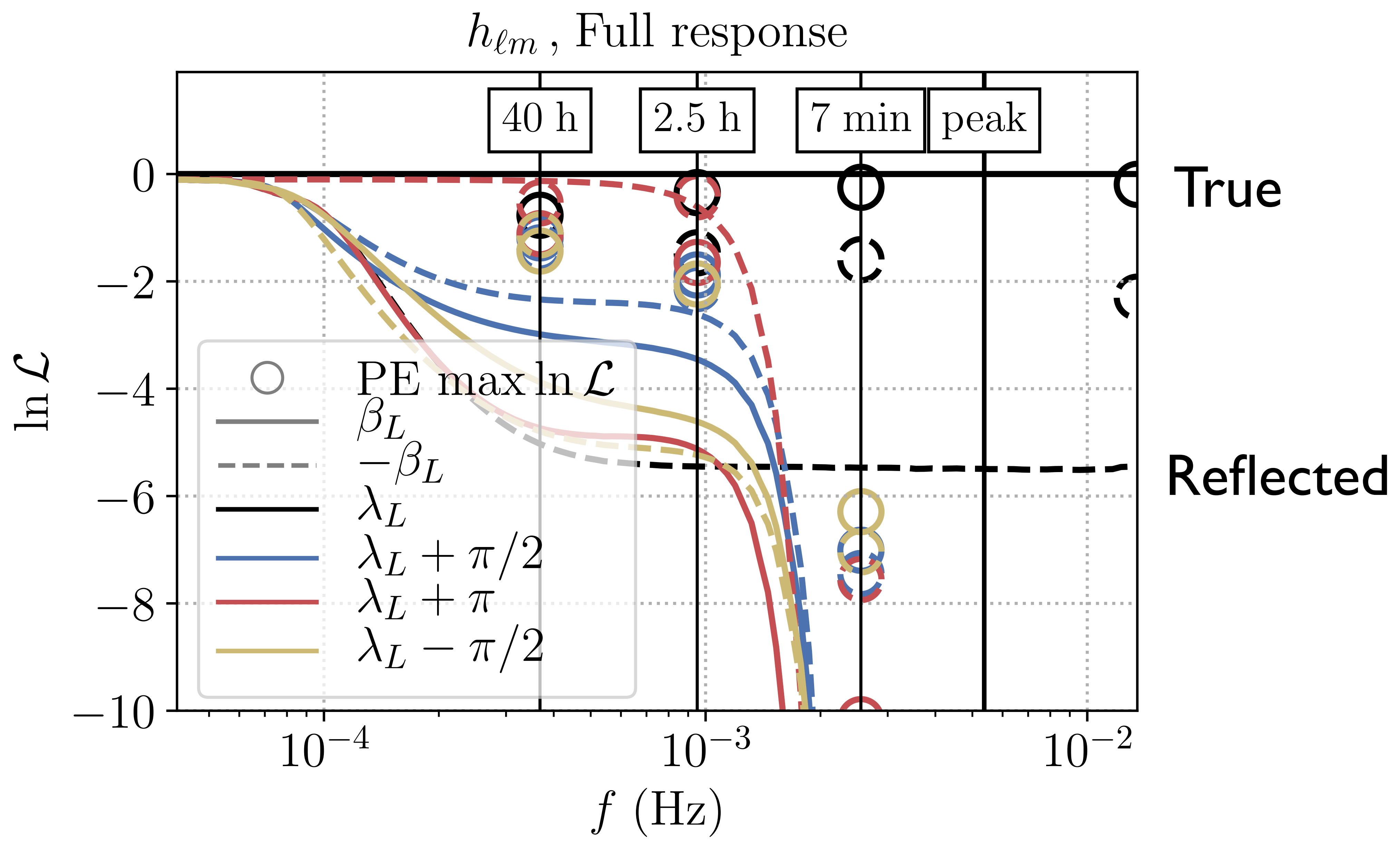
Multimodality pattern:





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

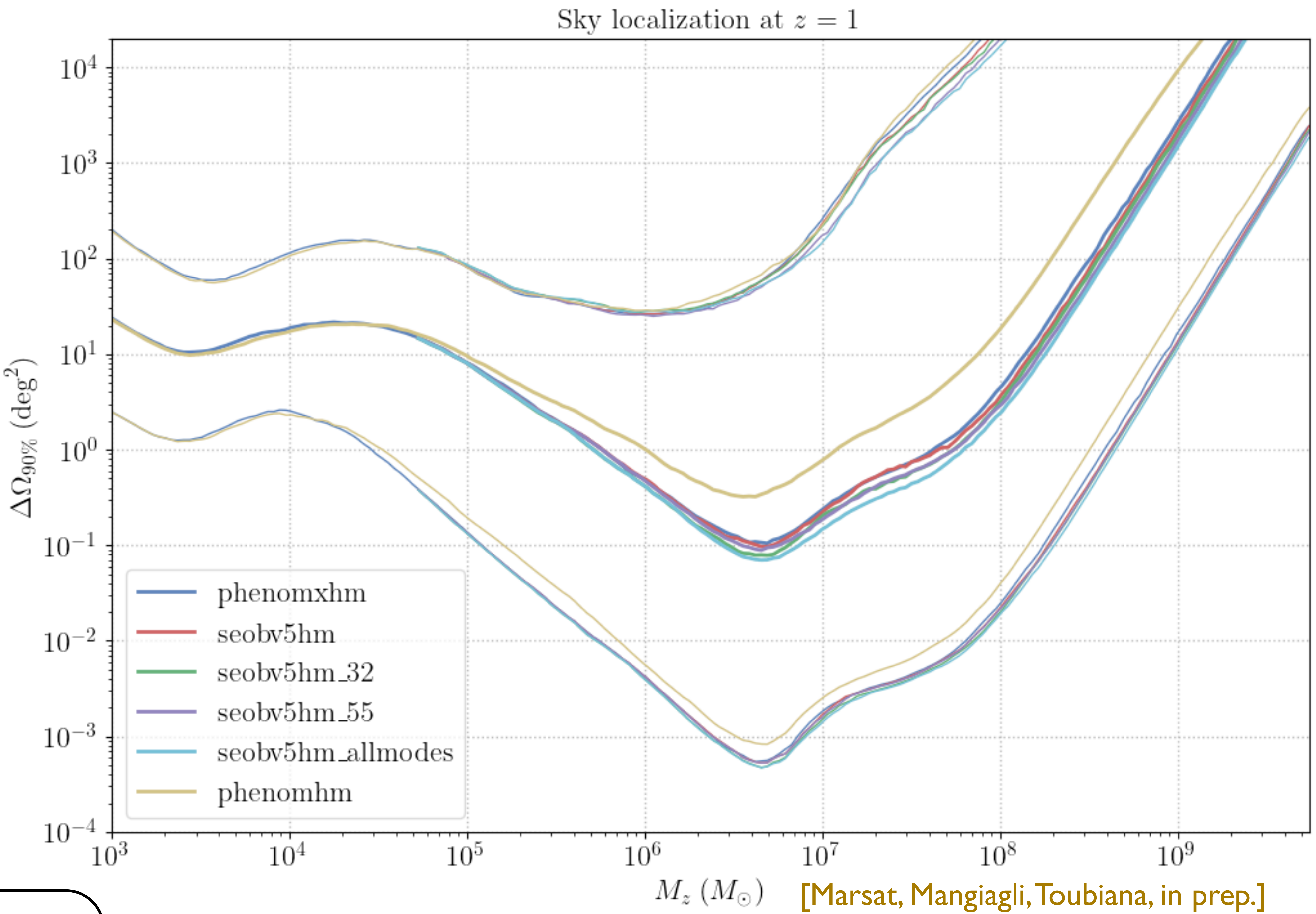


- Multimodality broken by subdominant effects in response (motion, high-f)

Fisher localization: impact of waveform model

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

Sky multimodalities for LISA MBHBs

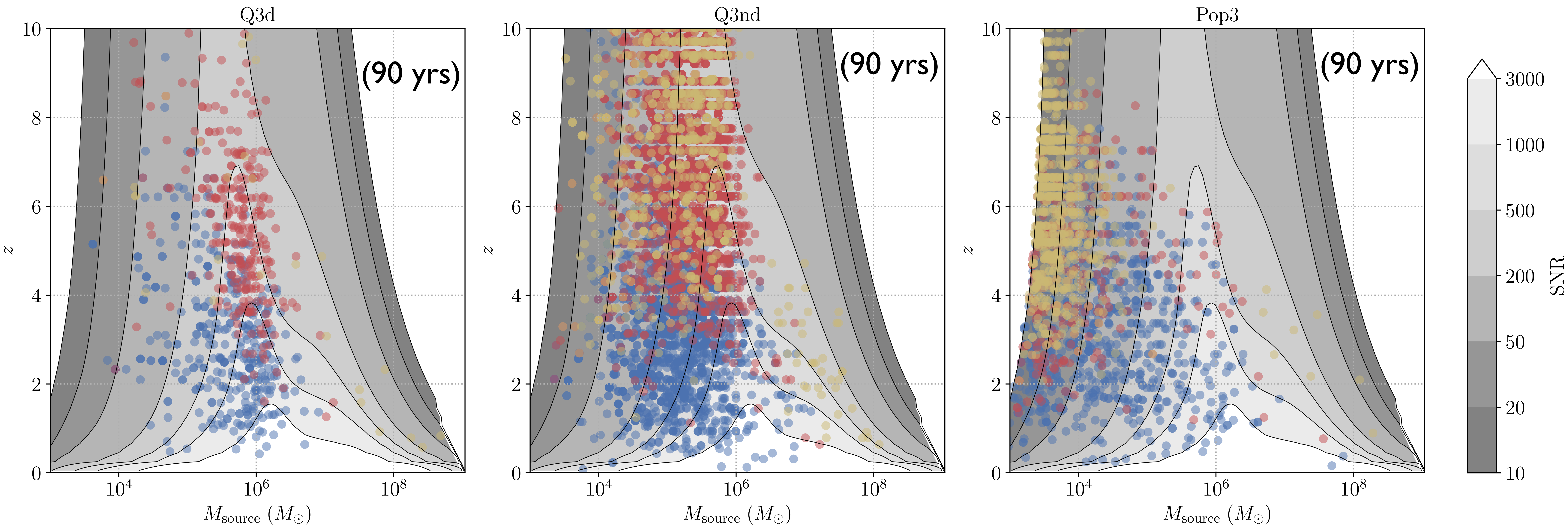
- **Bayesian PE** required to explore multimodal posteriors
- Simulation of 90yrs catalogs
- Custom proposals for degeneracies

Astrophysical models [Barausse 2012]:

- Heavy seeds - delay (Q3d)
- Heavy seeds - no delay (Q3nd)
- PopIII seeds - delay (Pop3)

o 1 mode
o 2 modes
o >2 modes

Threshold: 5% prob.



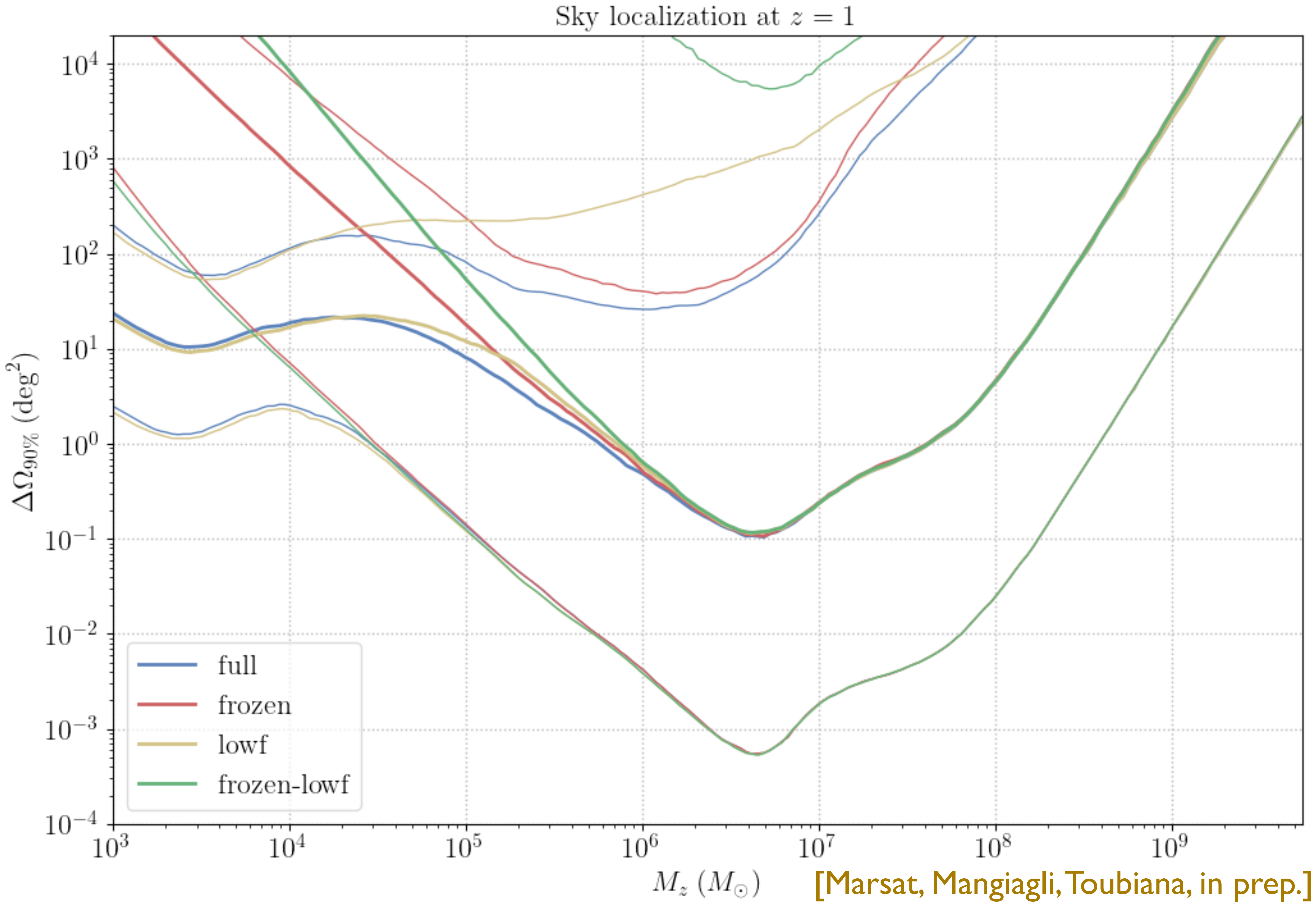
Applications: EM counterparts and cosmological inference
[Mangiagli&al 2022, Mangiagli&al 2023]

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

Fisher localization: impact of response approximation

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

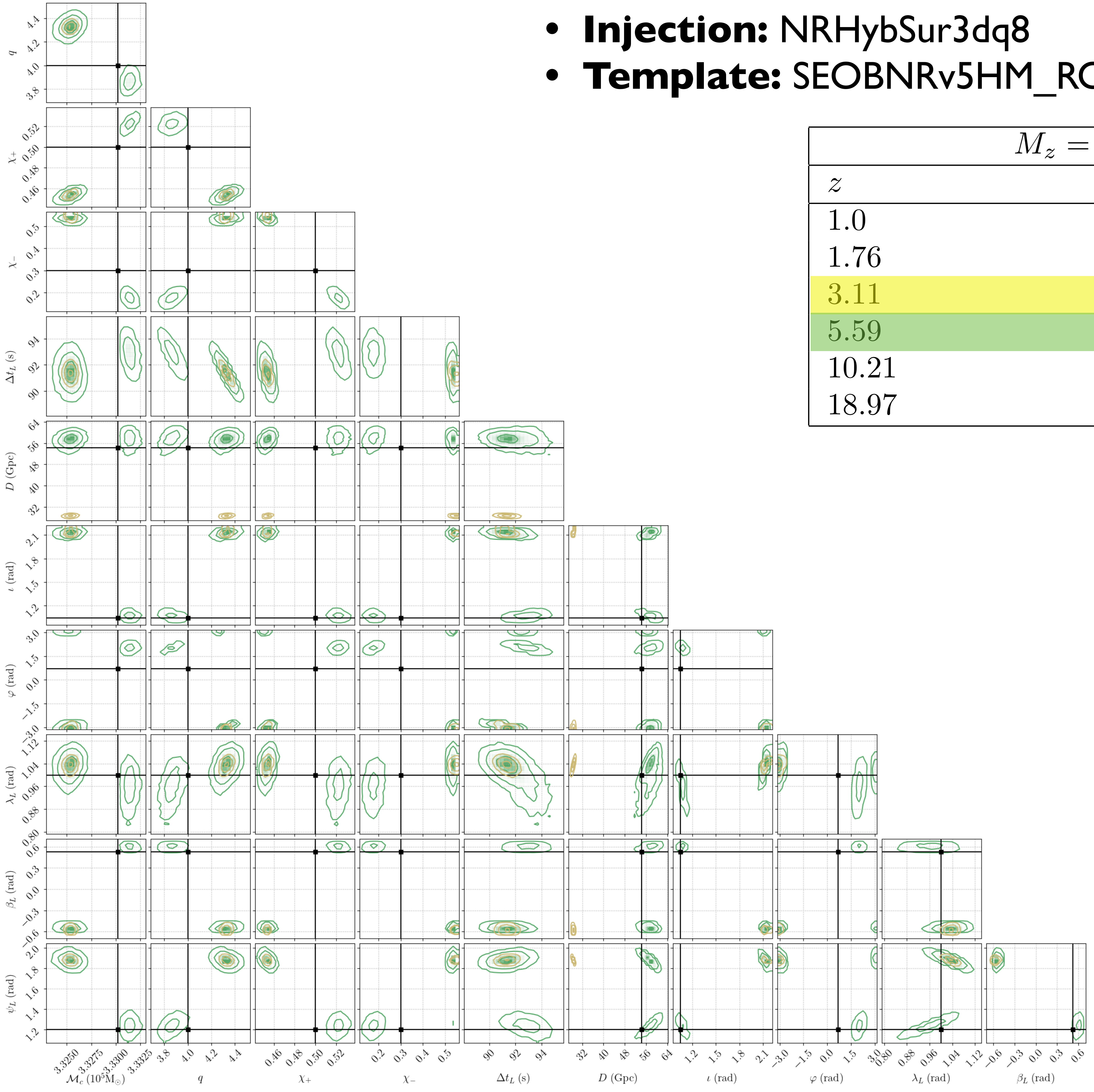


[Marsat, Mangiagli, Toubiana, in prep.]

- ‘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)

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

Example Parameter estimation with systematics III

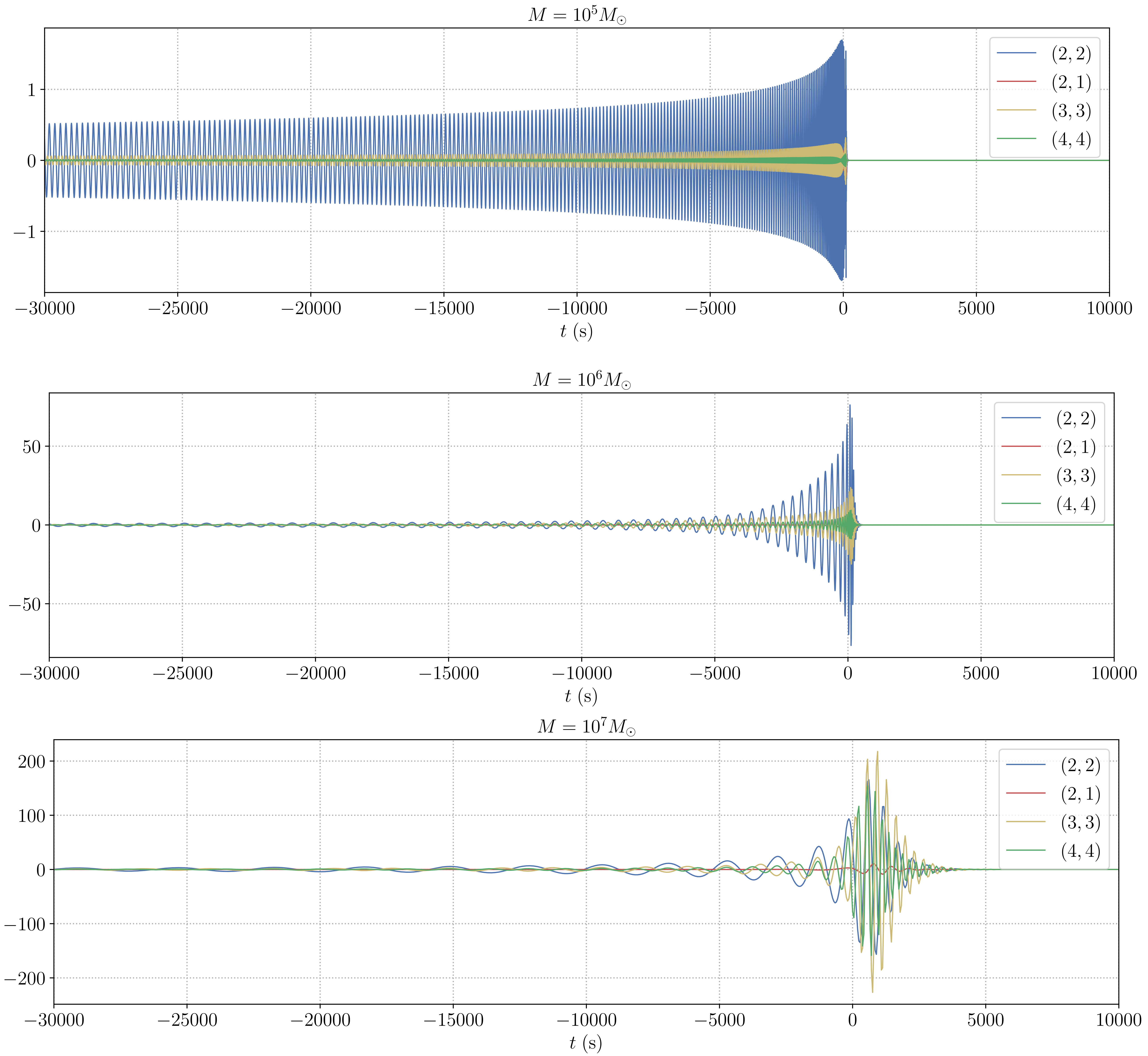


- **Injection:** NRHybSur3dq8
- **Template:** SEOBNRv5HM_ROM

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

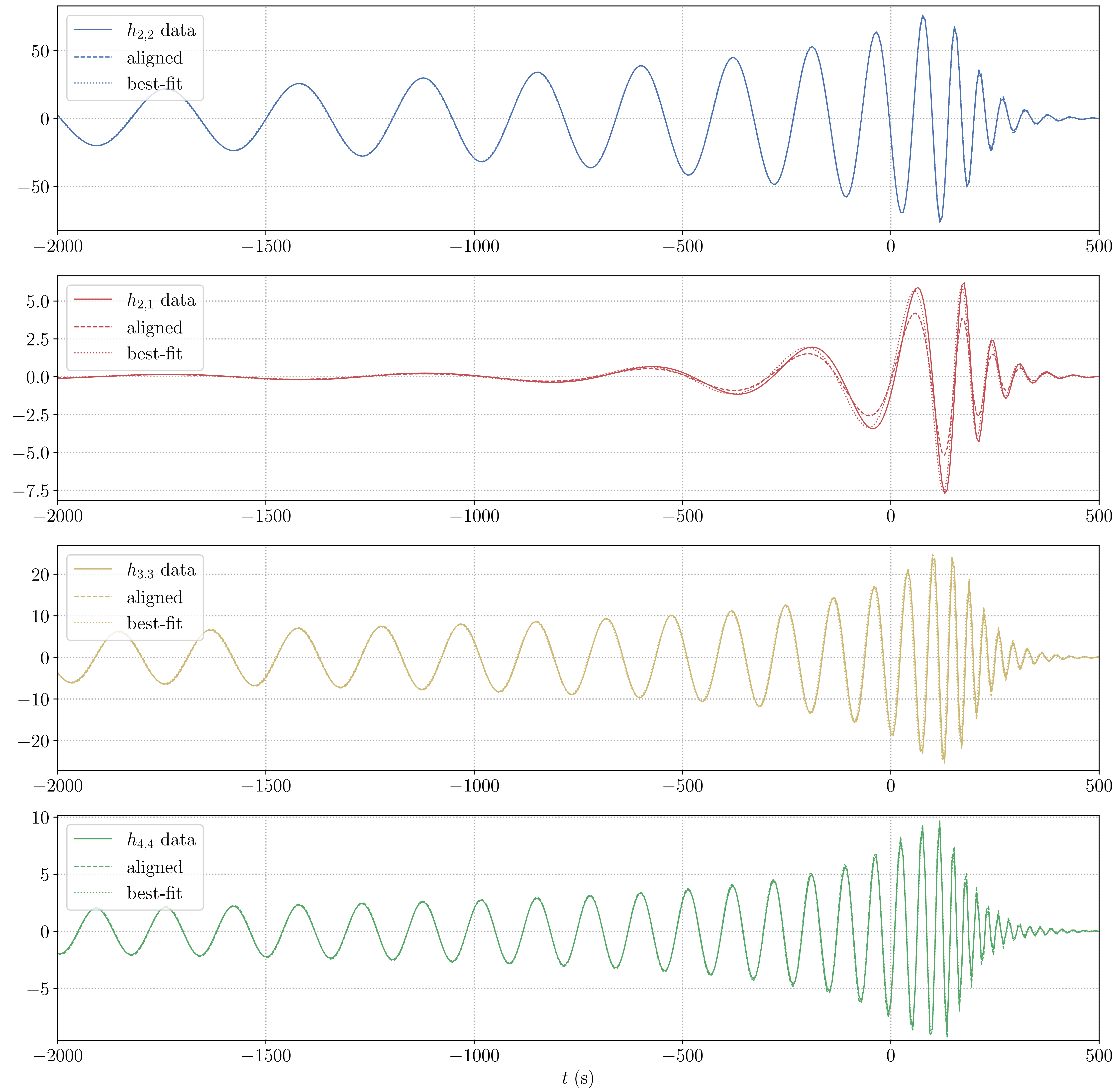
Example Parameter estimation with systematics III: TD signals and residuals

Whitened time-domain signals

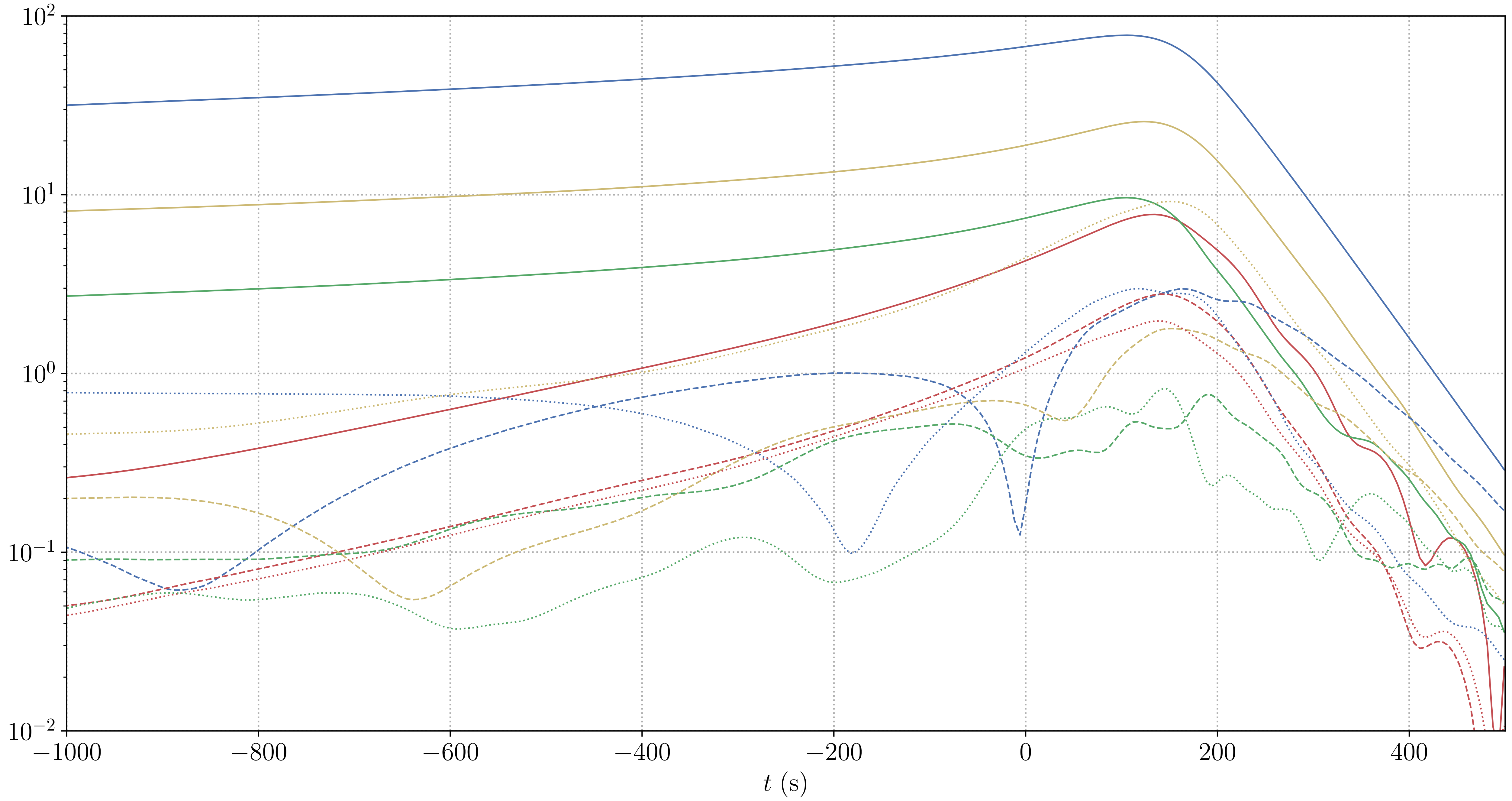


- Wide range of signal morphology
- High-mass signals are completely merger/HM dominated

Example Parameter estimation with systematics III: TD signals and residuals

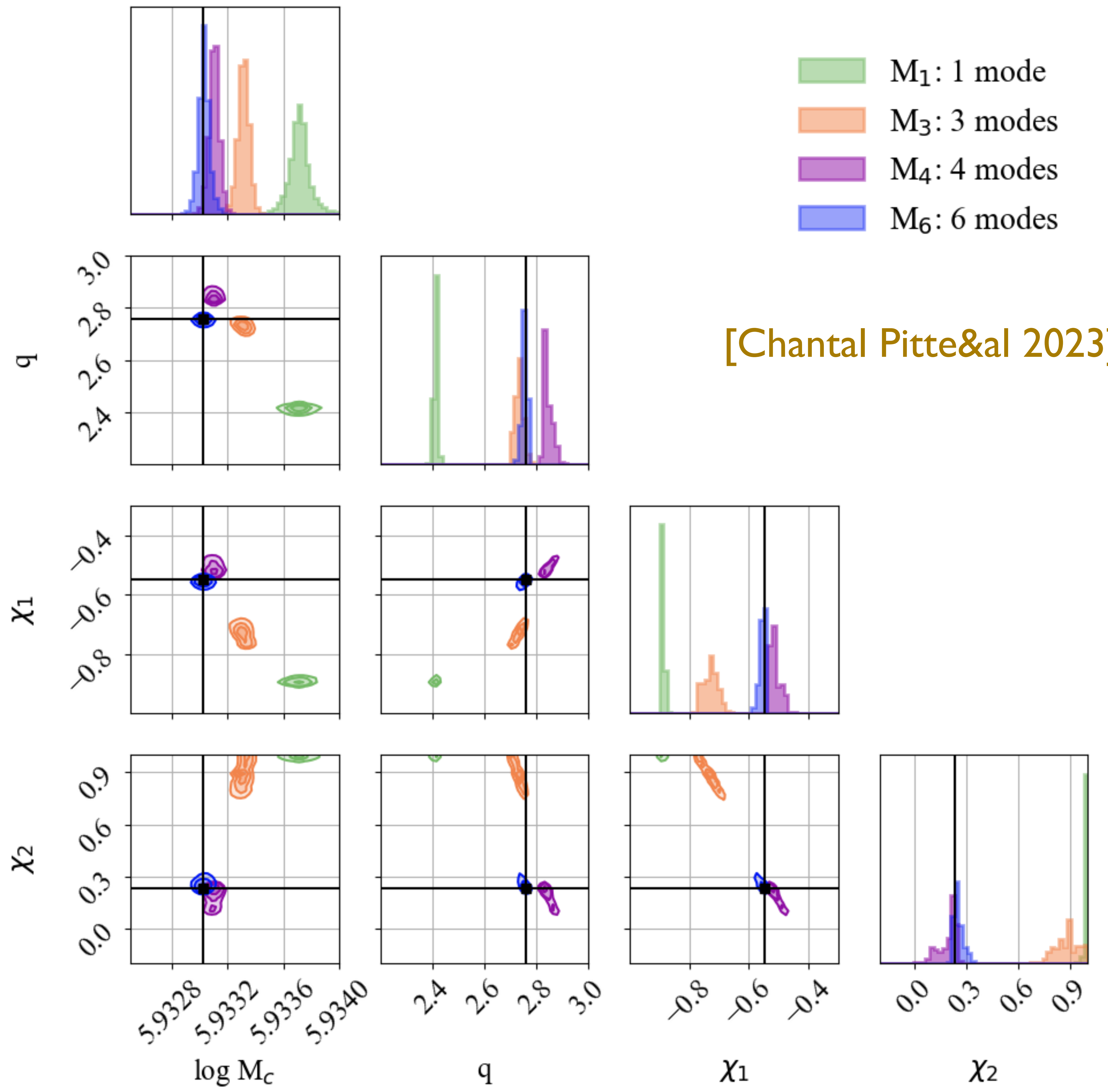


Residuals for the 'ugly' case

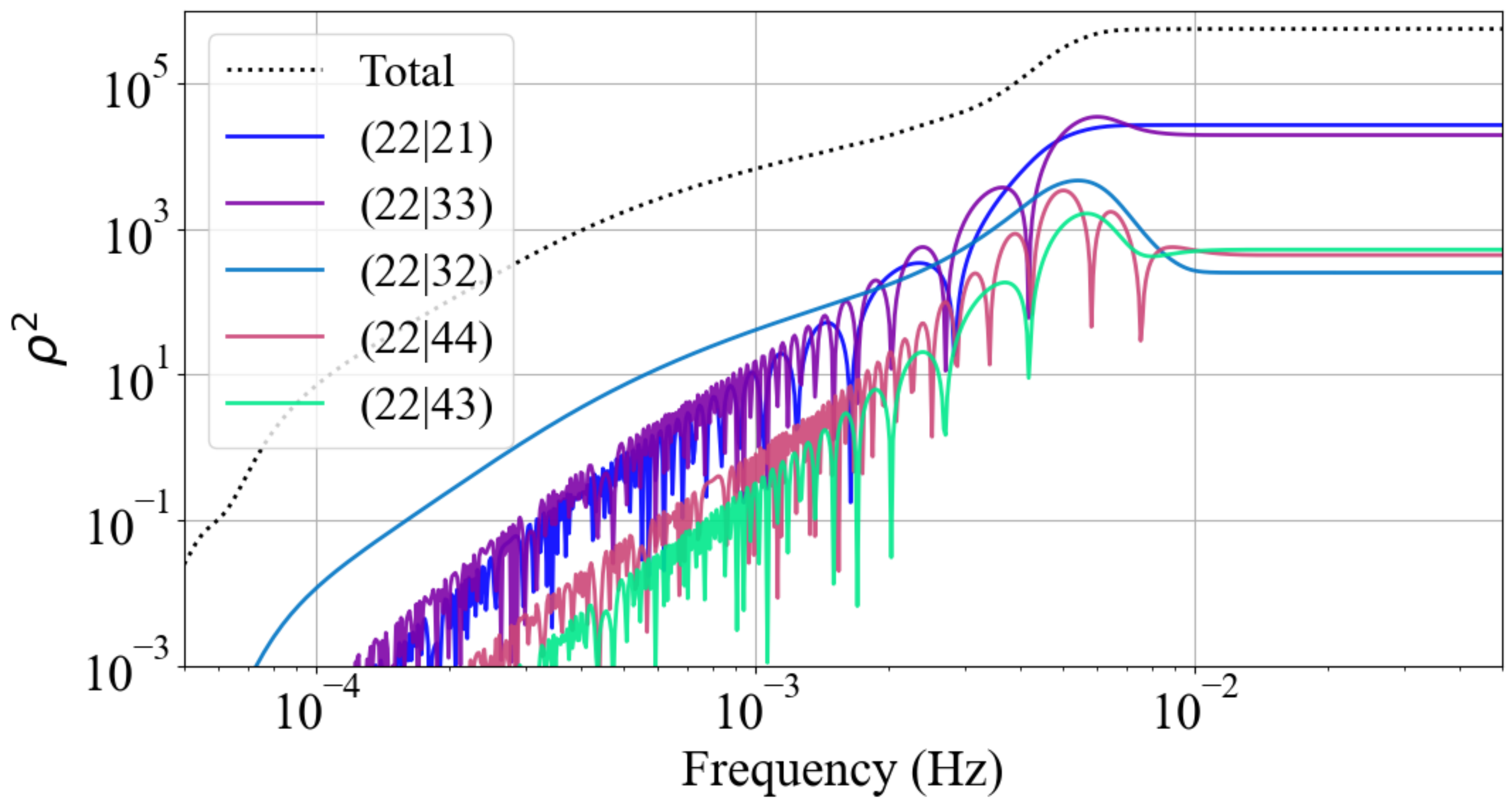


- Residuals are 'visually' small…
- Details of HM at merger are important

Biases caused by missing higher harmonics



[Chantal Pitte&al 2023]



‘Simpler’ waveform error: missing higher harmonic(s)

Missing higher harmonics can cause biases: cross-products between modes

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

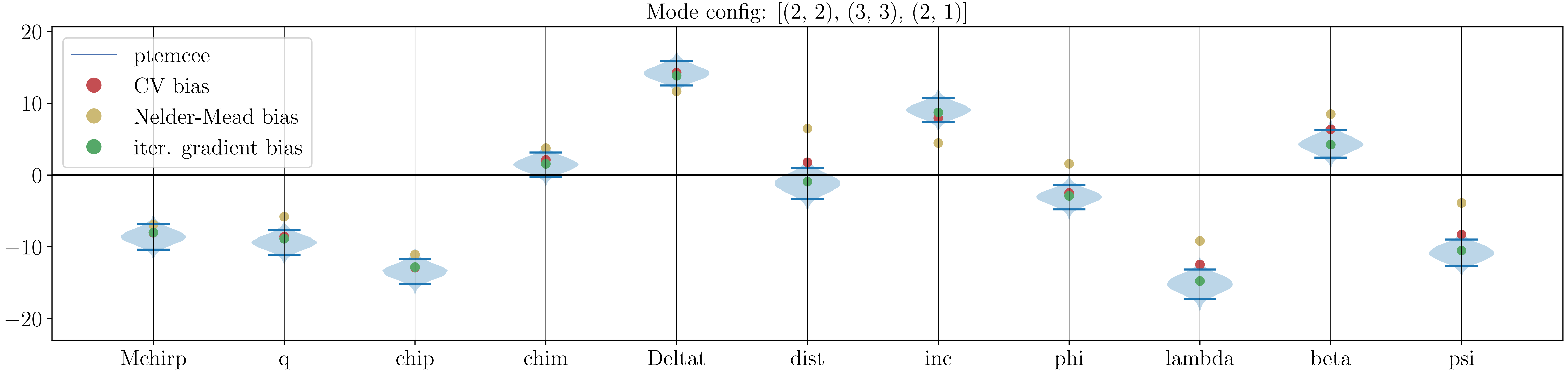
$$\Delta\theta = H^{-1} \cdot \nabla \ln \mathcal{L}$$
$$H_{ij} = \partial_i \partial_j \ln \mathcal{L} \sim (\partial_i h | \partial_j h)$$

- Are there better optimization algorithms ? (e.g. simplex method)

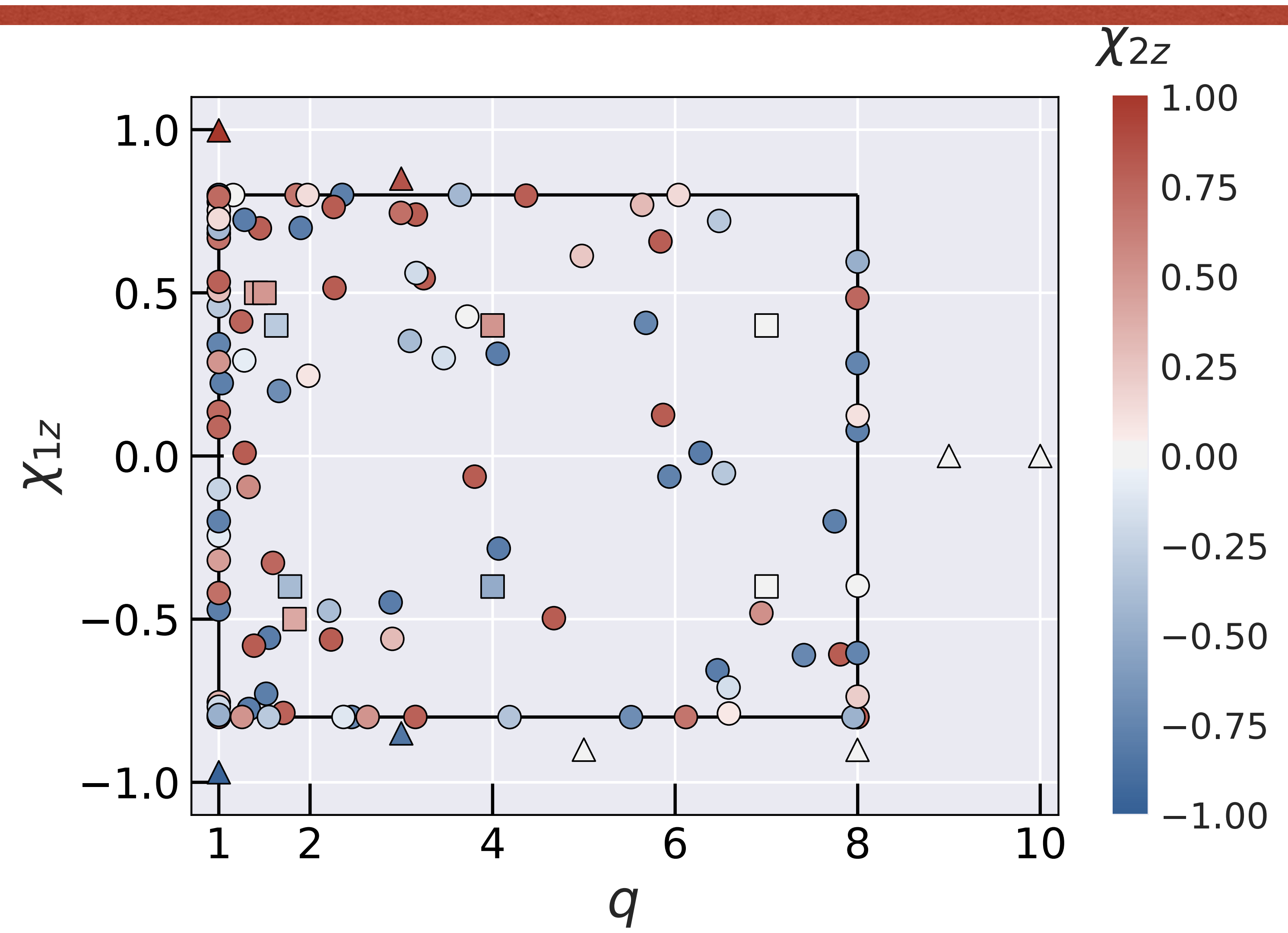
- One simple idea is to repeat CV as iterated gradient descent

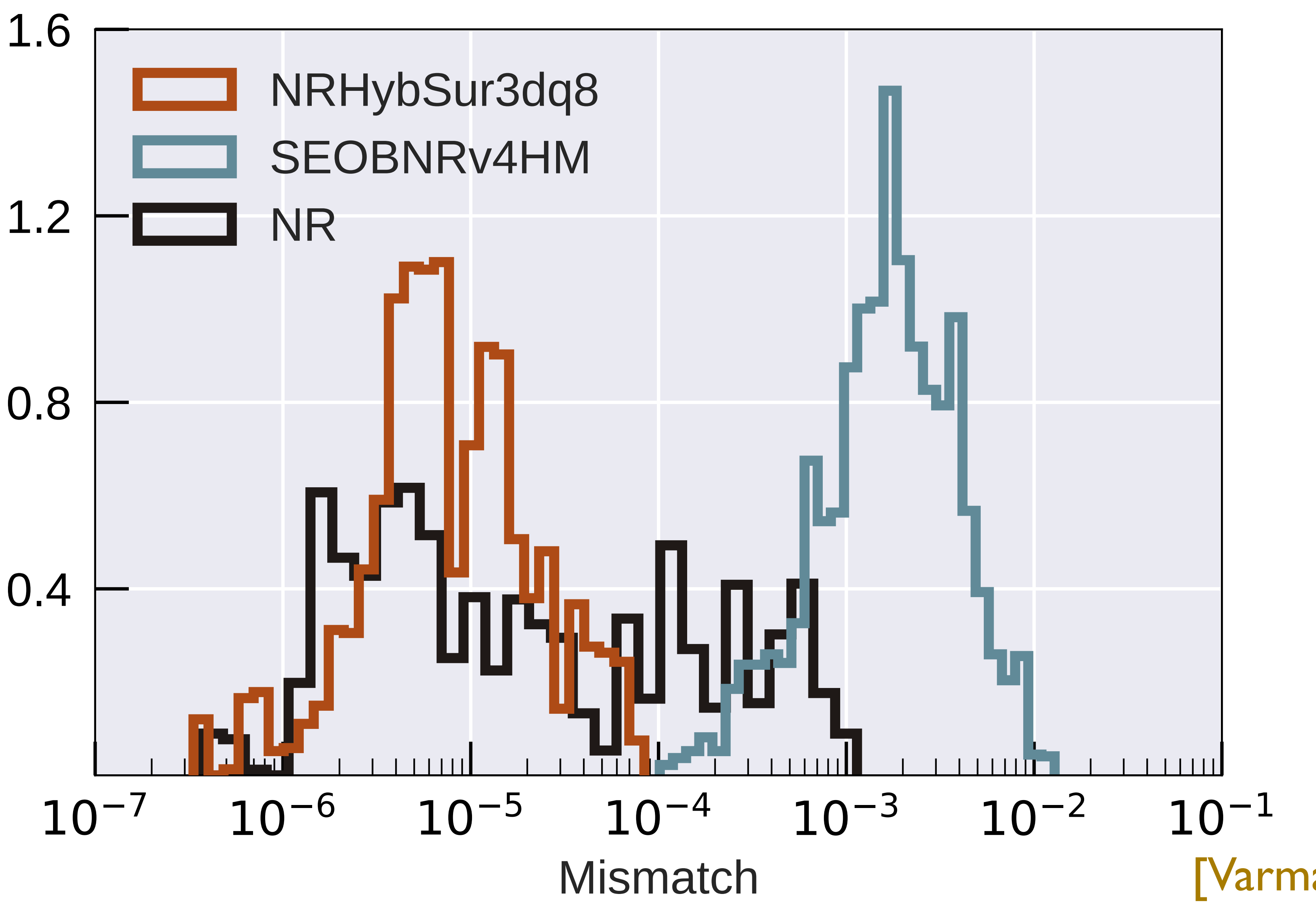
There should be better bias estimators than CV - robustness ?

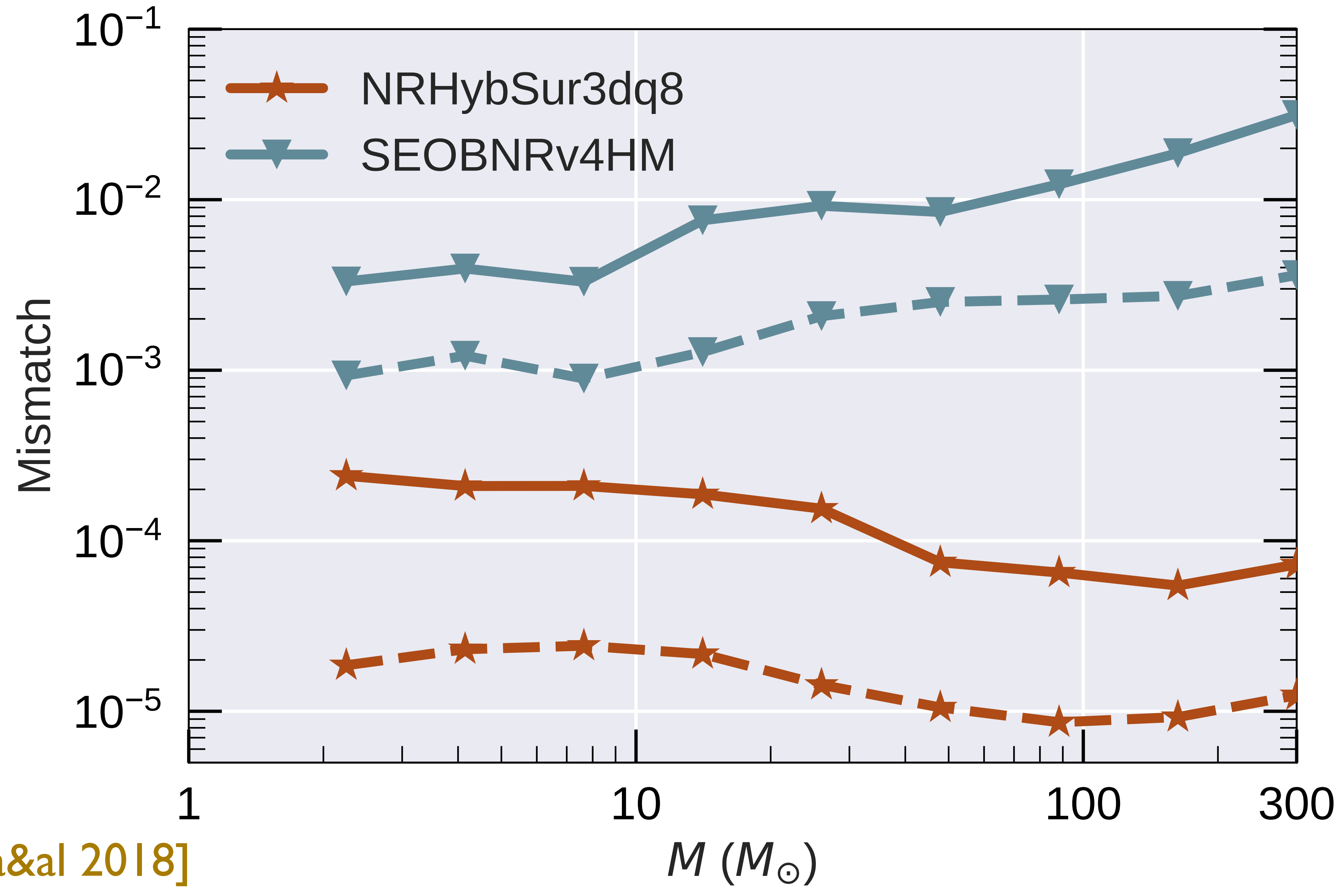
[Sophia Yi &al, in prep.]



Mismatches of NR surrogate

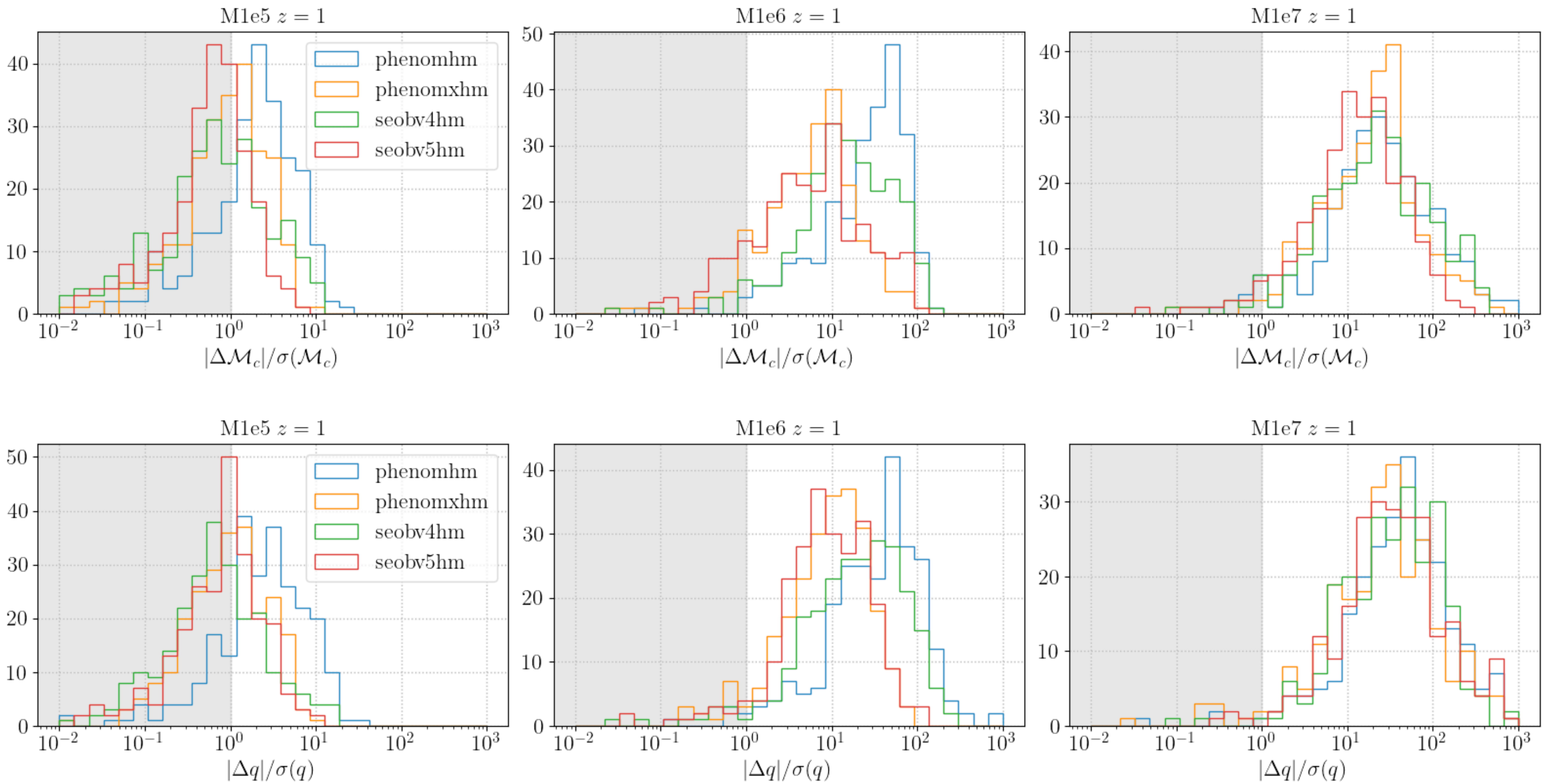




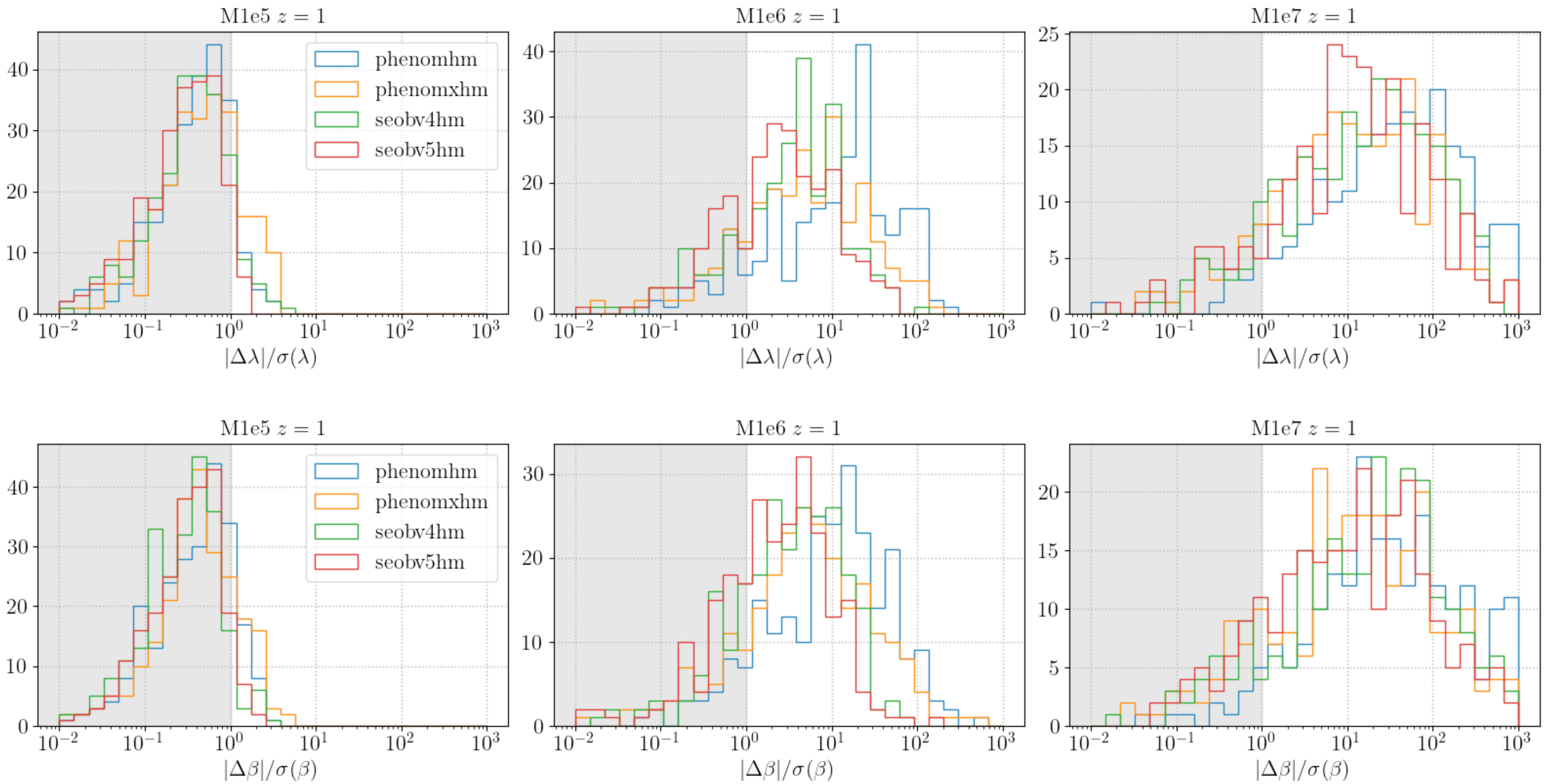


[Varma&al 2018]

Statistical significance of biases: intrinsic parameters



Statistical significance of biases: extrinsic parameters



Pre-merger analysis: likelihood with decomposed response

Degeneracy breaking for 8 sky maxima

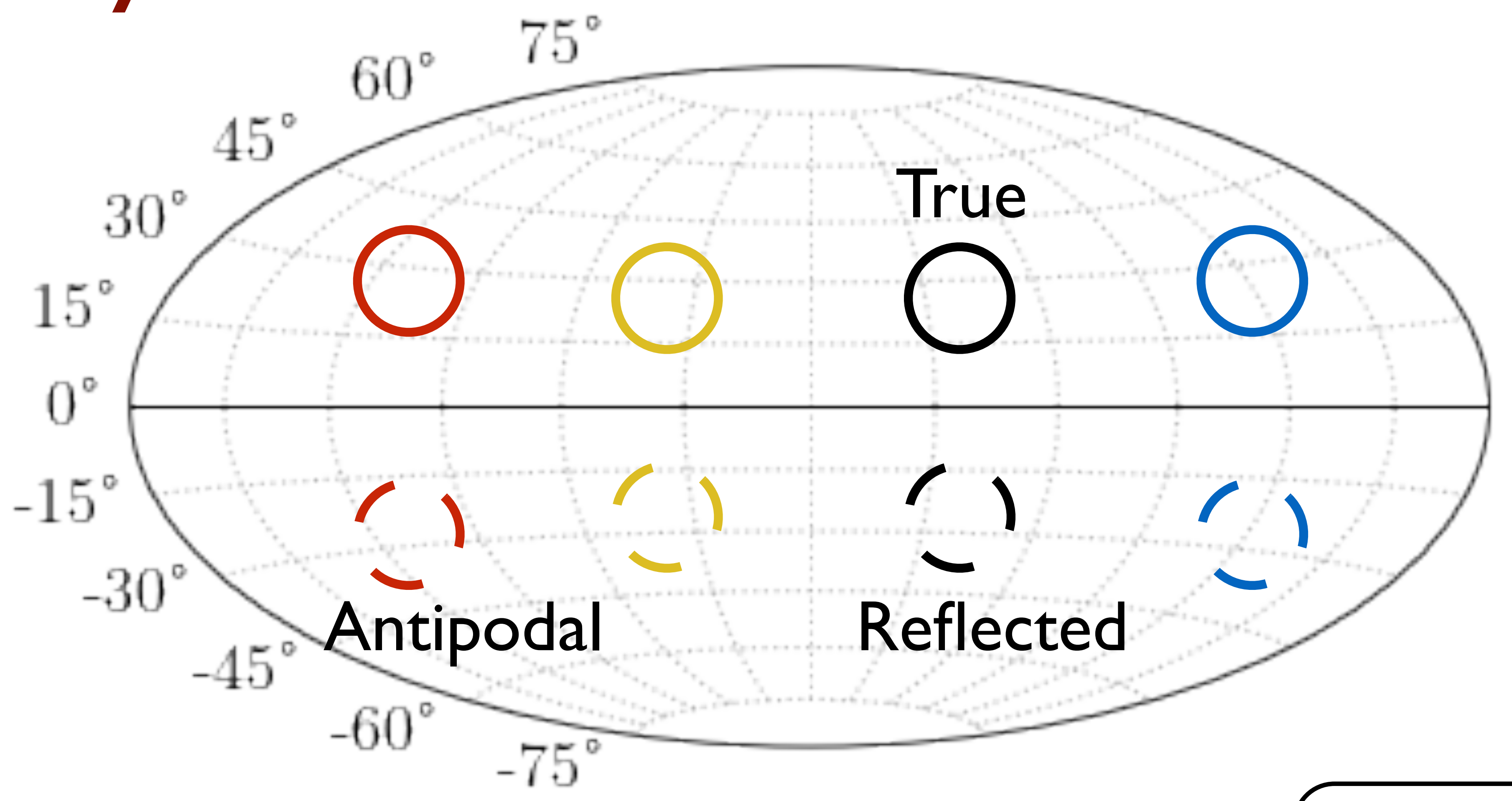
Instrument response:

- 'Full': keep all terms
- 'Frozen': ignore LISA motion
- 'Low-f': ignore f-dependency
- 'Frozen Low-f': ignore both

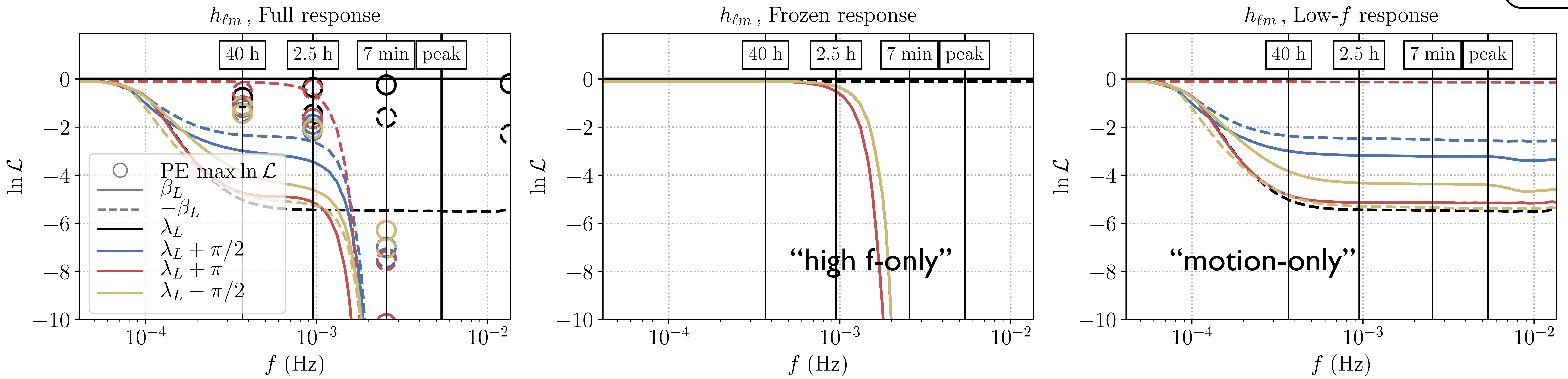
Likelihood: $\ln \mathcal{L}(d|\theta) = -\sum_{\text{channels}} \frac{1}{2}(h(\theta) - d | h(\theta) - d)$

Approximate degeneracy measure: likelihood at the other sky positions

Sky modes L-frame



High-f features crucial



[Marsat&al 2020]

Discussion: thoughts on biases

- Estimating biases is an SNR-independent deterministic optimization problem: best method ? Should waveform modellers evaluate parameter biases along mismatches ?

- The stochastic view on biases: introduce stochastic waveform errors, marginalize over waveform uncertainty

- The deterministic view on biases: model the bias as a parameter map (how quickly is it varying with parameters ?), for instance with dedicated NR injections in the vicinity of detected signals

- Different goals for waveform errors: evaluate width of posterior / likelihood surface simply connected through gradient descent to optimum / no significant biases

- Tests of GR: are modified-GR effects orthogonal to waveform errors ? Orthogonality better in the inspiral, not much at merger