Waveform systematics and sky localization for LISA signals of massive black hole binaries

Sylvain Marsat (L2IT, Toulouse)



in collaboration with A. Mangiagli (AEI Potsdam), A. Toubiana (U. Milano Bicocca)

ACME Workshop: The gravitational wave sky and complementary observations



Sylvain Marsat

2025-04-09

•MBHBs and LISA response

- Sky localization: main mode
- Sky localization: sky degeneracies
- Sky localization: galaxy counts
- Pre-merger sky localization
- Waveform systematics

2

Massive black hole binaries for LISA



The LISA instrumental response

The low-frequency response



The full response



Single-link response:

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

+ 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:

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

constant prefactors dependent on geometry





4



LISA response and multimodality in the sky

Multimodality pattern:



Degeneracy breaking:

• MBHBs and LISA response

•Sky localization: main mode

- Sky localization: sky degeneracies
- Sky localization: galaxy counts
- Pre-merger sky localization
- Waveform systematics

6

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
 - '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



MBHB sky localization at merger



See also: [McGee&al2018, Mangiagli&al 2020]

- 10 sq. deg.: LSST field of view
- 0.4 sq. deg.: Athena Wide Field Imager

Sky area: which parameters are the most important?



- MBHBs and LISA response
- Sky localization: main mode

•Sky localization: sky degeneracies

- Sky localization: galaxy counts
- Pre-merger sky localization
- Waveform systematics

MBHB catalogs: full parameter estimation

Astrophysical models [Barausse 2012]:

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

LISA detection rates from 90 yrs simulated:

- Q3d: 30 / 4yrs
- Q3nd: 471 / 4 yrs
- Pop3: 129 / 4yrs

MBHB catalogs: sky multimodality

- MBHBs and LISA response
- Sky localization: main mode
- Sky localization: sky degeneracies

•Sky localization: galaxy counts

- Pre-merger sky localization
- Waveform systematics

13

Luminosity distance determination

For distance determination, higher harmonics are **crucial**, breaking the distance-inclination degeneracy.

Looking for a host by converting dL to z, weak lensing and peculiar motions have to be taken into account

MBHB catalogs: sky localisation and galaxy counts

Sky areas

Sky area and error volume computed from Bayesian PE (main mode), with lensing

Catalogs: [Barausse 2012]

Simplistic, simulated catalog cut in mass with no consideration of completeness, EM emissions...

Simulated galaxy catalog courtesy of [D. lzquierdo-Villalba&al]

- MBHBs and LISA response
- Sky localization: main mode
- Sky localization: sky degeneracies
- Sky localization: galaxy counts

Pre-merger sky localization

• Waveform systematics

16

Pre-merger localization

- MCMC: 100 PE runs
- Here, sky area of the main mode

[Piro&al 2022]

[See also Mangiagli&al 2020]

Advance localization challenging, much better post-merger

Large dispersion in sky area, ~4 orders of magnitude

17

Pre-merger localization: degeneracies

Bayesian PE: sky localization cutting at different times

- 'Platinum': M3e5, z=0.3
- 'Gold': M3e6, z=1
- 'Heavy': MIe7, z=1

- Wide range of multimodalities dep. on parameters
- Post-merger localization unimodal here

- MBHBs and LISA response
- Sky localization: main mode
- Sky localization: sky degeneracies
- Sky localization: galaxy counts
- Pre-merger sky localization
- •Waveform systematics

19

Waveform systematics and parameter estimation

Are current waveform models accurate enough for LISA ? Can the sky localization be biased ?

Parameter space exploration:

- $M_z = [10^5, 10^6, 10^7] M_{\odot}$
- $z_{\min} = 1$
- N = 240 PE runs
- uniform q, χ_1, χ_2
- randomize orientations

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

Injections:

NRHybSur3dq8

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

Templates:

Efficient Fourier-domain models:

- PhenomHM
- PhenomXHM
- SEOBNRv4HM_ROM
- SEOBNRv5HM_ROM

• **Template:** PhenomXHM

• **Template:** PhenomXHM

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

The good:

- converges on the true parameters
- mild bias at z = 1, SNR = 317

• **Template:** PhenomXHM

• **Template:** PhenomXHM

• **Template:** PhenomXHM

• **Template:** PhenomXHM

+3.33

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

 β_L

Statistical significance of biases: intrinsic parameters

Bias in chirp mass:

Bias in longitude (on corrected skymode):

Wrong skymode recovered:

Highlights

- MBHB signals are merger-dominated, post-merger localisation can be very good
- Main mode localization: from pattern response at high mass, inclination/latitude dominated
- Pre-merger localisation can be challenging, except for the very best events
- Degeneracies in the sky position can occur, worse pre-merger
- Systematics: possibly strong for high-mass signals, can also mislead towards the wrong sky mode

Outlook

- More realistic waveforms: precession and eccentricity
- More realistic analysis: proper time-domain analysis, superposition of multiple signals, realistic noise, data gaps, glitches...

• LISA localisation capabilities for MBHBs crucial for multimessenger science

Galaxy counts in the LISA + weak lensing error box

Simplistic estimate: simulated catalog cut in mass with no consideration of completeness, EM emissions...

Simulated catalog courtesy of [D. Izquierdo-Villalba&al]

[see also Lops&al 2022]

Fitting factors in parameter space

Waveform systematics and parameter estimation

Indistinguishability criterion:

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

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

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

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

- Constant *D*: dimension, approximate
- Scaling SNR² robust

Mismatch (unfaithfulness):

Mismatch, optimization over time/phase/polarization:

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

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

Linearized biases (Cutler-Vallisneri): [Flanagan-Hughe

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

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

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

Can we assess biases with efficient tools ?

Linking mismatches and biases

From indistinguishability criterion:

From bias measured in PE:

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

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

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

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

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

easured in PE: $\Delta \theta$ $\overline{(\theta)}$

Relation between mismatch and bias unclear

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

Both

42

Massive black hole binaries

43

eccentricity

MBHB signals are merger-dominated in SNR

Most of the SNR accumulates in the last hours before coalescence

The length of MBHB signals

MBHB detected signals:

Astrophysical models [Barausse 2012]:

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

Fisher vs MCMC localisation

Early detection for 'golden' sources

Injections:

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

48

Pre-merger analysis: accumulation of information with time

Method

 Represent a cut in time-tomerger by a cut in frequency, becomes inaccurate at merger

SNR-based time cuts:

SNR	DeltaT
10	40h
42	2.5h
167	7min
666	-

8-maxima sky degeneracy only broken shortly before merger

2-maxima sky degeneracy

survives after merger ('Reflected')

LDC-2: source superposition and global analysis

LISA Data analysis challenges

- Superposition of many sources, with a population of GBs also forming a stochastic background -> Global fit
- High SNR for MBHBs, waveform systematics important
- EMRI waveform models
- Data gaps
- Glitches, instrumental non-stationarity

LISA Data Challenge 2 'Sangria'

- ~10 massive black holes
- Population of galactic binaries (~10000 resolved)
- Unkown noise level

 10^{-3} \tilde{X} -TDI, 1/Hz 10^{-} 10^{-5}

-2.0

-2.5 [ZH] And Hz -3.0

-3.5

 -4.0^{-1}

The SNR of higher harmonics

52

Understanding degeneracy breaking by higher harmonics

The role of higher harmonics

$$h_{+} - ih_{\times} = \sum_{-2} Y_{\ell m}(\iota, \varphi) h_{\ell m}$$

$${}_{-2}Y_{\ell m}(\iota,\varphi) = {}_{-2}Y_{\ell m}(\iota,0)e^{im\varphi}$$

When measuring several modes $h_{\ell m}$:

- Distance/inclination degeneracy broken
- Phase independently measured
- Better sky localization (caveat: edge-on, see [Katz&al 2020])

Sky area: impact of WD background

The WD confusion background matters Sky localisation will be updated as the GB analysis is refined

Degradation of median sky localization at z = 1 due to WD

Pre-merger localization: role of instrumental response for 'golden' systems

Here: main mode sky area

Response (signal with HM here):

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

For low masses, best candidates for advance localization:

- Localization from the LISA motion saturates reaching merger
- Localization from high-f effects dominates at merger

For high masses, HM at merger convey most of the information

Pre-merger localization: role of HM for 'golden' systems

Pre-merger analysis: likelihood with decomposed response

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

[Marsat&al 2020]

LISA/Athena candidates

'Platinum'

$$M_{\rm source} = 3 \times 10^5 M_{\odot}, z = 0.3$$

- Very long: > lyr
- Localization unimodal early on, no sky degeneracies

'M3e6'

• Observable for ~2w

$$\times 10^6 M_{\odot}, z = 1$$

'Mle7'

$$M_{\rm source} = 3 \times 10^7 M_{\odot}, z = 1$$

• Observable for ~2d

59

Multimodality of the sky localization: astrophysical catalogs

$$\mathcal{L} > -20$$

Multimodality of the sky localization: a likelihood estimator ?

