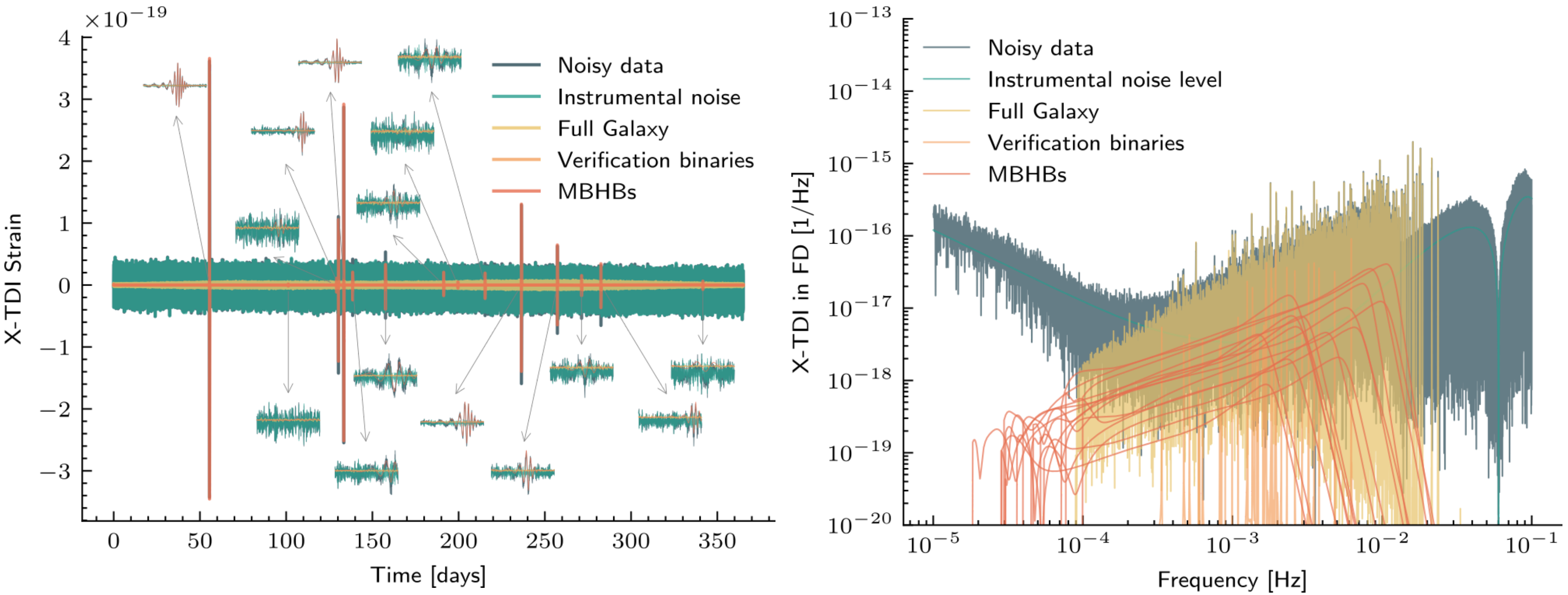
# Fast detection and reconstruction of merging Massive Black Hole Binary signals

Sen-wen Deng (APC)

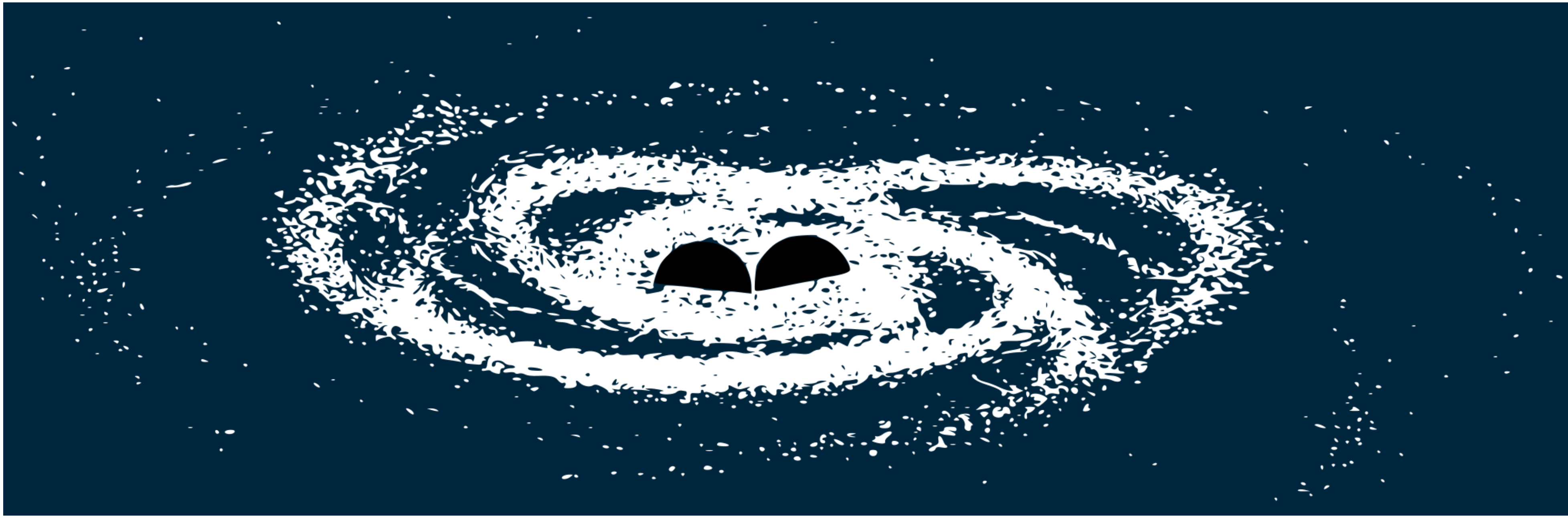
14th May 2025, SISSA

# Profile of LISA Data

- Dominated by GW signals all over the sky and the observation time
- Many signals are long-lived and overlapping in time
  - GBs, EMRIs, etc.
- MBHBs are loud, broadband and "transient"
- Unresolved sources contribute to the confusion noise



# MBHBs in LISA

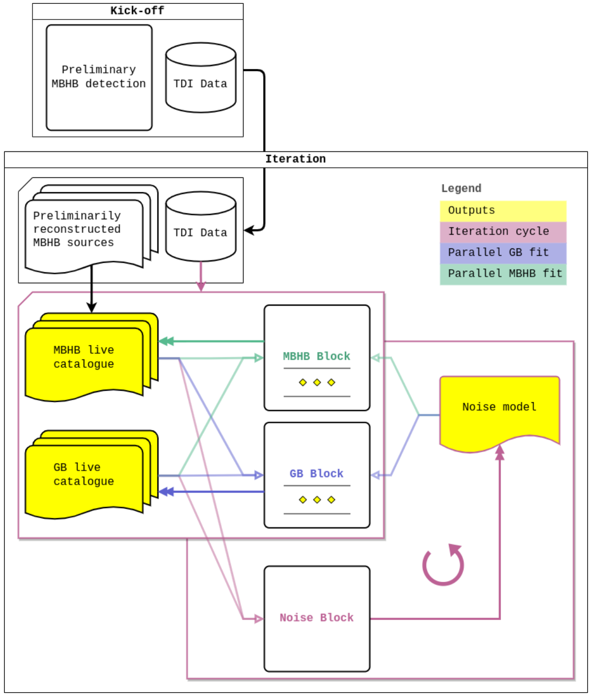


Reproduction / Credit: ESA

- Masses: $10^4$ - $10^7$ $M_\odot$
- Inspiral-merger-ringdown (IMR) waveform
  - Identical to LVK source waveforms (if we scale the amplitude and the frequency)
  - ! More complicated LISA response
- Signal-to-noise ratio (SNR) to several hundreds or even thousands ⇒ easy to detect
  - Mostly contributed by the merger part
  - Detectable inspiral-only

# Dealing with the MBHBs

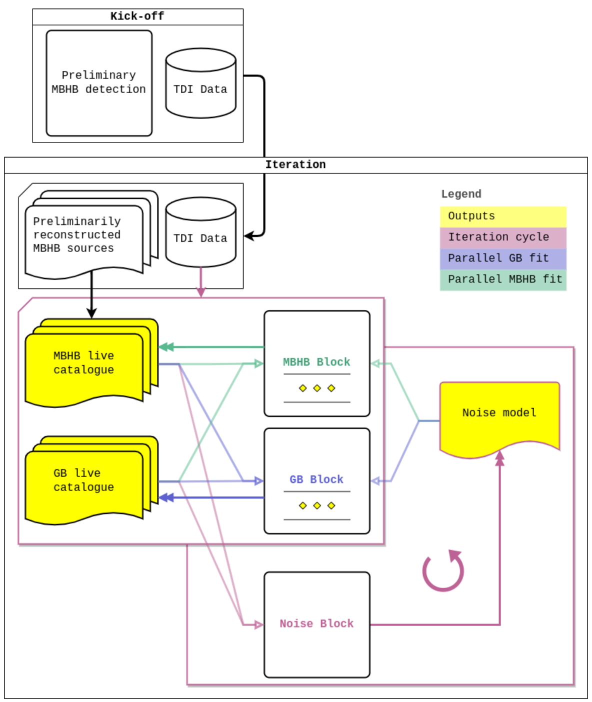
- In a global-fit pipeline
  - Loud MBHB signals hinder the proper estimation of the noise model
    - We aim to subtract the loud MBHB signals from the data as a kick-off step




Architecture of a global-fit pipeline

# Dealing with the MBHBs

- In a global-fit pipeline
  - Loud MBHB signals hinder the proper estimation of the noise model
    - We aim to subtract the loud MBHB signals from the data as a kick-off step




Architecture of a global-fit pipeline

- In a low-latency pipeline
  - The merger of MBHBs carries the most information but is transient
    - We want to protect the mergers from the scheduled maintenance and antenna repointing
    - We aim to predict the time of upcoming mergers
  - Predicting enough in advance helps the multimessenger observations

# Fast detection and reconstruction scheme

## Strategy

1. Crudely simplify the response
2. Reduce the dimension of the parameter space
3. Run an optimiser $\Rightarrow$ detection and reconstruction

# Fast detection and reconstruction scheme

## Strategy

1. Crudely simplify the response
    - Assume the LISA constellation is frozen
    - Push the long-wavelength regime outside its region of validity

$$\tilde{h}_{\mathrm{A}}^{\ell m}, \tilde{h}_{\mathrm{E}}^{\ell m} = i\sqrt{2}\sin(2\pi f L) e^{-2i\pi f L}(-6i\pi f L) \times e^{2i\pi f \mathbf{k}\cdot\mathbf{p_0}} F_{\mathrm{A,E}}^{\ell m} \tilde{h}^{\ell m}$$

- For frozen LISA, $F_{\mathrm{A,E}}^{\ell m}$ are constant complex numbers
- $\tilde{h}_{\mathrm{A,E}}^{\ell m} = a_{\mathrm{A,E}}^{\ell m} \tilde{H}^{\ell m}(f) + b_{\mathrm{A,E}}^{\ell m} i \tilde{H}^{\ell m}(f)$ with $a_{\mathrm{A,E}}^{\ell m}, b_{\mathrm{A,E}}^{\ell m} \in \mathbb{R}$

# Fast detection and reconstruction scheme

## Strategy

1. Crudely simplify the response
   - $\tilde{h}^{\ell m}_{\mathrm{A,E}} = a^{\ell m}_{\mathrm{A,E}} \tilde{H}^{\ell m}(f) + b^{\ell m}_{\mathrm{A,E}} i \tilde{H}^{\ell m}(f)$ with $a^{\ell m}_{\mathrm{A,E}}, b^{\ell m}_{\mathrm{A,E}} \in \mathbb{R}$
2. Reduce the dimension of the parameter space
   - Log-likelihood ratio (data containing signal vs. noise-only data)
   - $\langle k|g\rangle := 4\Re \int_{f_{\mathrm{min}}}^{f_{\mathrm{max}}} \frac{\tilde{k}(f)\tilde{g}^*(f)}{S_{\mathrm{n}}(f)} \mathrm{d}f$
   - $\ln \mathcal{L} = \langle d|h\rangle - \frac{1}{2}\langle h|h\rangle$
   - If $\tilde{h} = a\tilde{H} + bi\tilde{H}$ with $a, b \in \mathbb{R}$, then $\ln\mathcal{L}$ can be maximised by $a = \frac{\langle d|H\rangle}{\langle H|H\rangle}, b = \frac{\langle d|iH\rangle}{\langle H|H\rangle}$

$$\mathcal{F}\text{-statistic} := \ln \mathcal{L}_{\mathrm{max}} = \frac{1}{2}\left( \frac{\langle d|H\rangle^2}{\langle H|H\rangle} + \frac{\langle d|iH\rangle^2}{\langle H|H\rangle} \right)$$

Only the intrinsic parameters (masses and spins) and the coalescence time $t_{\mathrm{c}}$ enter the $\mathcal{F}$-statistic.

# Fast detection and reconstruction scheme

## Strategy

1. Crudely simplify the response
2. Reduce the dimension of the parameter space

$$\mathcal{F}_{\mathrm{A,E}} = \frac{1}{2}\left( \frac{\langle d_{\mathrm{A,E}}|H\rangle^2}{\langle H|H\rangle} + \frac{\langle d_{\mathrm{A,E}}|iH\rangle^2}{\langle H|H\rangle} \right)$$

When the noise is Gaussian

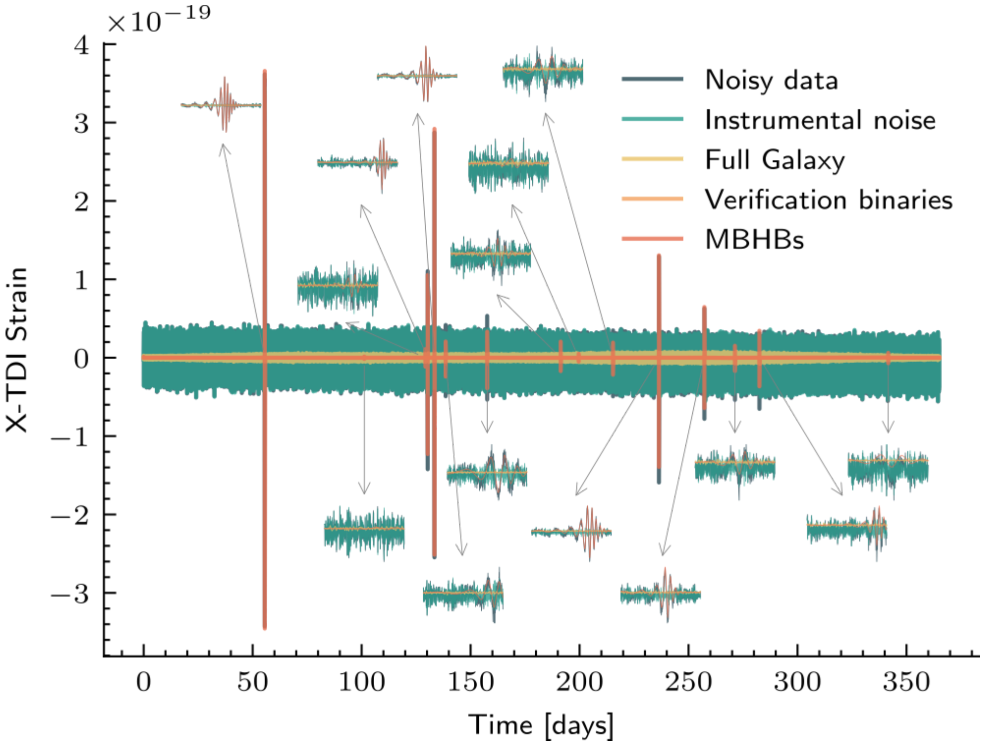
- $2\mathcal{F} := 2(\mathcal{F}_A + \mathcal{F}_E)$ follows the central $\chi^2$ distribution with $\mathrm{d.o.f} = 4$ if no signal is present
- $P_{\mathrm{FA}} = 1 - \mathrm{CDF}(2\mathcal{F}; 4)^N$ for $N$ independent trials

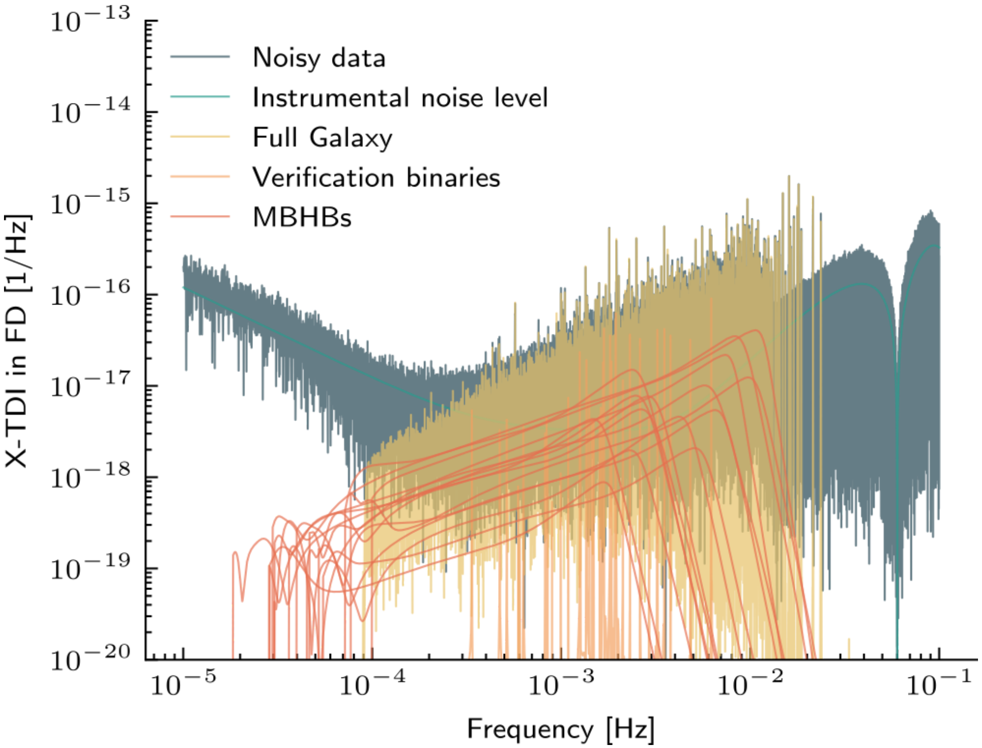
# Fast detection and reconstruction scheme

## Strategy

1. Crudely simplify the response
2. Reduce the dimension of the parameter space
   - $\mathcal{F}_{\mathrm{A,E}} = \frac{1}{2}\left(\frac{\langle d_{\mathrm{A,E}}|H\rangle^2}{\langle H|H\rangle} + \frac{\langle d_{\mathrm{A,E}}|iH\rangle^2}{\langle H|H\rangle}\right)$
3. Run an optimiser ⇒ detection and reconstruction
   - Parameters (masses, spins, $t_{\mathrm{c}}$) enter $\mathcal{F}$ through $H$
   - IFFT finds the optimal $t_{\mathrm{c}}$
   - Good stochastic optimisers do the trick for masses and spins
     - Differential Evolution (DE)
     - Adapted Particle Swarm Optimisation (APSO)
     - ...
   - Detection: $\mathcal{F} \geqslant \mathcal{F}_{\mathrm{th}}$
   - Reconstruction: $a_{\mathrm{A,E}} = \frac{\langle d_{\mathrm{A,E}}|H\rangle}{\langle H|H\rangle}, b_{\mathrm{A,E}} = \frac{\langle d_{\mathrm{A,E}}|iH\rangle}{\langle H|H\rangle}$

# Applied to "Sangria"

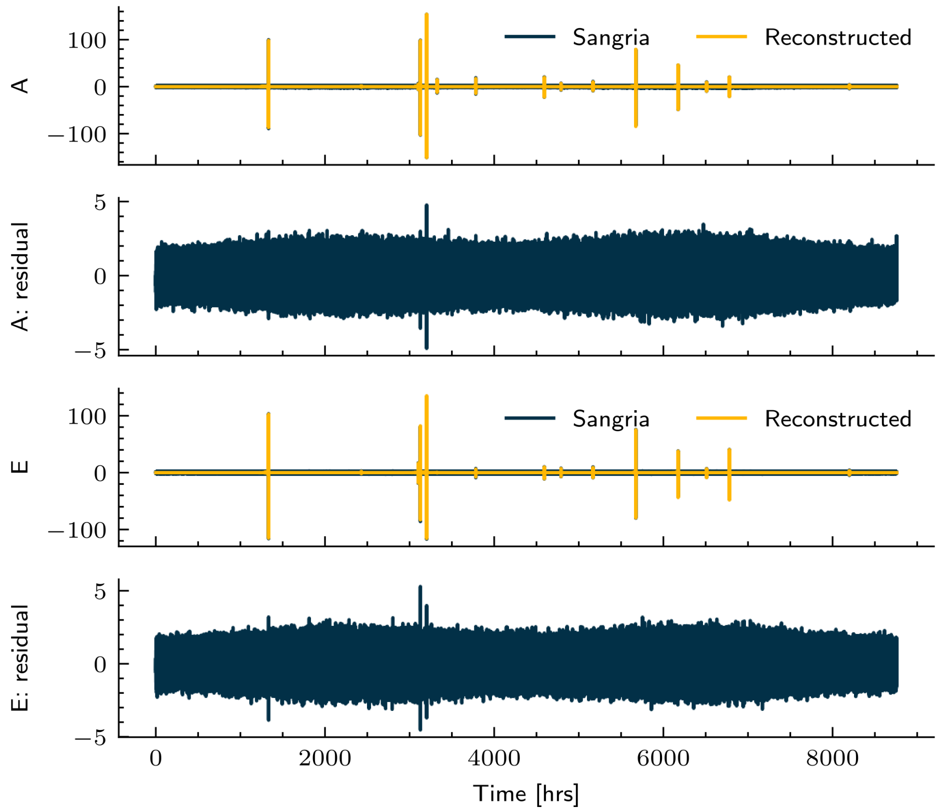


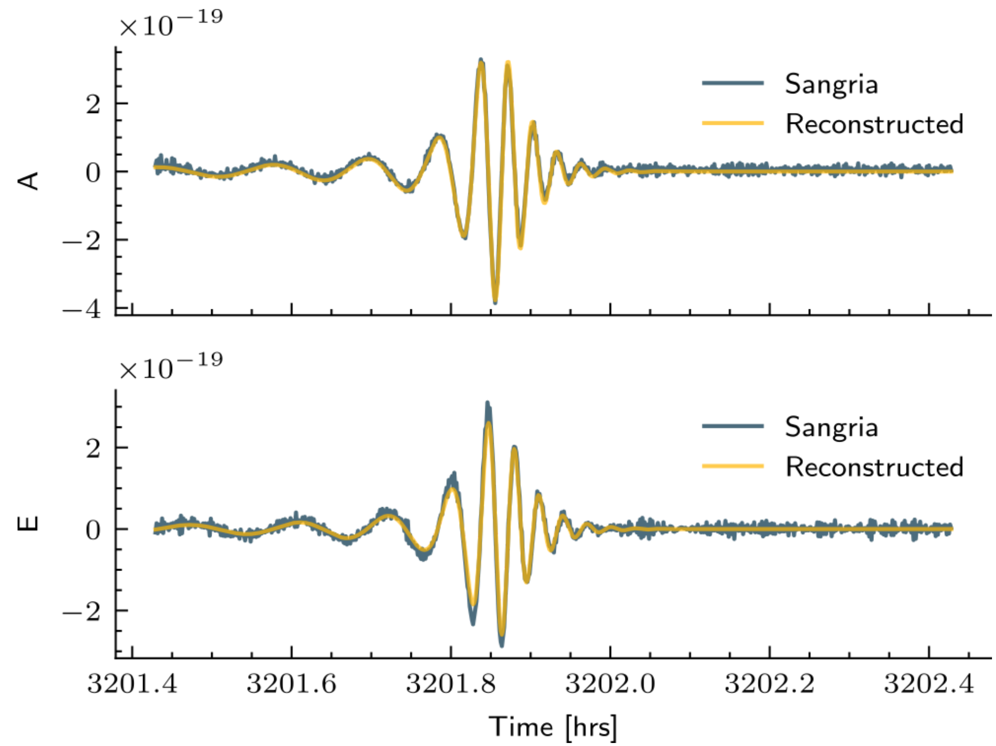




- As the global-fit kick-off
- In the low-latency merger prediction

# "Sangria" - global-fit kick-off

- Mergers are easy to identify
- Analyse two-weeks-long data segments around the mergers

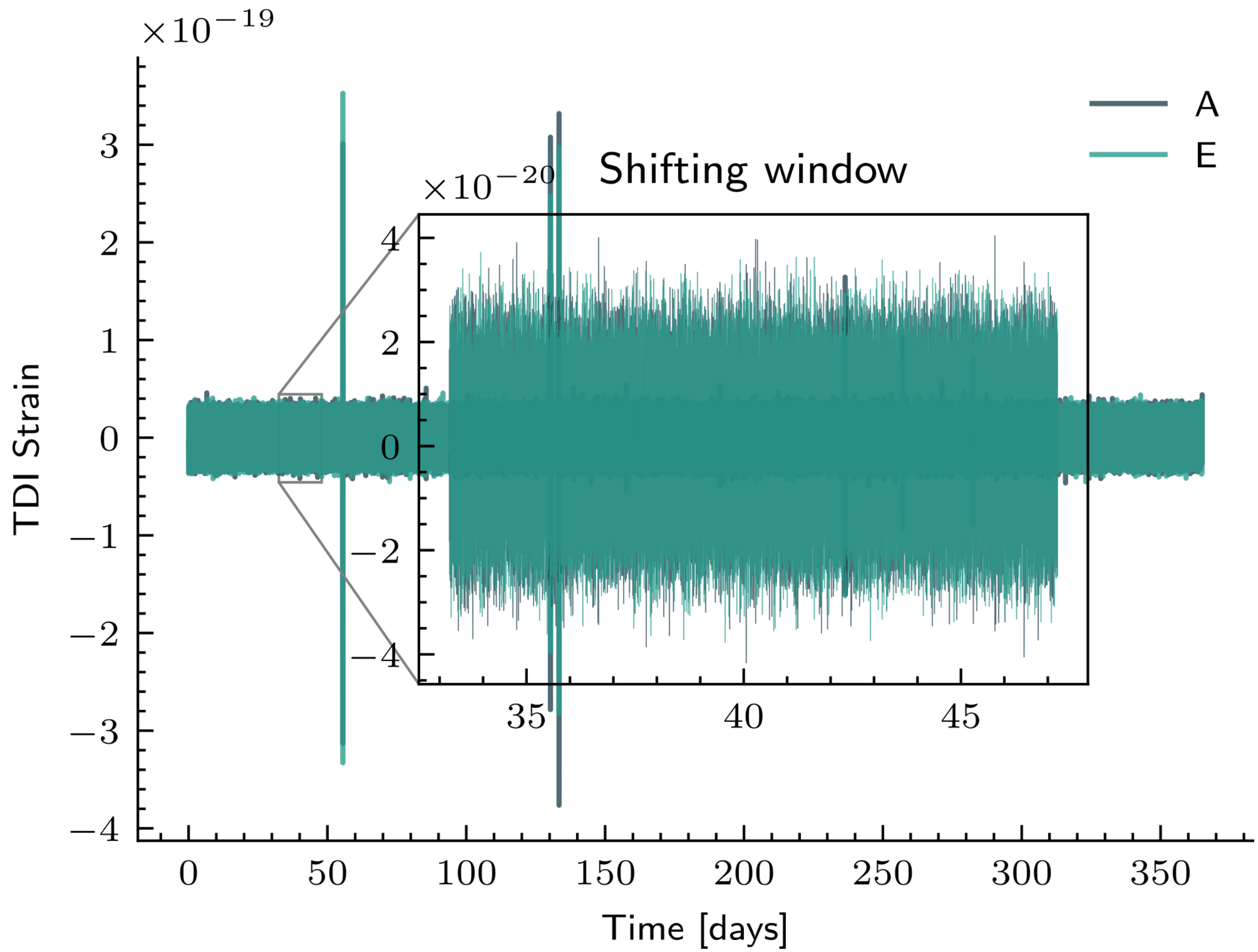



Zoom-in around MBHB 4

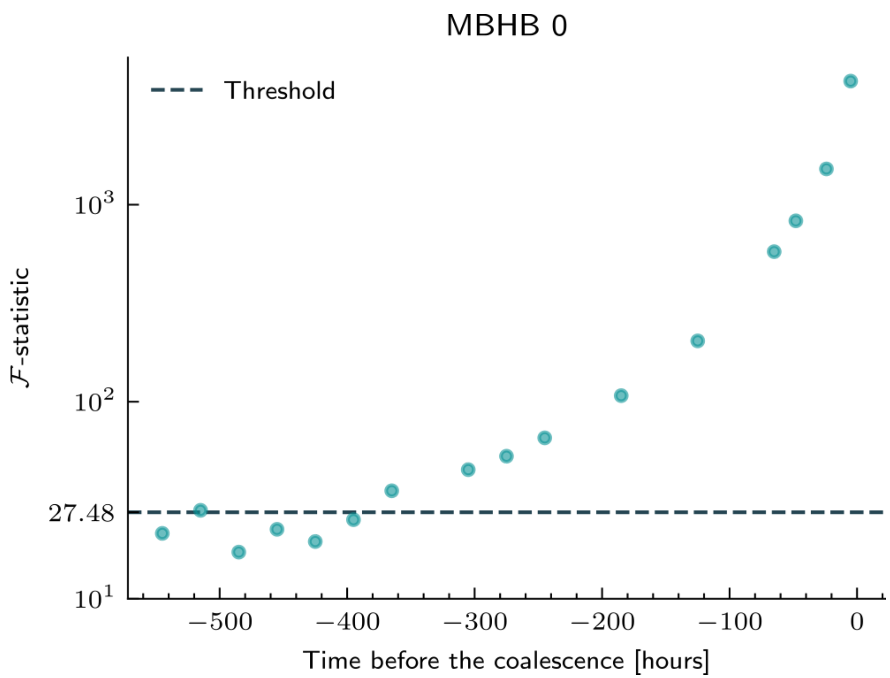
# "Sangria" - merger prediction

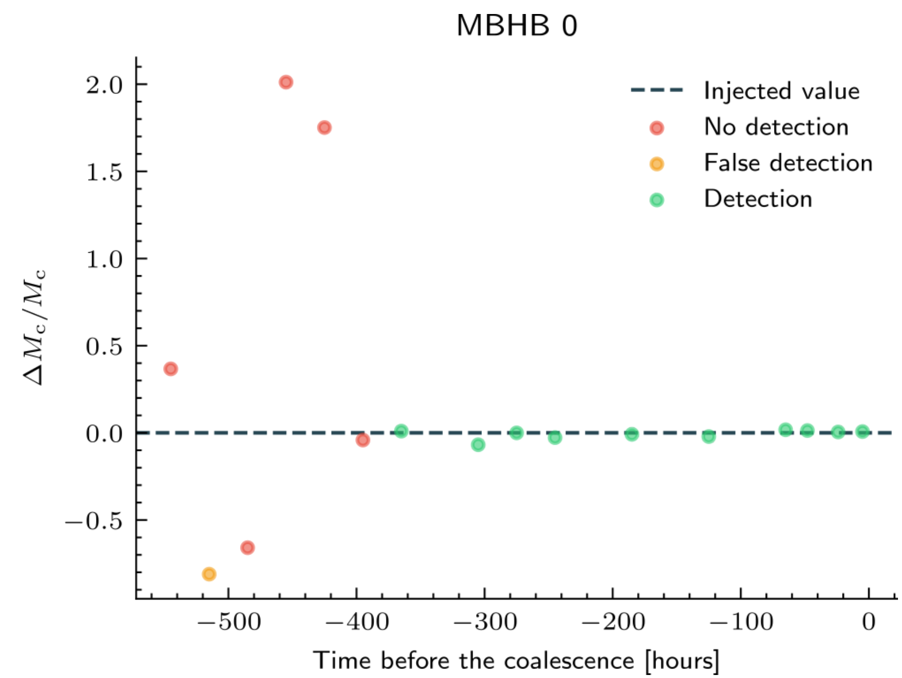
- Remove the GBs resolved by the global-fit pipeline (to satisfy the Gaussianity)

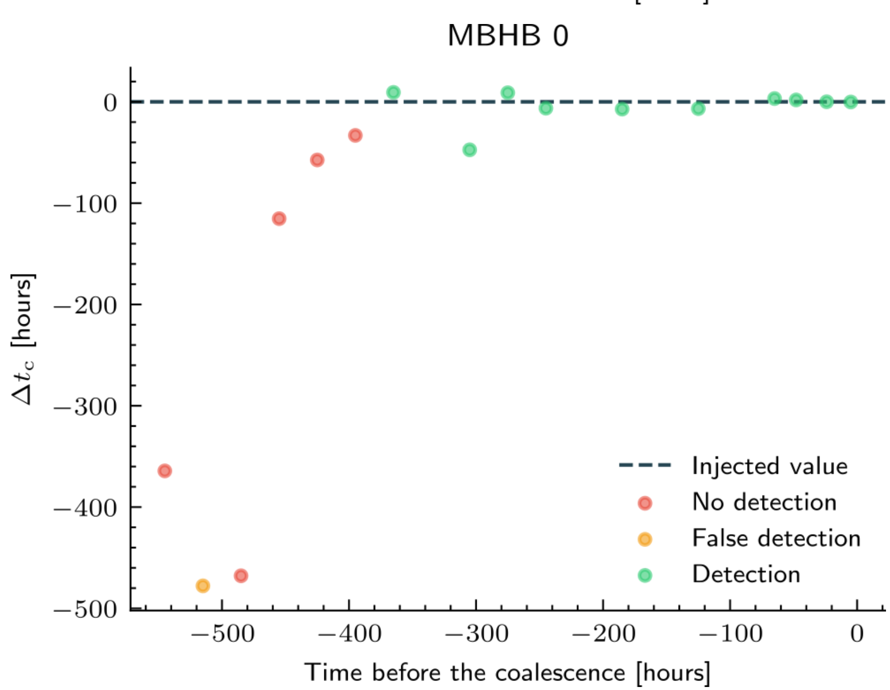



End of the window at 200 hours before the merger of MBHB 0

# "Sangria" - merger prediction




- The threshold corresponds to SNR about 7.5
- Accuracy on $t_c$ comparable to existing machine learning methods

# "Sangria" - merger prediction

| Dist. [h] | $\mathcal{F}$-statistic | $\Delta M_c/M_c$ | $\Delta t_c$ [h] |
|---|---|---|---|
| MBHB 2<br>$M_c$=1 199 474.45 $M_\odot$, $t_c$=3102.17 h<br>SNR=339.6, E.D.H.=23.13 h | | | |
| 545.0 | 36.51 | -39.46% | 54.49 |
| 455.0 | 35.82 | -35.16% | 135.74 |
| 365.0 | 53.40 | -39.26% | 54.25 |
| 275.0 | 80.19 | -34.67% | 115.34 |
| 245.0 | 106.80 | -39.10% | 62.87 |
| 185.0 | 135.87 | -39.64% | 59.40 |
| 125.0 | 217.44 | -34.98% | 102.81 |
| 65.0 | 406.94 | -34.12% | 106.60 |
| 48.0 | 443.37 | -34.11% | 107.58 |
| 24.0 | 582.82 | -41.68% | 26.03 |
| 5.0 | 838.17 | -41.22% | 26.10 |
| MBHB 3<br>$M_c$=696 728.57 $M_\odot$, $t_c$=3127.44 h<br>SNR=2098.6, E.D.H.=719.59 h | | | |
| 545.0 | 41.59 | 2.59% | 54.77 |
| 455.0 | 45.90 | 8.98% | 97.54 |
| 365.0 | 57.78 | 1.45% | 15.73 |
| 275.0 | 87.01 | 4.70% | 34.92 |
| 245.0 | 120.65 | 12.99% | 84.22 |
| 185.0 | 152.34 | 7.64% | 61.25 |
| 125.0 | 262.34 | 12.79% | 83.86 |
| 65.0 | 527.15 | 9.44% | 70.98 |
| 48.0 | 618.30 | 12.28% | 78.95 |
| **24.0** | **1031.57** | **1.70%** | **1.21** |
| **5.0** | **2357.42** | **0.64%** | **-0.07** |
| MBHB 4<br>$M_c$=772 462.86 $M_\odot$, $t_c$=3201.86 h<br>SNR=3356.0, E.D.H.=1173.19 h | | | |
| **545.0** | **44.10** | **-0.96%** | **40.28** |
| **455.0** | **79.46** | **1.54%** | **13.34** |
| **365.0** | **102.94** | **0.08%** | **16.68** |
| **275.0** | **181.50** | **0.76%** | **6.91** |
| **245.0** | **189.30** | **-4.03%** | **-21.33** |
| **185.0** | **348.68** | **2.17%** | **9.13** |
| **125.0** | **590.48** | **1.08%** | **2.99** |
| **65.0** | **1226.59** | **3.02%** | **5.42** |
| **48.0** | **1876.36** | **-0.12%** | **-1.34** |
| **24.0** | **3237.47** | **0.99%** | **0.04** |
| **5.0** | **6758.56** | **0.35%** | **0.01** |

- E.D.H. stands for "Expected Detection Horizon"
  - How well in advance do we accumulate SNR to 7.5?
- MBHB 2, MBHB 3 and MBHB 4 are close to each other
  - MBHB 4 is low-mass and loud
  - In many shifting windows, its inspiral part has the highest SNR → it is detected
- Strategy for the shortly separated mergers
  - Detect the loudest one → reconstruct and subtract it → loop over the residual

# "Sangria" - merger prediction

| Dist. [h] | $\mathcal{F}$-statistic | $\Delta M_c/M_c$ | $\Delta t_c$ [h] |
|---|---|---|---|
| MBHB 2<br>$M_c$=1 199 474.45 $M_\odot$, $t_c$=3102.17 h<br>SNR=339.6, E.D.H.=23.13 h | | | |
| 545.0 | 28.54 | -41.97% | 42.80 |
| 455.0 | 30.48 | -42.02% | 37.12 |
| 365.0 | 48.38 | -43.70% | -15.61 |
| 275.0 | 70.21 | -41.06% | 43.83 |
| 245.0 | 74.25 | -40.93% | 37.33 |
| 185.0 | 96.27 | -40.66% | 35.89 |
| 125.0 | 175.40 | -43.15% | 18.25 |
| 65.0 | 334.54 | -41.74% | 24.65 |
| 48.0 | 413.57 | -40.43% | 30.72 |
| 24.0 | 580.53 | -34.86% | 103.69 |
| 5.0 | 682.74 | -34.58% | 103.37 |
| MBHB 3<br>$M_c$=696 728.57 $M_\odot$, $t_c$=3127.44 h<br>SNR=2098.6, E.D.H.=719.59 h | | | |
| **545.0** | **28.94** | **-3.34%** | **-45.32** |
| **455.0** | **40.70** | **-3.68%** | **-21.32** |
| **365.0** | **53.45** | **9.49%** | **111.08** |
| **275.0** | **66.06** | **0.46%** | **4.66** |
| **245.0** | **71.63** | **-2.91%** | **-8.89** |
| **185.0** | **116.49** | **-5.17%** | **-21.25** |
| **125.0** | **214.68** | **3.10%** | **8.42** |
| **65.0** | **476.61** | **1.49%** | **2.24** |
| **48.0** | **590.19** | **0.25%** | **0.25** |
| 24.0 | / | / | / |
| 5.0 | / | / | / |

After subtracting MBHB 4 (reconstructed independently for each window)

- E.D.H. stands for "Expected Detection Horizon"
  - How well in advance do we accumulate SNR to 7.5?
- MBHB 2, MBHB 3 and MBHB 4 are close to each other
  - MBHB 4 is low-mass and loud
  - In many shifting windows, its inspiral part has the highest SNR → it is detected
- Strategy for the shortly separated mergers
  - Detect the loudest one → reconstruct and subtract it → loop over the residual

# "Sangria" - merger prediction

| Dist. [h] | $\mathcal{F}$-statistic | $\Delta M_c/M_c$ | $\Delta t_c$ [h] |
|---|---|---|---|
| MBHB 2<br>$M_c$=1 199 474.45 $M_\odot$, $t_c$=3102.17 h<br>SNR=339.6, E.D.H.=23.13 h | | | |
| 545.0 | 21.07 | -97.14% | -452.51 |
| 455.0 | 17.45 | -76.29% | -398.62 |
| 365.0 | 16.64 | -74.63% | 127.08 |
| 275.0 | 16.04 | -74.85% | 75.72 |
| 245.0 | 16.35 | -87.62% | -217.19 |
| 185.0 | 17.54 | -73.58% | 414.94 |
| 125.0 | 16.08 | -95.82% | 101.08 |
| 65.0 | 19.05 | 100.78% | -45.89 |
| 48.0 | 20.91 | 1.39% | 2.10 |
| **24.0** | **36.29** | **-9.21%** | **-8.05** |
| **5.0** | **44.48** | **-3.75%** | **-2.07** |

After subtracting MBHB 4 and then MBHB 3 (reconstructed independently for each window)

- E.D.H. stands for "Expected Detection Horizon"
  - How well in advance do we accumulate SNR to 7.5?
- MBHB 2, MBHB 3 and MBHB 4 are close to each other
  - MBHB 4 is low-mass and loud
  - In many shifting windows, its inspiral part has the highest SNR → it is detected
- Strategy for the shortly separated mergers
  - Detect the loudest one → reconstruct and subtract it → loop over the residual

# "Sangria" - merger prediction

| Dist. [h] | $\mathcal{F}$-statistic | $\Delta M_c/M_c$ | $\Delta t_c$ [h] |
|---|---|---|---|
| MBHB 4<br>$M_c$=772 462.86 $M_\odot$, $t_c$=3201.86 h<br>SNR=3356.0, E.D.H.=1173.19 h | | | |
| 545.0 | 17.87 | -94.63% | 43.27 |
| 455.0 | 19.39 | -59.90% | 46.71 |
| 365.0 | 16.61 | -97.94% | -271.81 |
| 275.0 | 16.25 | -94.20% | -19.34 |
| 245.0 | 16.76 | -94.21% | -18.53 |
| 185.0 | 16.12 | -90.90% | 103.71 |
| 125.0 | 17.64 | -91.06% | 98.52 |
| 65.0 | 18.29 | 110.99% | 135.76 |
| **48.0** | **36.61** | **208.54%** | **-45.16** |
| **24.0** | **60.57** | **489.50%** | **-22.50** |
| **5.0** | **554.51** | **-1.83%** | **-1.71** |

After subtracting MBHB 4 (reconstructed independently for each window)

- E.D.H. stands for "Expected Detection Horizon"
  - How well in advance do we accumulate SNR to 7.5?
- MBHB 2, MBHB 3 and MBHB 4 are close to each other
  - MBHB 4 is low-mass and loud
  - In many shifting windows, its inspiral part has the highest SNR → it is detected
- Strategy for the shortly separated mergers
  - Detect the loudest one → reconstruct and subtract it → loop over the residual
  - !! Residuals could also be detected
    - Need for more accurate reconstruction
    - Deep analysis ⇒ sky localization

Thanks for your attention!