ETH zürich

cWB Space : Towards a morphology independent transient search pipeline for LISA*

*Are you worried! That 2-body problem is not solvable in GR, given that binaries in strong field is our main observational target? There can be beyond GR features, environmental effects or third body interactions as well .... can we detect them with LISA? There can be surprises and totally new events in LISA data.... how would we find these things ?

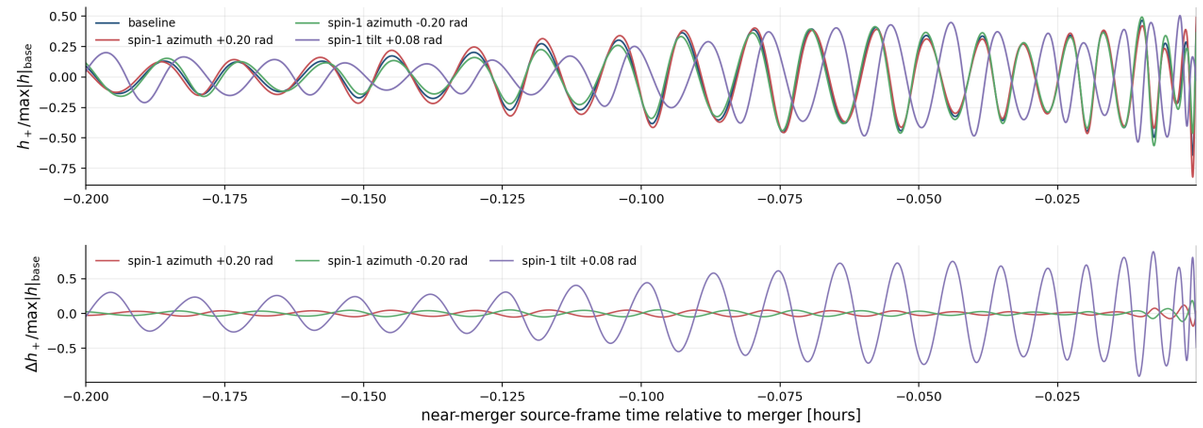
Shubhanshu Tiwari, UIB Palma, LISA Spain meeting, 2026

Based on *cWB-space: A time-frequency unmodelled transient-search pipeline for LISA* Shubhanshu Tiwari, Yumeng Xu, Giovanni A. Prodi e-Print: 2609.12986

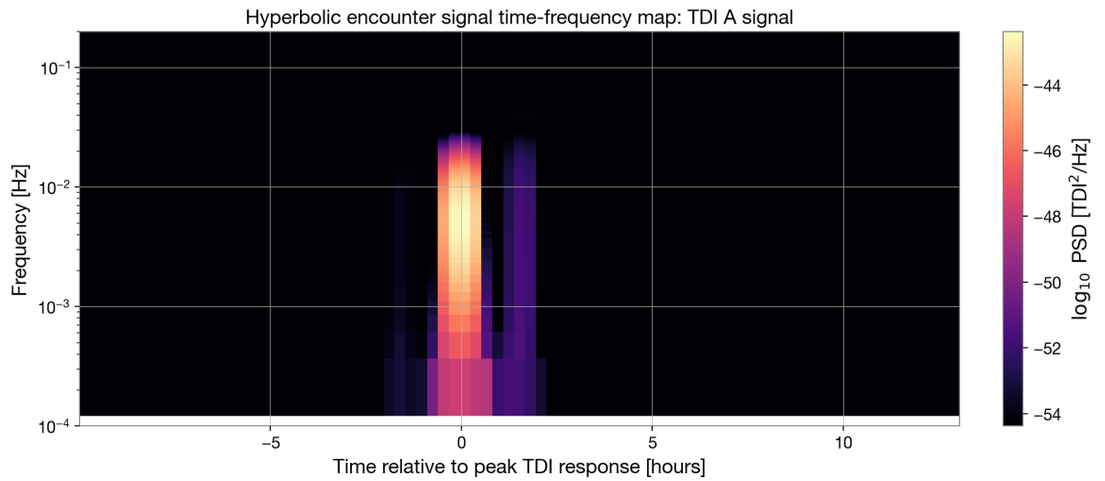
Why/what is a morphology independent pipeline ?

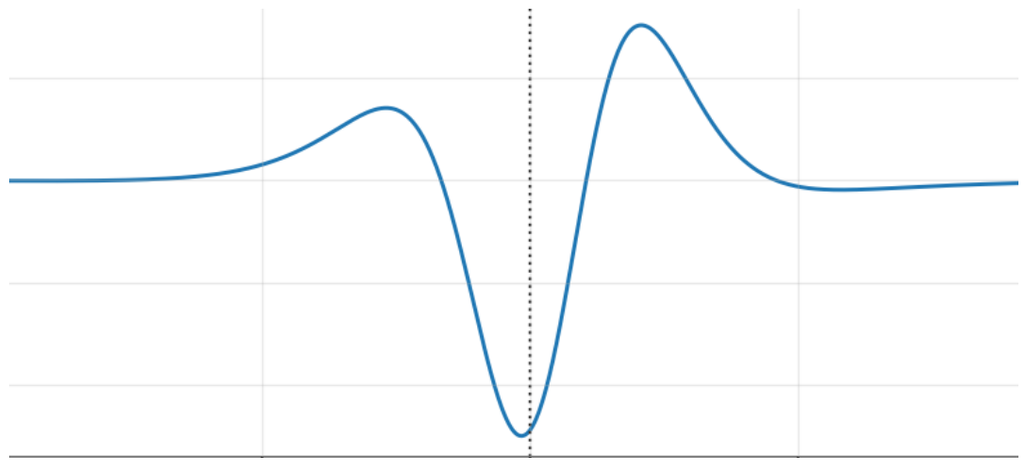
If you were not worried yet.....

- LISA will observe a large diversity of transient sources, the loudest of them will be massive black hole binaries (MBHBs)
  - MBHB waveforms are build from carefully patching various approaches and they show remarkable accuracy and speed but still they are not fully devoid of systematics
  - MBHBs will be very loud with SNRs reaching 1000s, current LVK detection are limited to SNRs below 100, so this its a new regime where there can be “extra effects”
- LISA can also observe the extreme mass ratio inspirals with one component massive black hole and the other stellar or intermediate mass black hole
  - The systems are notorious to model/analyze and a small perturbations due to environment or beyond GR or systematics can cause a very different phase evolution of the system
- LISA will also be sensitive to a class of cosmic strings interactions, hyperbolic encounters etc



IMRPhenomTPHM via phenomxpy; same M=1e5 Msun, eta=0.18, DL=6230 Mpc, iota=1.35 rad, dt=0.1 s. Only the spin-vector precession angles are perturbed; no smoothing is applied.

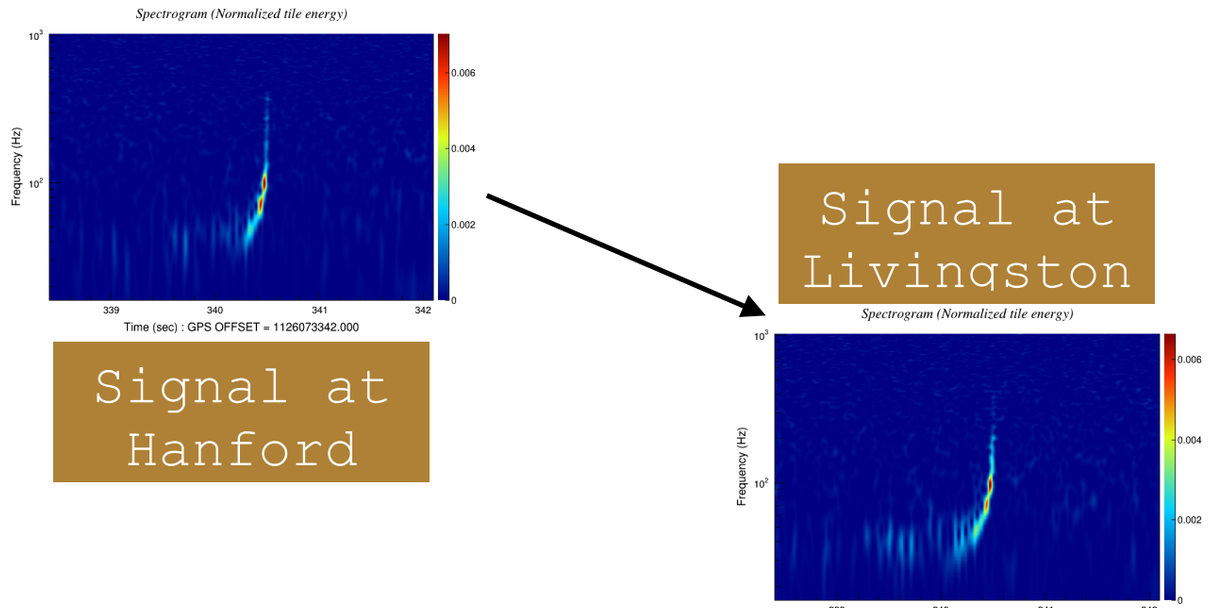


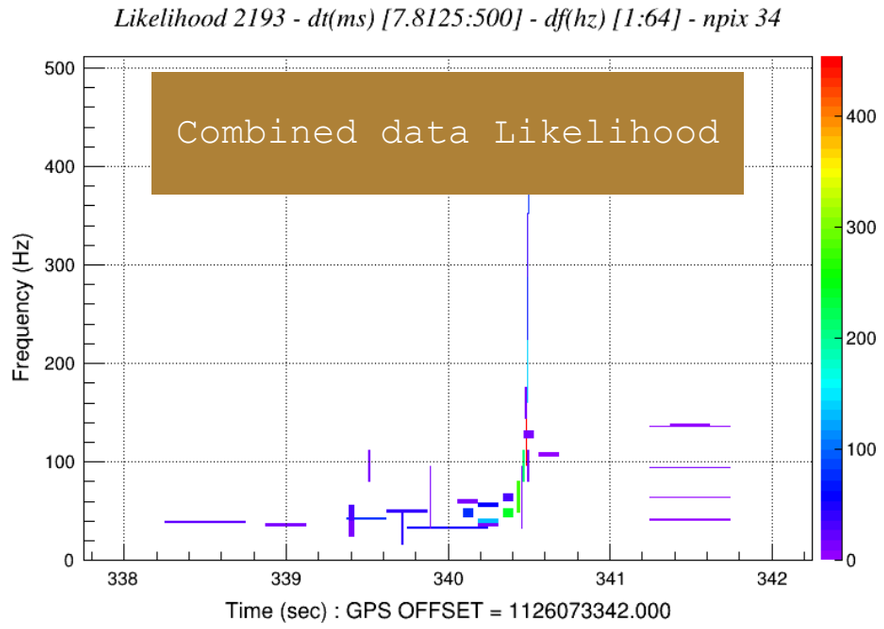


Why/what is an morphology independent pipeline ?

If you were not worried yet.....

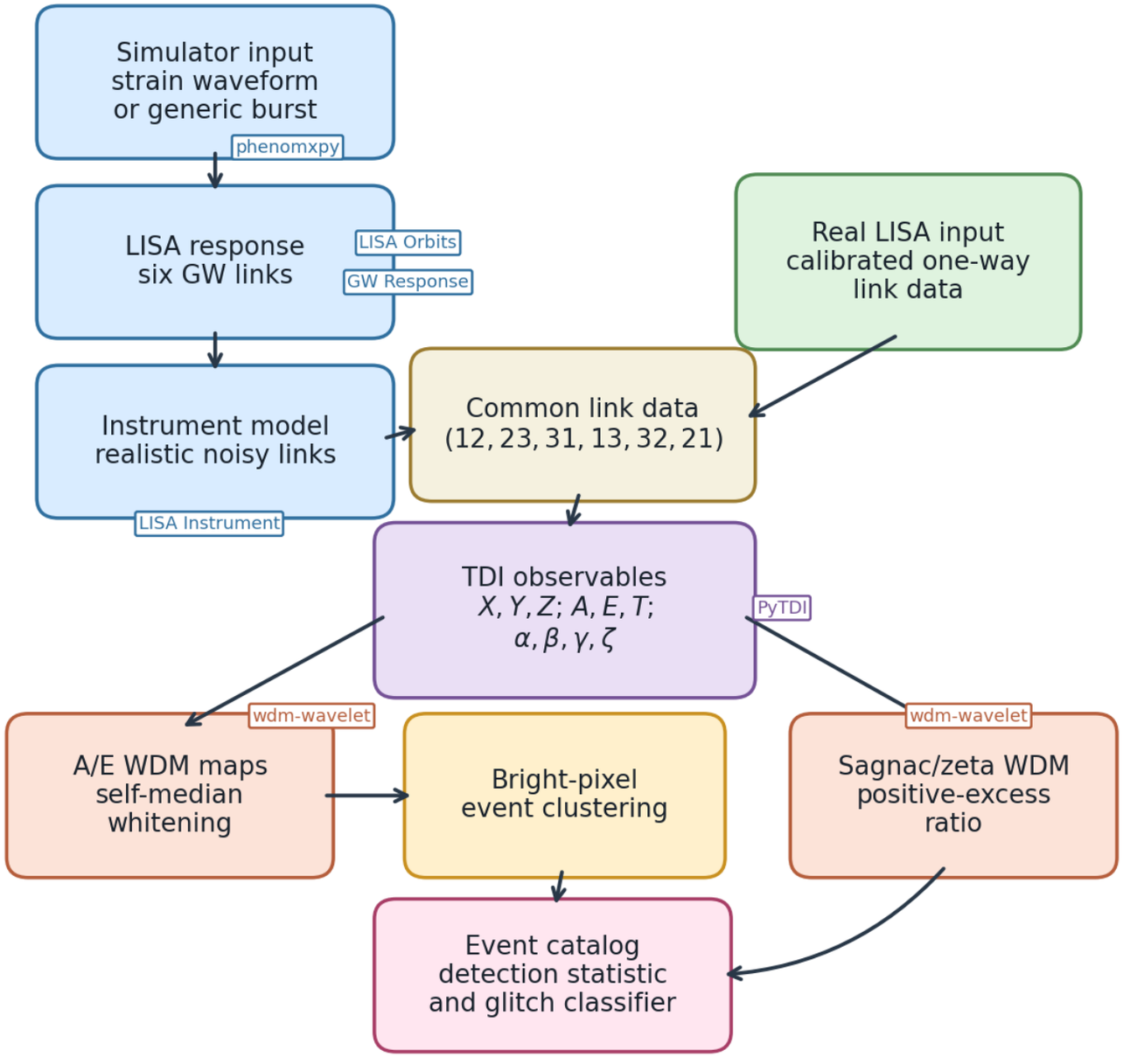
- Indeed the first gravitational wave observation GW150914 was detected first by a morphology independent (also known as bursts) pipeline
  - Broadly bursts pipelines work by not relying on the assumption of the underlying morphology* of the incoming signal, sky direction, polarisation etc
  - They build a detection statistics based on excess power above the noise floor, which is “coherent” in different data channels
- For e.g. coherent WaveBursts in LVK work by noting the excess power above the detector noise floor and compute the coherence as a function of sky direction using multiple detectors as different data channels





cWB-space : building cWB for LISA

cWB-space accepts either simulated link products or calibrated real one-way LISA links.

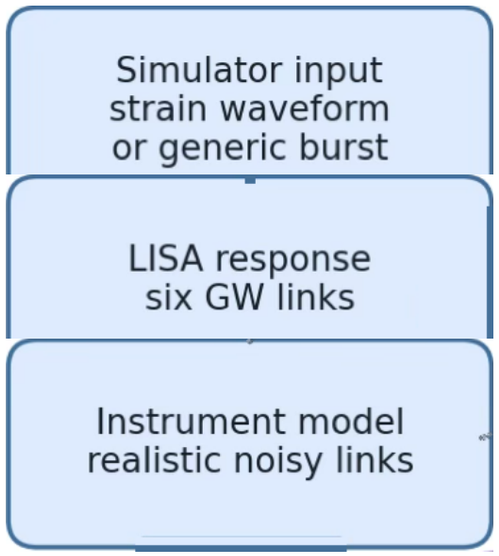


Search
Find concentrations of power in time and frequency, then rank the candidates.

Glitch study
Compare a candidate across combinations of LISA measurements to investigate its origin.

Reconstruction
Retain the candidate region and invert the wavelet transform to recover a waveform.

cWB-space : building cWB for LISA



- **Simulator**: Generate common time series grid with pre defined burn-in (currently 600 sec); create MBHB polarizations h_+,h_x using phenomxpy, LISABurstsWaveform (can accept any waveform generator); peak-align, interpolate, and zero-pad each source onto its injection day (remove artifacts of simulated data generation)

- **LISA response**: Project h_+,h_x through lisagwresponse using ESA trailing lisaorbits, sky position (ra,dec), and six one-way links convert y_{ij} to Hz via the relation $2.816 \times 10^{14}\ y_{ij}$.

- **Instrument/noise model**: Generate link-level measurements with lisainstrument using same orbit/grid, fixed seeds, noise floors, optional GW response file, and optional lisaglitch readout injections.

Simulator/instrument: *Bayle & Hartwig (2023); Baghi et al. (2022); Spadaro et al. (2023).*

MBHB waveforms: *Pratten et al. (2021); García-Quirós, Tiwari & Babak (2025).*

Xu, Tiwari et al LISA Bursts Waveform generator

cWB-space : building cWB for LISA

TDI observables
$X, Y, Z; A, E, T;$
$\alpha, \beta, \gamma, \zeta$

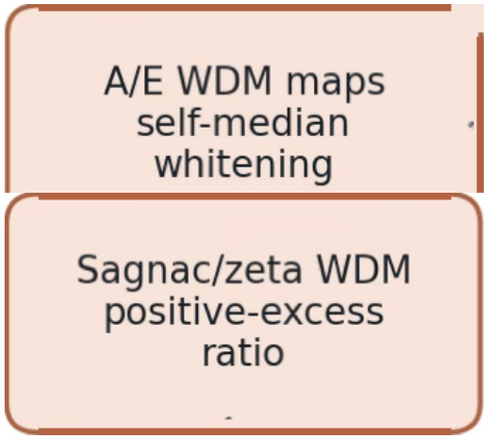
- **Michelson XYZ:** Build second-generation X2,Y2,Z2 with pytdi from either noisy instrument HDF5 streams or GW-only response files, using orbit delays, sampling rate, interpolation order and optional delay derivatives

- **A/E/T:** Rotate Michelson channels into A2=(Z2-X2)/sqrt(2), E2=(X2-2Y2+Z2)/sqrt(6), and T2=(X2+Y2+Z2)/sqrt(3), then crop burn-in/edge samples and band-limit to 1e-5 <= f <= 1e-1 Hz.

- **Sagnac/Zeta:** Compute alpha2, beta2, gamma2 and zeta21, zeta22, zeta23 from the same delayed link measurements with the consistent interpolation, cropping, and band-limiting conventions.

Based on : Tinto & Dhurandhar (2005); Dhurandhar, Nayak & Vinet (2002); Prince et al. (2002); Muratore et al. (2022); Hartwig & Muratore (2022); Tinto, Dhurandhar & Malakar (2023).

cWB-space : building cWB for LISA



- **WDM maps:** Transform each TDI streams $d_c(t)$ into complex Wilson-Daubechies-Meyer coefficients using wdm-wavelet, then analyze pixel power in the band, here we use two resolutions $W_c(t_i, f_j) = \mathscr{W}[d_c](t_i, f_j), Q_c(t_i, f_j) = |W_c(t_i, f_j)|^2$

- **Whitening / self-normalization:** Normalize each WDM frequency row by a median noise floor $B_c(f_j) = \text{median}_i |W_c(t_i, f_j)|^2, P_c(t_i, f_j) = \frac{|W_c(t_i, f_j)|^2}{B_c(f_j)}$. So the data is whitened with signal in it, since the assumption is that the signal is transient the presence of rare loud event in data stream does not effect the noise floor computation, this is also the reason why we use the median and not mean to avoid large outliers.

- **Zeta-WDM (R_ζ) :** Compute WDM power in the channels alpha2, beta2, gamma2, zeta21, zeta22, zeta23. Then we subtract the noise WDM power in each channel, keeping only positive whitened residual power, then compare the total normalized excess in the three zeta channels with the total normalized excess in the three alpha/beta/gamma channels. Low R_ζ is GW-like, while order-unity R_ζ indicates link-localized instrumental structure. In summary we are trying to capture if the loud signal is appearing only in one link or is showing in all the links.

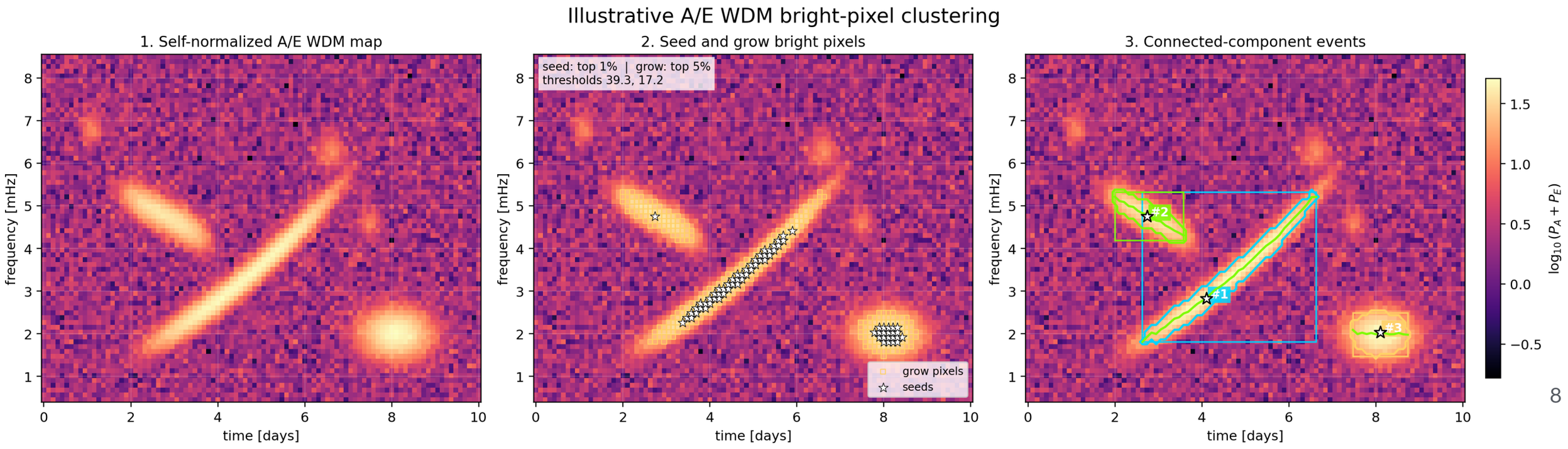
WDM introduced by : Klimenko, Necula, Vedovato, Prodi et al for the coherent wave bursts pipeline

Package used here is the wdm-wavelet

cWB-space : building cWB for LISA

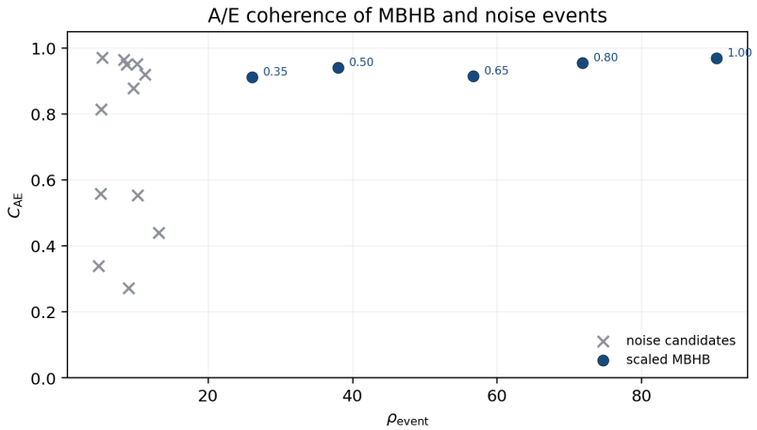
Bright-pixel event clustering

- **Bright-pixel selection:** Use the self-normalized A/E WDM discovery map; pixels above the 99th percentile seed events, and connected pixels above the 95th percentile are allowed to grow around each seed.
- **Connected-component growth:** Start from the loudest unused seed and grow through the 95th-percentile mask with model independent clustering: for the chirping case we have the rule which connects nearby frequency bins at the same time and allow larger frequency jumps between adjacent time bins to preserve chirp-like tracks.
- **Event filtering and ranking:** Reject components with fewer than eight pixels or fewer than two occupied time bins, then rank surviving events using total A+E excess power, A/E balance, and chirp-like centroid evolution.



cWB-space : building cWB for LISA

Event catalog
detection statistic
and glitch classifier



- **Detection statistic:** For each surviving event mask, sum the self-normalized A and E powers to get total A+E excess, penalize events that are bright in only one channel through the A/E balance factor, and multiply by a chirp-morphology factor based on whether the event's frequency centroid mostly increases with time $\rho_{\mathrm{event}} = \sqrt{E_{\mathrm{AE}}}\sqrt{B_{\mathrm{AE}}}\left(\frac{1+m_{\mathrm{chirp}}}{2}\right)$

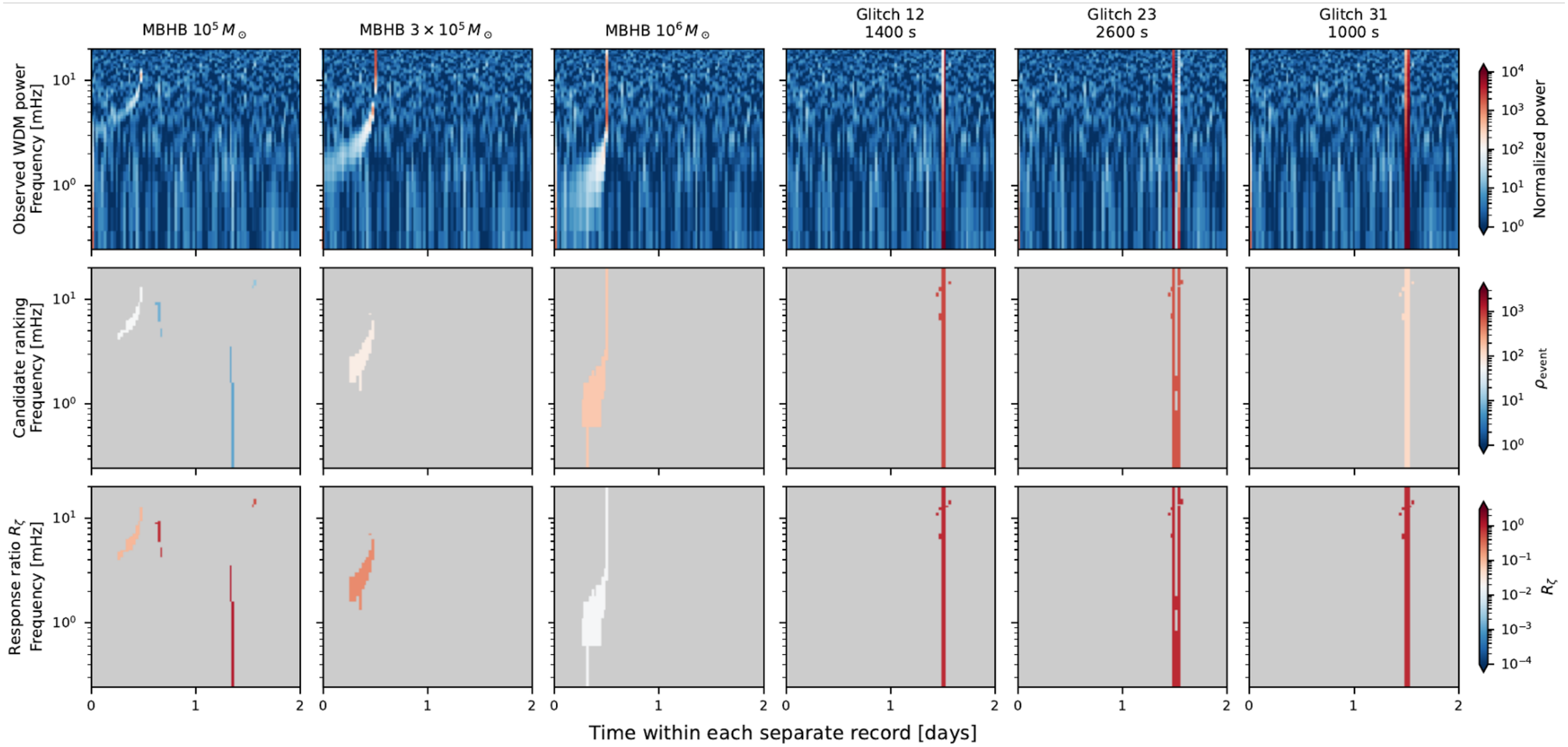
  here $(E_{\mathrm{AE}} = E_A + E_E), (B_{\mathrm{AE}} = 4E_A E_E/(E_A + E_E)^2)$ and mchirp is a measure of the quality of chirp

- **A/E coherence:** Compute complex A/E coherence on the same selected event pixels using the whitened WDM coefficients; high coherence means the event has a phase-consistent excess in both A and E, while low coherence suggests a less GW-like or less consistently projected structure.

cWB-space : injection set and recovery

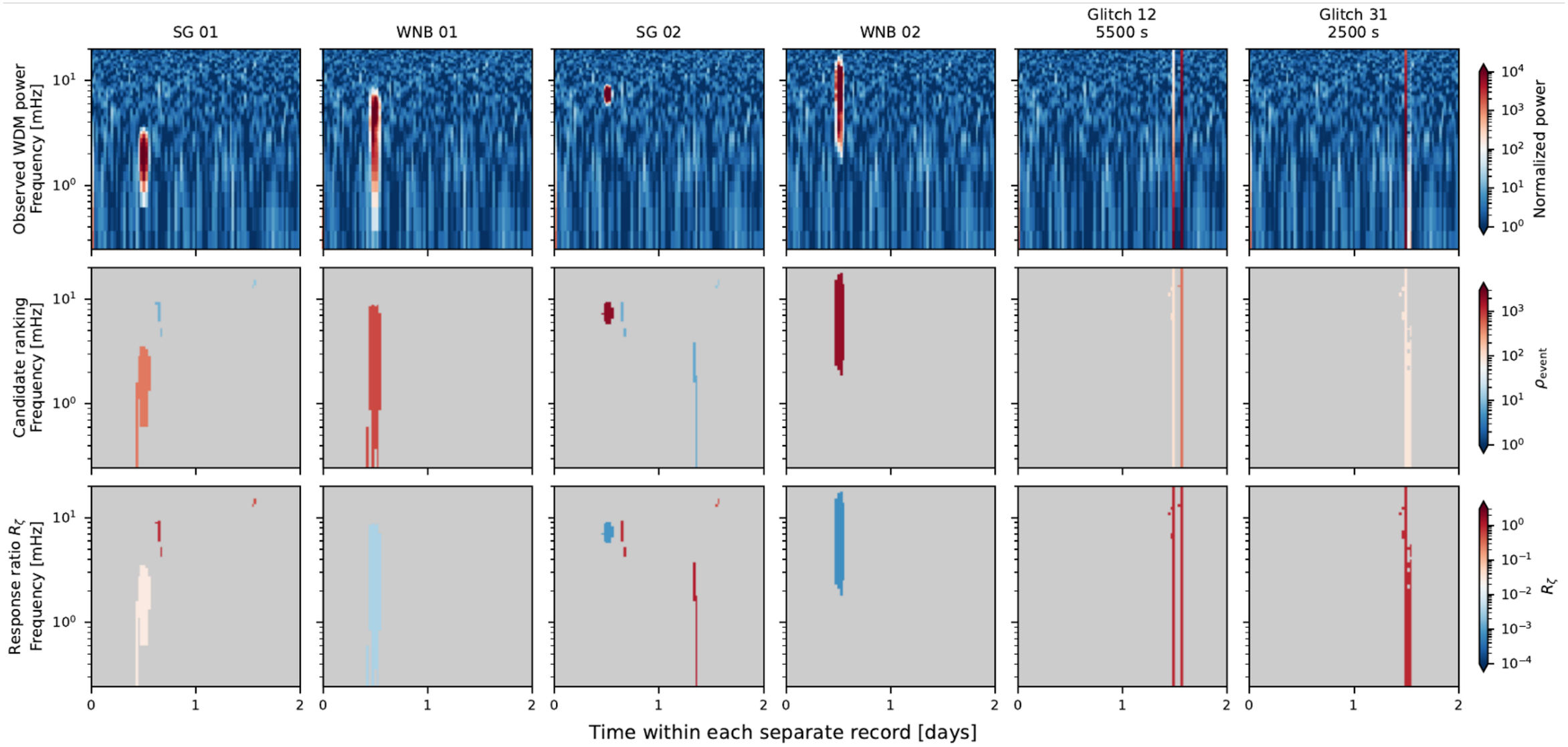
Injection	**Mass/ frequency**	**Shape / width**	**Distance/amplitude**
MBHB, q = 1	10^5 $M_\odot$	Aligned spins	6.23 Gpc
MBHB, q = 0.6	3×10^5 $M_\odot$	Aligned spins	6.23 Gpc
MBHB, q = 0.3	10^6 $M_\odot$	Aligned spins	6.23 Gpc
Sine-Gaussian 01	2.2 mHz	9000 s	8×10^{-18} strain
White-noise burst 01	1.2-6 mHz	11000 s	7×10^{-18} strain
Sine-Gaussian 02	7.5 mHz	6500 s	8×10^{-18} strain
White-noise burst 02	3-12 mHz	8000 s	7.5×10^{-18} strain
Glitch, link 12	Primary study	1400 s	1.295×10^{-3} Hz
Glitch, link 23	Primary study	2600 s	6.297×10^{-3} Hz
Glitch, link 31	Primary study	1000 s	2.186×10^{-3} Hz
Glitch, link 12	Burst study	5500 s	1.294×10^{-3} Hz
Glitch, link 31	Burst study	2500 s	5.124×10^{-3} Hz

MBHBs injections



- The first MBHB appears alongside weaker noise clusters. Its response ratio is smaller than those of the surrounding fluctuations.
- MBHB R_ζ: 0.0967, 0.201, 0.0165. The displayed link glitches give 0.802-0.815.

Bursts injections



- Different transient shapes show the expected response contrast in these examples. The GW-burst ratios span 0.000725-0.0272.

- The two displayed link glitches have Rζ = 0.824 and 0.829.

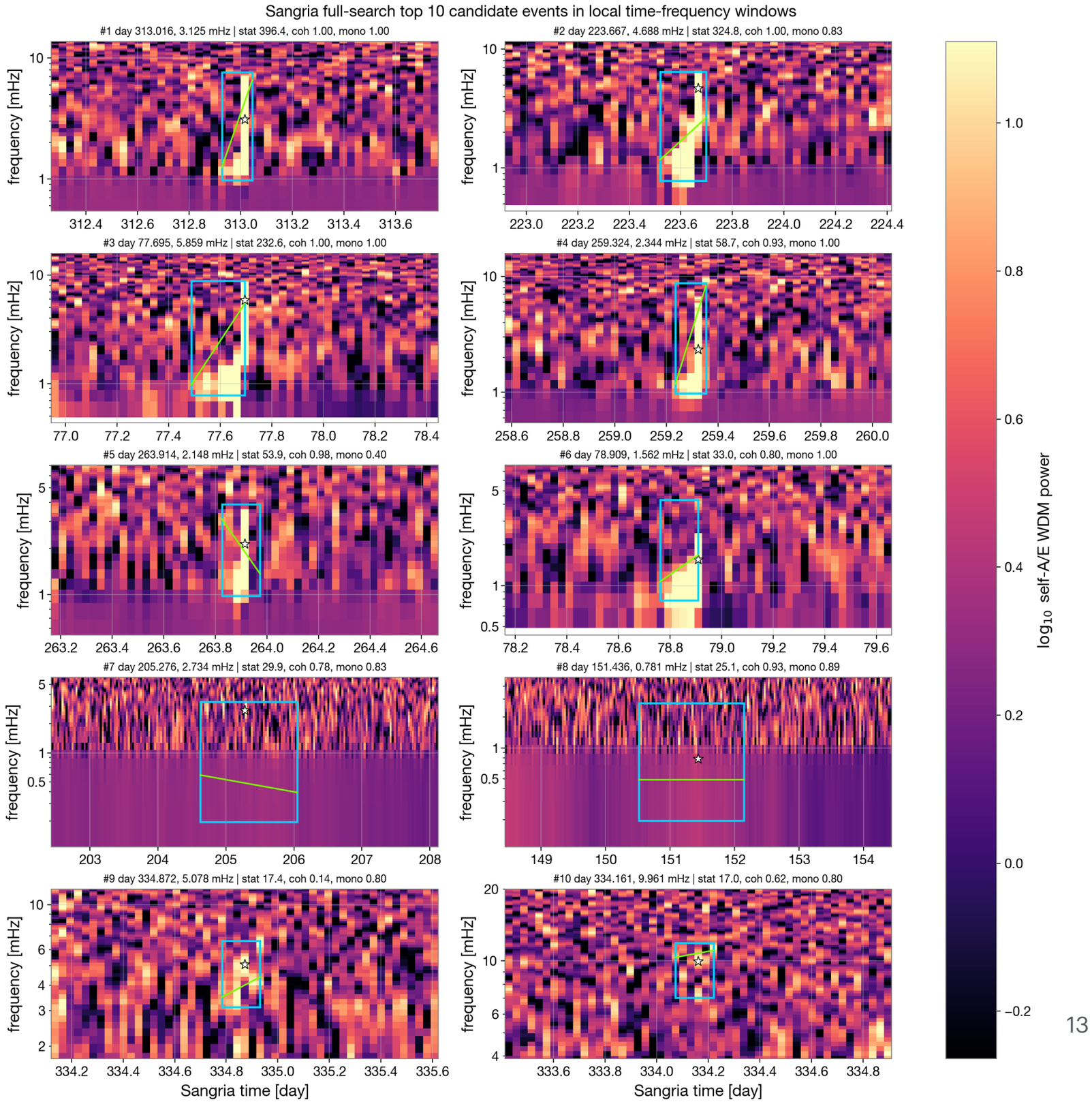
Search in LISA MDC : Sangria

- Below is the ranking of the top 8 events found in the blind search, top 6 are confirmed MHBHs in the unblinded Sangria data,

- We can not compute the R_{\zeta} since the data for this is only available for the AET channels

Rank	Peak day	(\rho_{\rm event})
1	313.016	396.44
2	223.667	324.85
3	77.695	232.61
4	259.324	58.70
5	263.914	53.93
6	78.909	32.96
7	205.276	29.86
8	151.436	25.07

The cyan outline marks the connected event's time-frequency extent.

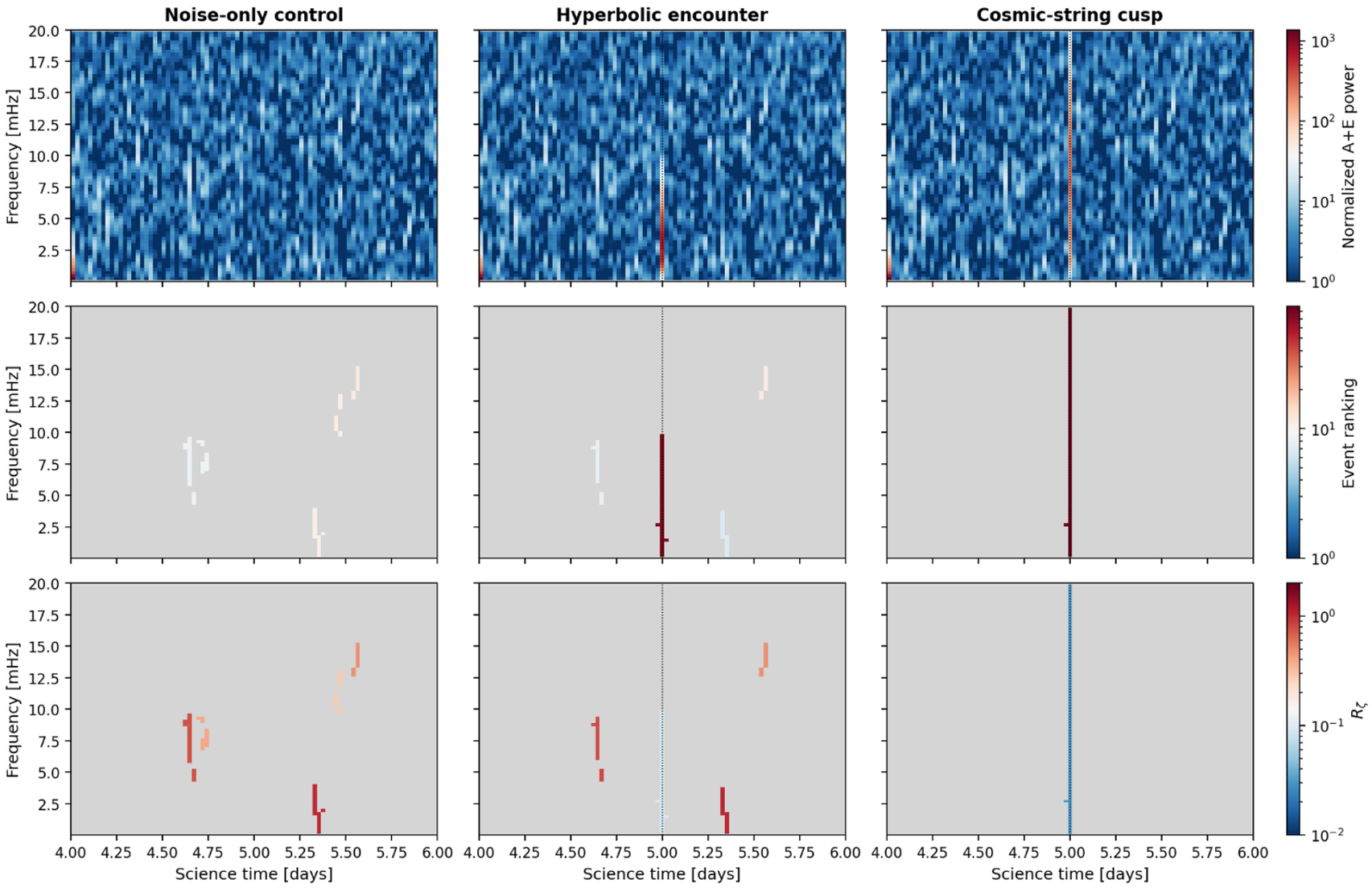


Cosmic Strings and Hyperbolic encounters

- We have a bursts waveform generator as part of the CU-WAV bursts subunit
- It has two source models hyperbolicTD and Cosmic Strings (cusps and kinks)
- Using this we also inject and recover the following systems with cWB-space

Parameter	**Hyperbolic encounter**	**Cosmic-string cusp**
Waveform	Nonspinning unbound encounter	Linearly polarized cusp
Component masses	$m_1 = m_2 = 50{,}000\ M_\odot$	Not applicable
Encounter geometry	Initial e = 1.5, b = 100 GM/c²	Not applicable
Orientation	Inclination = 0.75 rad	$h_\times = 0$ in the source basis
Injected strength	Distance = 58.86 Mpc	$A = 4.9024 \times 10^{-20}$ strain $s^{-1/3}$
Frequency scales	-	0.3 mHz turnover 10 mHz rolloff
Peak strain	2.216×10^{-18}	2.182×10^{-18}

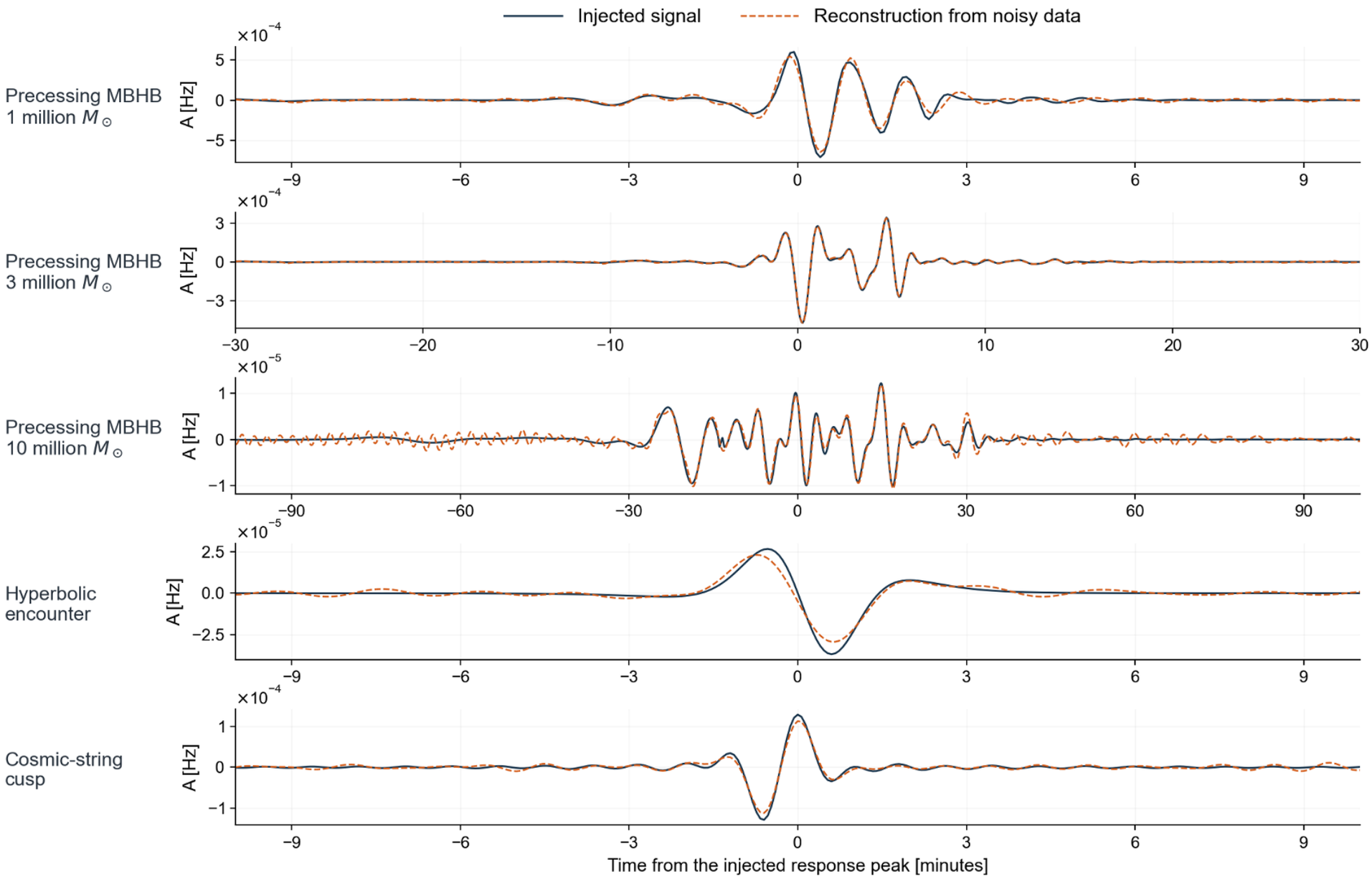
Cosmic Strings and Hyperbolic encounters



- Hyperbolic: ρ = 75.56, Rζ = 0.0934. Cusp: ρ = 87.62, Rζ = 0.0328.

Reconstruction

Some examples of reconstruction of the signal with cWB-space, after detecting the signal we go for higher resolution. Search: K = 32. Waveform follow-up: K = 256.



Conclusion and future work

- cWB-space does transient searches, waveform reconstruction and glitch classification without requiring source waveform templates.

- All six injected Sangria MBHBs are recovered as the six highest-ranked candidates,

- Reconstructions recover waveform structure from aligned-spin and precessing binaries. Detector-response comparisons help distinguish many glitches, although some correlated disturbances can resemble gravitational waves.

- We are working on multi resolution analysis - See Yumeng's talk

- *Establish sensitivity and detection significance*: broaden injection and background studies, including changing noise levels and unresolved foregrounds.

- *Extend the scientific targets*: adapt clustering and ranking to the time-frequency structure of eccentric binaries and EMRIs.

- *Develop practical applications*: test low-latency follow-up and transient screening before Galactic-binary searches within global-fit analyses.

- *Investigate source physics*: compare reconstructed waveforms with model predictions to study unexplained structure, environmental effects and possible departures from general relativity, while assessing instrumental and modeling uncertainties.

Thanks!!