A test for LISA foreground Gaussianity and stationarity
Extreme mass-ratio inspirals

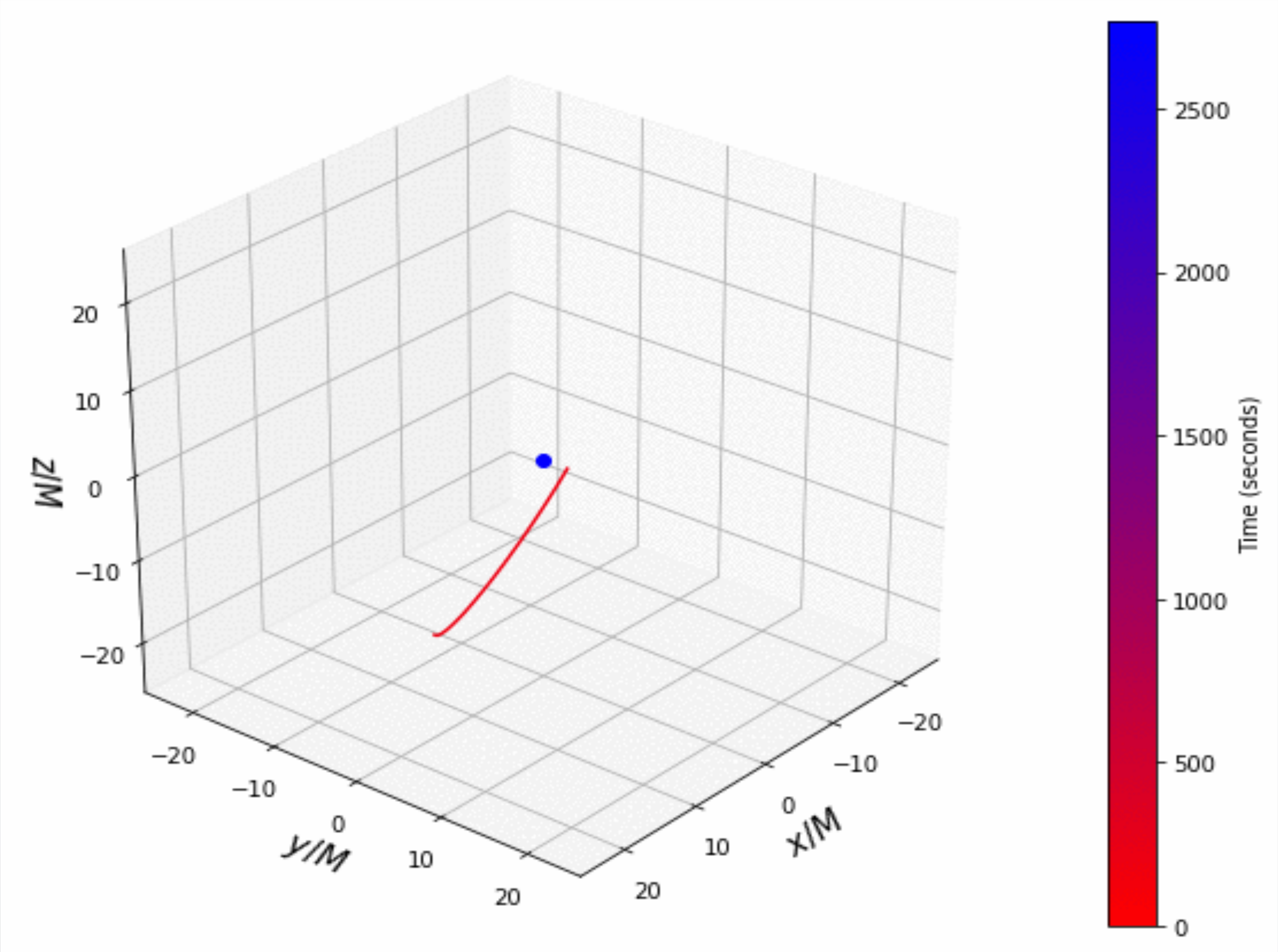
Manuel PIARULLI
Riccardo BUSCICCHIO, Federico POZZOLI,
Ollie BURKE, Matteo BONETTI, Alberto SESANA

Piarulli et al. (2024) [arXiv:2410.08862]

JUN 24, 2025

EMRI Search and Inference within the LISA Global Fit, Paris

Extreme Mass Ratio Inspirals (EMRIs)



Credits: Ollie Burke

not all EMRIs are expected to be loud enough to be individually detected

$\Downarrow$

GWB

Pozzoli et. al 2023 [arxiv:2302.07043]
Bonetti, Sesana 2020 [arxiv:2007.14403]

Kerr MBH mass

$m_1 \sim 10^4 - 10^7 M_\odot$

small CO mass

$\mu \sim 1 - 100 M_\odot$

extreme mass ratio:

$q = \dfrac{m_2}{m_1} \leq 1 \quad q \sim 10^{-3} - 10^{-6}$

highly eccentric:

$e \sim 0.1 - 0.9$

huge number of cycles:

$n_{cyc} \approx 10 \left(\frac{M_{\rm tot}}{10^6 M_\odot}\right)^{-5/3} \left(\frac{f}{1\rm mHz}\right)^{-5/3} \frac{(1+q)^2}{q}$

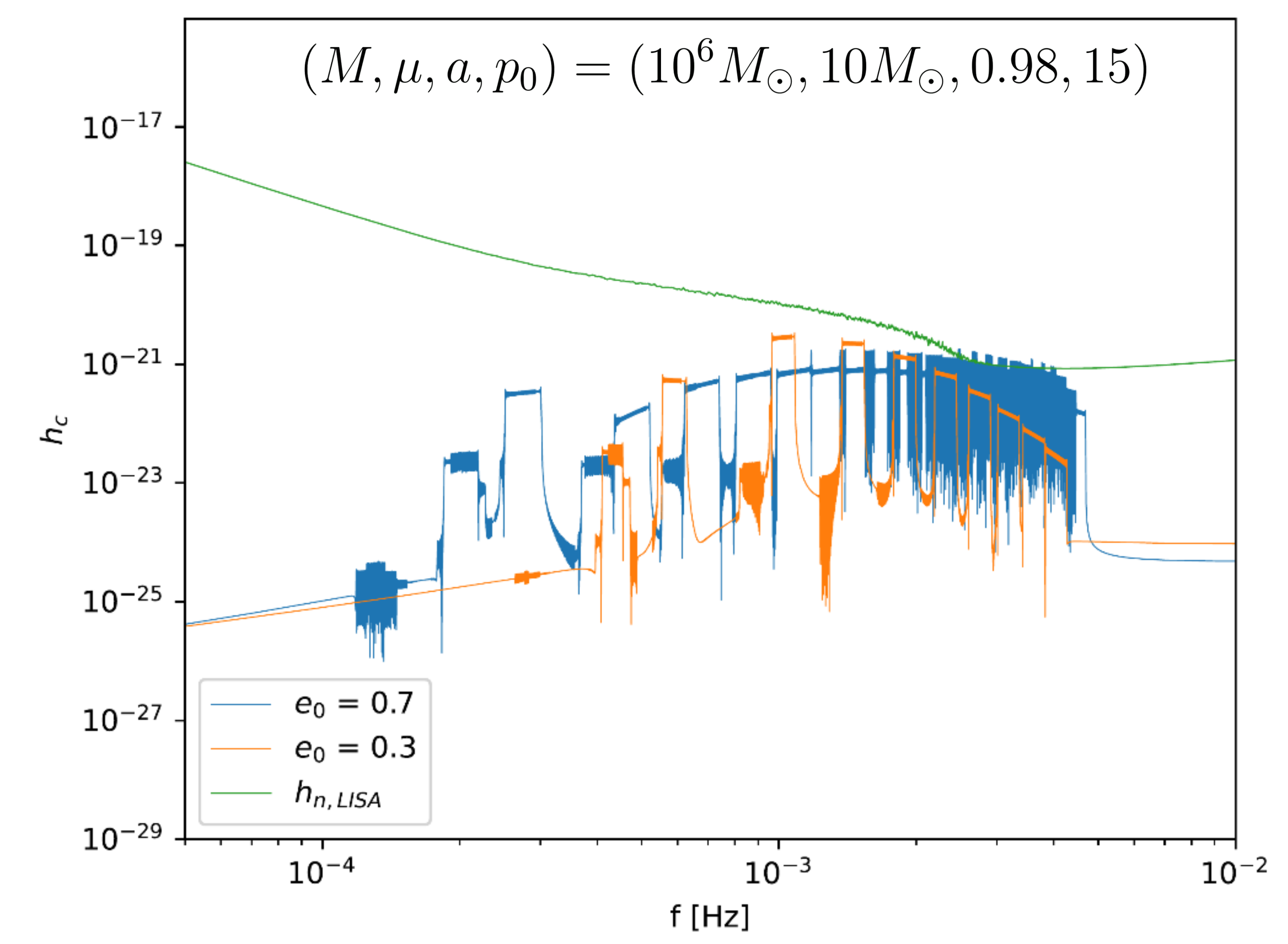
EMRI waveforms

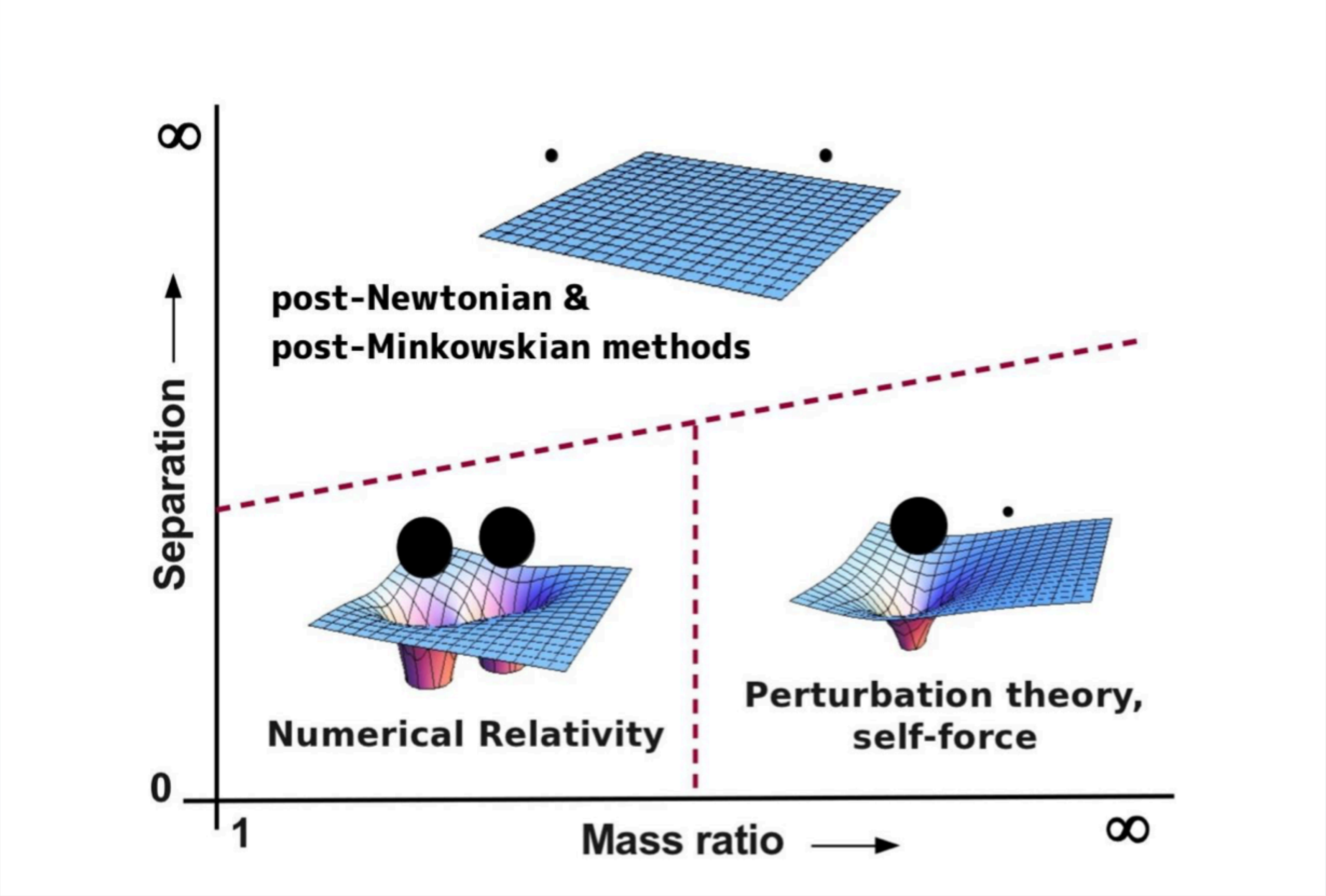
we require fully generic EMRI waveforms, capturing:

- high MBH spin ~ 0.98
- high eccentricities $\sim 0.1 - 0.99$
- inclined orbits

$$h_+ - ih_\times = \sum_{lmnk} A_{lmnk}(t)\ \exp\Big[-i\Phi_{mnk}(t)\Big]$$

up to ~1000 different harmonics





Credit: Leor Barack

GSF formalism is very accurate $\Longrightarrow$ computationally expensive

specific family of approximate generic-orbit EMRI waveforms

Kludge waveforms (AK - NK - 5PN AAK)

FastEMRIWaveforms package
[arXiv:2104.04582]

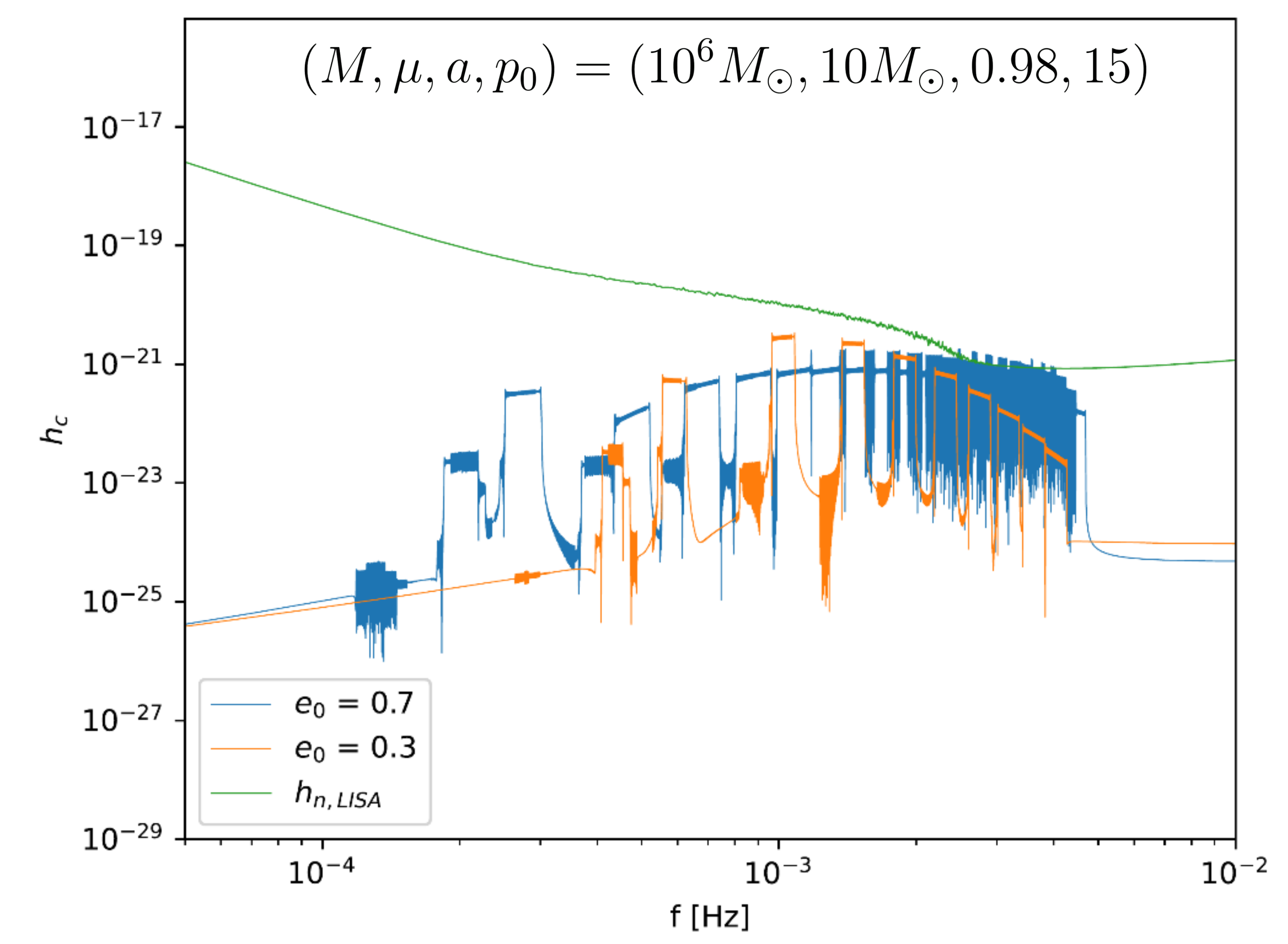
EMRI waveforms

we require fully generic EMRI waveforms, capturing:

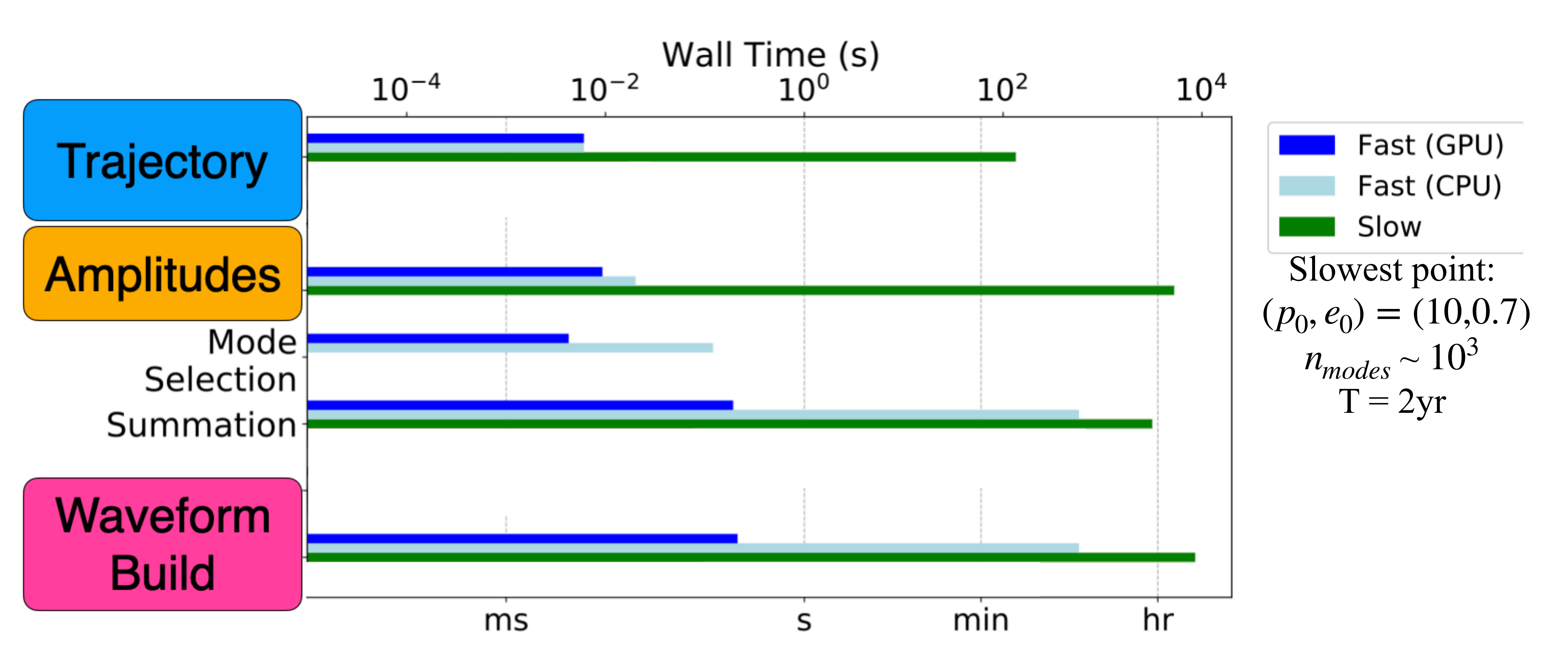
- high MBH spin ~ 0.98
- high eccentricities $\sim 0.1 - 0.99$
- inclined orbits

$$h_+ - ih_\times = \sum_{lmnk} A_{lmnk}(t) \;\exp\Big[-i\Phi_{mnk}(t)\Big]$$

up to ~1000 different harmonics



FastEMRIWaveforms package
Katz et al. 2021 [arXiv:2104.04582]



Slowest point:
$(p_0, e_0) = (10{,}0.7)$
$n_{modes} \sim 10^3$
$T = 2yr$

specific family of approximate generic-orbit EMRI waveforms

$\boxed{\text{Kludge waveforms (AK - NK - 5PN AAK)}}$

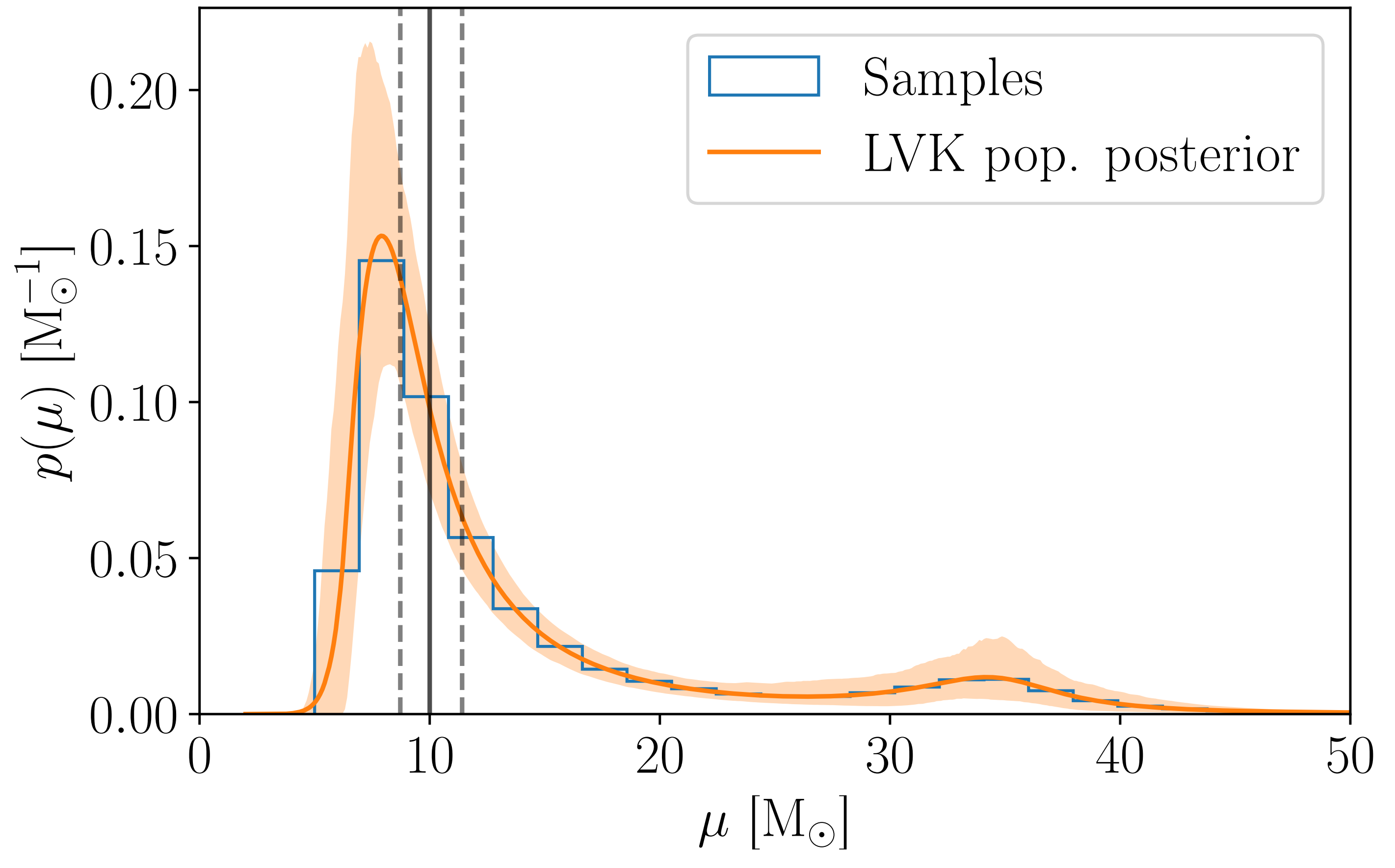
EMRI populations

EMRI event rates are uncertain
expected between ~ 10 and ~ 20000 per year

Model	Mass function	MBH spin	Cusp erosion	M–σ relation	N_p	CO mass $[M_\odot]$	Total	EMRI rate $[\mathrm{yr}^{-1}]$ Detected (AKK)	Detected (AKS)
M1	Barausse12	a98	yes	Gultekin09	10	10	1600	294	189
M2	Barausse12	a98	yes	KormendyHo13	10	10	1400	220	146
M3	Barausse12	a98	yes	GrahamScott13	10	10	2770	809	440
M4	Barausse12	a98	yes	Gultekin09	10	30	520 (620)	260	221
M5	Gair10	a98	no	Gultekin09	10	10	140	47	15
M6	Barausse12	a98	no	Gultekin09	10	10	2080	479	261
M7	Barausse12	a98	yes	Gultekin09	0	10	15800	2712	1765
M8	Barausse12	a98	yes	Gultekin09	100	10	180	35	24
M9	Barausse12	aflat	yes	Gultekin09	10	10	1530	217	177
M10	Barausse12	a0	yes	Gultekin09	10	10	1520	188	188
M11	Gair10	a0	no	Gultekin09	100	10	13	1	1
M12	Barausse12	a98	no	Gultekin09	0	10	20000	4219	2279

Babak et al. 2017 [arxiv:1703.09722]

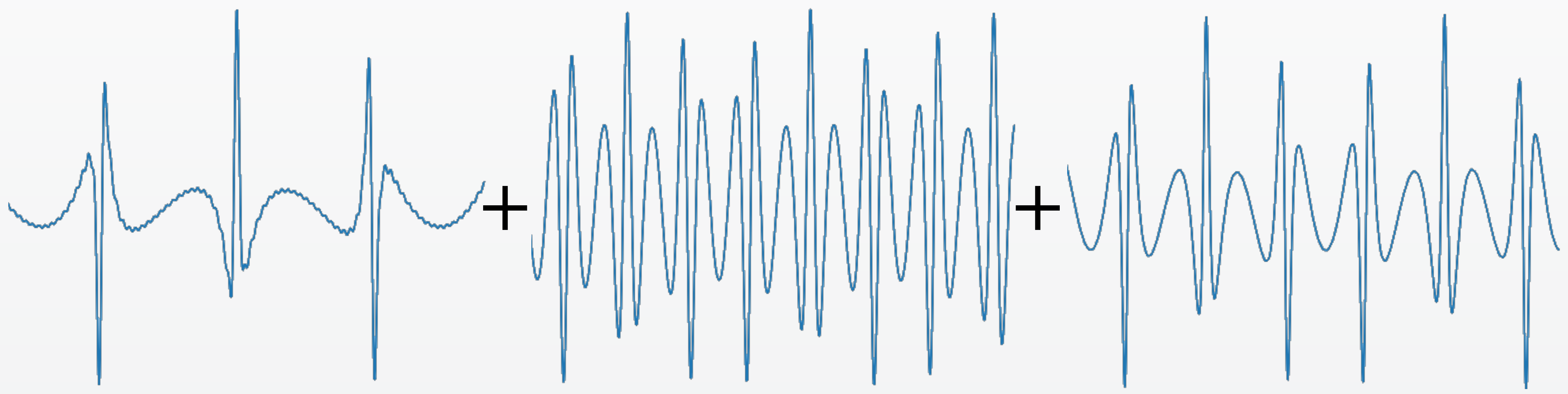
we allow the mass of the small CO to vary:
population of compact binary mergers
observed by ground-based detector



LVK Population analysis with GWTC-3 (2022)
[arxiv:2111.03634]

we focus on 3 of the 12 catalogs
M1 - M8 - M12
intermediate - pessimistic - optimistic

Statistical Properties of a GWB



$$h_1 \quad + \quad h_2 \quad + \ \dots \ + \quad h_n$$

$$h_{ij}(t,x) = \sum_{A=+,\times} \int_{-\infty}^{\infty} \mathrm{d}f \int \mathrm{d}^2\hat{n}\, \tilde{h}_A(f,\hat{n}) e_{ij}^A(\hat{n}) e^{i2\pi f(t-\hat{n}\cdot x/c)}$$

- **Stationary**: the statistical properties of the GWB do not depend on time

$$\langle \tilde{h}_A^*(f,\hat{n})\tilde{h}_{A'}(f',\hat{n}') \rangle \propto \delta(f-f')C_{AA'}(\hat{n},\hat{n}')$$

- **Gaussian**: central limit theorem VS finite source density

uniquely characterized by the spectral density $S_h(f)$

- **Isotropic and sky-uncorrelated**: extragalactic origin

$$C_{AA'}(\hat{n},\hat{n}') = \frac{1}{4\pi}\delta^2(\hat{n},\hat{n}')C_{AA'}$$

- **Unpolarized** $\quad C_{AA'}(\hat{n},\hat{n}') = \frac{1}{4\pi}\delta^2(\hat{n},\hat{n}')\delta_{AA'}$

IS IT GAUSSIAN AND STATIONARY AS FREQUENTLY ASSUMED?

Characterize the statistical properties of time-series

a toy model to capture the relevant features

ergodicity of the signal

replace averages over statistical ensemble
with averages over time

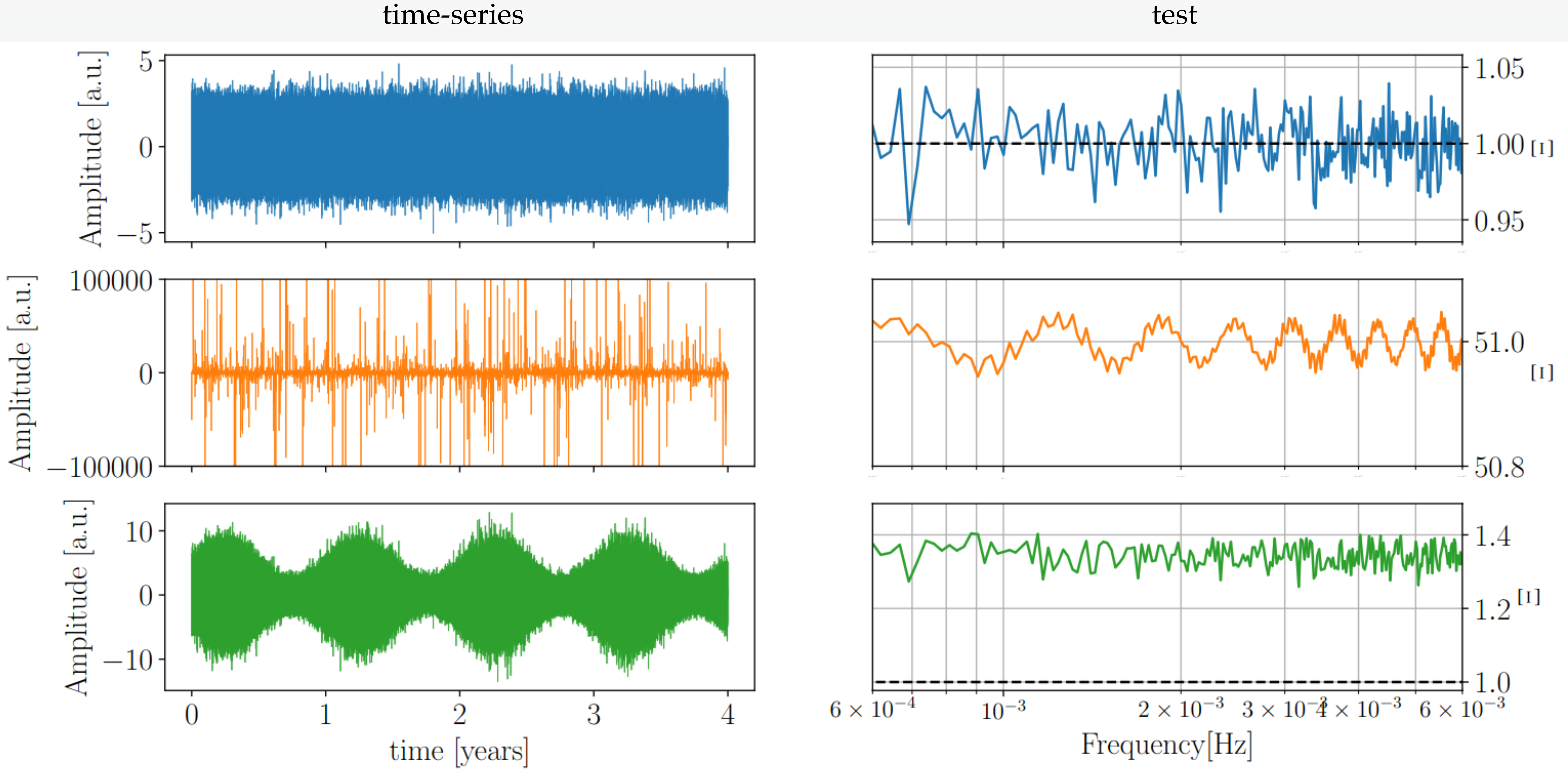
$\Downarrow$

split each time series into N_{chunks}

Rayleigh test

$$\Theta = \frac{\sigma(|\tilde{s}(f)|^2)}{\mu(|\tilde{s}(f)|^2)} \to 1$$ in the infinite sample limit

Acernese et. al 2023 [arXiv:2210.15634]



Analysis breakdown

Steps to characterize the GWB from EMRIs in LISA:

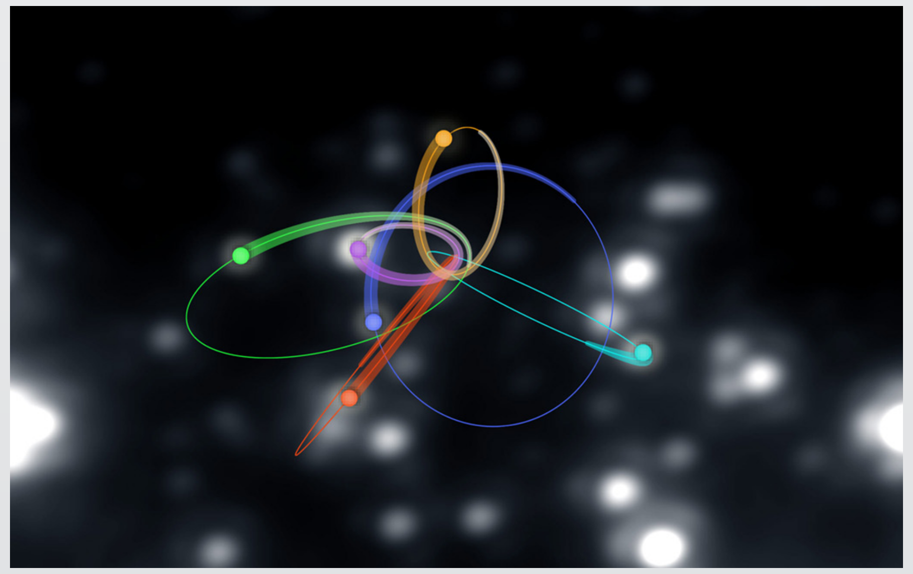
i) build EMRI populations → Models by Babak et al. 2017 + LVK population + Bonetti, Sesana 2020 [arXiv:2007.14403]

ii) compute the gravitational wave signal → Fast EMRIs Waveform, Augmented Analytic Kludge (5PN AAK)

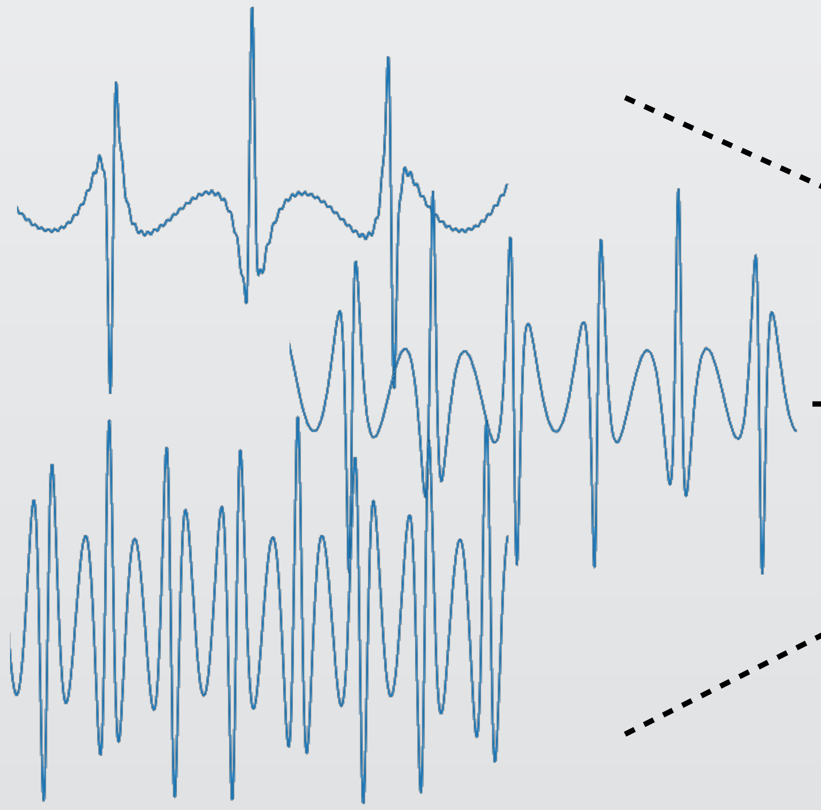
iii) inject it inside LISA → 1st Gen. Time Delay Interferometry

iv) characterize statistical properties of the GWB → Rayleigh test

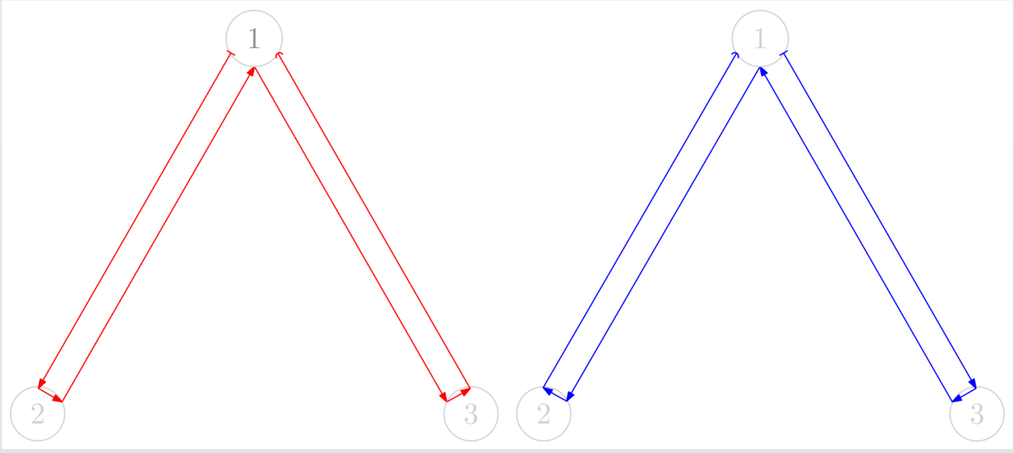
i)



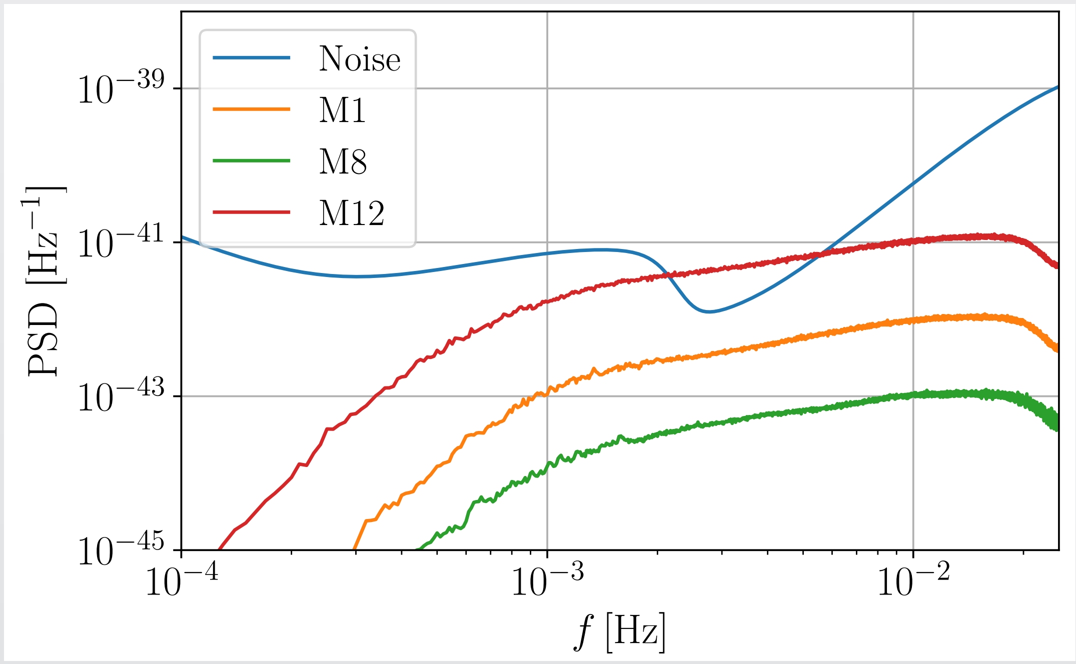
ii)



iii)



iv)



EMRI catalogs definition

LISA launch: 2035
me waiting for M12: 2085

4-year LISA observation time, dt = 10s

$h_{+,\times}(t) + TDI \sim 45s$ on CPUs

Model	N_{start}
M1	1217952
M8	124968
M12	21315202

speed up

GPU

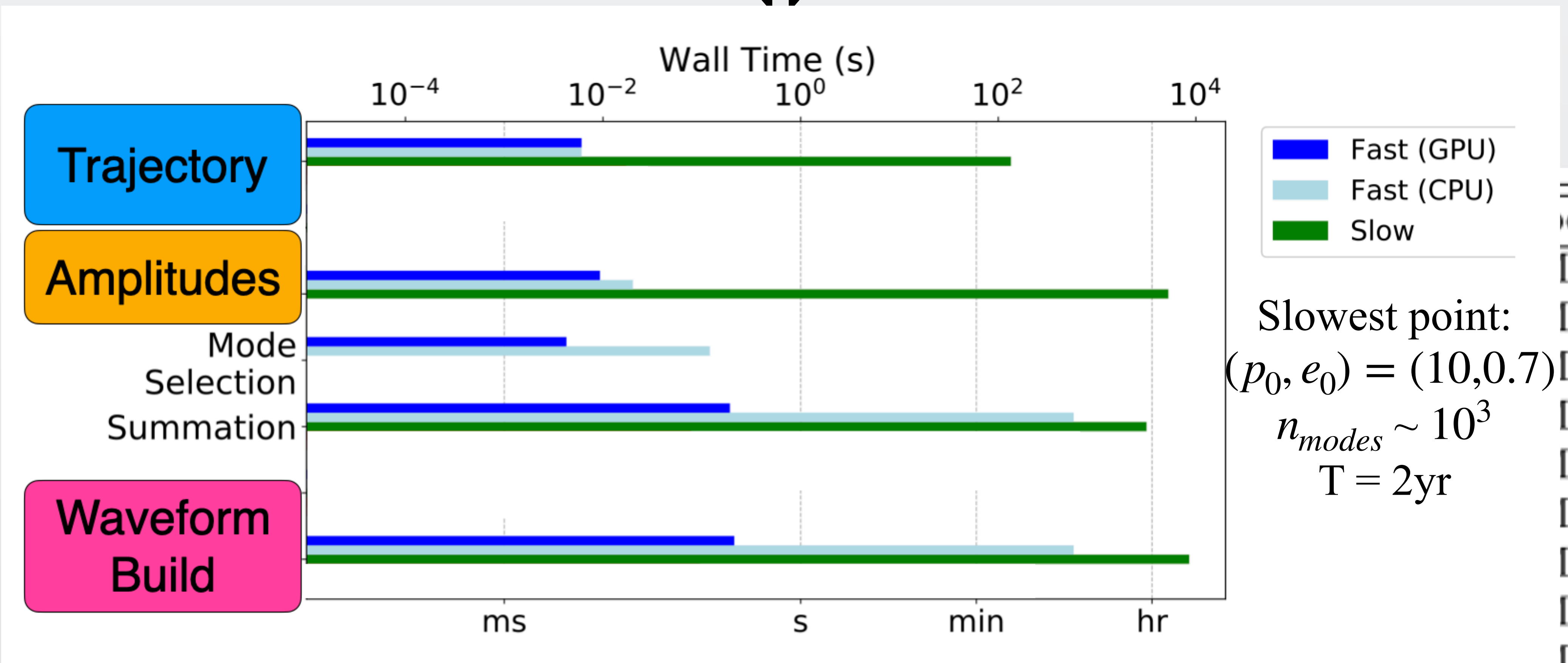
~ 0.3 ÷ 10 s
per single source

FastEMRIWaveforms
[arXiv:2204.06633]
+
FastLISAResponse
[arXiv:2104.04582]

cut in the population
remove sources with
SNR < 1 using an
inclination-polarization
averaged version of the
AK waveform
$\rho_{AK,th} = 1$

cut ~ 97% of the sources
losing only about 5%
of the total SNR

$\Downarrow$



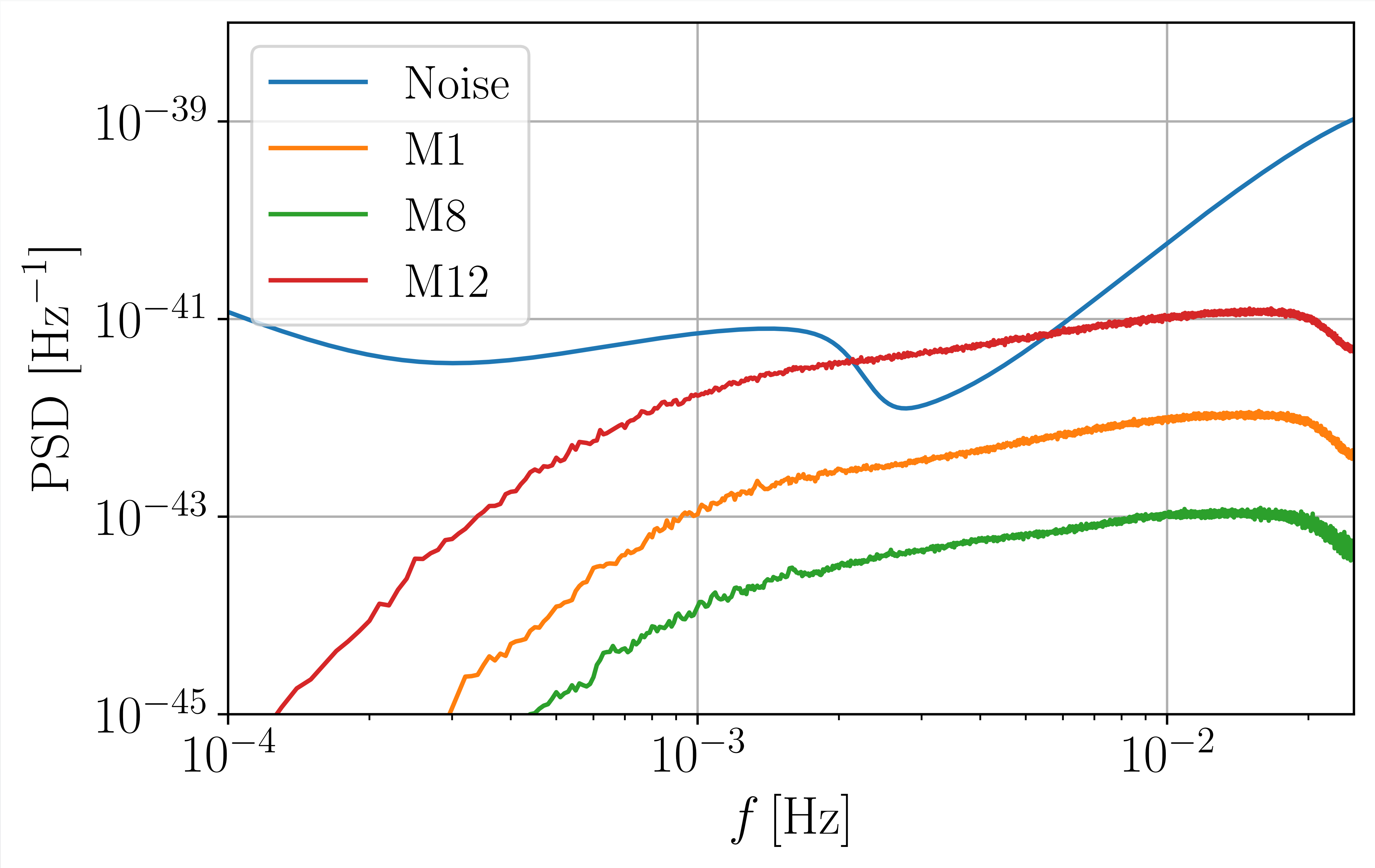
Katz et al. 2021 [arXiv:2104.04582]

del	N_{tot}	N_f	$\mathrm{SNR}_{\mathrm{tot}}$	$\mathrm{SNR}_{\rho>1}$	SNR_f	Detections
I1	1225158	31764	571	534	460	366
I2	580149	17030	434	418	352	386
I3	2030059	68100	1174	1121	839	1123
I4	43607	11166	960	926	750	1713
I5	1772409	44011	727	695	535	622
I6	396800	2486	48	44	27	35
I7	8218425	303348	4940	4748	3960	4619
I8	89010	3620	63	61	51	53
I9	872231	31356	501	480	419	426
M10	823589	30597	478	459	408	406
M11	34724	287	4.53	4.11	2.32	0
M12	16547658	394583	6475	6200	4553	5580

Pozzoli et al. 2023 [arxiv:2302.07043]

EMRI background spectra and SNR

A Channel



Model	$N_{\rm final}$	Detections	$\rho_{\rm gwb}$
M1	26932	522	311
M8	3209	64	38
M12	319309	5909	3684

$$\rho_{\mathrm{gwb},i} = \sqrt{T_{\mathrm{OBS}} \int_0^\infty \mathrm{d}f \left(\frac{S_{\mathrm{gwb},i}}{S_{\mathrm{n},i}} \right)^2}$$

resolvable sources with $\rho_i > 20$

when computing the single source ρ_i

$$S_n(f) = S_{instr}(f) + S_{WD}(f)$$

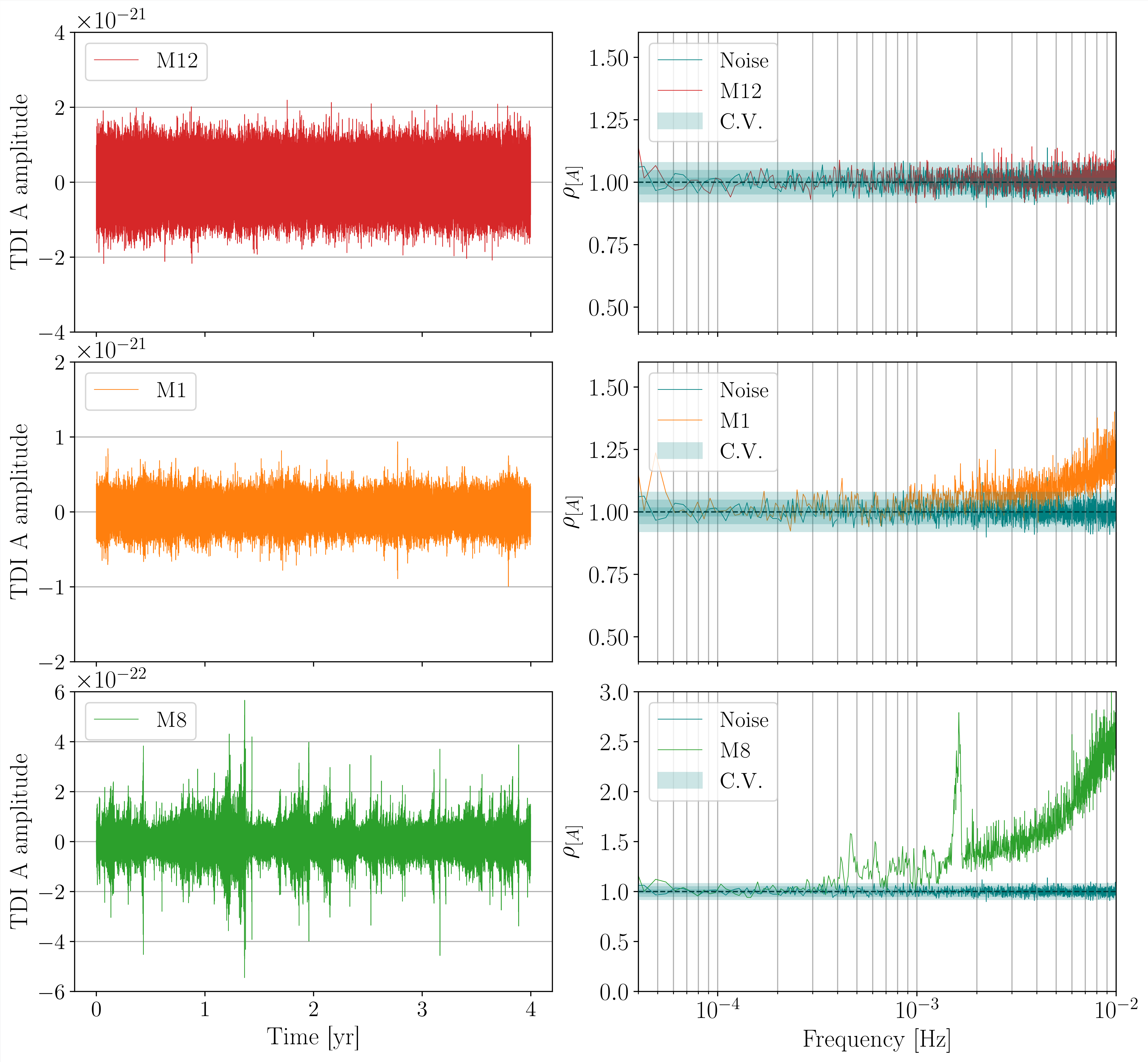
A more realistic approach would be to perform an Iterative Foreground Estimation (IFE) including also the EMRI GWB in the $S_n(f)$

see: Karnesis et al. 2021 [arXiv:2103.14598]

pile-up of the unresolvable sources

3 possible scenarios
between an upper and lower limit

Is it Gaussian and Stationary?



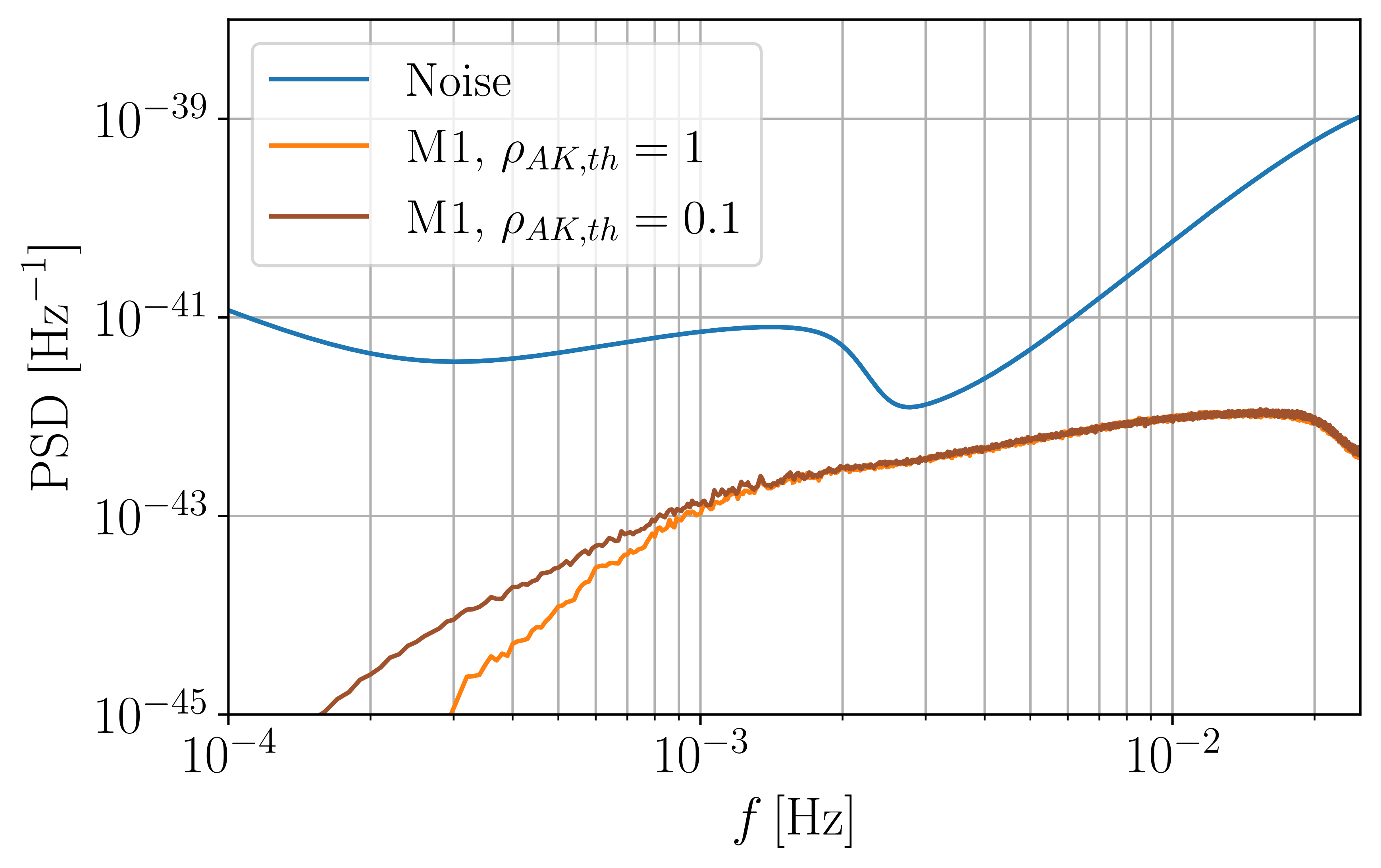
Model	$N_{\rm final}$	Detections	$\rho_{\rm gwb}$
M1	26932	522	311
M8	3209	64	38
M12	319309	5909	3684

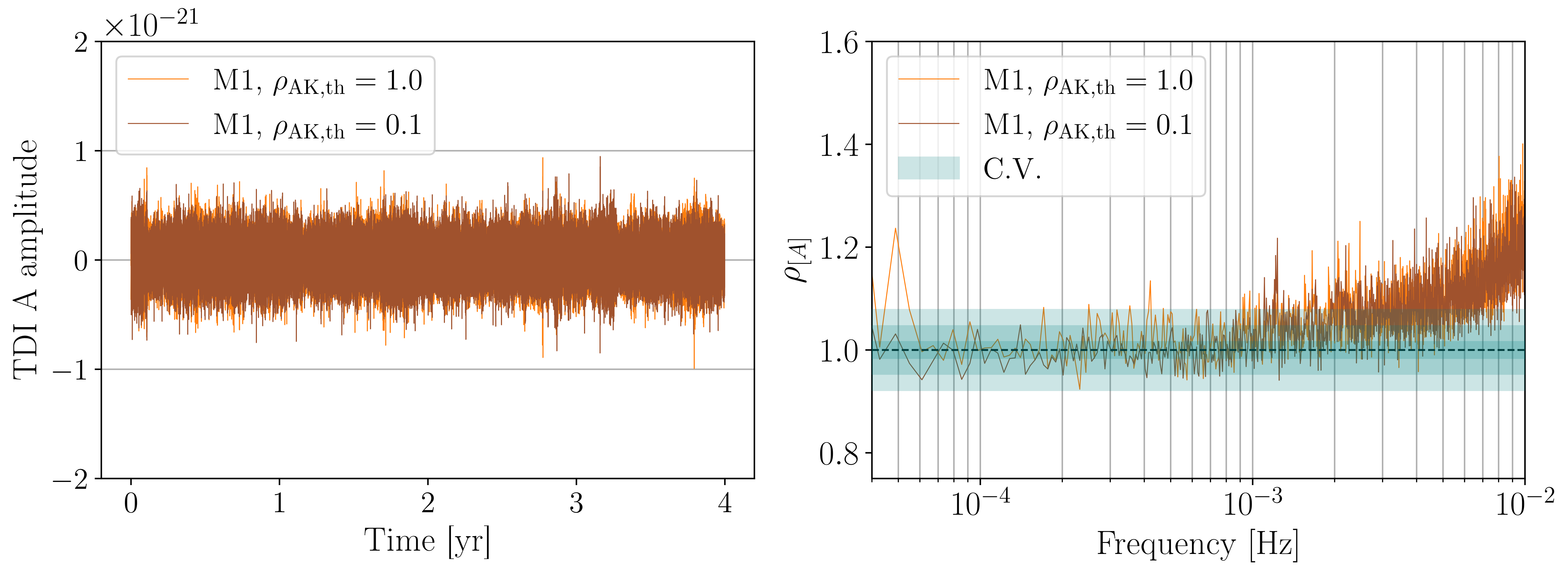
M12: we do not reject the hypothesis of stationarity and Gaussianity.
$N_{final} = 319309$

M1: presence of either non-Gaussianities or non-stationarities for frequencies exceeding 1mHz.
$N_{final} = 26932$

M8: rejects the hypotheses of stationarity and Gaussianity, due to the low number of sources in the catalog.
$N_{final} = 3209$

Is our choice of excluding faint EMRIs robust?





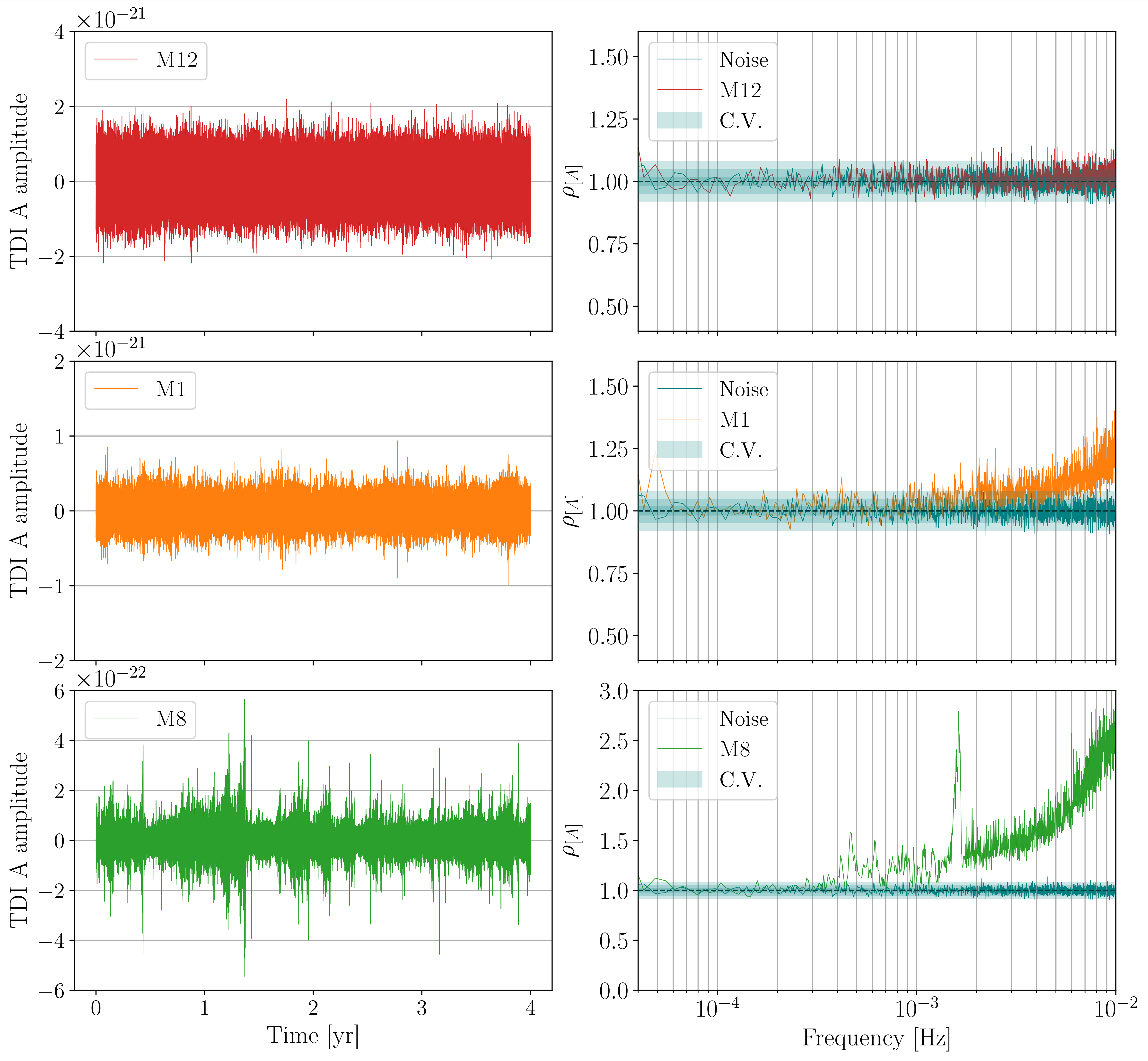
relax the threshold for source removal for model M1

$$\rho_{AK,th} = 1 \to 0.1 \quad \Longrightarrow \quad N_{final} = 26932 \to 209072 \quad \Longrightarrow \quad \rho_{gwb} = 311 \to 319$$

the Rayleigh test remains largely unchanged

the choice of considering only $\rho_{AH} > 1$ suffices to support our conclusions.

Is it Gaussian and Stationary?



Consequences

- **Gaussian-likelihood could be only approximately valid**
- **Global fit couples SGWB detection, estimation, and resolvable source PE**

  inadequate modeling could potentially introduce biases in global inference results, affecting foreground estimation, background detection, and individual source parameter reconstruction.

More work is needed to assess the impact of such biases.

Updating EMRI detection rates and parameter uncertainties

Manuel PIARULLI,
Danny LAGHI, Ollie BURKE, Shubam KEJRIWAL
Christian CHAPMAN-BIRD, Federico POZZOLI, Nikos KARNESIS

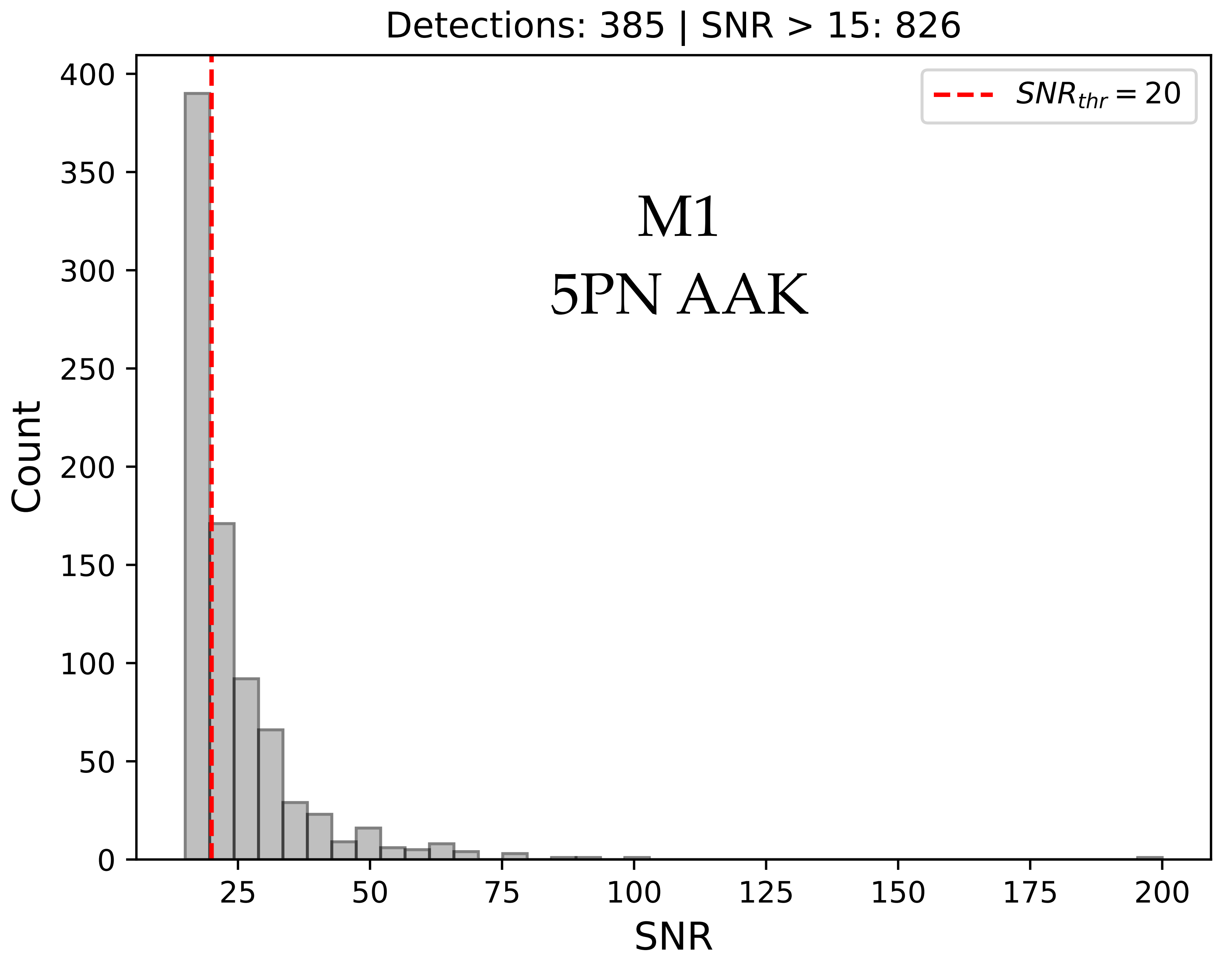
We focus on 5 EMRI catalogs
M1-M6 - M5-M11 - M7
intermediate - pessimistic - optimistic

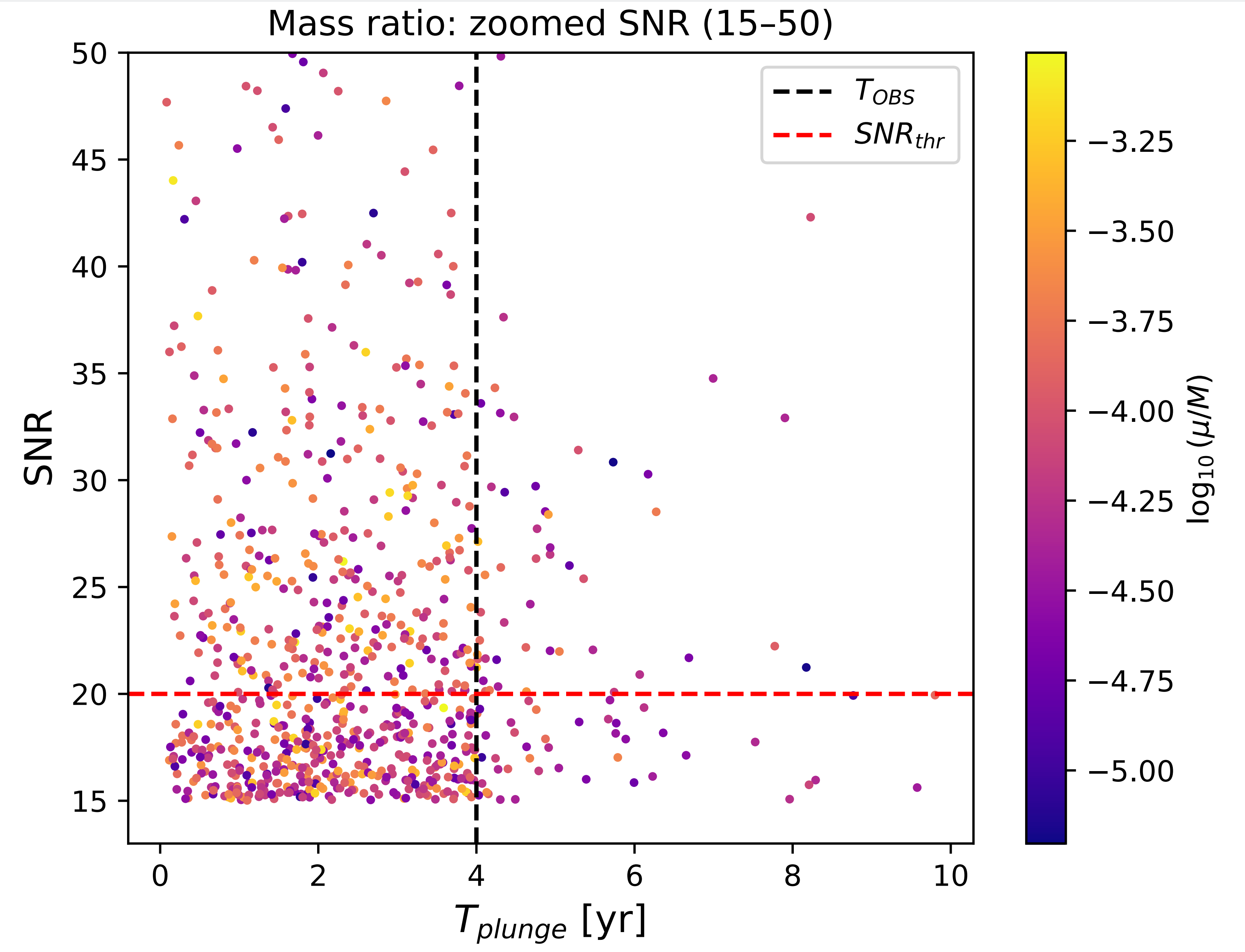
Updates from previous work:

2ndGenTDI
IFE for subtraction of resolvable sources

Model	# before IFE	# after IFE
M1	420	385
M5	26	26
M6	387	352
M7	3228	RUNNING
M11	1	1

Number of resolvable sources
(before and after IFE)
$T_{LISA} = 4$ years
5PN-AAK waveform





We focus on 5 EMRI catalogs

M1-M6 - M5-M11 - M7

intermediate - pessimistic - optimistic

Updates from previous work:

Catalog definition

2ndGenTDI

IFE for subtraction of resolvable sources

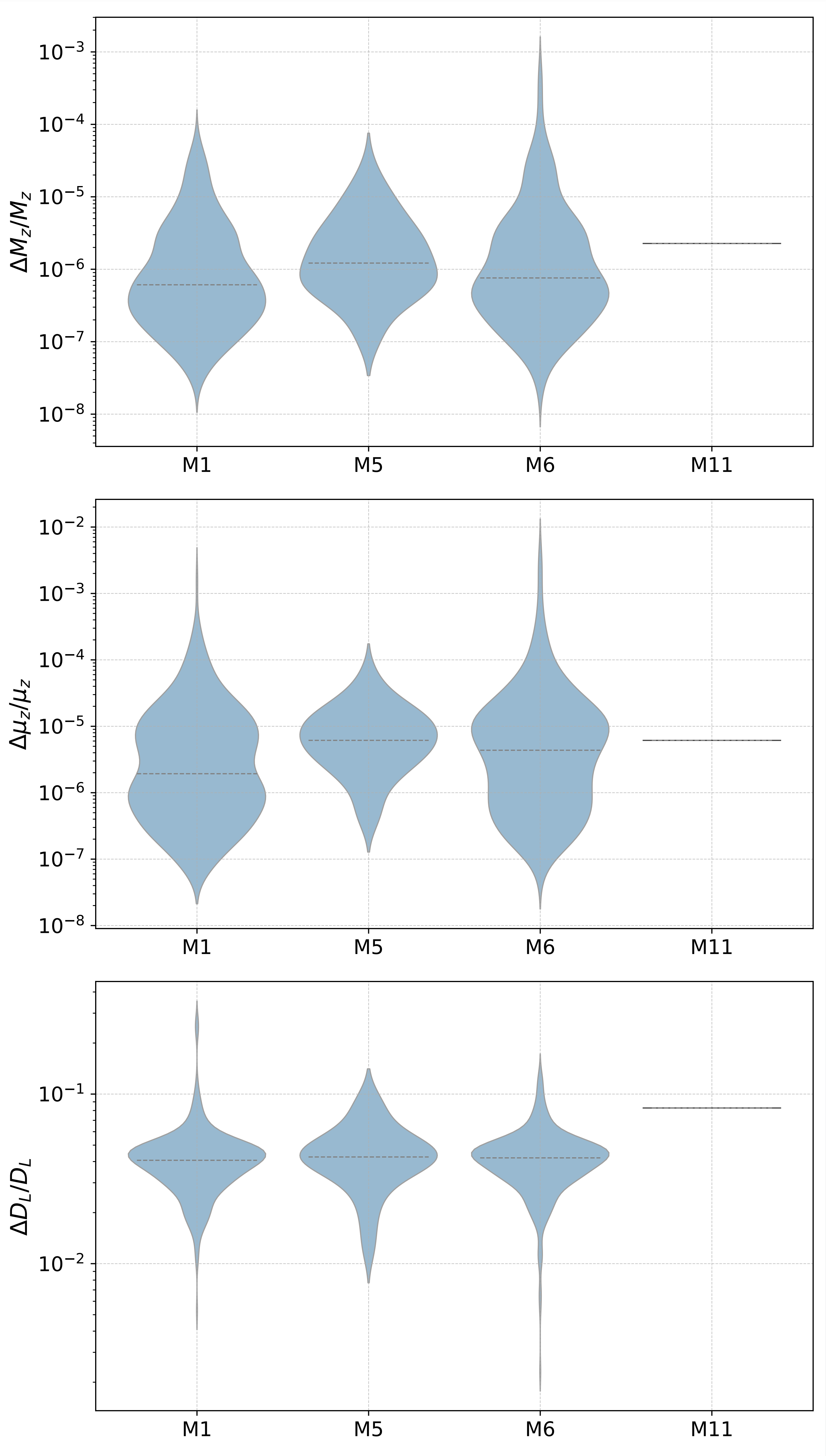
Model	# before IFE	# after IFE
M1	420	385
M5	26	26
M6	387	352
M7	3228	Running
M11	1	1

Resolvable sources (before/after IFE)

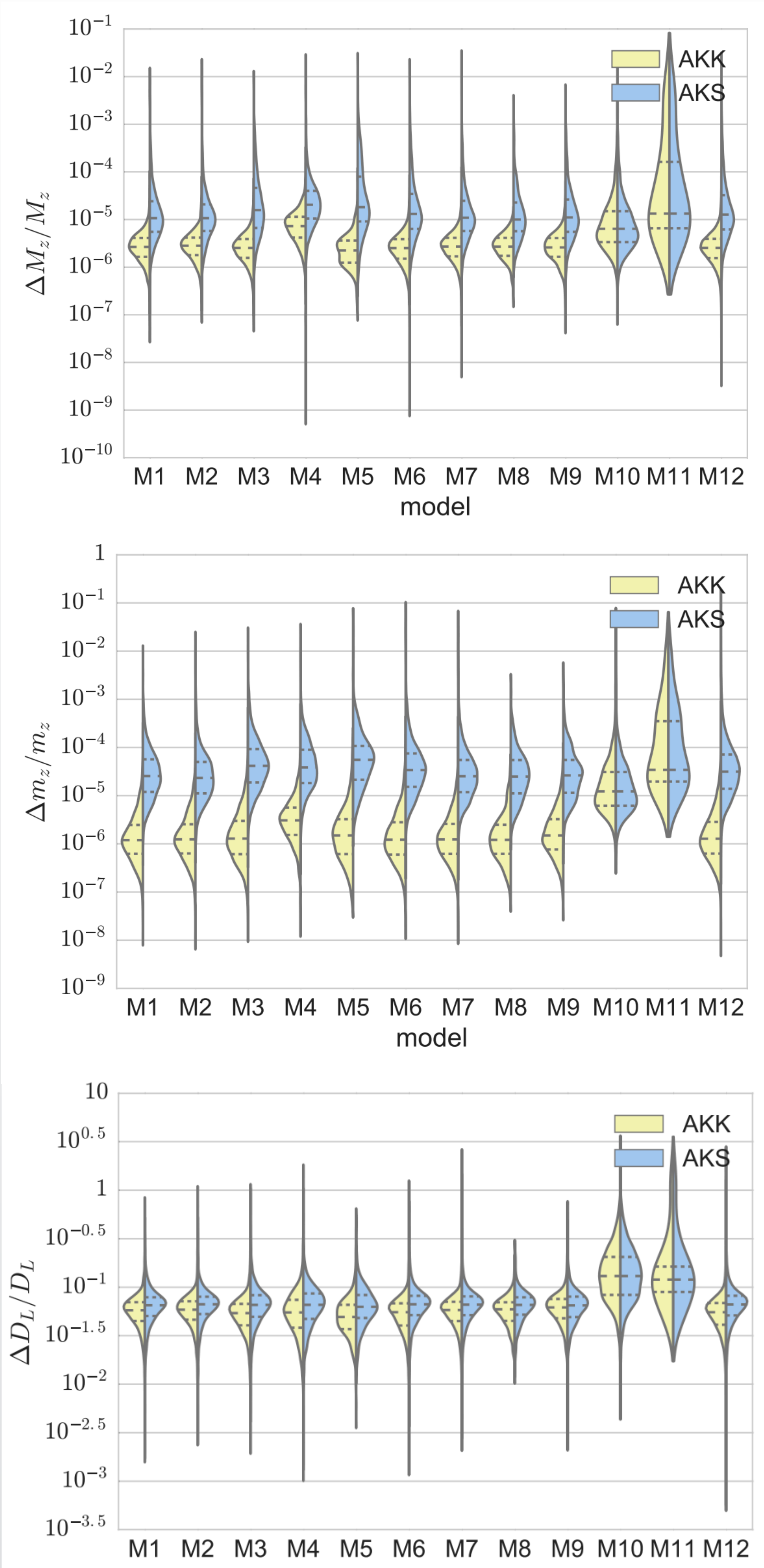
$T_{LISA} = 4$ yrs

AAK waveform

AAK



Babak et al. 2017 [arxiv:1703.09722]



We focus on 5 EMRI catalogs

M1-M6 - M5-M11 - M7

intermediate - pessimistic - optimistic

Updates from previous work:

Catalog definition

2ndGenTDI

IFE for subtraction of resolvable sources

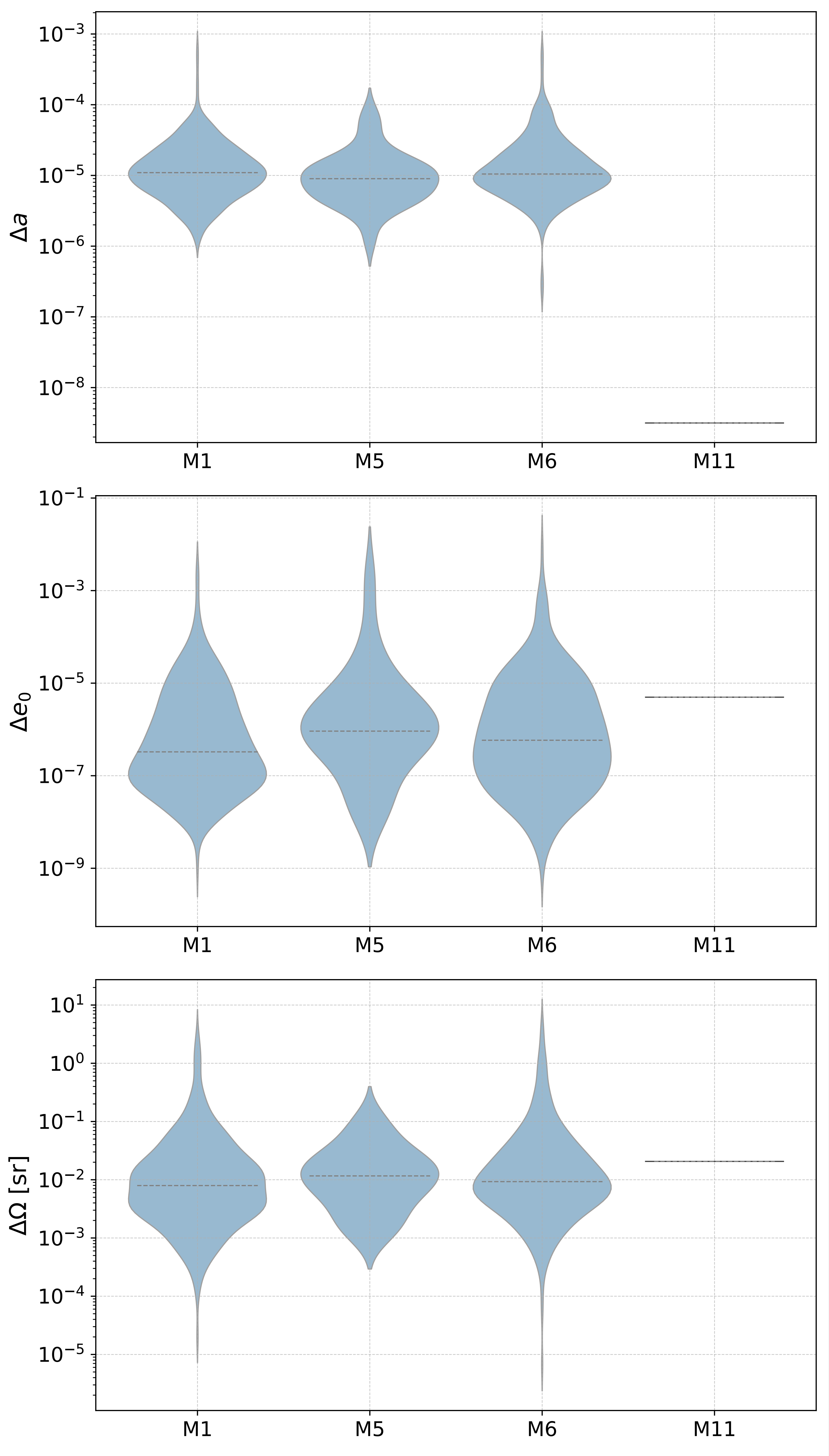
Model	# before IFE	# after IFE
M1	420	385
M5	26	26
M6	387	352
M7	3228	RUNNING
M11	1	1

Resolvable sources (before/after IFE)

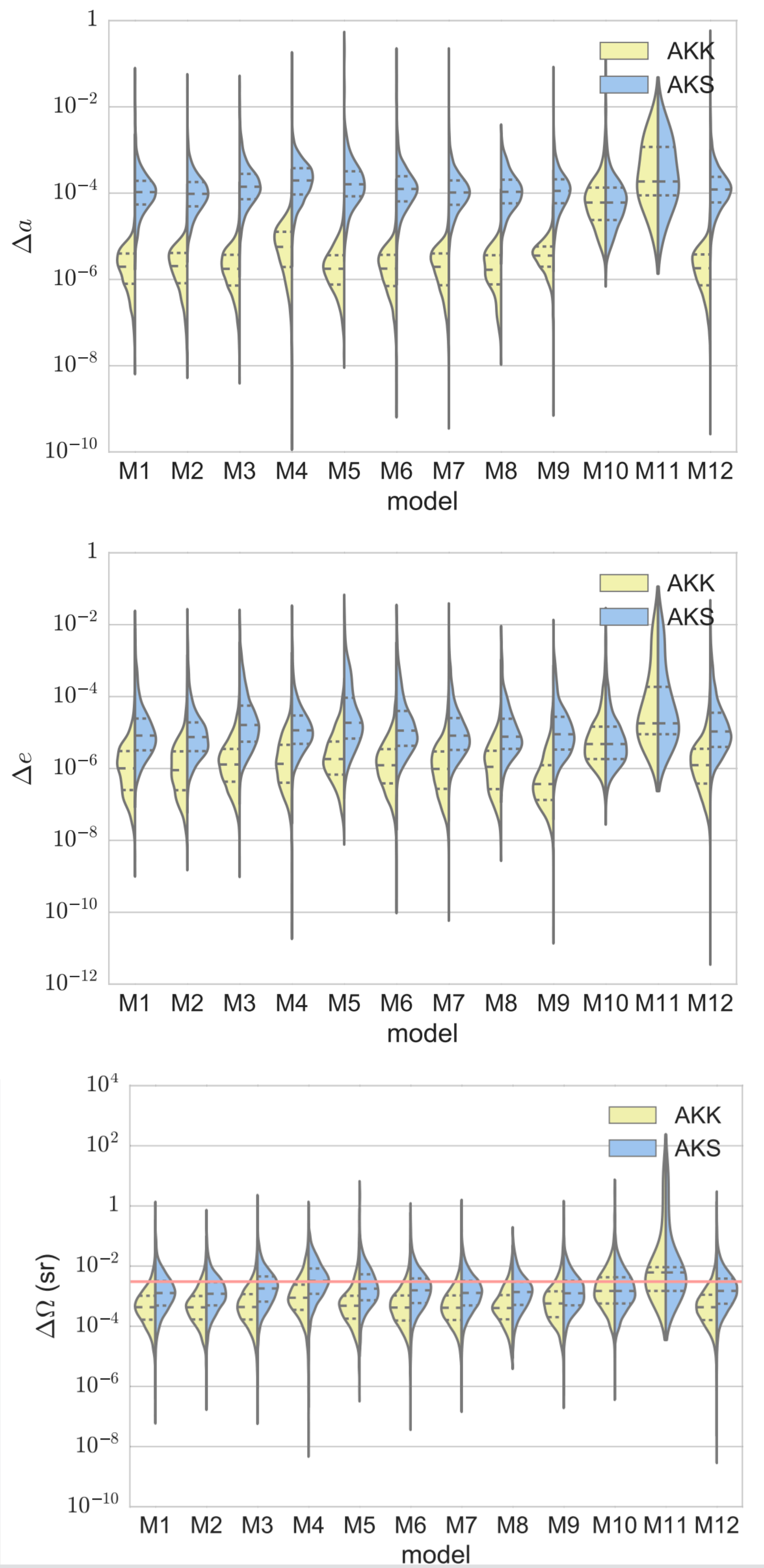
$T_{LISA} = 4$ yrs

AAK waveform

AAK



Babak et al. 2017 [arxiv:1703.09722]

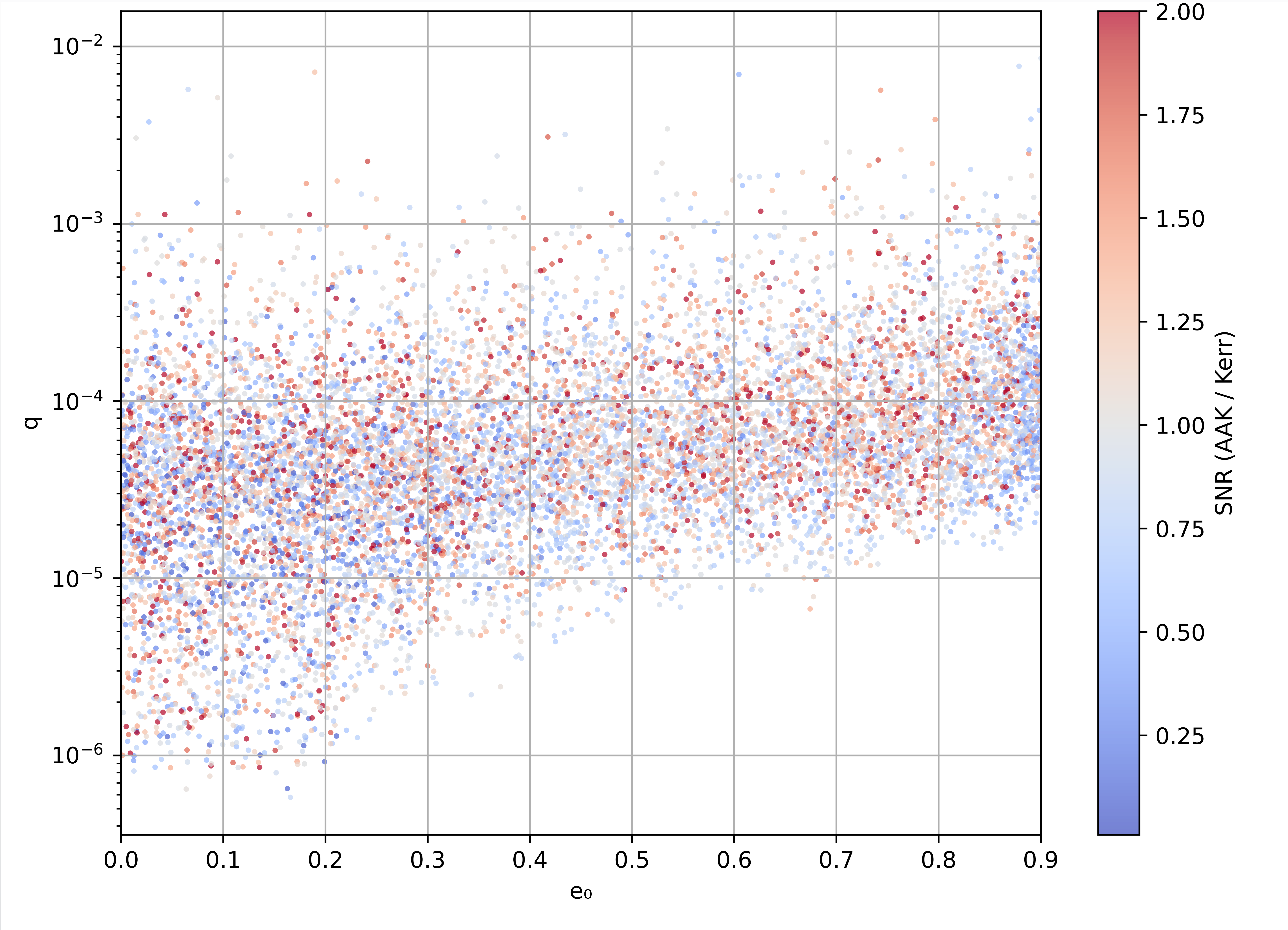


<u>Next steps:</u>

- KerrEccentricEquatorial
to asses SNR accuracy, and so possible changes in the detection rate.
Limited parameter space
(Equatorial and $e_0 < 0.9$)

- Full Bayesian PE for Fisher validation on a few detections

$e_0 < 0.9$ the only difference is Y_0



Thanks for the attention!

happy to take questions