APS APRIL MEETING 2024

IMPACT OF NOISE TRANSIENTS ON GRAVITATIONAL-WAVE BURST DETECTION EFFICIENCY OF THE *BAYESWAVE* PIPELINE WITH MULTI-DETECTOR NETWORKS

3RD OF APRIL 2024

YI SHUEN (CHRISTINE) LEE

DR. MEG MILLHOUSE

PROF. ANDREW MELATOS

OzGrav

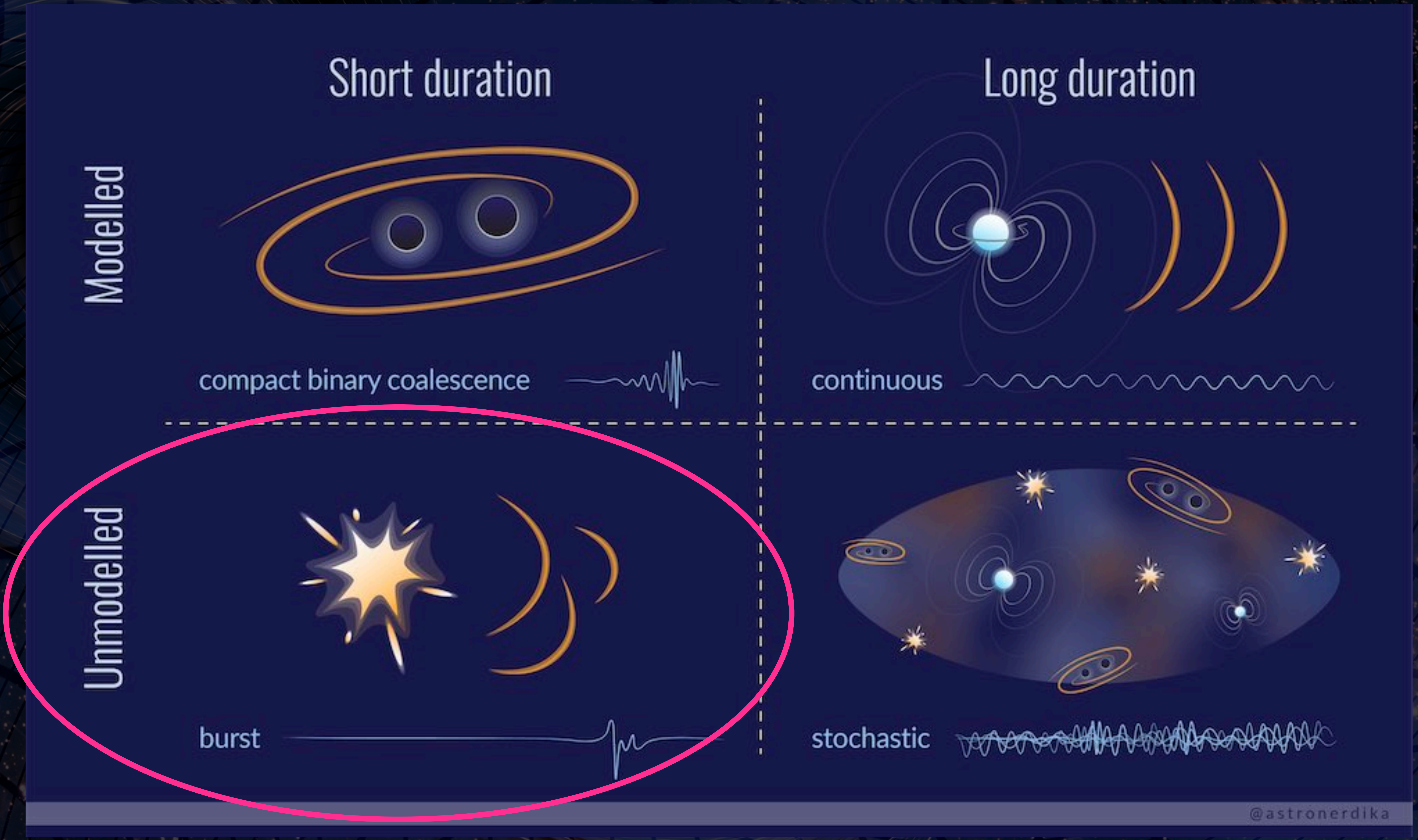
ARC Centre of Excellence for Gravitational Wave Discovery

LIGO VIRGO KAGRA

POSTERA CRESCAM LAUDE

THE UNIVERSITY OF MELBOURNE

TYPES OF GW SOURCES

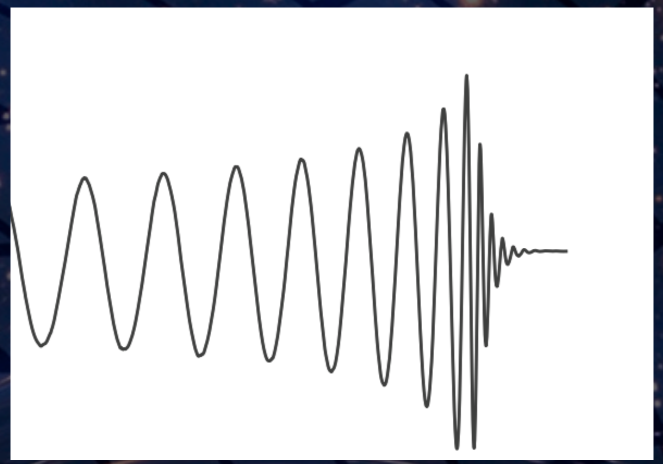


[Image credits: Shanika Galaudage]

WHAT IS IN THE DETECTOR DATA?

Data = Signal (maybe) + Noise

$\vec{s}$



$h(\vec{\theta})$

$\vec{n}$

Noise components:

$$\vec{n} = \overrightarrow{n_G} + \vec{g}$$

Stationary, Gaussian noise

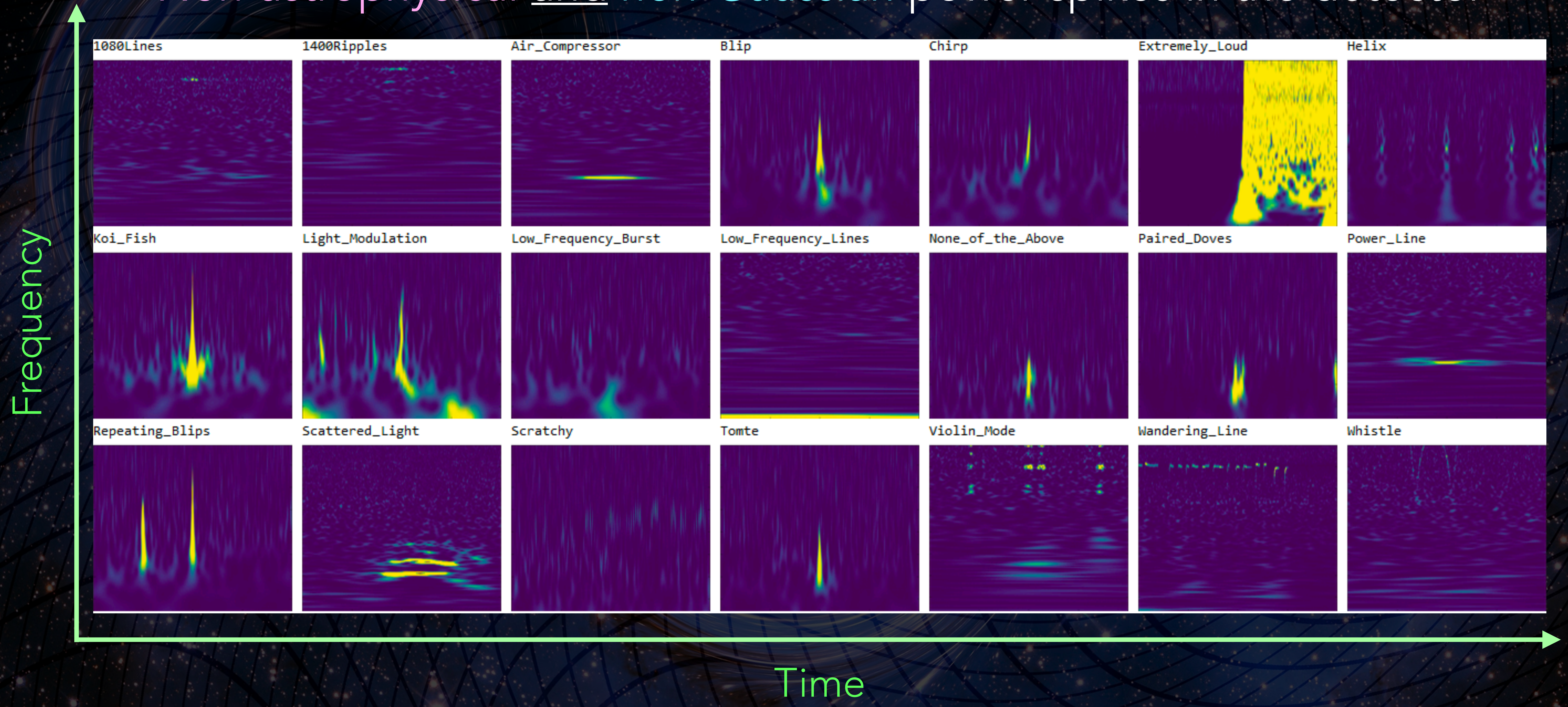
Instrumental glitches (non-stationary, non-Gaussian)

X

[Adapted from Meg Millhouse's slide]

THE PROBLEM: *INSTRUMENTAL GLITCHES*

Non-astrophysical and non-Gaussian power spikes in the detector



[Image from Gravity Spy]

BAYESWAVE: *ALGORITHM OVERVIEW*

- Unmodelled transient gravitational wave (burst) analysis algorithm
- Enables joint detection of instrumental glitches and GW bursts
- Reconstructs transient, non-Gaussian features in the data by summing a set of sine-Gaussian wavelets, with no *a priori* assumptions

Images courtesy of Meg Millhouse

BayesWave publications:
Cornish + Littenberg, Class. Quant. Grav 32, 130512 (2015)
Cornish et al., Phys. Rev. D 103, 044006 (2021)

BAYESWAVE: MODEL SELECTION

- Three independent models:
  - Signal plus Gaussian-noise model, $\mathscr{S}$
  - Glitch plus Gaussian-noise model, $\mathscr{G}$
  - Gaussian-noise only model, $\mathscr{N}$
- Compare models using the **Bayes Factor**

Bayesian evidence ratio between two models

$$\text{e.g. } \mathscr{B}_{\mathscr{S},\mathscr{G}} = \frac{p(\vec{s}\,|\,\mathscr{S})}{p(\vec{s}\,|\,\mathscr{G})}$$

If $\ln \mathscr{B}_{\mathscr{S},\mathscr{G}} > 0 \Rightarrow$ $\mathscr{S}$ is more strongly supported by data than $\mathscr{G}$

BAYESWAVE: ROLE IN GW BURST SEARCHES

- GWTC-1, -2 and -3:
  Assess consistency with matched-filter (model-based) searches
- O1, O2 and O3 all-sky burst searches:
  Follow-up trigger events found by coherent WaveBurst (cWB)
- Glitch subtraction
- PSD estimation

GWTC-1: Phys. Rev. X 9, 031040 (2019), **All-sky search O1:** Phys. Rev. D 95, 042003 (2017),
GWTC-2: Phys. Rev. X 11, 021053 (2021), **All-sky search O2:** Phys. Rev. D 100, 024017 (2019),
GWTC-3: arXiv:2111.03606 (2021). **All-sky search O3:** Phys. Rev. D. 104, 122004 (2021).

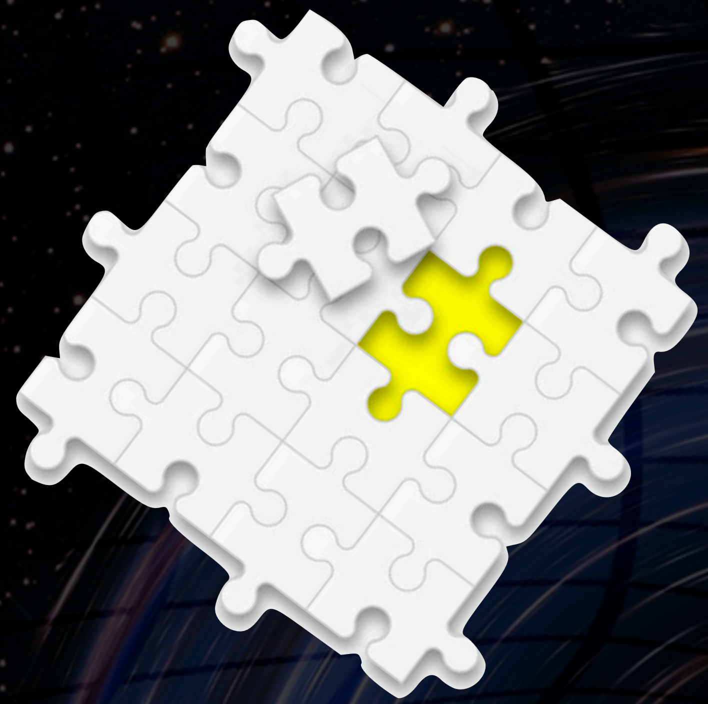
PREVIOUS WORK - PHYS. REV. D **103**, 062002

- *BayesWave's* detection statistics: $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$
- Global detector network is expanding
- How does the size of the detector network affects $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$?
- Analytically and empirically showed that $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ scales as
  - $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}} \sim \mathscr{I} N \ln \mathrm{SNR}_{\mathrm{net}}$

    (i) $\mathscr{I}$: number of detectors

    (ii) N : number of wavelets (i.e. model complexity)

    (iii) $\mathrm{SNR}_{\mathrm{net}}$: signal-to-noise ratio



THE MISSING PUZZLE PIECE

- This work studies the improvement in $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ of astrophysical triggers with expanded detector networks
- BUT…
  Larger detector networks are more susceptible to instrumental glitches i.e. noisier detector backgrounds!

AIM OF THIS WORK

- To study the overall performance of *BayesWave* by accounting for the noise background
- Compare performance between:
  - Hanford-Livingston (HL), 2-detector network
  - Hanford-Livingston-Virgo (HLV), 3-detector network

FALSE ALARM PROBABILITY, P_{FA}

/measure of detection significance/

a function of detection statistics ($\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$)

Probability that a trigger with a given $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ is a false alarm

Lower P_{FA} = Higher astrophysical significance

DETECTION EFFICIENCY P_{det}

/figure of merit for *BayesWave's* overall performance/

Probability of detecting an astrophysical event with a given significance (P_{FA})

EFFICIENCY CURVES

To characterise

DETECTION EFFICIENCY P_{det}

as a function of

FALSE ALARM PROBABILITY P_{FA}

EFFICIENCY CURVE "RECIPE"

2 COMPONENTS:

(1) **NOISE BACKGROUND MEASUREMENTS** → P_{FA}

- Distribution of *BayesWave*'s detection statistics $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ for **non-astrophysical triggers**

(2) **ASTROPHYSICAL MEASUREMENTS** → P_{det}

- As above, but for **astrophysical triggers**

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT

STEP 1: Obtaining detection statistics for background triggers

- Down-select triggers from the Coherent WaveBurst (cWB) trigger lists of O3a time-slide data
- Selected triggers satisfy the cWB nominal significance threshold: $\rho \geq 7$
- Analyse triggers using *BayesWave* to get $\ln \mathcal{B}_{\mathcal{S},\mathcal{G}}$

Method adapted from Kanner et al, Phys. Rev. D 93, 022002 (2016)

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT

STEP 1: Down-selecting background triggers

- From Coherent WaveBurst (cWB) trigger lists of O3a time-slide data

Stage 1 down-selection:

Triggers satisfying $\rho \geq 7$

- $\sim 2 \times 10^3$ triggers for HL
- $\sim 7 \times 10^3$ triggers for HLV

Stage 2 down-selection:

Keep only a fraction of $\rho \geq 7$ triggers (randomly selected)

- 45% of HL → 1008 triggers
- 15% of HLV → 1134 triggers

Method adapted from Kanner et al, Phys. Rev. D 93, 022002 (2016)

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT

STEP 2: Identifying Gaussian-noise-like background triggers

- Events more consistent with the Gaussian-noise model, $\mathscr{N}$ than the Gaussian-noise plus signal model, $\mathscr{S}$:

$$\ln \mathscr{B}_{\mathscr{S},\mathscr{N}} - \Delta \ln \mathscr{B}_{\mathscr{S},\mathscr{N}} \leq 0$$

- Meaningless to compare odds between signal and glitch models via $\ln \mathscr{B}_{\mathscr{S},\mathscr{G}}$
- Cannot remove from dataset as they still satisfy $\rho \geq 7$
- Assign arbitrarily low detection statistics: $\ln \mathscr{B}_{\mathscr{S},\mathscr{G}} = -500$
- **THIS APPLIES TO ALL EVENTS IN THIS STUDY!**

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT

STEP 3: Calculate P_{FA} as a function of $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$

P_{FA}
/probability of false alarm/

Fraction of non-astrophysical triggers detected above a given $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ threshold

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT

STEP 3: Calculate P_{FA} as a function of $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$

P_{FA}
/probability of false alarm/

Per-trigger probability (i.e. fraction) of non-astrophysical triggers detected above a given $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ threshold

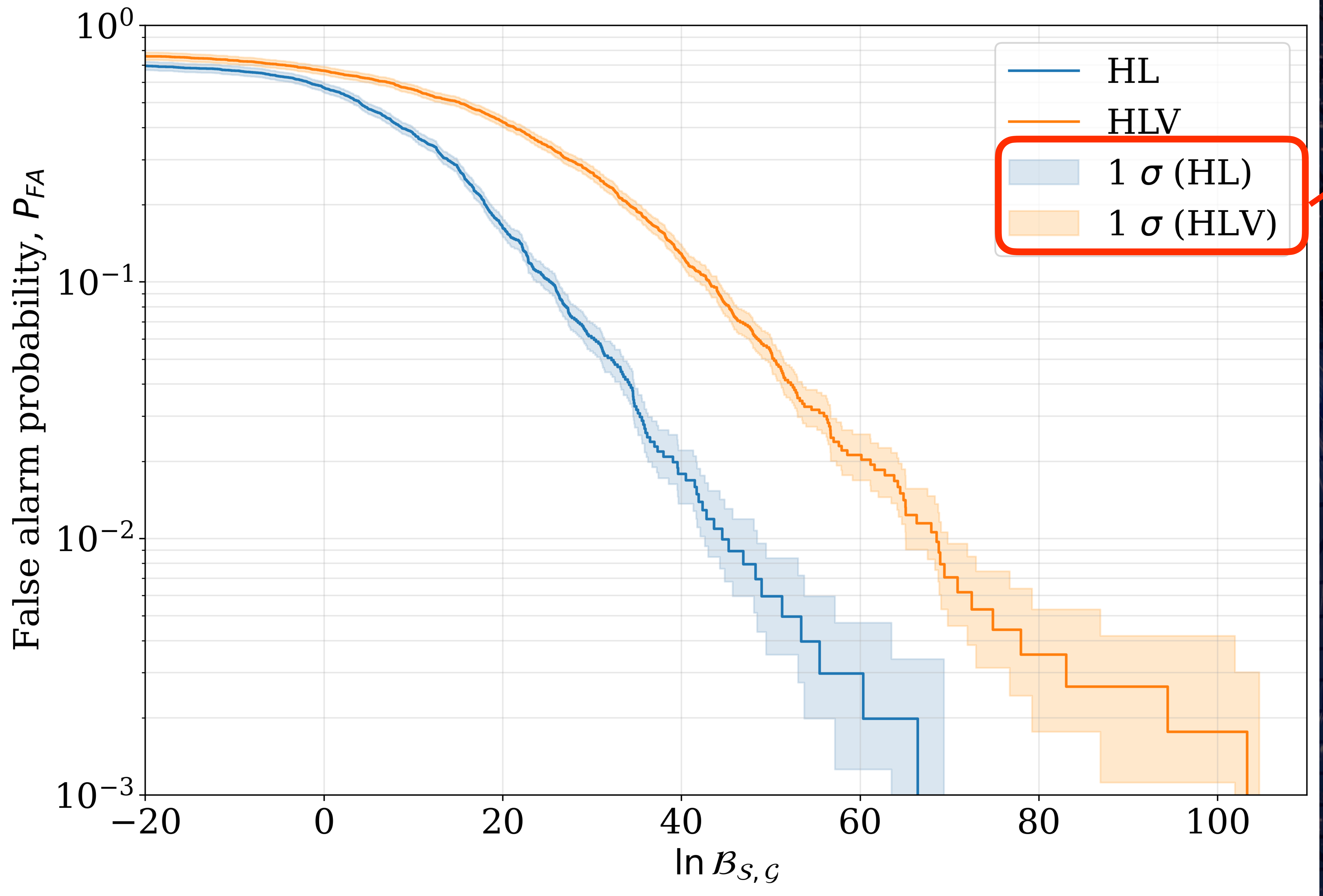
Why not False alarm RATE (FAR)?

- FAR is a time-averaged quantity
- The timescales of HL and HLV time-slide background data are different

- P_{FA} is time-independent and marginalises over the number of triggers analysed is more suitable for comparison between the two networks

X

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT



STEP 3: Error bars

Assume Poisson process

s.t.

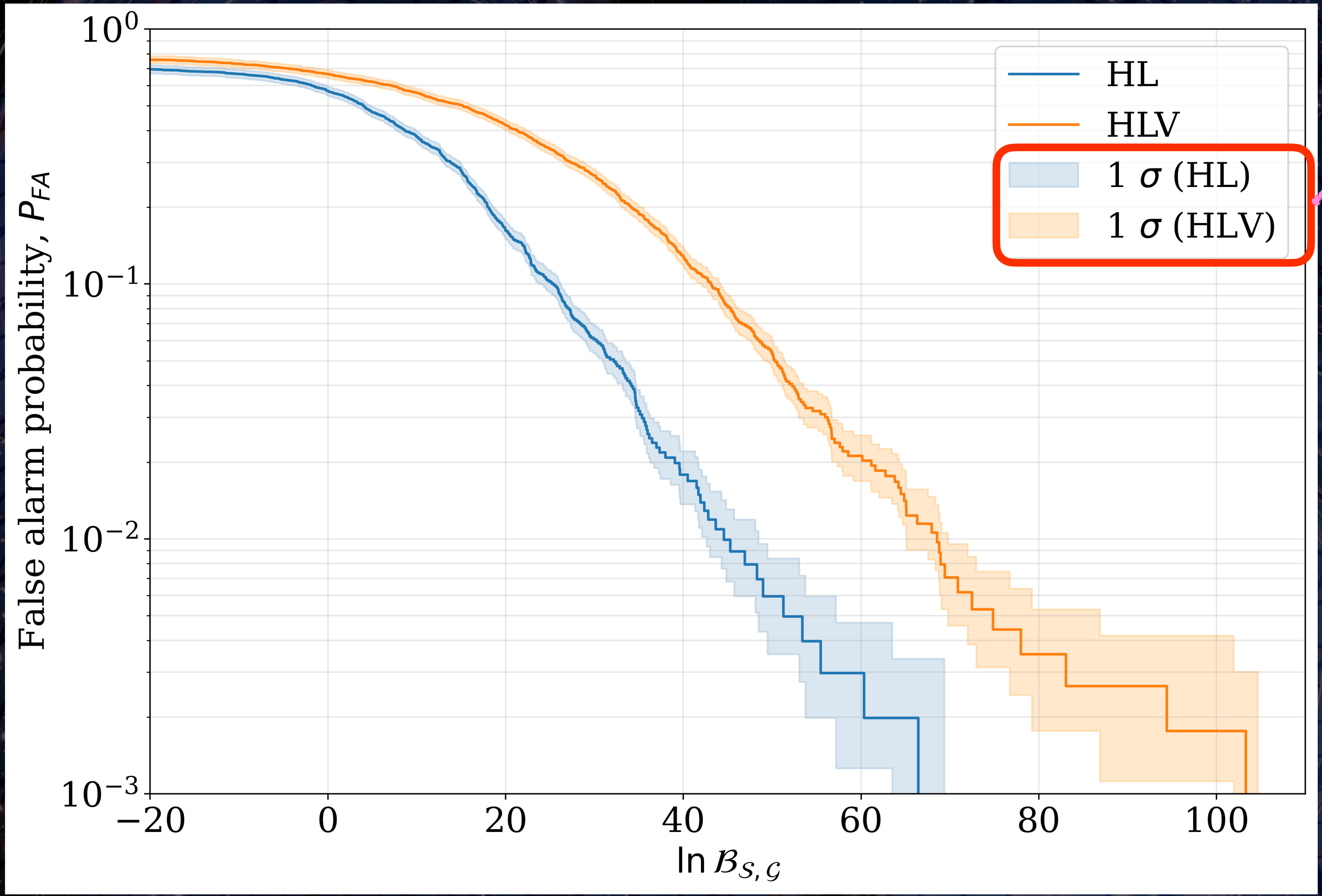
1-σ in background event counts, $n = \sqrt{n}$

$\therefore$

1-σ in P_{FA} is $\dfrac{\sqrt{n}}{N}$

N: Total number of triggers (1008 in HL, 1134 in HLV)

EFFICIENCY CURVE:
(1) NOISE BACKGROUND MEASUREMENT



Poisson uncertainty region

Assuming background noise is a Poisson process

P_{FA} in HLV is <u>higher than HL</u> for all $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$

i.e. HLV background is NOISIER than HL

EFFICIENCY CURVE:
(2) ASTROPHYSICAL MEASUREMENTS: PHENOMENOLOGICAL INJECTIONS

EFFICIENCY CURVE: (2) ASTROPHYSICAL MEASUREMENTS

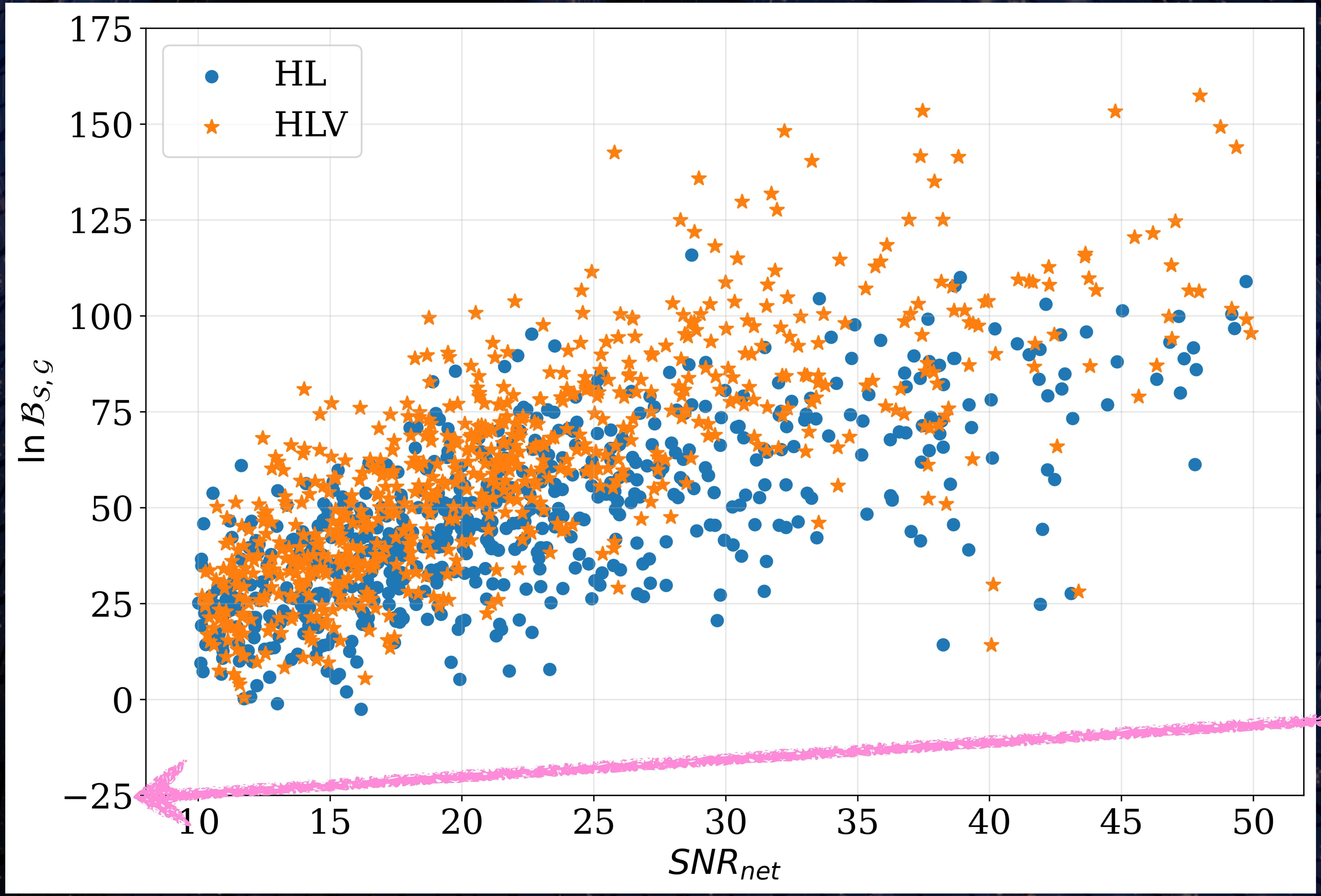
Properties of Injection Set 1 (IS1)

- **Phenomenological binary black hole (BBH) waveforms**
- **Injected into** randomly selected segments throughout all of **O3a**
- 790 injections, all satisfying $\mathrm{SNR}_{\mathrm{net}} \geq 10$

  c.f. Only followed-up cWB (background) events with $\rho \geq 7$

EFFICIENCY CURVE: (2) ASTROPHYSICAL MEASUREMENTS

Distribution of IS1 injections



Consistent with previous work:

$$\ln \mathscr{B}_{\mathcal{S},\mathcal{G}} \sim \mathcal{I} N \ln \mathrm{SNR}_{\mathrm{net}}$$

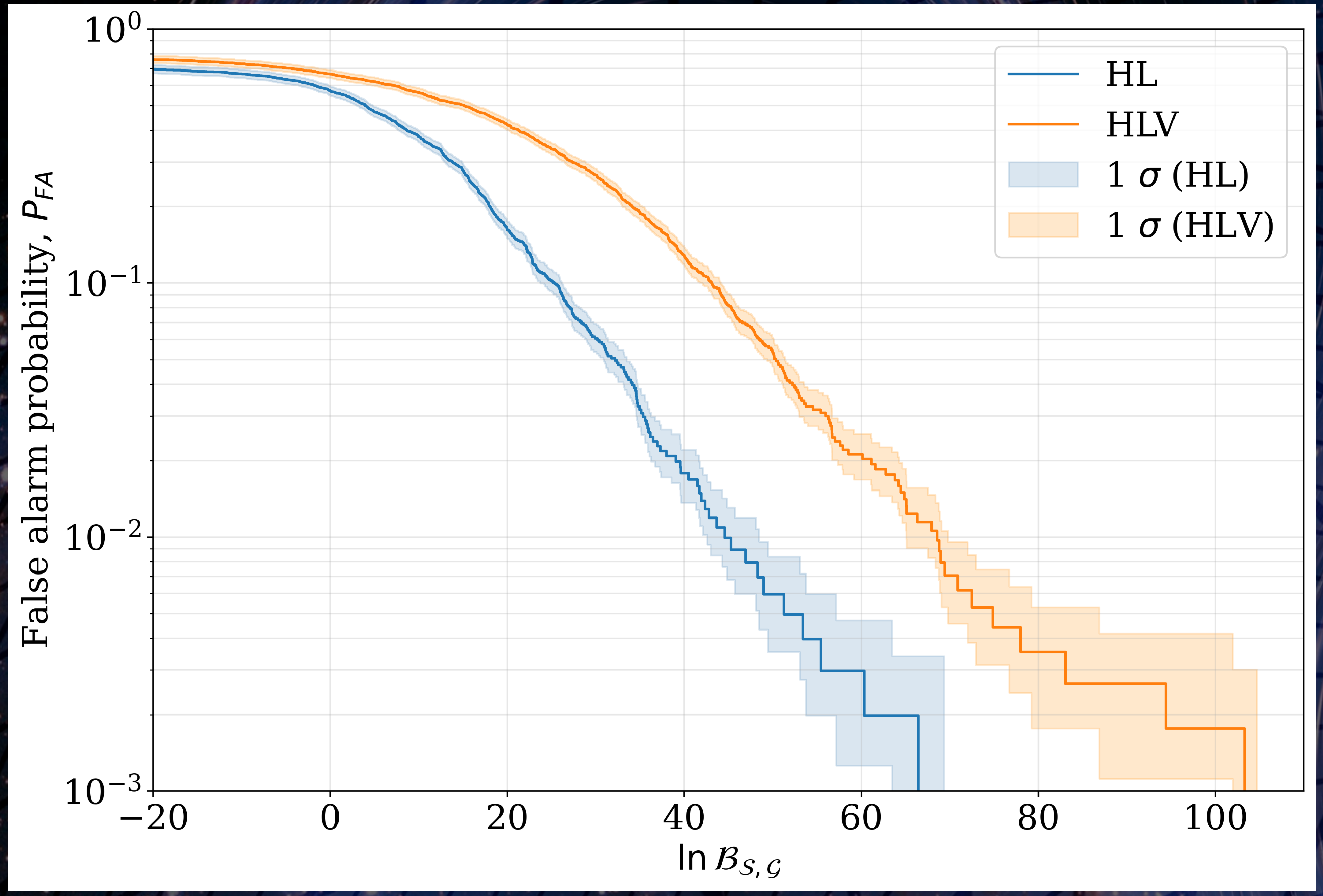
$\ln \mathscr{B}_{\mathcal{S},\mathcal{G}} = -500$ events not shown

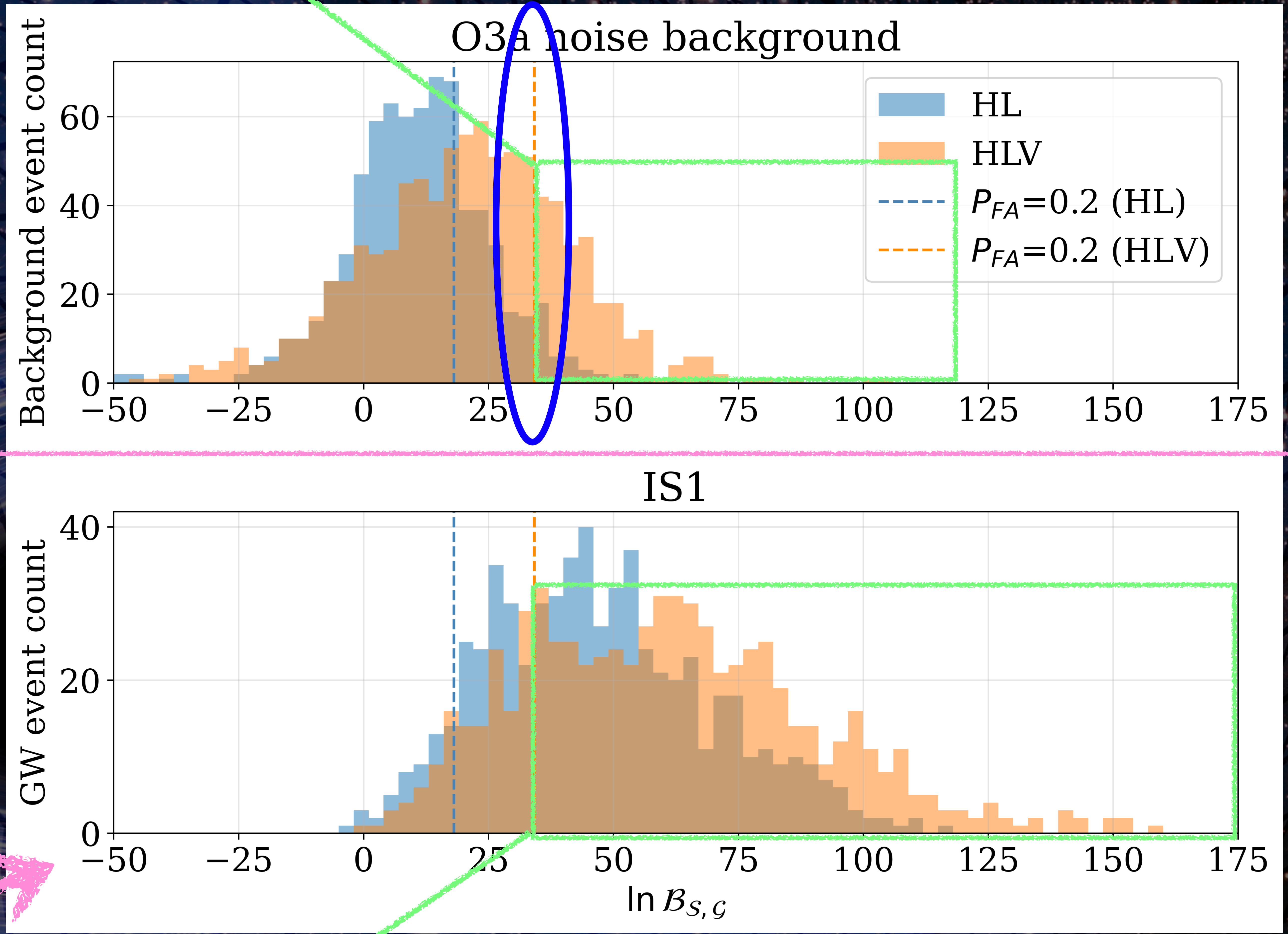
CONSTRUCTING EFFICIENCY CURVE

$P_{\mathrm{FA}} = 0.2$

/probability of false alarm/

Fraction of non-astrophysical triggers detected above a given $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ threshold





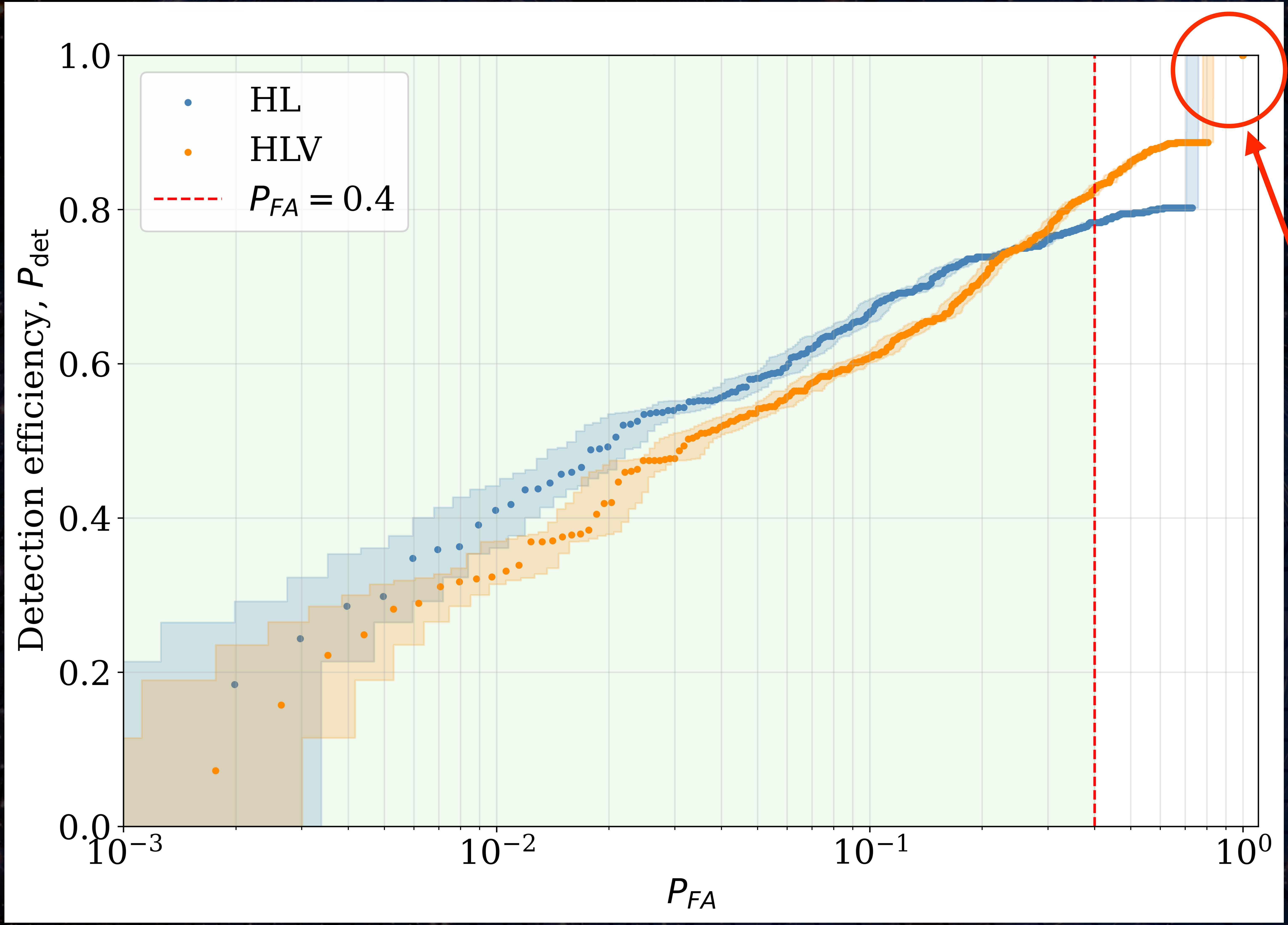
Again, $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}} = -500$ events not shown

P_{det}

/Detection efficiency/

Fraction of astrophysical events detected with significance i.e. $P_{\mathrm{FA}} \geq 0.2$

EFFICIENCY CURVES OF HL AND HLV



1-σ P_{FA} **POISSON UNCERTAINTY REGION**

same as background measurement

Cluster of Gaussian-noise-like events with

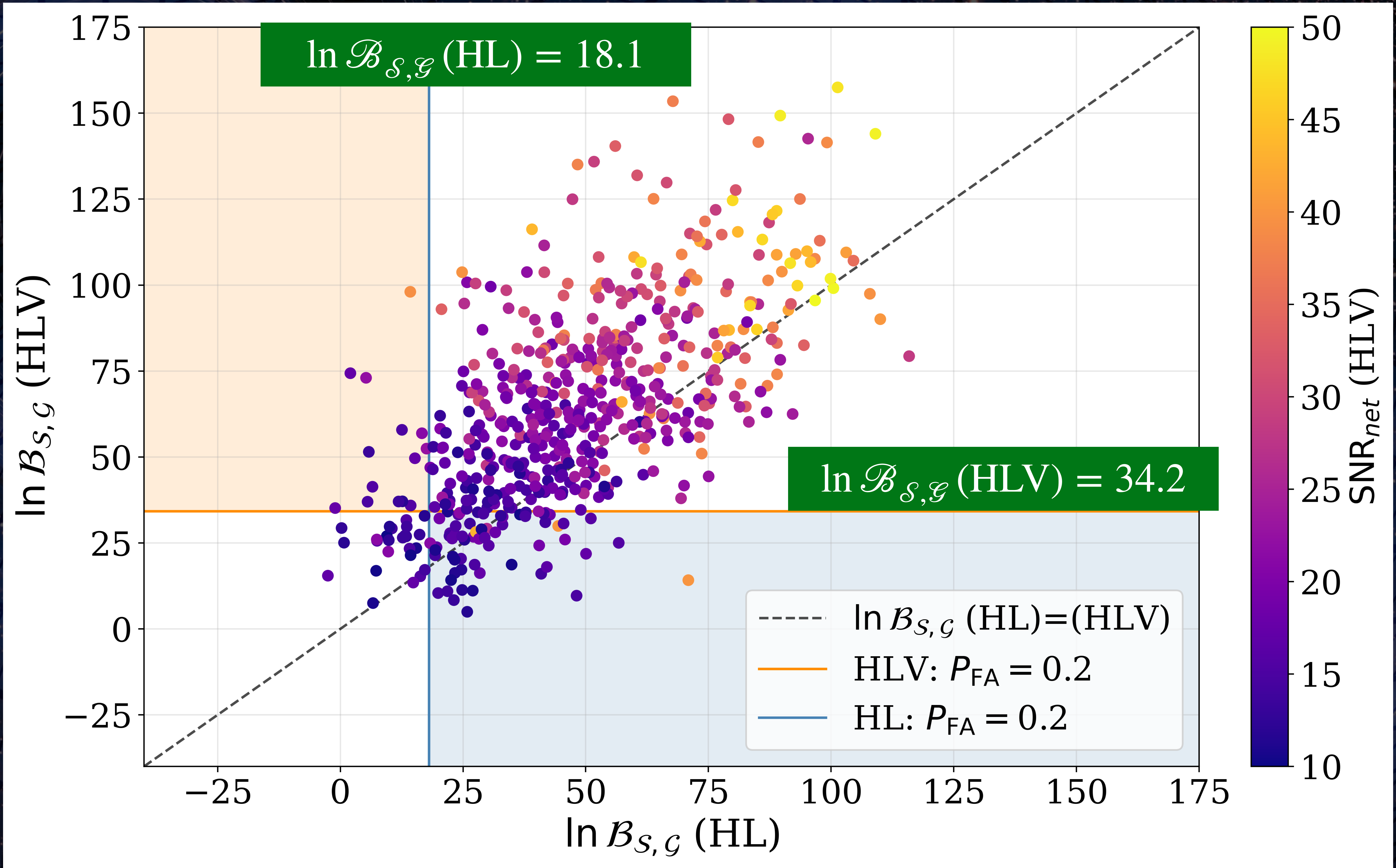
$\ln \mathscr{B}_{\mathscr{S},\mathscr{G}} = -500$ and $P_{\mathrm{FA}} = 1$

(i.e. minimal detection significance)

Practically meaningful regime $P_{\mathrm{FA}} \leq 0.4$

No significant difference between P_{det} of HL and HLV

P_{FA} AND $\ln \mathscr{B}_{\mathscr{S},\mathscr{G}}$ COMPARISON



X

INJECTION SET 2 (IS2)
O3-LIKE CBC INJECTIONS

(1) A consistency check

(2) To measure *BayesWave*'s detection significance of O3 GW candidates

Using the HL and HLV backgrounds measured in this study

O3-LIKE CBC INJECTIONS

Properties of Injection Set 2 (IS2)

- **Off-source waveforms:**
  - Waveform parameters sampled from matched-filter posteriors of O3 GW events
  - Injected in the proximity of the actual event
- **18 GW events, each with 50 off-source waveforms**
  - Only follow-up injections with $\mathrm{SNR}_{\mathrm{net}} \geq 10$

O3-LIKE CBC INJECTIONS

Properties of Injection Set 2 (IS2)

- **Off-source waveforms:**
  - Waveform parameters sampled from matched-filter posteriors of O3a/O3b candidate GW events, injected in the proximity of the actual event
  - Consistency test candidates in GWTC-2/GWTC-3
- 17 candidates from O3a, 5 candidates from O3b
  ** O3b waveforms are injected into random segments of O3a data
- 50 off-source waveforms per GW candidate
  i.e. (17+5) x 50 = 1100 injections in IS2

X

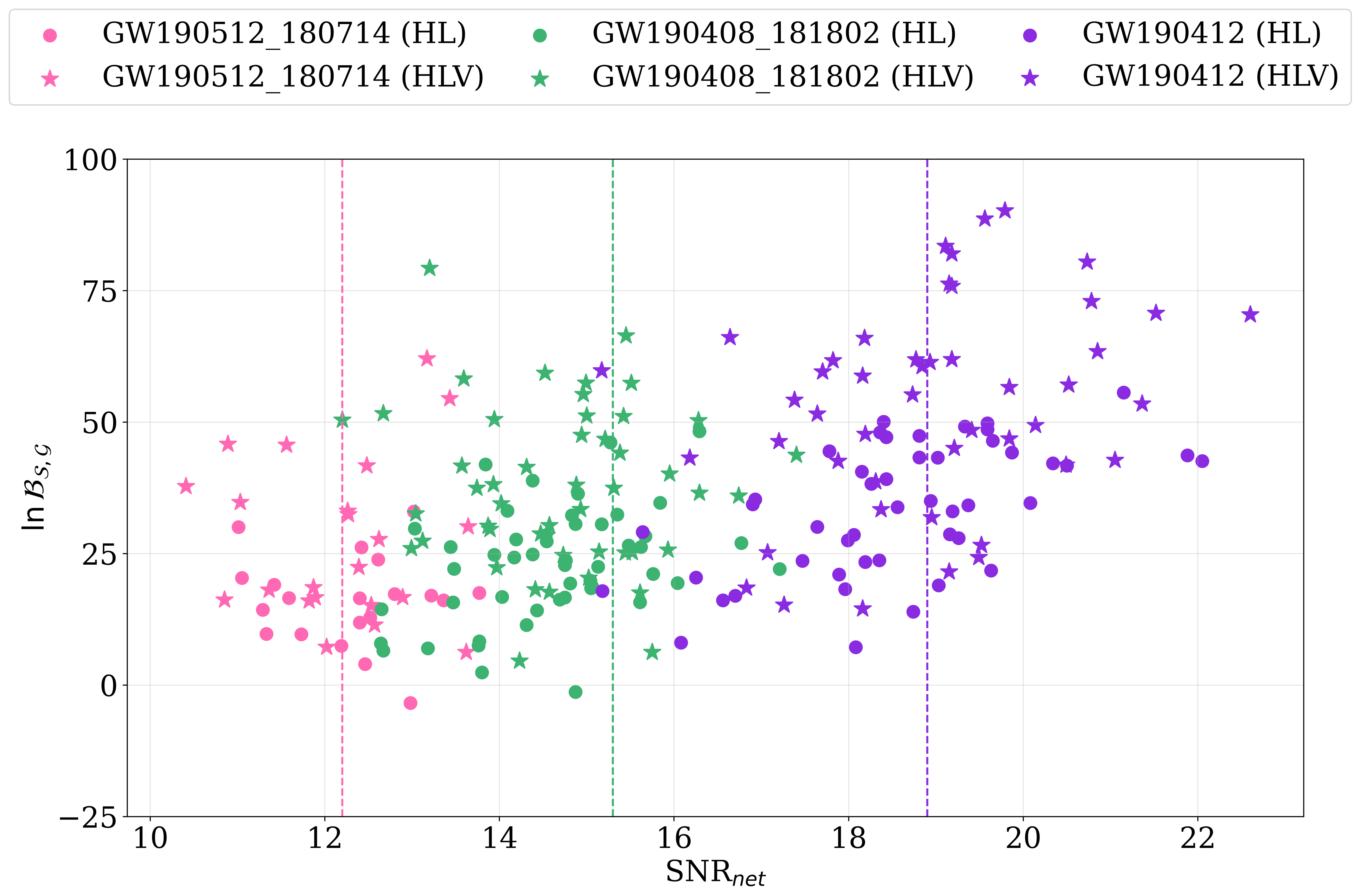
O3-LIKE CBC INJECTIONS

Filtering Injection Set 2 (IS2)

- **Non-detections:**
  - Only follow-up injections with $\mathrm{SNR}_{\mathrm{net}} \geq 10$
  - 4 (low match-filter SNR) GW candidates removed $\because \leq 25$ (less than half of) off-source waveforms satisfy *BayesWave*'s detection criteria
- **Gaussian-noise-like injections**
  - Assign arbitrarily low detection statistics: $\ln \mathscr{B}_{\mathscr{S},\mathscr{G}} = -500$

O3-LIKE CBC INJECTIONS

Visualising distributions of off-source injections



Median match-filter SNR

GW190512_180714 → 12.2

GW190408_181802 →15.3

GW190412 → 18.9

X

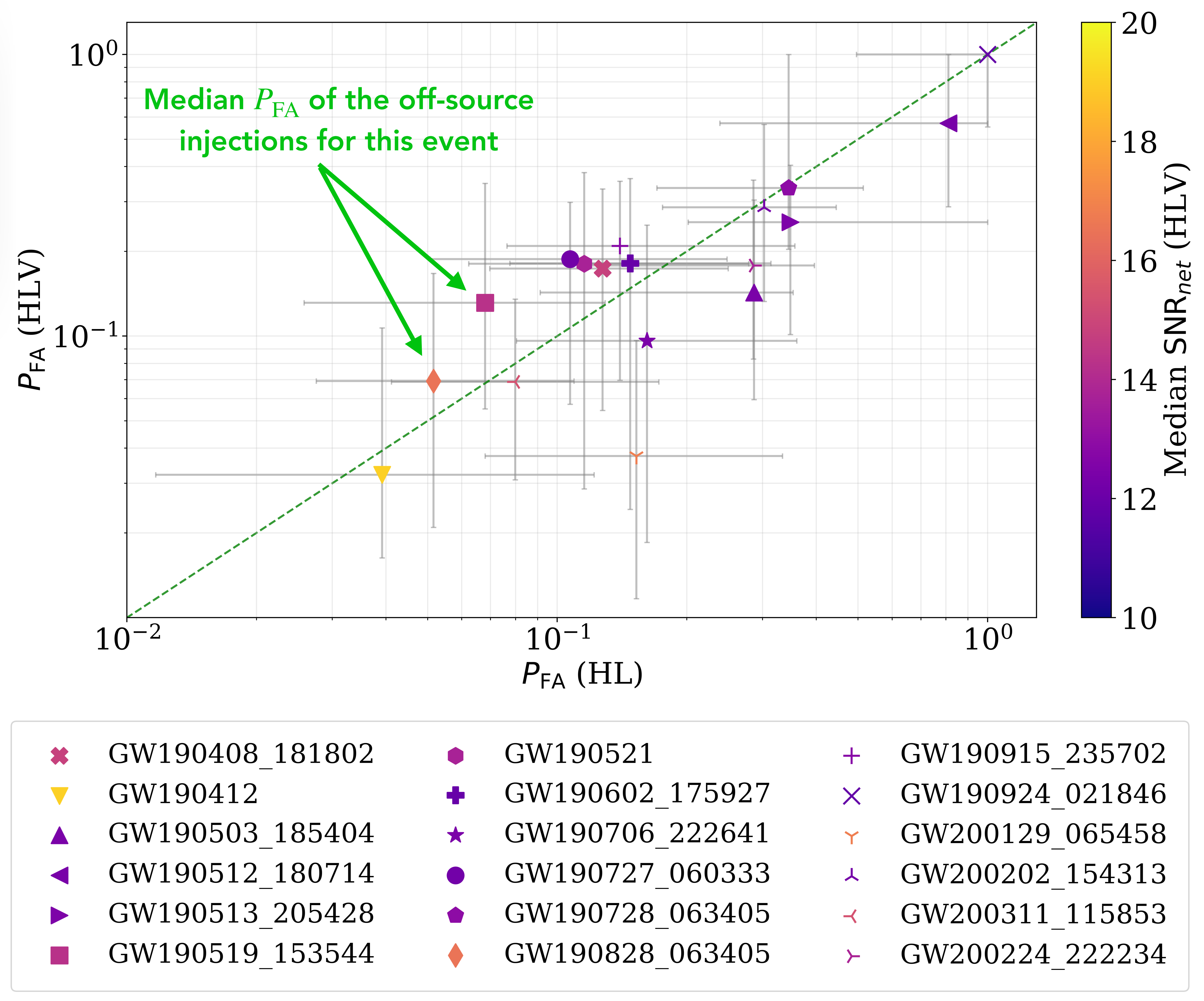
P_{FA} comparison - HLV vs. HL

18 O3-LIKE CBC INJECTIONS

Error bars

IQR of P_{FA} within each GW candidate

All events are detected with comparable significance between HL and HLV



SUMMARY OF RESULTS

- **Both the phenomenological BBH injections (IS1) and O3-like events (IS2) show comparable *BayesWave* performance with HL and HLV**
  - Although $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ increases with more detectors,
    false alarm probability P_{FA} is also higher (i.e. noisier detector backgrounds)
    ∴ Larger detector networks need to attain higher $\ln \mathscr{B}_{\mathcal{S},\mathcal{G}}$ to achieve the same significance (P_{FA}) as smaller networks

- **Findings consistent with:**
  - LVK O3 all-sky burst search (Phys. Rev. D **104**, 122004)
  - ML-enhanced cWB burst search, Szczepańzyk et al. (2023)

@AstroOrca

KEY TAKEAWAY

BayesWave's detection efficiency is comparable for two- and three-detector networks

because

more frequent glitches in expanded detector networks offsets the advantage of higher detection statistics ($\ln \mathscr{B}_{\mathscr{S},\mathscr{G}}$)

Link to arXiV

Accepted PRD - arXiv:2403.16837 (2024)