Detecting Gravitational Waves using the Public LIGO-VIRGO Data

Barak Zackay
Weizmann institute of Science

A little bit about me

- Main tools: Algorithms and Statistical methods
- Goal: Making new discoveries.

- In every modern project in astronomy, data analysis plays a key role.
- Improvement in analysis methods directly translates to better sensitivity, **to all similar surveys**.
- Many times data is under utilized.

Example - Image Subtraction

- Problem: $\frac{\#\mathrm{FalsePositives}}{\#\mathrm{SN}} \approx 10$ (AFTER cleanup efforts)
  - Human candidate scanning required (No automatic followup)
    - Does not scale.
- Science depending on image subtraction:
  - GW counterparts, Tidal disruption events, supernovae, Microlesning, Novae,…

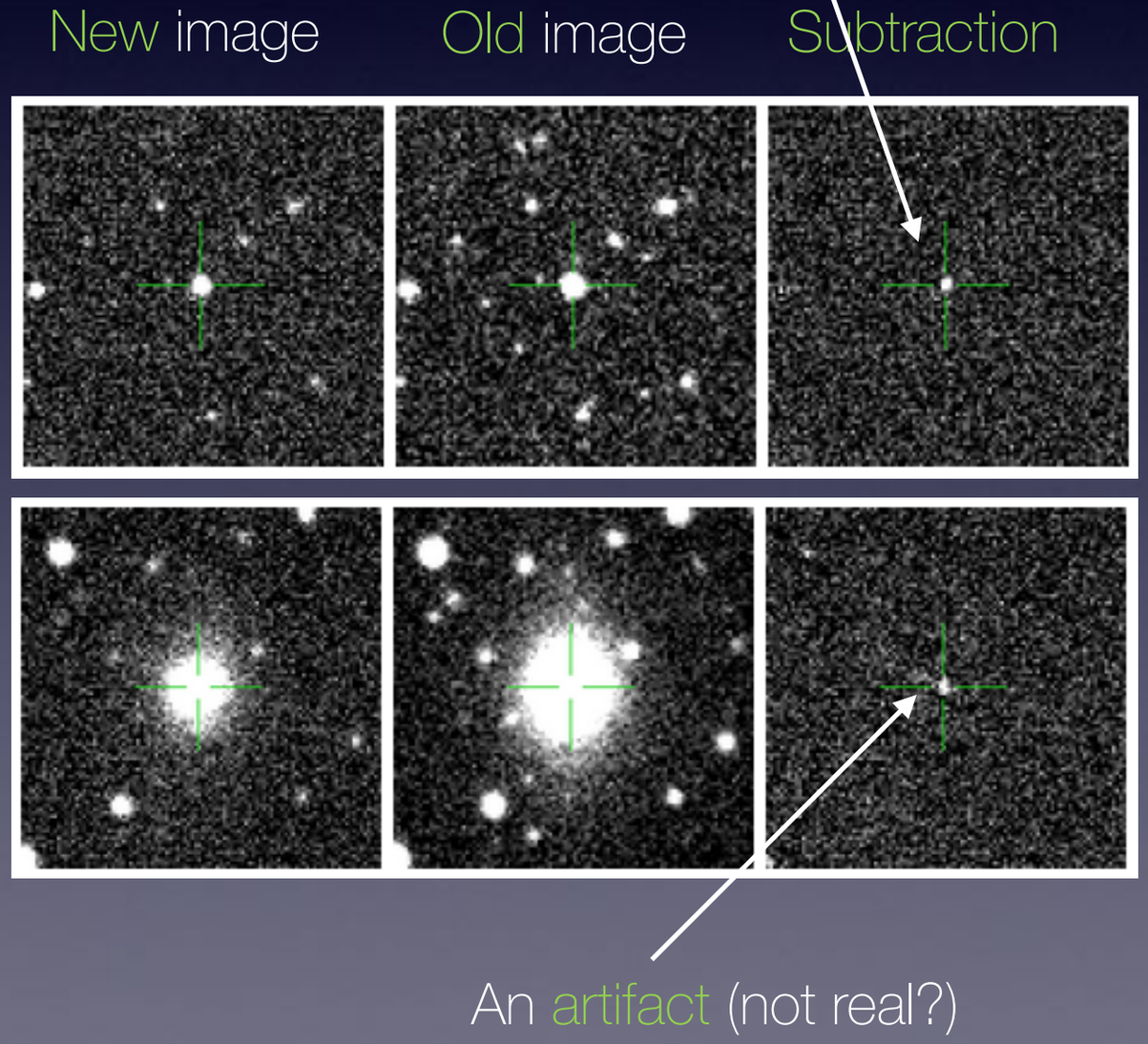


Image Subtraction: The real problem

- No actual derivation of used method
- Noise in reference image ignored.
- When PSFs don't match, effective deconvolution happens.

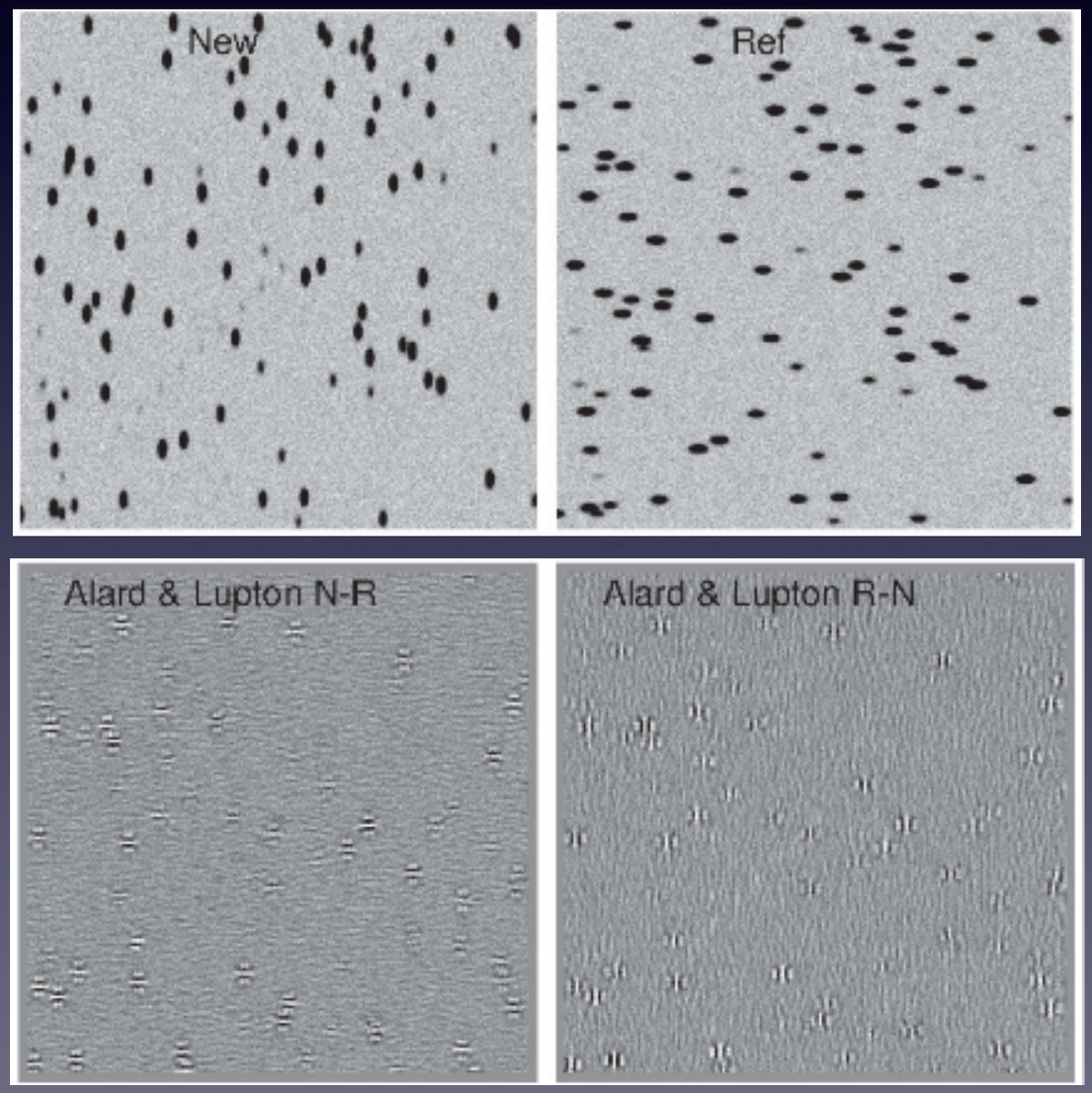
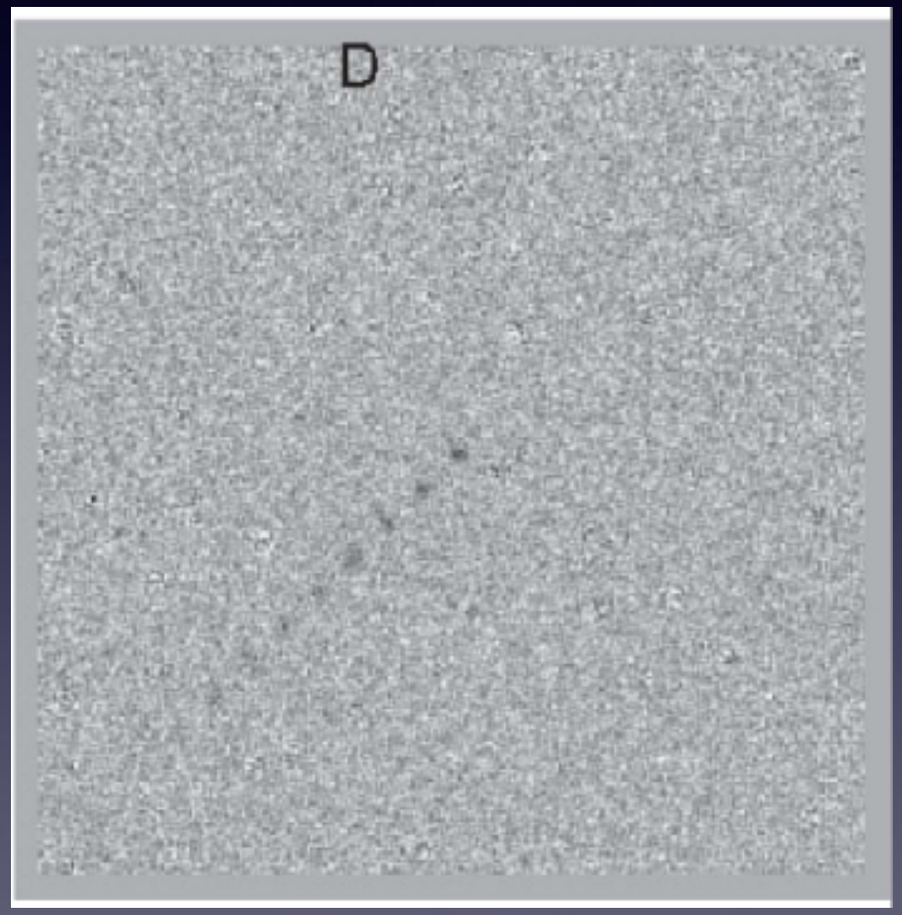


Image subtraction - Solution

- Identify the problem.
- Write statistical model encompassing all relevant effects
- Solve it from first principles (BZ, Ofek & Gal-Yam, 2017)
- Results:
  - False positives gone.
  - Automatic followup on ZTF.
  - 20-50% more survey speed.
- Many (>6!) independent implementations



More examples

- Image addition [+10-20% survey speed for ground based surveys (BZ & Ofek 2017 a,b)]
- Detecting FRBs [x100 reduction in computational cost (BZ & Ofek 2017c), ~20% increase in sensitivity (upcoming, used by Law+ 2020)]
- Fast parameter estimation for Gravitational Waves (BZ, Dai, Venumadhav 2018 and Dai, Venumadhav, BZ 2018).[$\times 10 - \times 10^5$ reduction in computational cost]

In the making (don’t hold your breath)

- **Pruning algorithm** - Detecting Pulsars in binary systems [Vishnu Balakrishnan (will start soon)].
    - Also 10 other problems, for example: continuous gravitational waves, merging black holes with LISA/ET/CE, planet 9] - (**Hiring!**)
- Solving for periodicity in very sparse time of arrival measurements. (**Hiring!**)
  [FRB periodicity, gamma-ray pulsars, other surprising applications in physics]
- Reducing the “technichal noise” in advanced LIGO Data [with IAS team] **(hiring)**

I am hiring multiple students and postdocs

Detecting Gravitational Waves

Barak Zackay, Tejaswi Venumadhav, **Jonathan Mushkin,** Liang Dai,
Javier Roulet, Seth Olsen, Horng-Sheng Chia and Matias Zaldarriaga

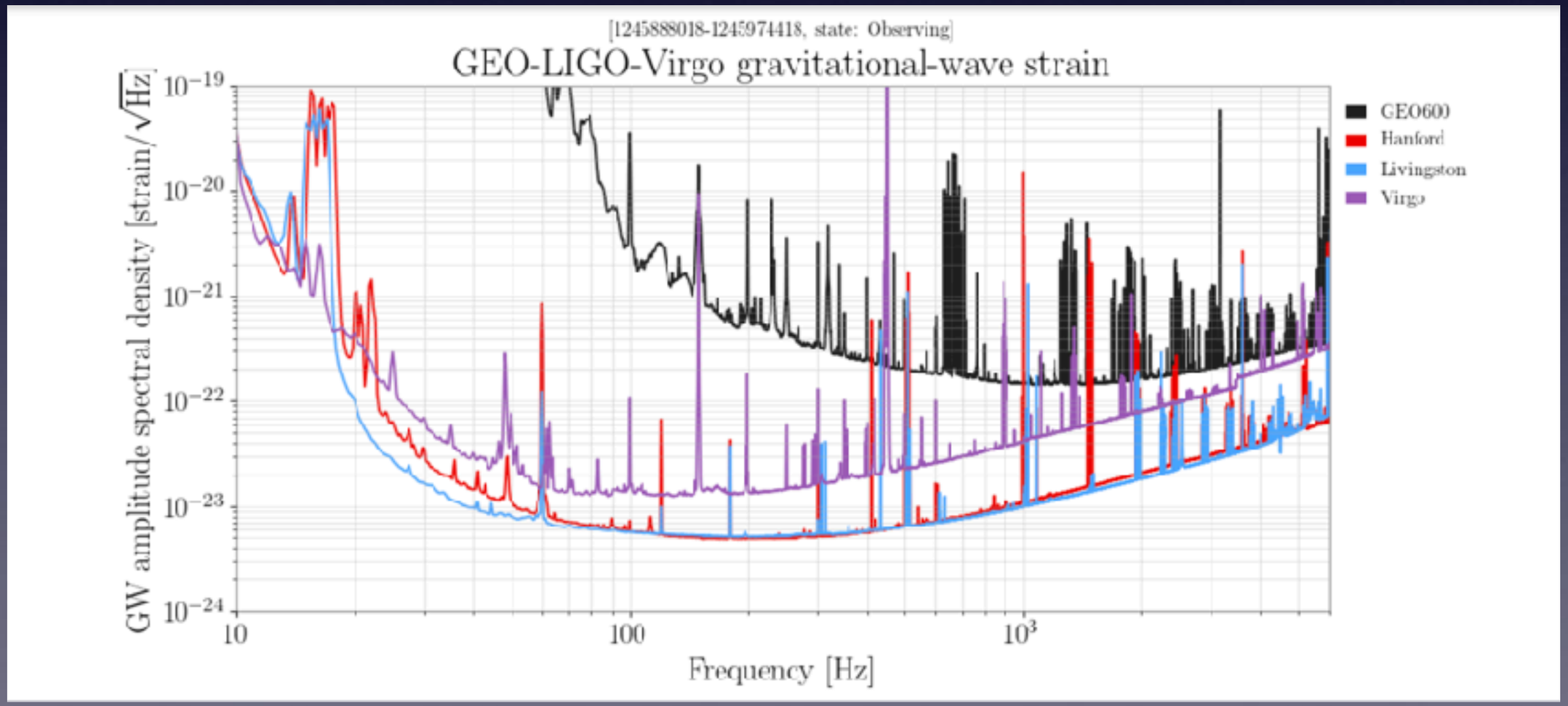
Gravitational waves detected

- Ten Binary Black Holes [GWTC-1].
- Binary Neutron star merger (along with EM counterparts across all the EM spectrum).
- We constructed an independent pipeline and discovered additional 8 Binary Black Hole signals in O1 and O2.
  (BZ+ 2019a,b,c and Venumadhav, BZ + 2019a,b)
- Population analysis (Roulet, Venumadhav, BZ+ 2020)
- Lensing search (Dai, BZ+ 2020)
- Another independent group had recently adopted a few of our improvements, and detected the LVC events + 3 of our newly reported events.
- Inferring population models from O3a data (Roulet+ 2021)
- A possible detection of precession in GW151226 (Chia+ 2021)

Intro to detecting merging compact objects in strain data.

Data Model

- Gaussian noise, with a high dynamic range power spectrum.



Signal Model

- System's intrinsic parameters are:
  - Total mass, mass ratio, effective spin (mass weighted average of out of plane spin).
  - In plane spins, tidal deformadibility.
- Extrinsic parameters - Distance, sky position, orbital orientation.

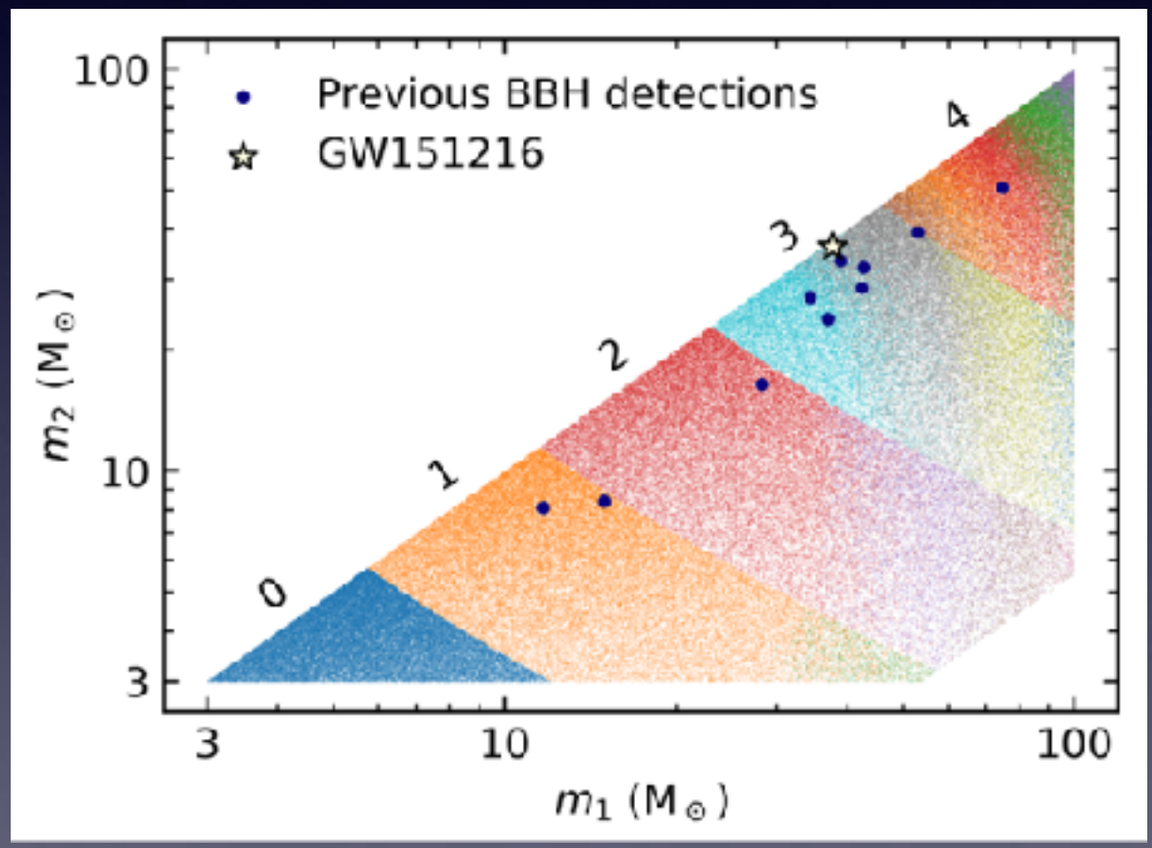
Single Detector Detection Statistic

- Fourier domain matched filtering - Simple inverse variance weighting

$$S = \frac{\sum_f \frac{h(f)T^*(f)}{S_n(f)}}{\sqrt{\sum_f \frac{|T(f)|^2}{S_n(f)}}}$$

Template Bank

- Templates are similar if using one to detect the other yields >97% of the S/N
- Should guarantee every physical waveform is represented.
- Every template is cross correlated with the data of all detectors.
- Times in which score is high (say SNR>4) are registered as “triggers”.



Joining Multiple Detectors

- Allowed time delay ~10ms

- Incoherent: Summing the single detector statistics (of the same template) in quadrature.

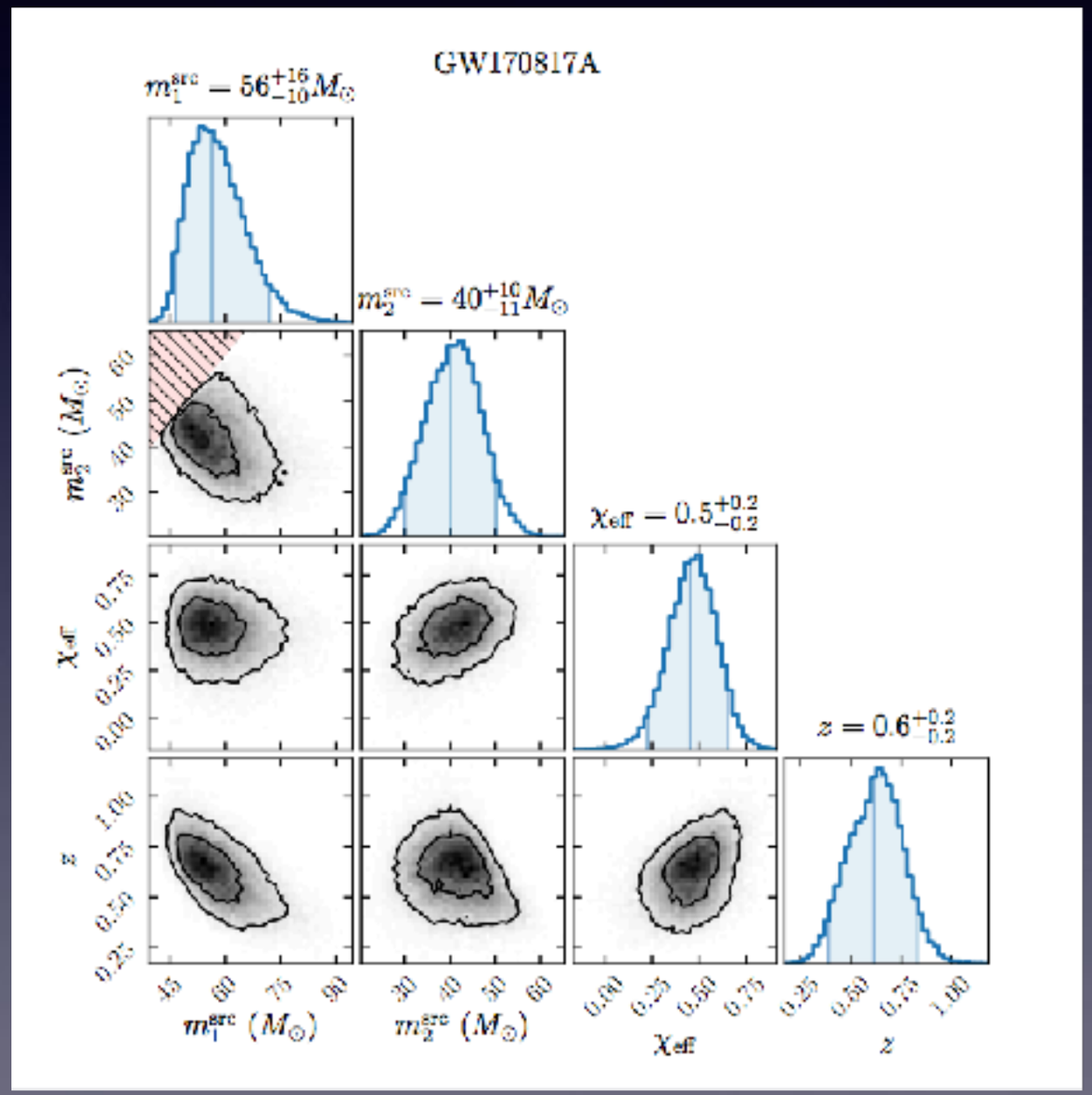
- Coherent: Detectors are “nearly” the same orientation in 3d => phases and responses of the LIGO detectors are expected to be close.

Computing False Alarm Rate "Time slides"

- Can exactly measure the background by repeating the multiple detector analysis with **unphysical** time delays
    - An important caveat - relevant when the signal in question is not in the "first place".
- A wonderful tool that keeps you honest. [Thank you LIGO]
    - In my view, many surveys can and should use it.

Measuring the parameters

- Likelihood assumes noise is Gaussian in vicinity of event.
- Define a prior distribution on physical parameters
- Sample the posterior



Why open the project? (July 2018)

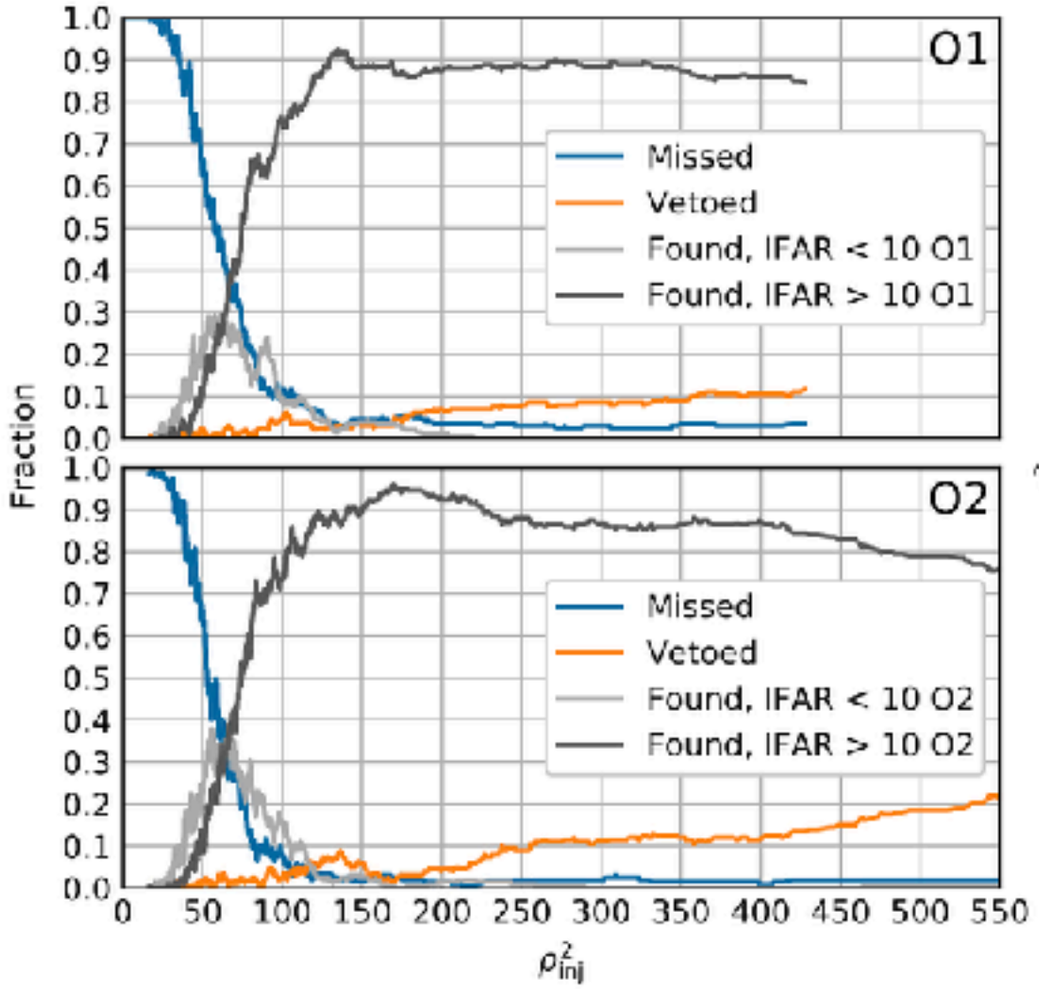
- Detection bar for LVC analysis was S/N ~ 10.
  - Barely detected GW151012 (S/N = 10).
  - One Trillion templates?
- Strategy for eliminating “glitches” seemed too simple.
  - One “silver bullet” chi2 test (unsafe to pure Gaussian noise).
- “Glitches” could not be the full story.
  - Inferred “glitch” rate is 1/ few sec? Is it really that bad?

List of improvements

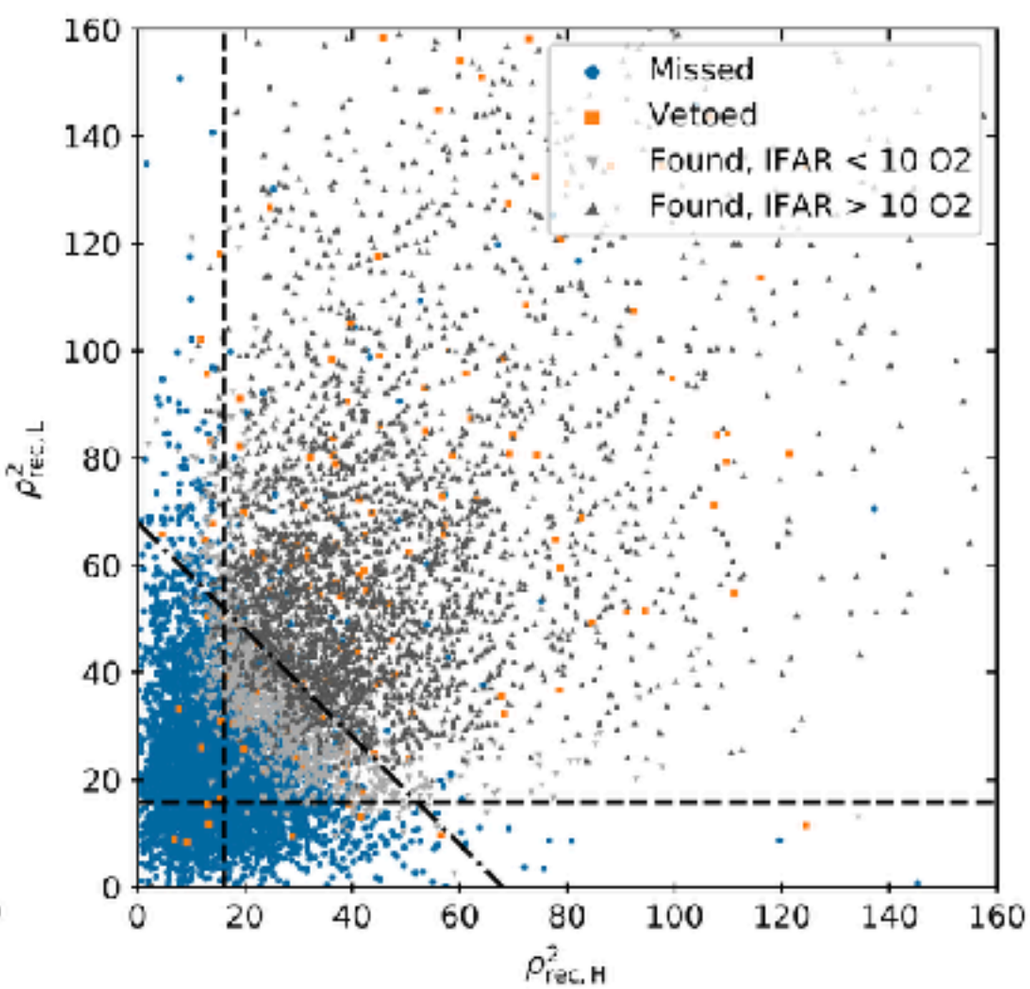
Analysis difference	ΔV(BBH)	ΔV(BNS)
Template prior that is flat in log mass	~30%	0%
Exact S/N determination	~30%	~30%
Agressive masking of bad data	0%	>10%
Inpainting of bad data	0%	>10%
Strong veto (vs chi2 renormalization)	~20%	>10%
Rank functions	~15%	0%
Treatment of disparate detector response	10-20%	10-20%
Total	~100%	?

Equivalent to reduction of ~4-5 orders of magnitude in background!

Injection study



(a) Injections in O1 (top) and O2 (bottom).



(b) Injections in O2.

A more “agnostic” prior

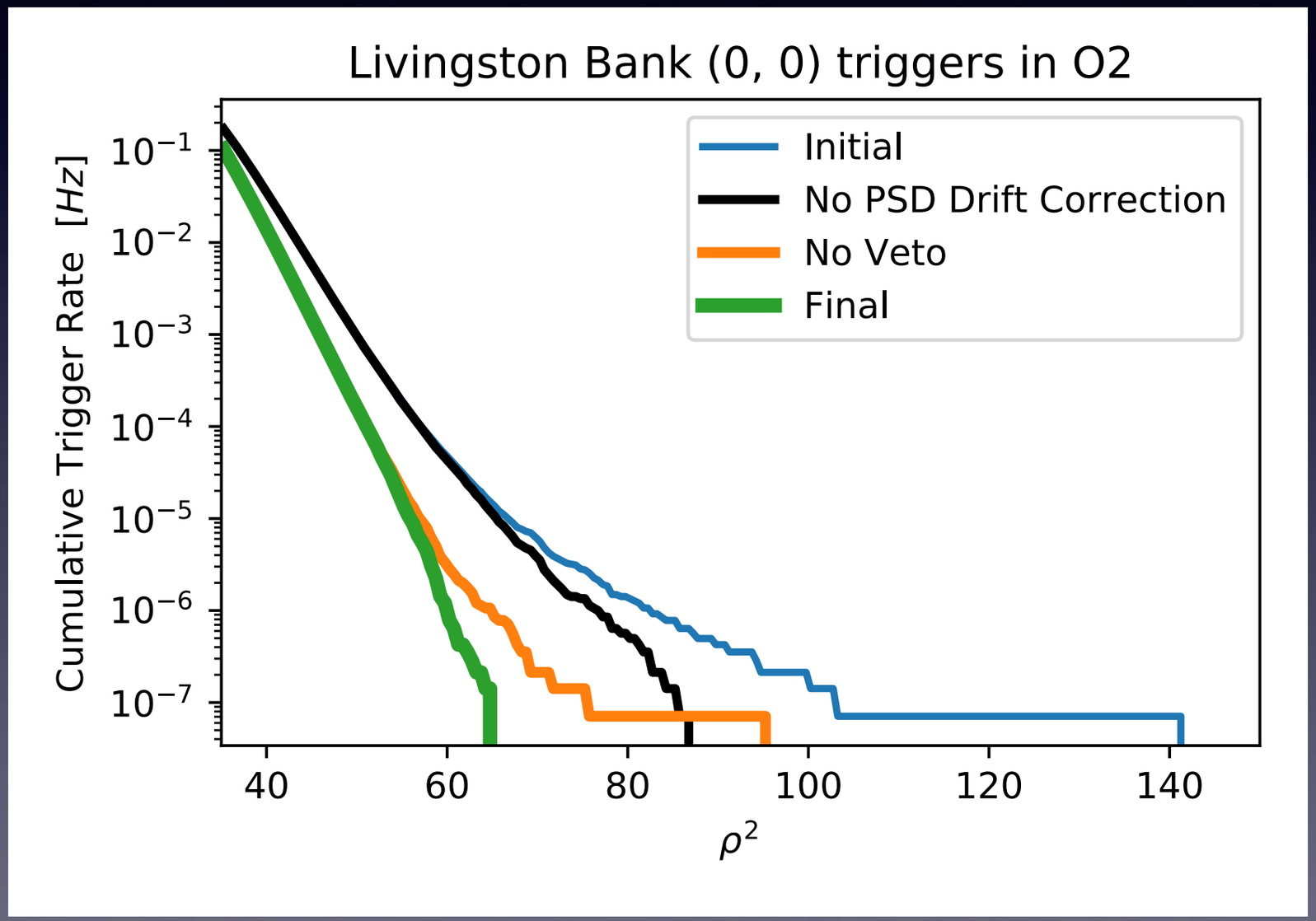
- LVC detection strategy:
  - All candidates from all templates treated equally
  - Do you believe that $R_{\mathrm{NSBH}} = 10^3 R_{\mathrm{BBH>20M_\odot}}$?

- Our detection strategy:
  - Divide to logarithmic banks depending on mass.
  - Phase-space within bank approximately constant.
  - Assumes rate is approximately constant between banks.

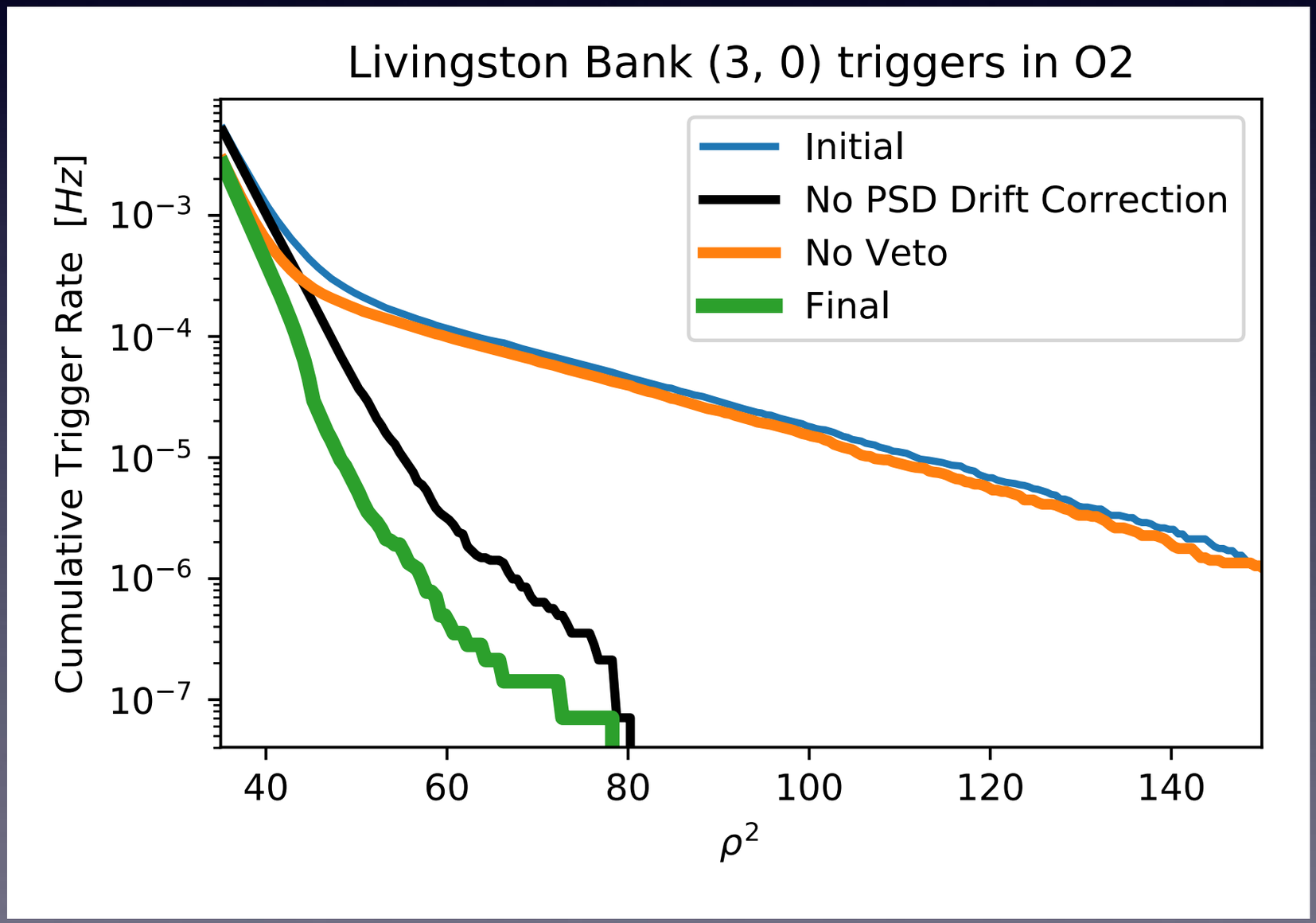
Exact SNR determination

- Problem: The PSD drifts faster than you can measure.
  - Manifestation: SNRs are mis-estimated to linear order => Tail in the trigger distribution
- Solution: linear order error is fully correctable with the measurement of one scalar.
  - Can precisely measure and track it's fast variability.
  - SNR measurable to 1% precision (and ~1% down from optimal).

Effect on detecting long, chirping events

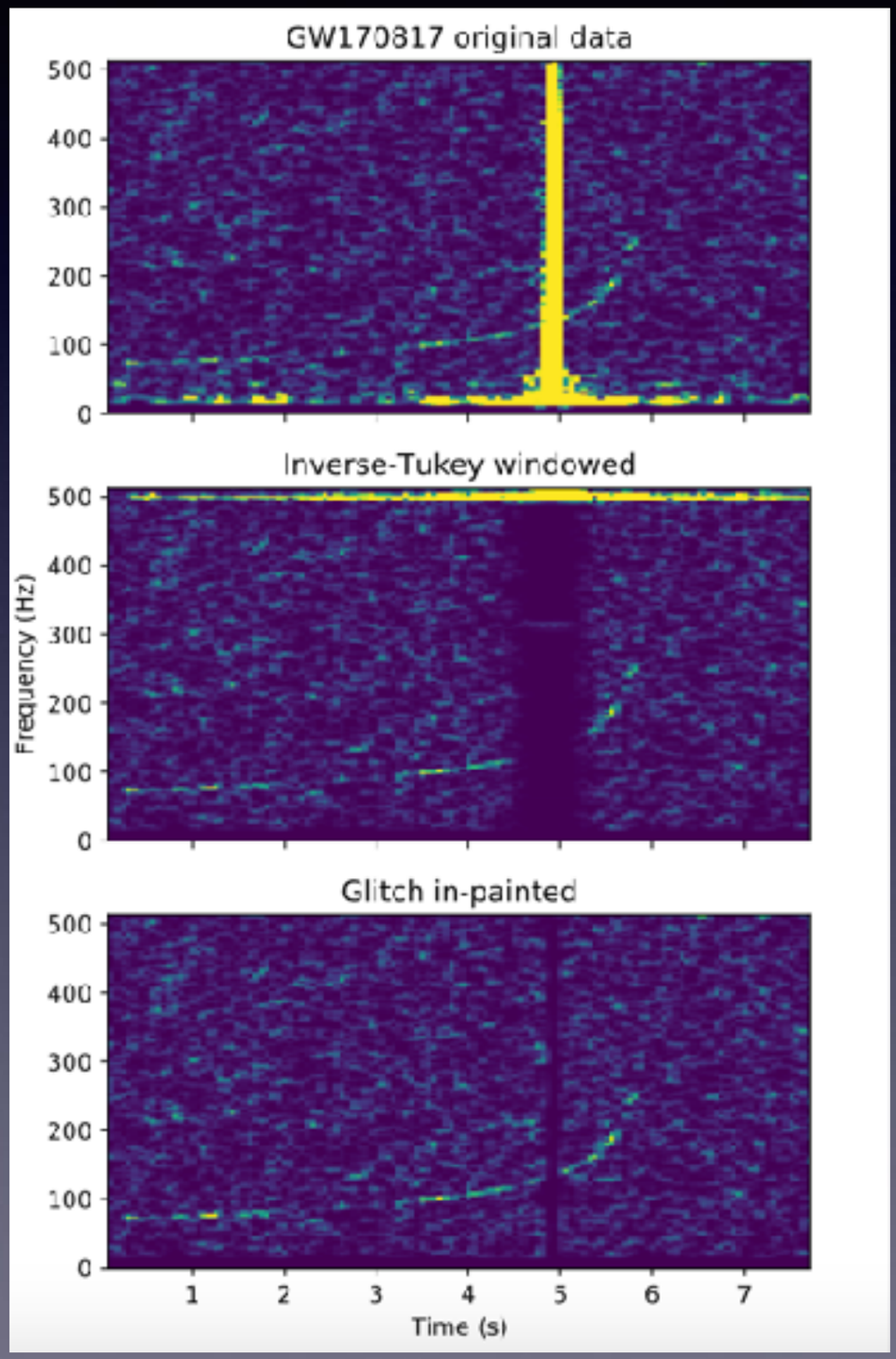


Effect on Heavy BBH



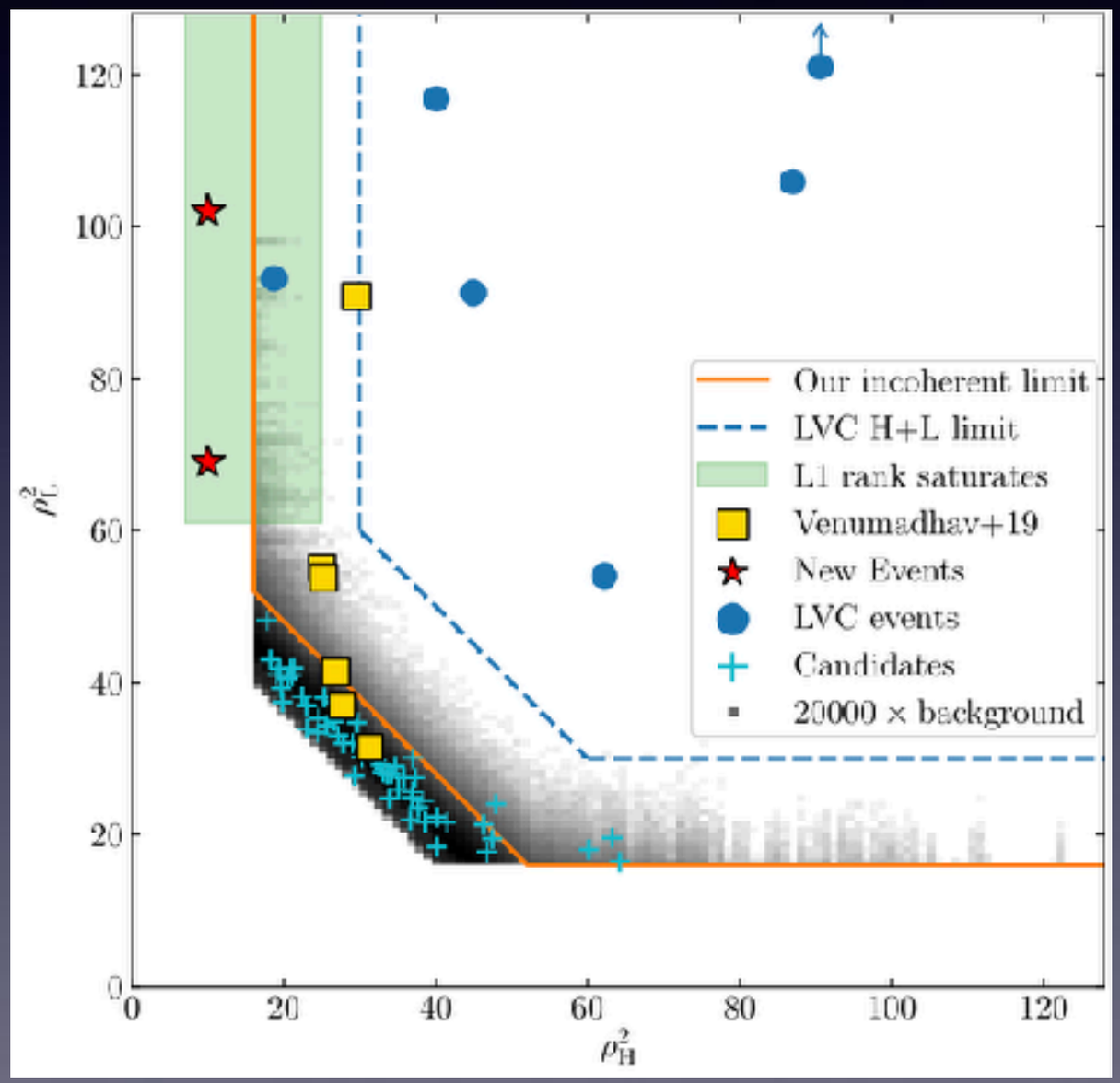
Inpainting

- Problem: Identified bad times are hard to remove.
- Solution: put an infinite noise term in “bad times” in the statistical model, solve again.
  - Effective data term emerges.
  - Replace bad data using linear predictor from surrounding samples.
  - Zero response to “delta function” events in bad times.

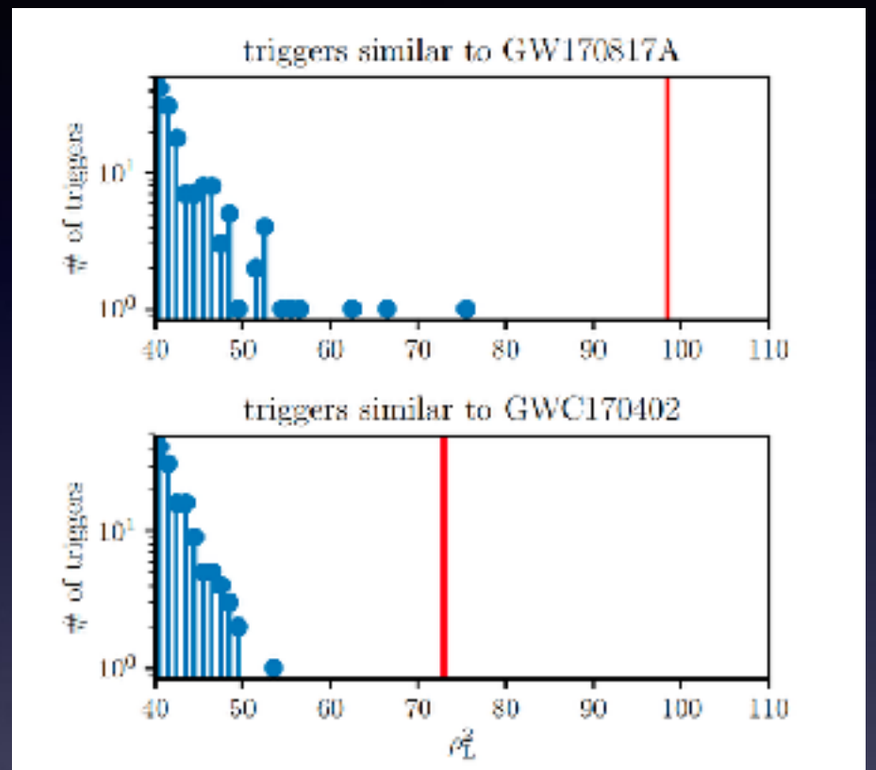


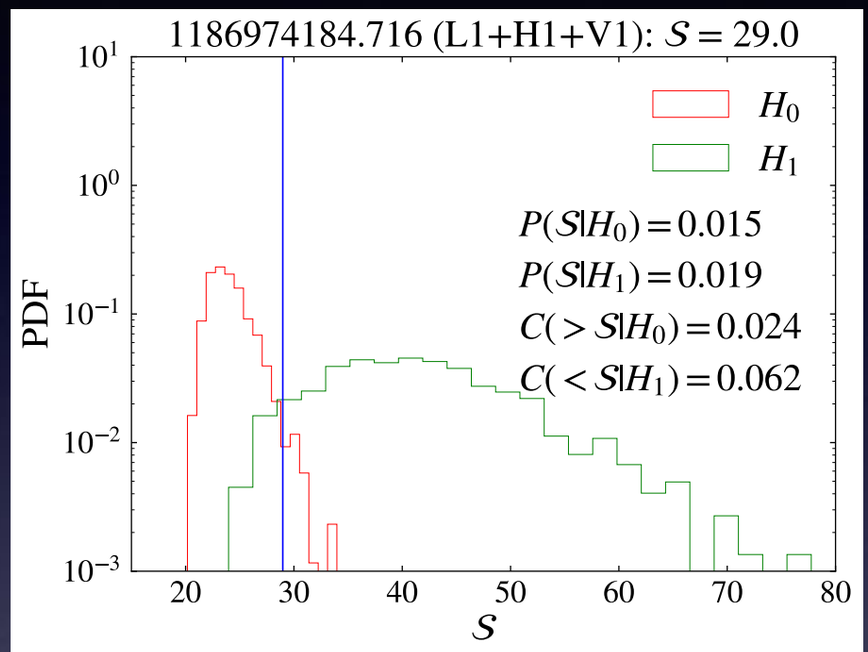
Improvement #7
“Fishing”

- Exact treatment of the case where detector response is highly disparate.
  - First stage:
    Separating Gaussian noise from “Glitch”/GW
  - Second stage:
    Global rank according to glitchiness of a template
  - Third stage:
    Scratch the bottom of the other detectors using exact marginalization over external parameters.



Fishing





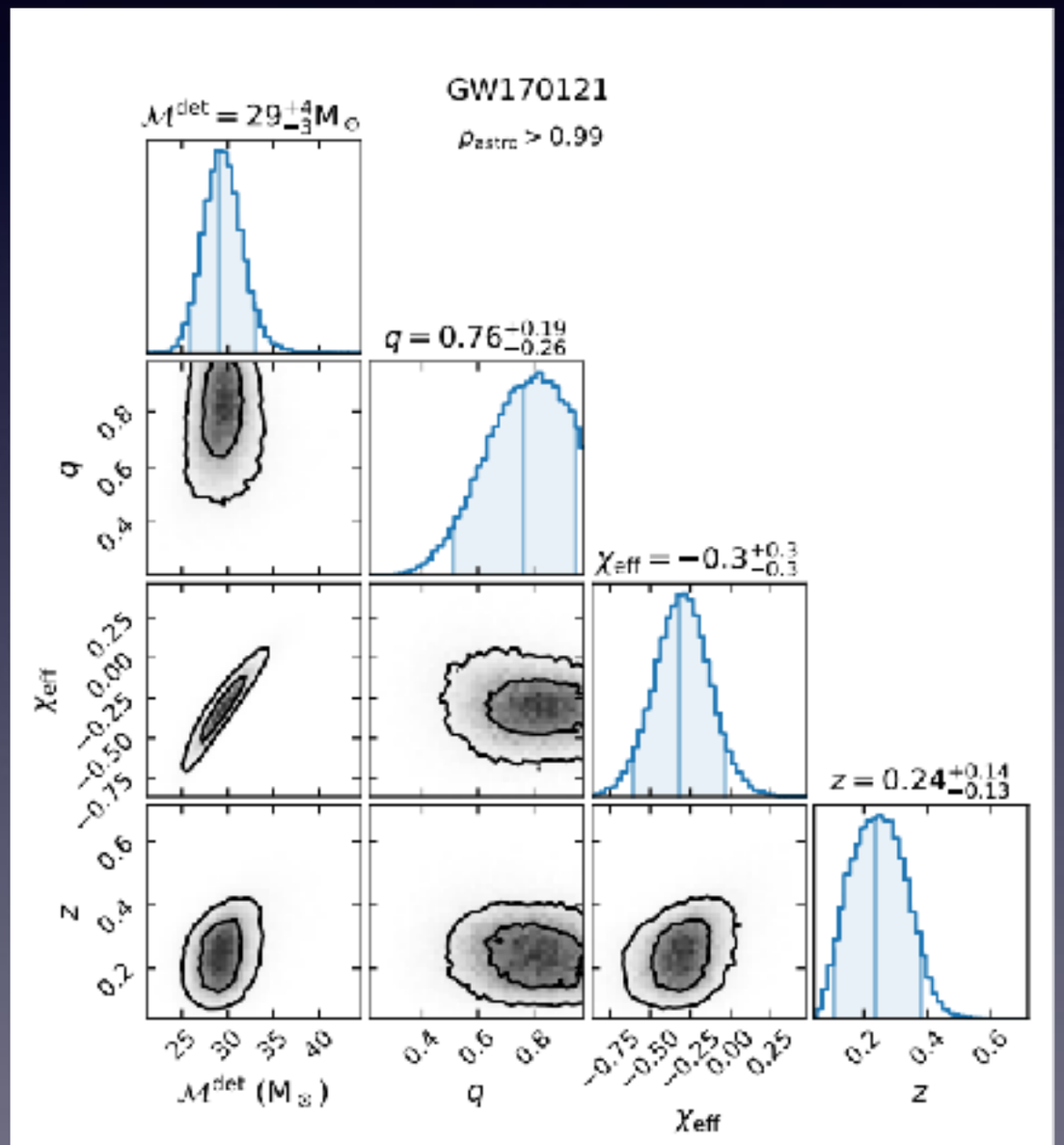
L1 rank	GPS time	ρ_L^2	# similar triggers	$C(S\|H_0)$	$C(S\|H_1)$	$\frac{P(S\|H_1)}{P(S\|H_0)}$	Comment
1	1187058327.068	93.1	0	$< 10^{-3}$	0.16	37	GW170818[a]
2	1187529256.504	92.1	0	-	-	-	GW170823
3	1169069154.564	90.8	0	-	-	-	GW170121
4	**1175205128.565**	**72.9**	**0**	**0.015**	**0.022**	**0.547**	**GWC170402**
5	1186741861.51	174.6	1	-	-	-	GW170814
6	1167559936.584	107.3	1	-	-	-	GW170104
7	1186302519.731	118.6	2	-	-	-	GW170809
8	**1186974184.716**	**98.5**	**5**	**0.028**	**0.055**	**0.98**	**GW170817A**
9	1174043898.842	75.7	9	0.36	0.001	0.008	Background
10	1170885005.109	66.4	16	0.49	0.003	0.013	Background
11	1178083239.592	74.4	22	0.34	0.003	0.016	Background
Removed[b]	1173477193.704	69.2	1	0.38	0.014	0.011	Artifacts present

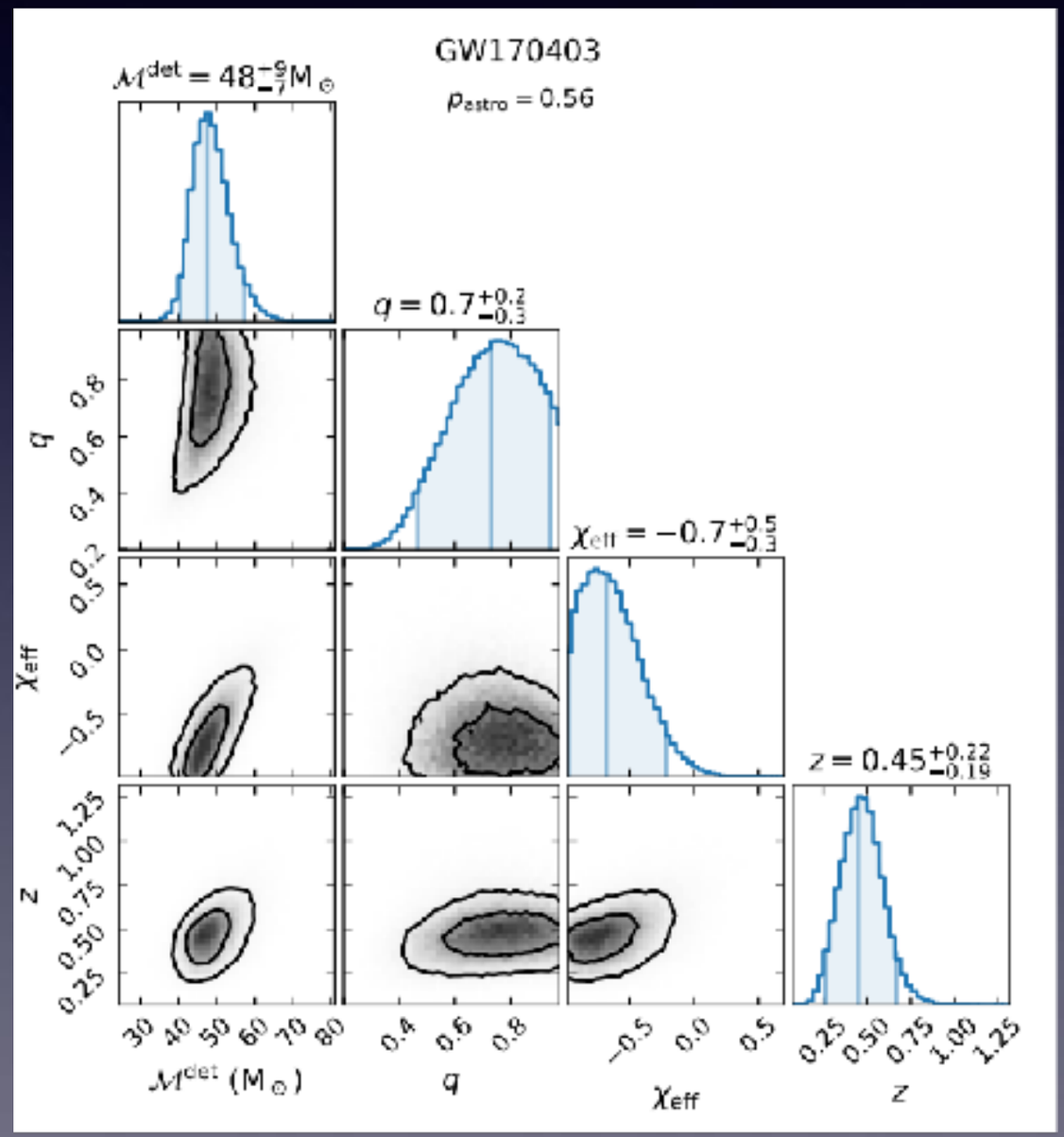
What does it tell us about astrophysics?

Quick recap on science questions concerning BBH

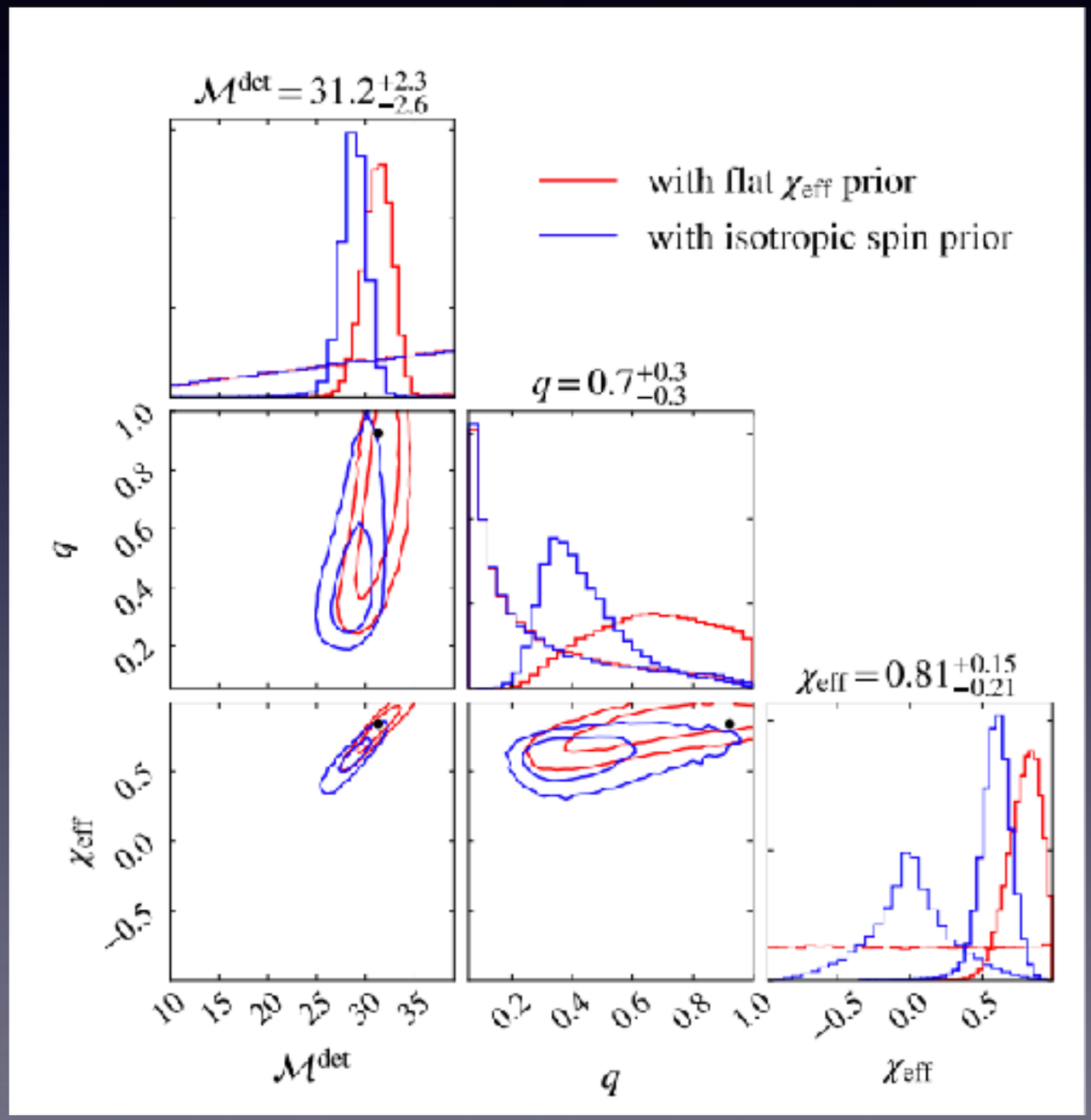
- How do binary black holes form?
  - Dynamic formation? [Key Predictions: Isotropic spin, few percent eccentric mergers]
    - Association with nuclear star clusters? [Prediction: located in the center of the host galaxy]
    - Binary Stellar Evolution [Key prediction: unisotropic spin distribution, possible prevalence of maximal spin]
- Upper mass gap?
  - (pulsational) pair instability supernovae ->, No black holes above ~45 solar mass.
- Lower mass gap?
  - Core collapse supernovae -> no black holes below ~5 solar mass?.
- Prompt coalescence? [Prediction: Should track star formation, star-forming hosts]
- Delay time distribution of 1/T? [Prediction: Rate should be smooth, some “Red and Dead” hosts]

Discoveries (#1) - Negative Spin?

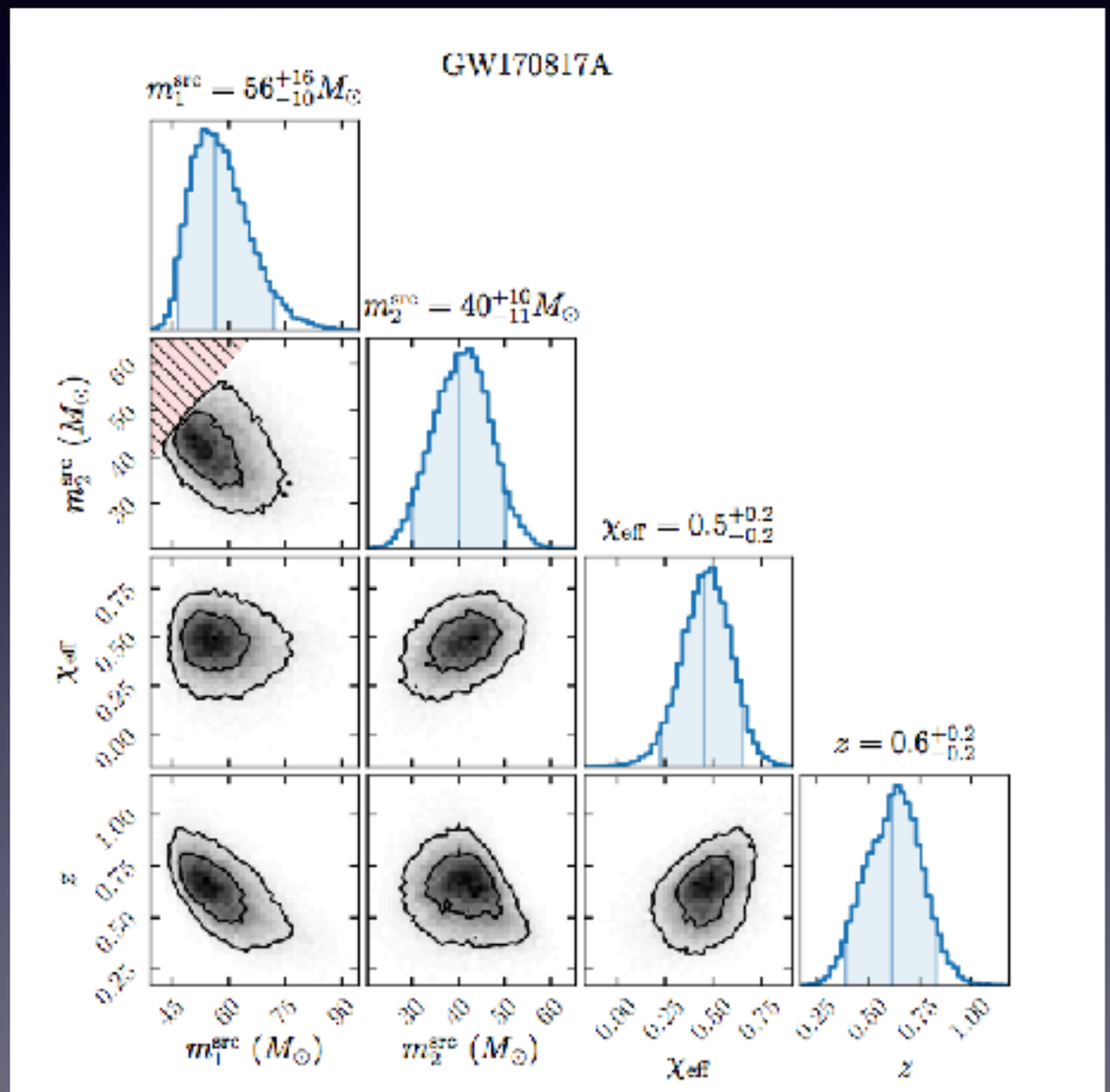




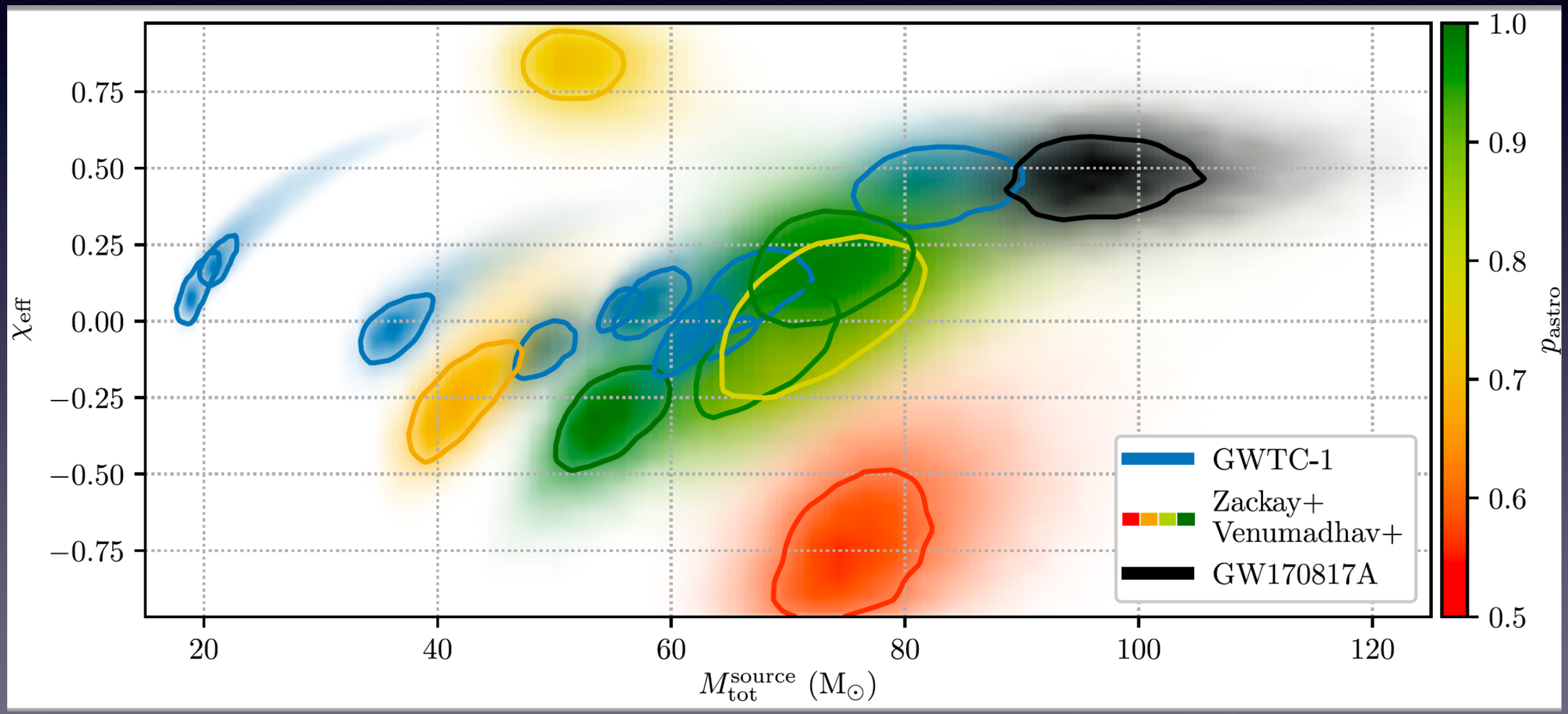
Discoveries #2 - High Spin



Discovery #3
High Mass



Discoveries #4
Correlation emerges?



Measured rate

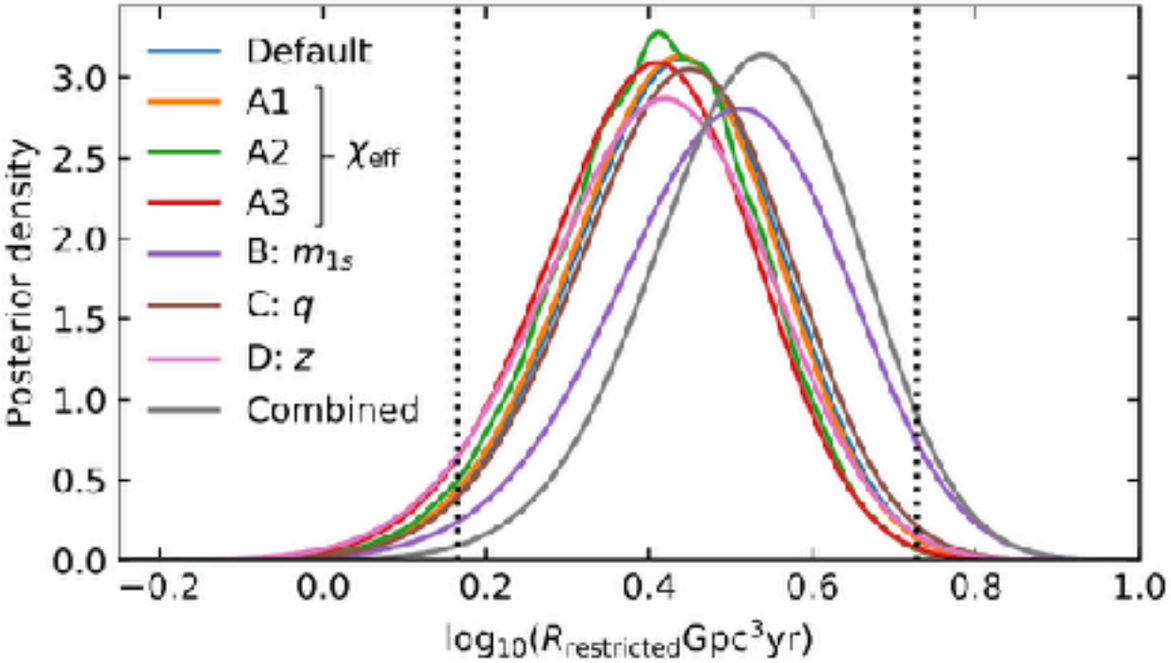
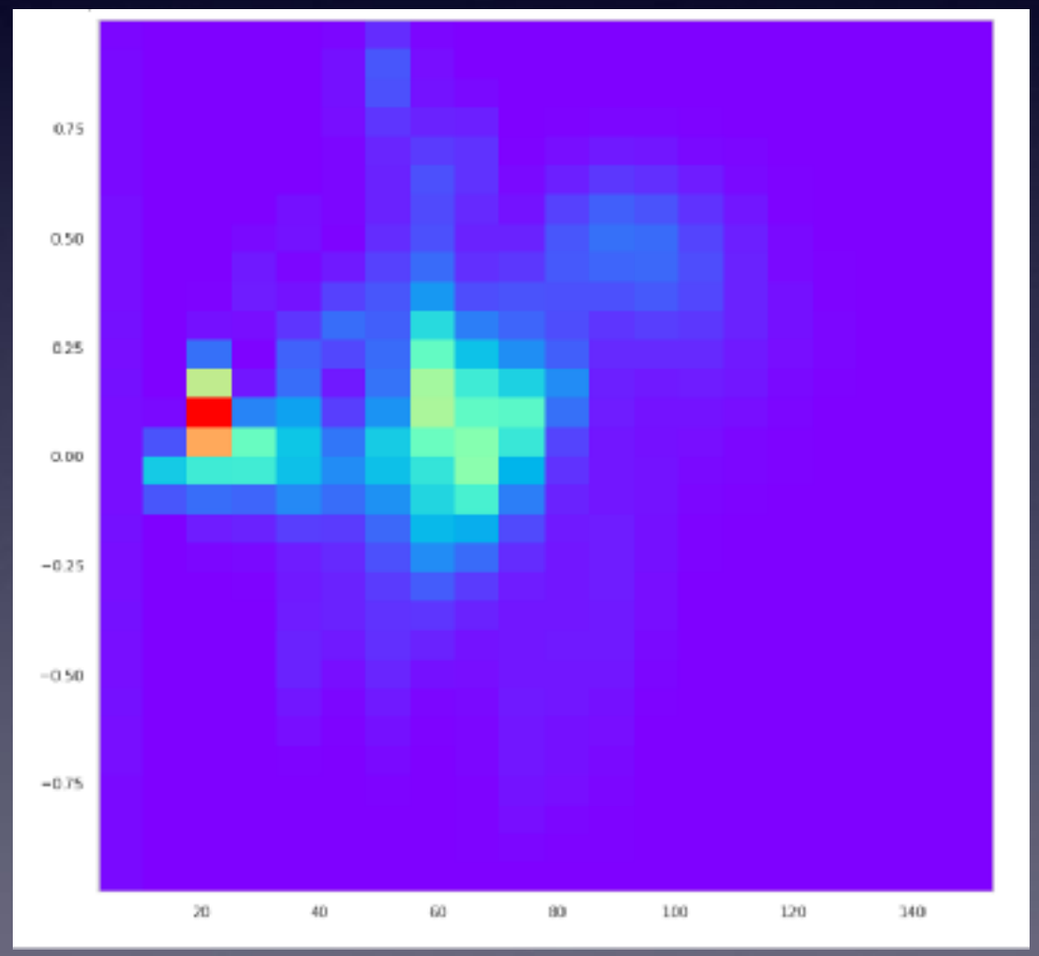


FIG. 9. Posterior for the comoving rate of mergers restricted to $20\,\mathrm{M}_\odot < m_{1s} < 30\,\mathrm{M}_\odot$, $q > 0.5$ and evaluated at redshift $z = 0.2$, under the various models we studied (§III A 1–III D, annotated with the variable each of them is intended to explore). We add a model that combines the maximum likelihood solutions of models A1, B and C. This quantity is better measured, and in better agreement across models, than the total local rate $\mathcal{R}$, cf. Figs. 2–8, Abbott *et al.* [3, fig. 12]. Dotted lines indicate the minimum over models of the 5% quantile, and the maximum of the 95% quantiles.

Charting new territory

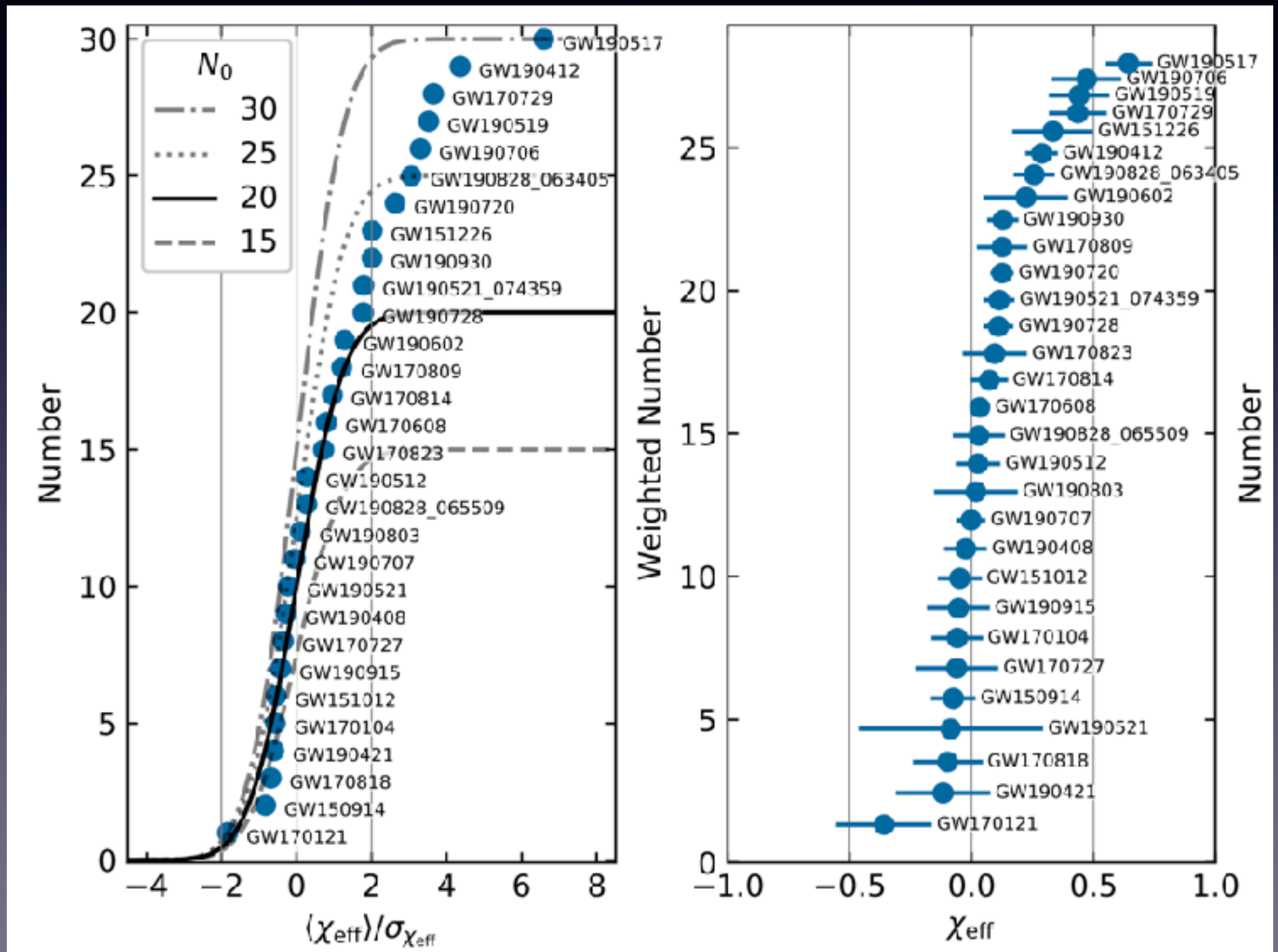
Number of detected BBH (KDE))
Uncorrected for selection effects

$$\chi_{\text{eff}} = \frac{M1\chi_{1,z} + M2\chi_{2,z}}{M1 + M2}$$

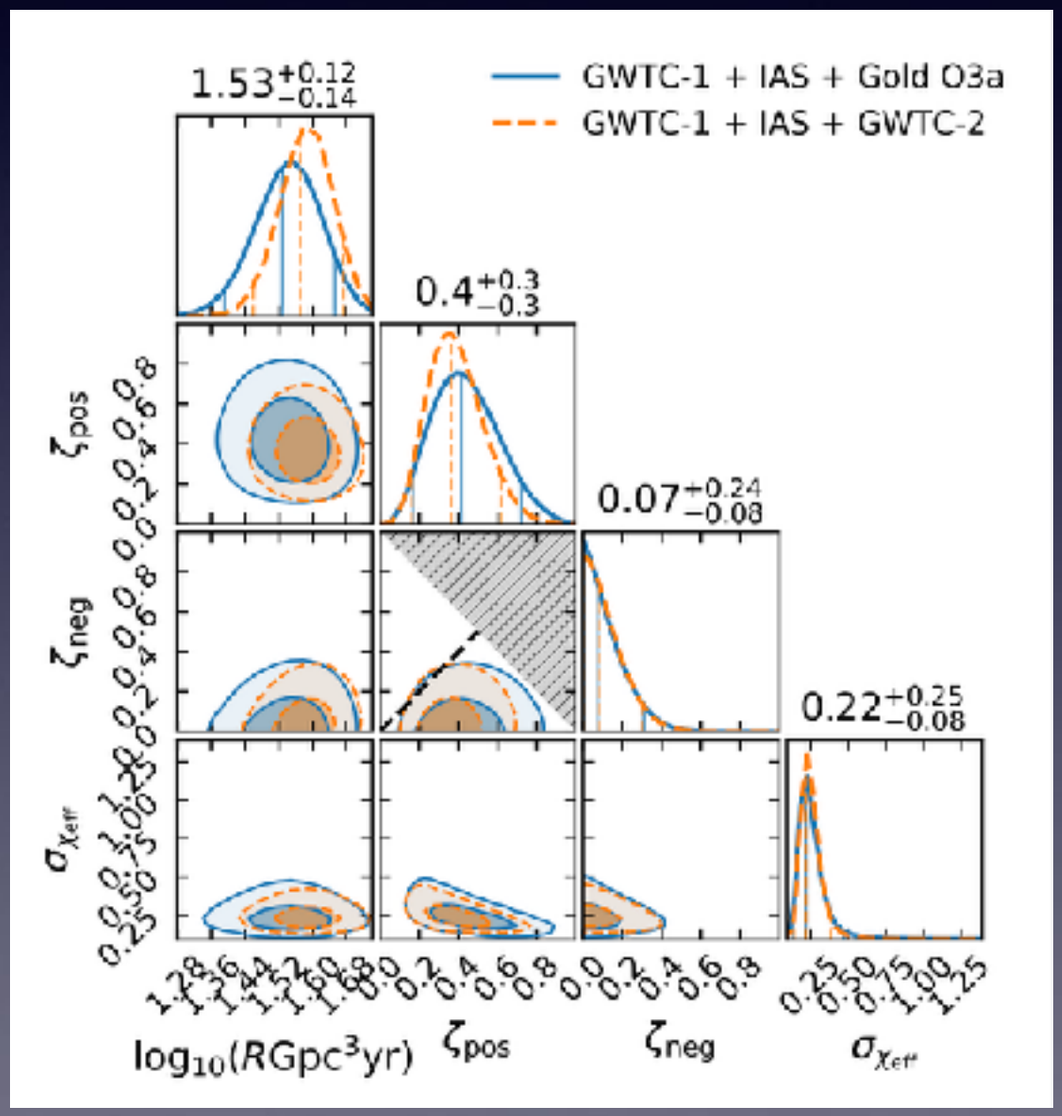


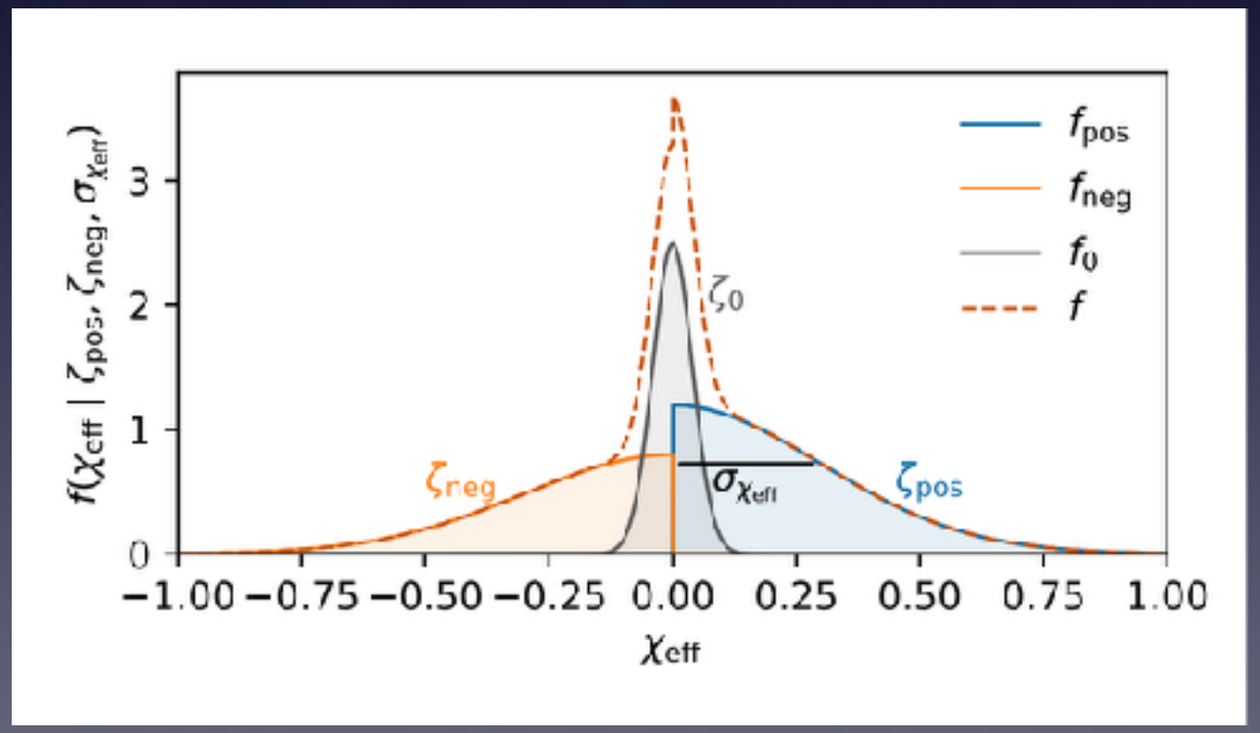
$M1 + M2$

Studying the symmetry of the effective spin



Studying the symmetry of the effective spin

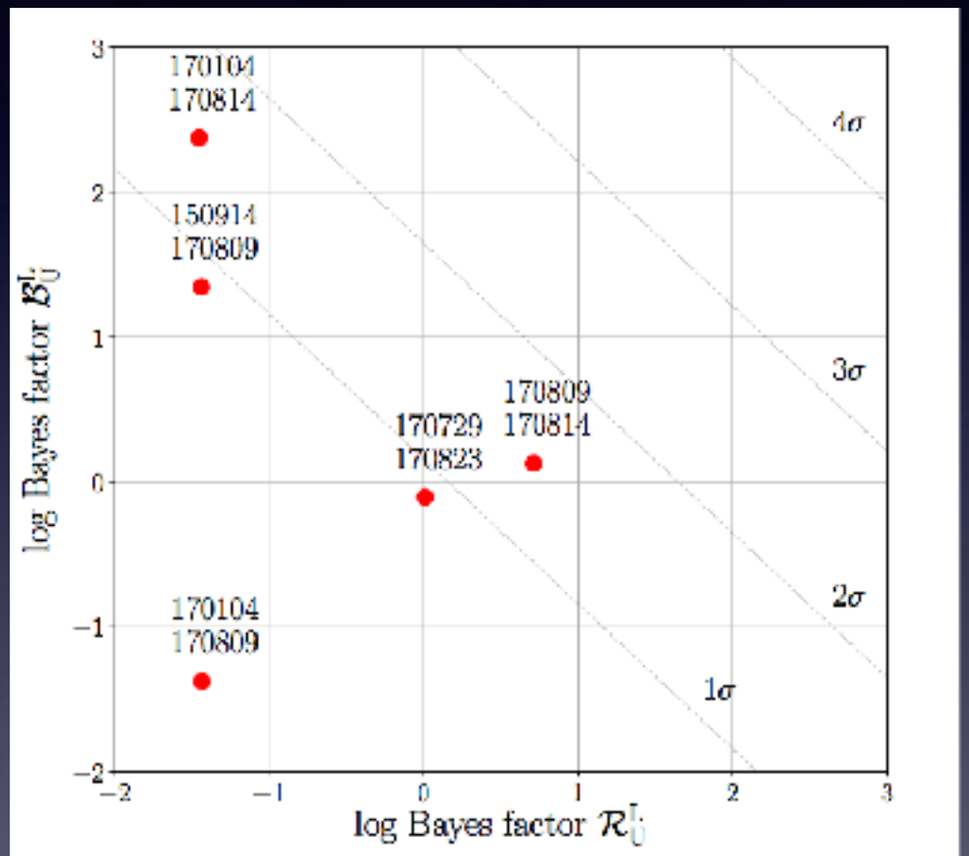




Vision for identifying the host galaxies of BBH?

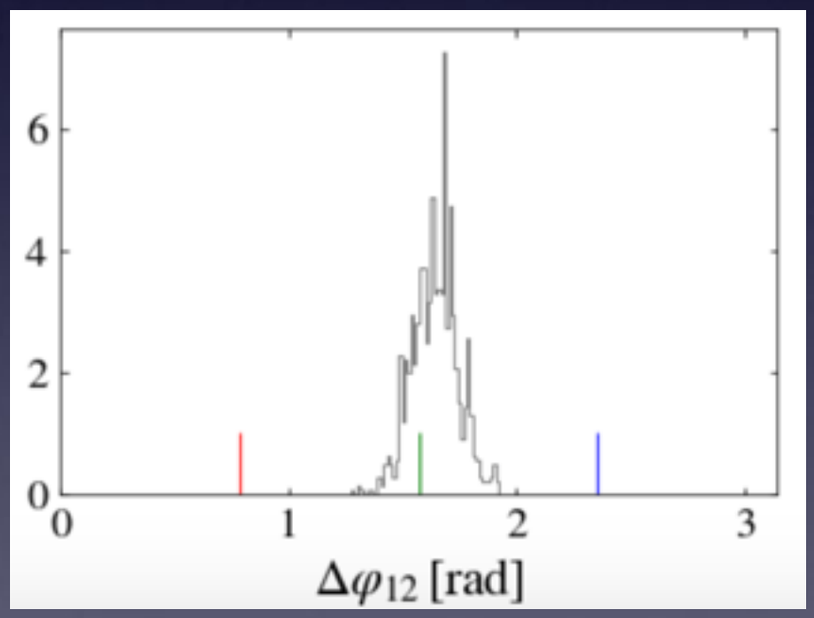
- General case: Not in advanced LIGO's band
  Yes, if we push for low frequencies (0.1-1 Hz)
- Yes if they are lensed. Measurable quantities:
  - Sky position error box - 10s of square degrees
  - Magnification ratio, Image types, Image time-delays.
  - A combination of magnification and redshift of host.
- => Only one lens with these parameters in the error-box
- Exact time delays allow identification of position within the host?
- Hubble constant measurements from these systems?

Curious candidate: GW170104, GW170814

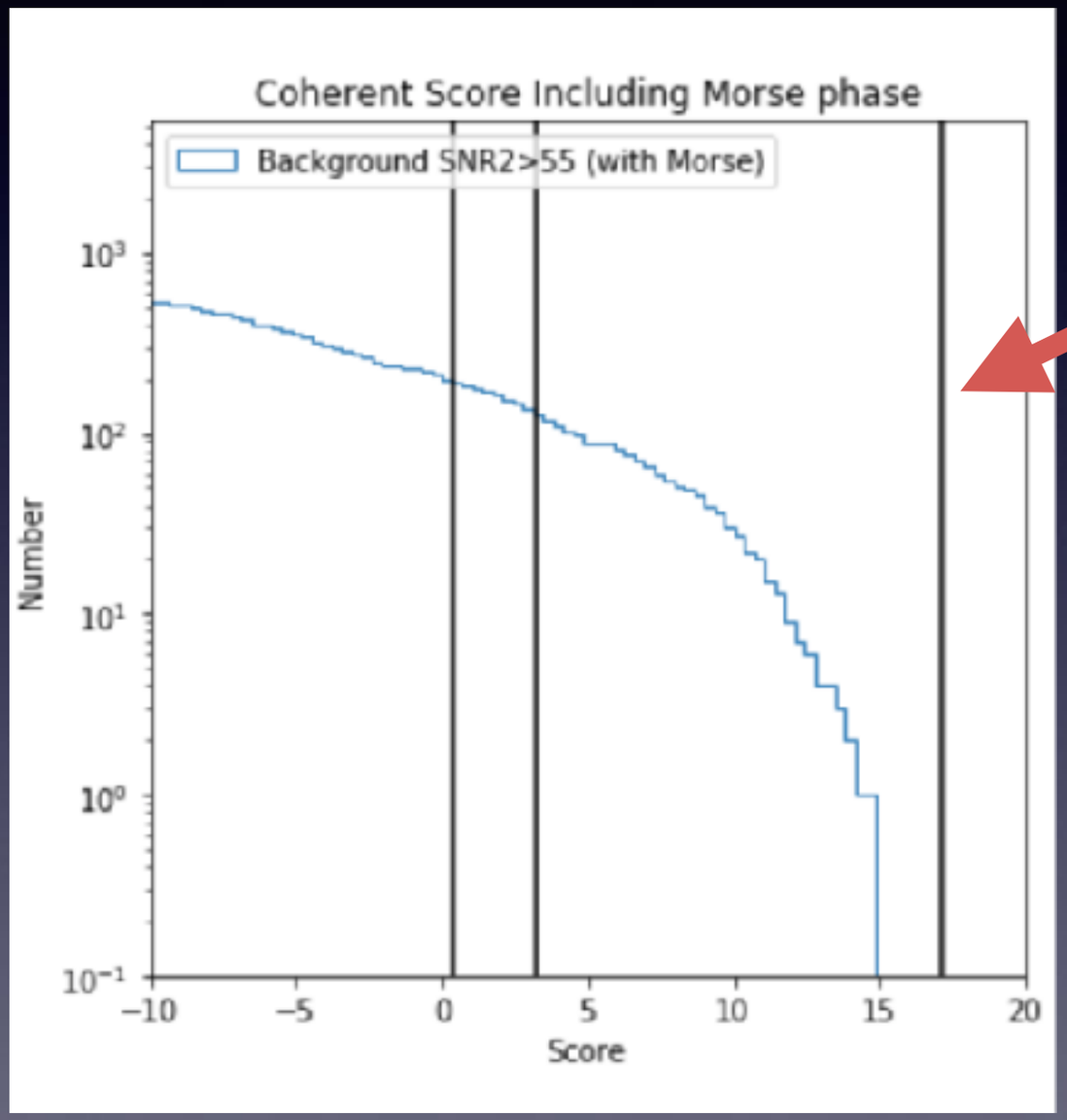


Hannuksela et al 2019

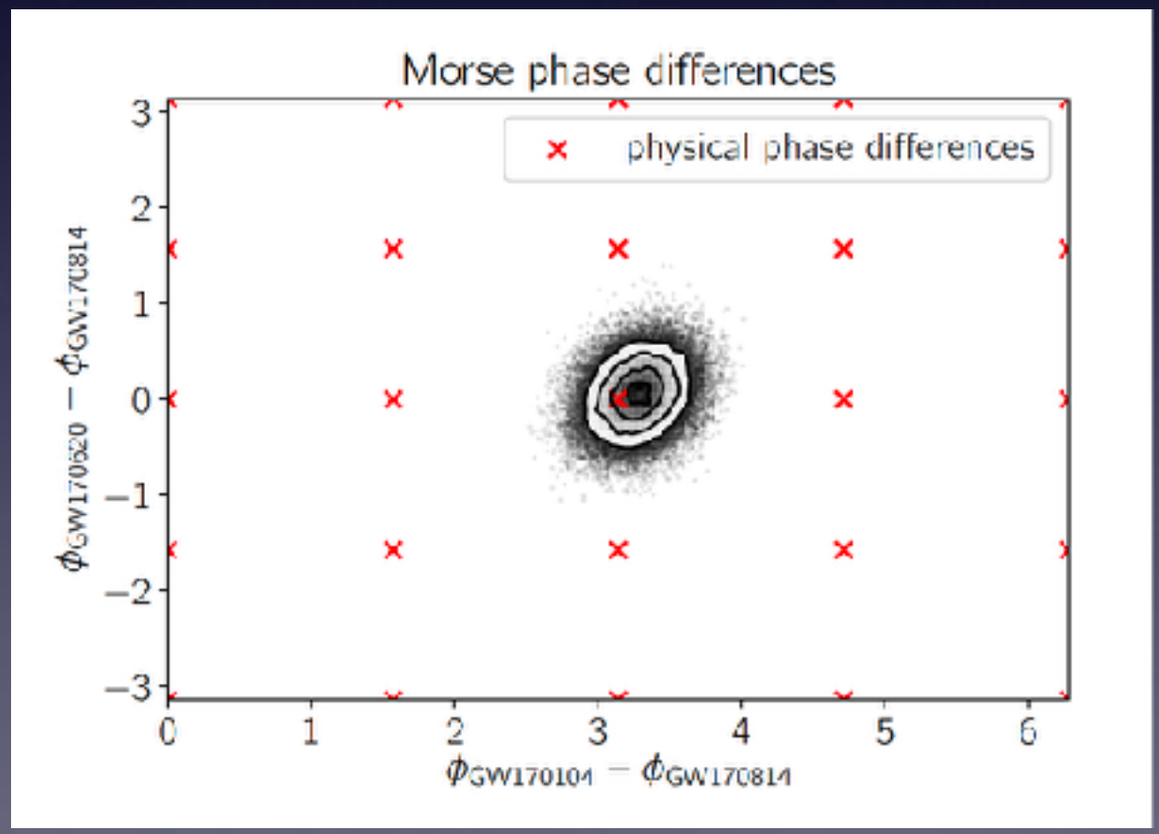
Morse phase neglected.
It actually works out!



Search for other images



GWC170620



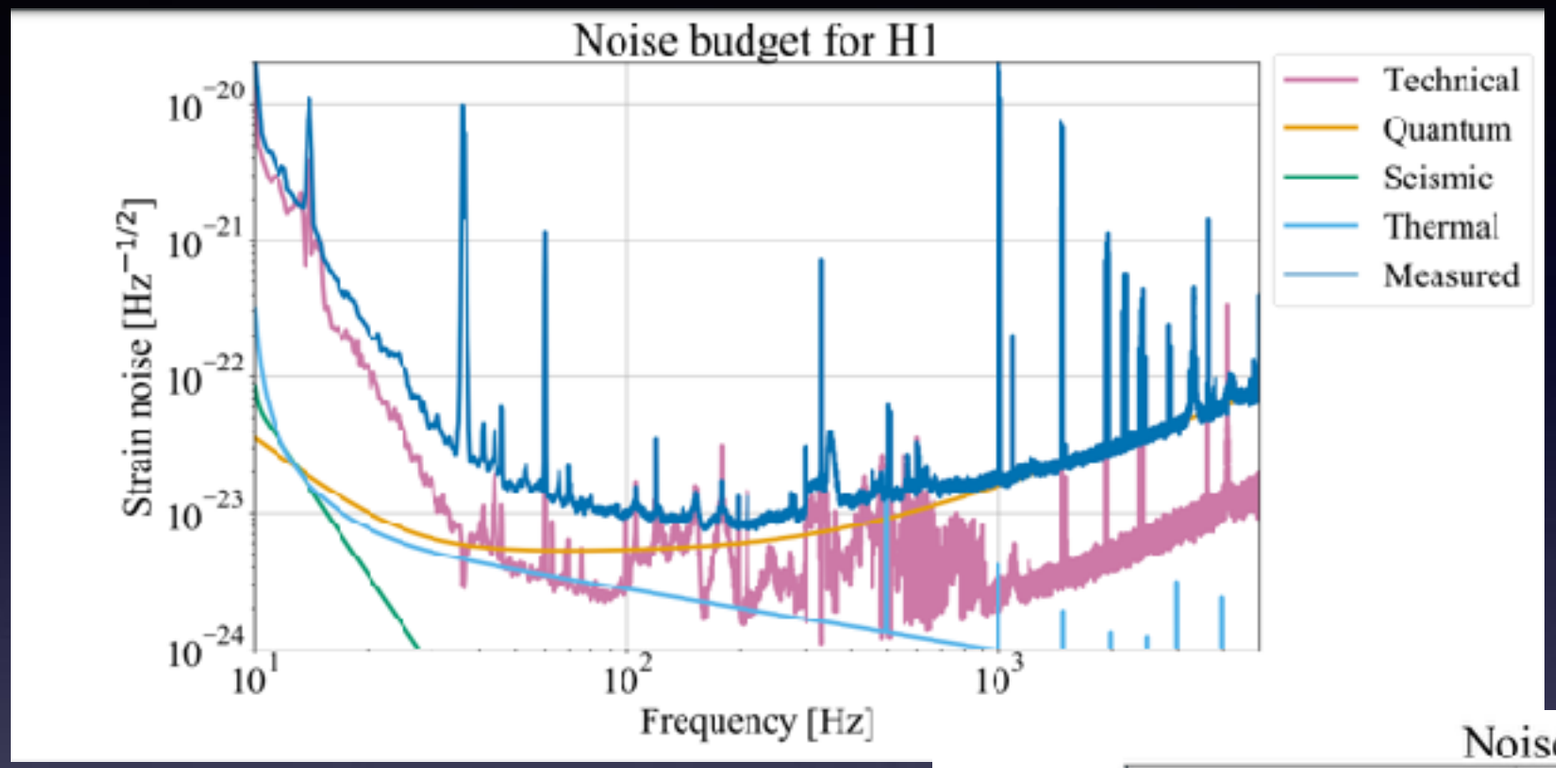
Evidence for and against lensing

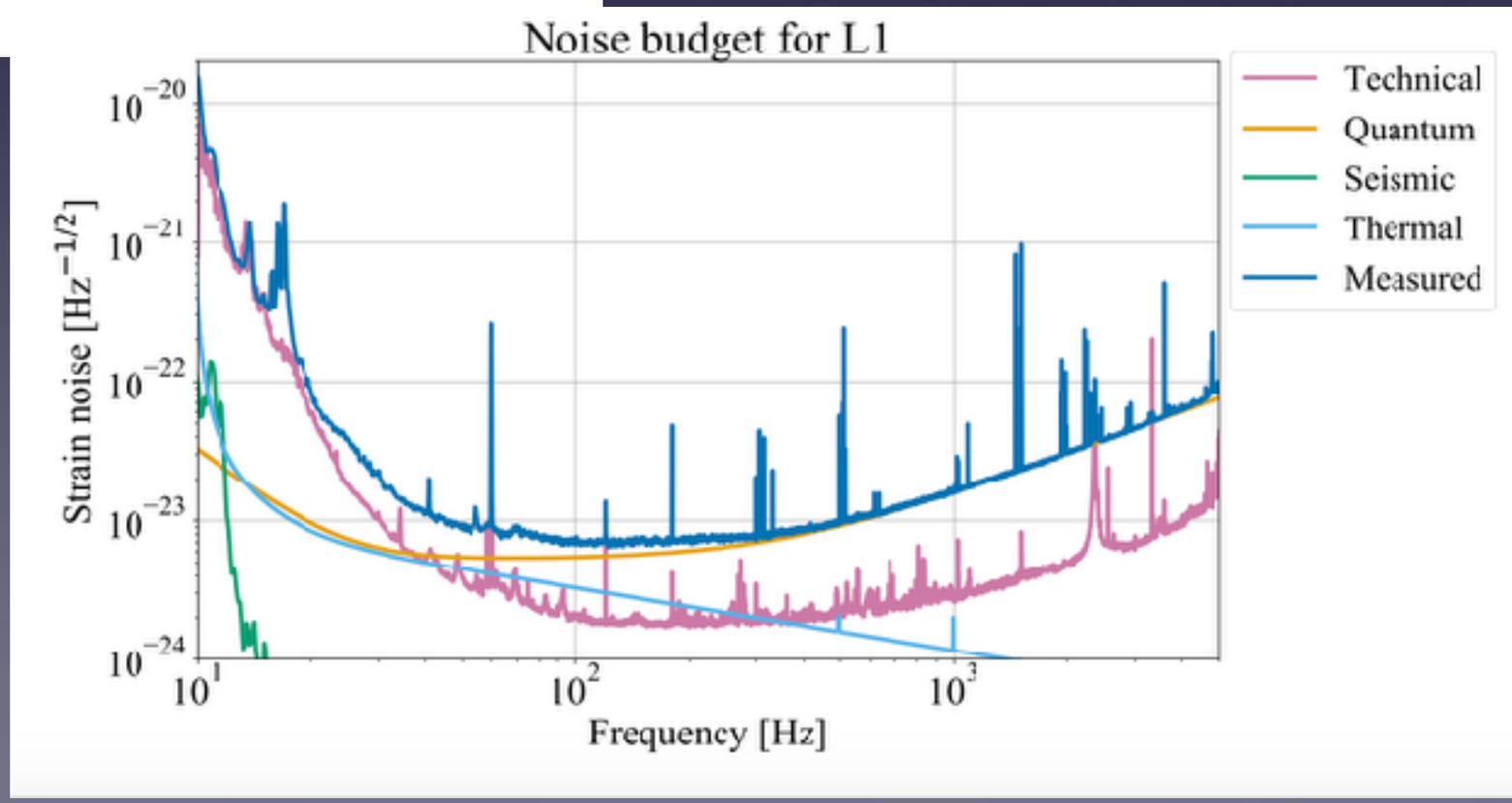
Expected Rate of Lensed events	$10^{-2} - 10^{-3}$/O2
Catalogue has pair GW170104 GW170814	$< 10^{-2}$
Existence of GWC170620	1.3×10^{-2}
Possible association of GW170202	0.02 (B~1)
Time delay not typical	Cluster or galaxy group?
Image types: H,L,L / L,H,H	H images are atypical
Magnification ratios 2.5:1:10	

Future projects in GW astronomy

- Using auxiliary channels to reduce low-frequency control noise.
- Veto (technique paper, led by Tejaswi)
- Search for heavier binary black holes and jet phenomena.
- Complete and report O3 (all sources) search (including “Fishing”).
- Fishing for short GRB counterparts.
- Construct a search that naturally includes precession and higher modes.
- What’s with GWC170402 ?

Control noise

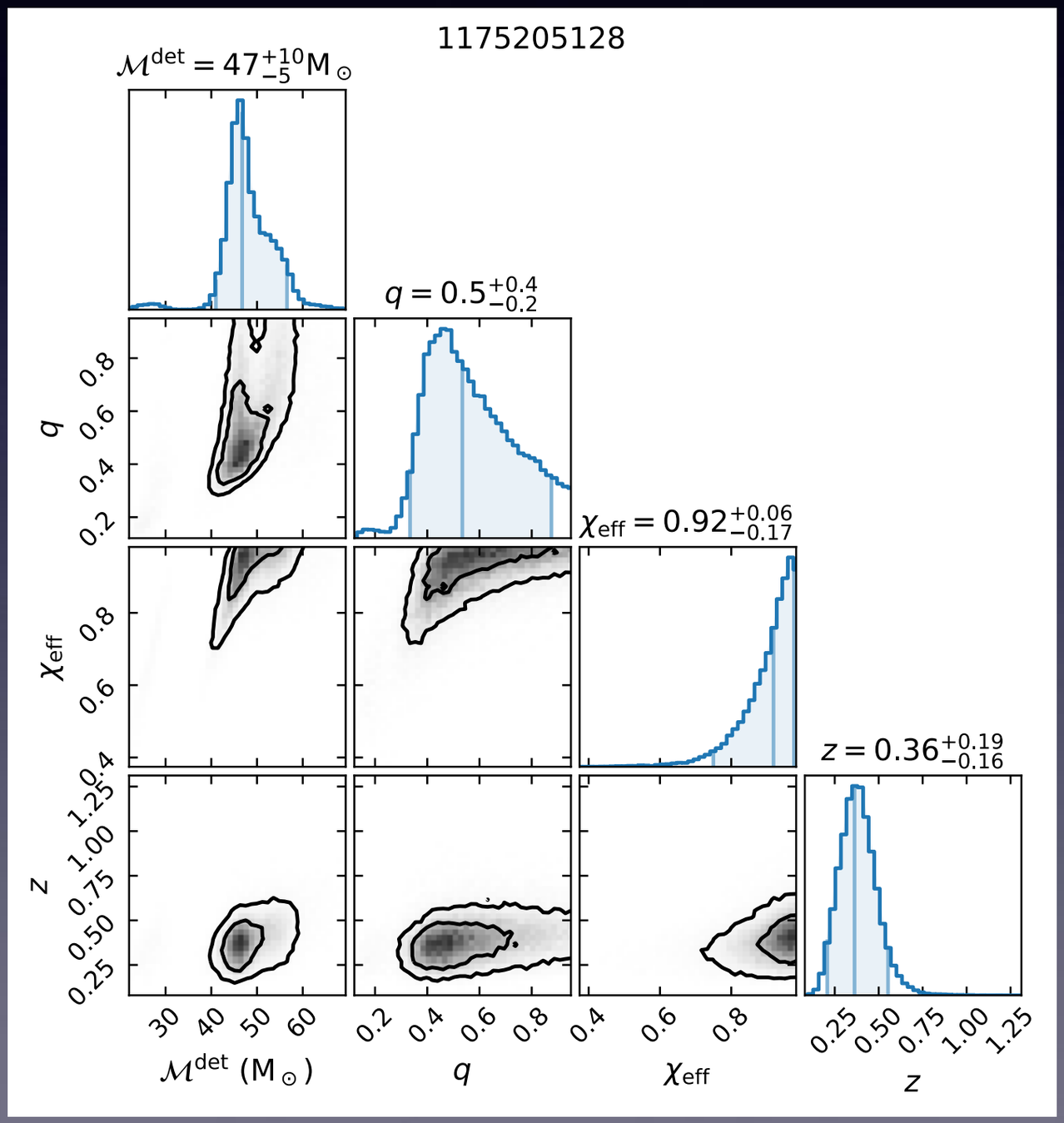


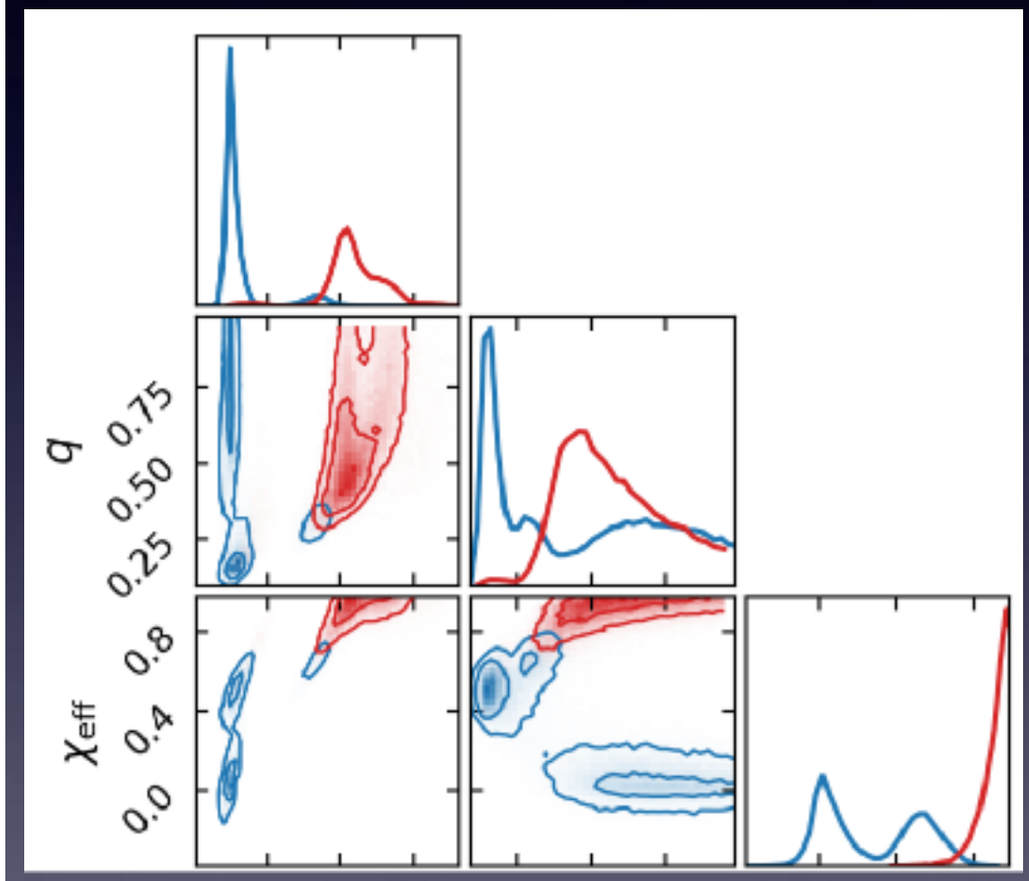


Questions?

I am hiring students and postdocs

What's with GWC170402 ?



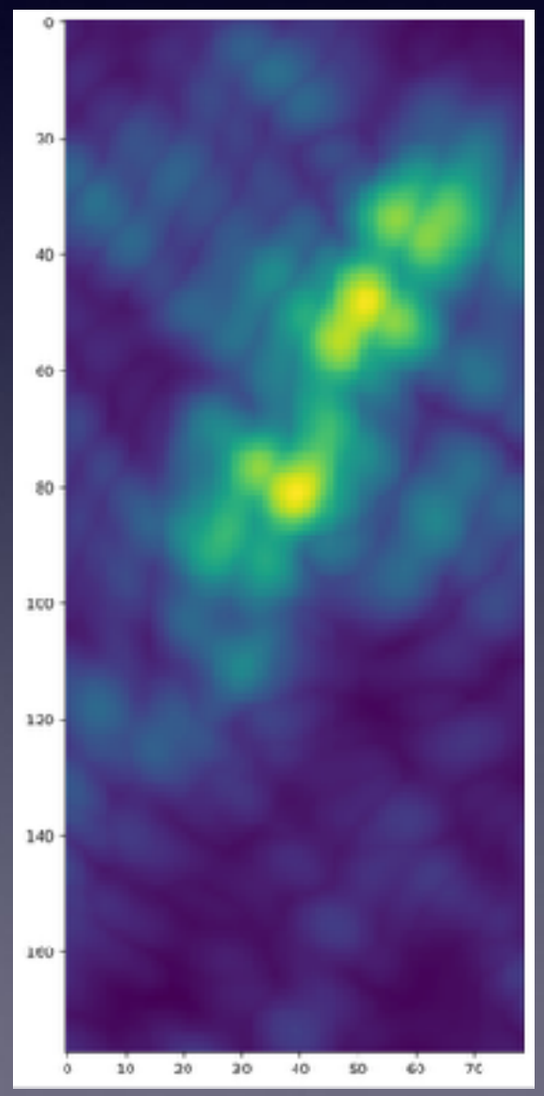


$\mathscr{M}_c$ q $\chi_{\rm eff}$

Are there really many solutions?

- YES
- Two separate solutions with equal response (chi squared of 71) in Livingston!
- Is that noise? NO (FAR ~ 1/10000)
- Is that physical?
  - Can reproduce a similar phenomenon using higher modes

Color - (S/N)^2



C2

C3

Including higher modes