Getting the most out of gravitational wave merger observations

Frans Pretorius

Princeton University

General Relativity – The Next Generation
Kyoto, February 23 2018

Outline I

- Motivation
  - understanding the dynamical, strong-field regime of gravity
  - binary black hole systems and the final state conjecture
- General relativity in the wake of LIGO's detections
  - first direct tests of this regime of GR
  - ability to constrain/rule-out alternatives limited by our lack of knowledge of gravity other than GR in this regime

Outline II

- Looking ahead
  - what kind of data can we expect in O3 and beyond, and how to maximize the science we can extract from this
  - going after faint signals common to a population of sources with coherent stacking
  - applications to quasi-normal ringdown, and the post-merger phase of binary neutron star mergers
- Conclusions

Strong Field Gravity

- This is the regime of general relativity (GR) where typical curvature scales are comparable to, or larger than other relevant scales in the problem
  - GR has no intrinsic length scale, so the scale where gravity becomes strong is always relative to some other physical scale in the problem
    - for compact objects (black holes and neutron stars) the radius of the object sets the scale
    - for the universe as a whole, the Hubble radius is the relevant scale

Strong Field Gravity

- The most extreme manifestation of strong field gravity is the presence of a *horizon*
  - general relativity then mandates than some form of singularity in the geometry is present somewhere in the spacetime
  - in a cosmological setting on scales of the Hubble radius there is not a horizon in the same sense as a black hole, nevertheless here the structure of spacetime is likewise markedly different from that of weak-field gravity (i.e. Minkowski spacetime)
- In dynamical situations the gravitational wave luminosity can approach a decent fraction of the Planck luminosity
  - the Planck luminosity $L_p=c^5/G$ does not dependent on h, but in some sense is a limiting luminosity even in classical GR

Why gather evidence for the GR description of strong-field gravity?

- GR itself has no intrinsic scale, and so one could argue the numerous existing confirmations of its weak-field properties should give confidence in all its predictions

- However, aside from basic scientific inquiry, there are reasons to be more cautious about blindly accepting GR's extreme gravity predictions

  – *the fundamental inconsistency with quantum mechanics*
    - ostensibly tensions should only manifest near the Planck scale, but some "firewall" proponents argue otherwise

  – *the existence of dark energy and dark matter*
    - the evidence for the latter does not rely on strong field gravity, but some have suggested the two phenomena are connected, e.g. Verlinde's emergent gravity proposal

Learning about gravity with binary black hole mergers

- *Binary black hole mergers* in general relativity are exquisite probes of dynamical, strong field gravity because the *Final State Conjecture* (Penrose) seems to be correct
  - *The generic, final state of all vacuum, 4D, asymptotically flat solutions of the Einstein field equations respecting cosmic censorship are a finite number of unbound black holes moving apart, together with gravitational waves streaming away to infinity*
  - *Each black hole asymptotes to a unique member of the 2-parameter (a, M) Kerr family of solutions*
- crucially, this is not "just" the no-hair theorem

No Hair Theorem

- *All* single, asymptotically flat, <u>*stationary*</u> black holes in 4D, vacuum GR (with no exterior naked singularities) are uniquely described by a member of the 2-parameter (a,M) Kerr family of solutions

- Taken by itself, this would suggest either

  (a) black hole solutions are sets of measure zero and not of astrophysical relevance at all

  (b) the Kerr family are "universal dynamical attractors" reached once gravitational collapse occurs

  - this option is essentially the FCS, and the important distinction compared to the no hair theorem alone is the FCS deals with the *dynamics* of BH spacetimes

The FSC and binary BH mergers

- Many profound consequences of the FSC; most relevant here are:
  - The *full structure* of spacetime exterior to the horizons of all *vacuum binary black hole spacetimes* allowed in GR, prepared in relative isolation sufficiently far to the past of coalescence, are essentially *uniquely* characterized by a *small, finite* set of numbers N
  - A merger waveform observed with large signal-to-noise ratio (SNR) will, from an information-theoretic perspective, require a correspondingly large set of numbers M to describe
  - For $M >> N$, can check for consistency with the FSC; an inconsistency indicates some assumption (pristine environment, cosmic censorship, GR, etc.) must be wrong

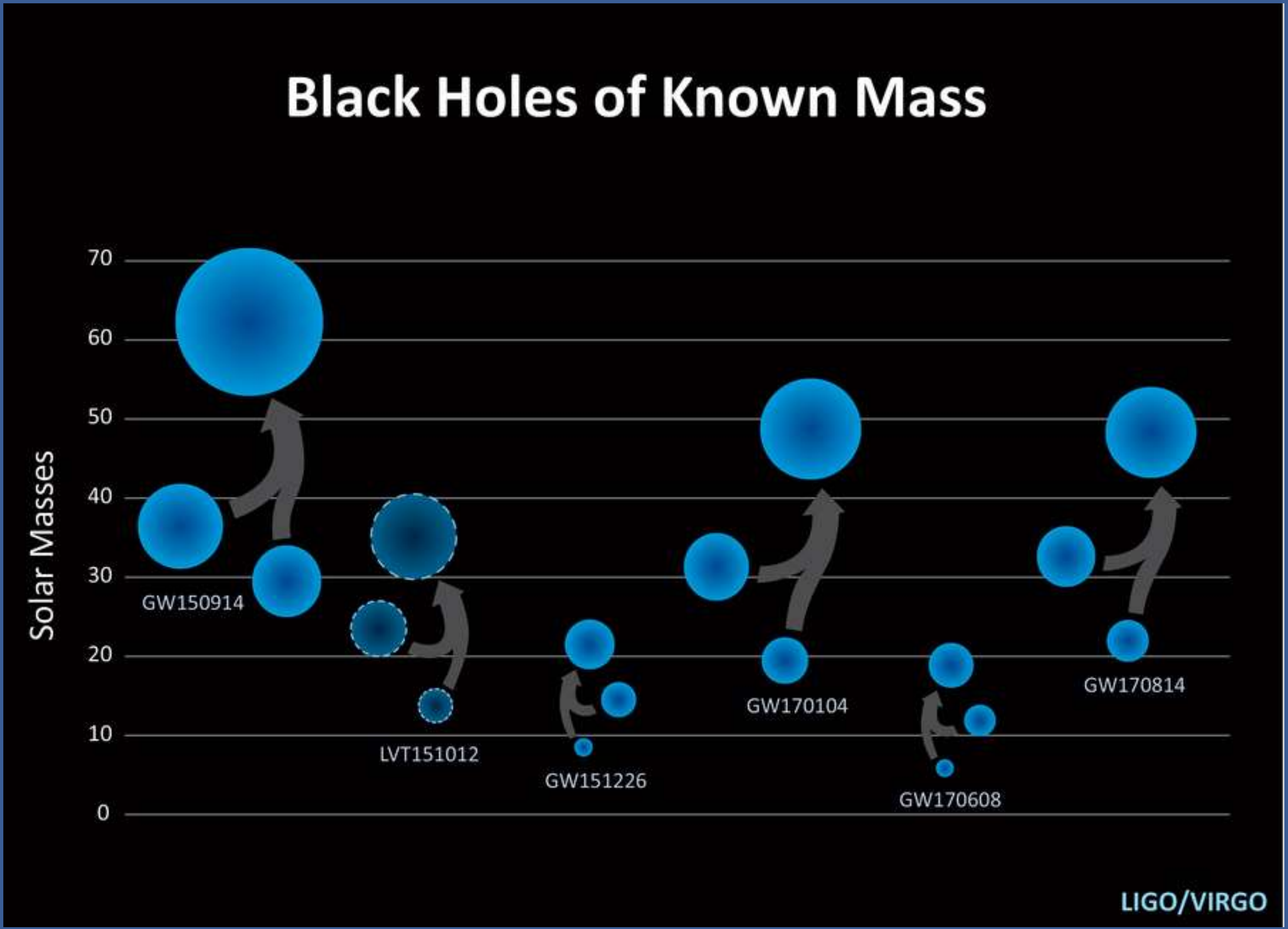


Image from LIGO website

LIGO/Virgo's set of GW events and the FSC

- All events so far consistent with GR, and are allowing us to begin making quantitative of the level of consistency
  - most "agnostic" test is the consistency of the residuals of the higher SNR events with noise
    - for GW150914, the data does not support more than a 4% modification from GR *[excluding classes of modification that would result in degeneracies with GR parameters, hence a larger inconsistency can get shuffled into a parameter estimation bias]*
    - this is implicitly a test of the FCS, as it limits the dimensionality of the template bank
  - other tests at present focus on the inspiral only portion, and consistency between parameters extracted from the inspiral vs ringdown portions of the waveform

Side comment : Beyond GR

- Constraining specific alternative theories (EDGB gravity, Chern-Simons gravity, …), or "exotic" compact object alternatives (gravastars, traversable wormholes, firewalls, etc.) is hamstrung at present by the following, or worse situation:

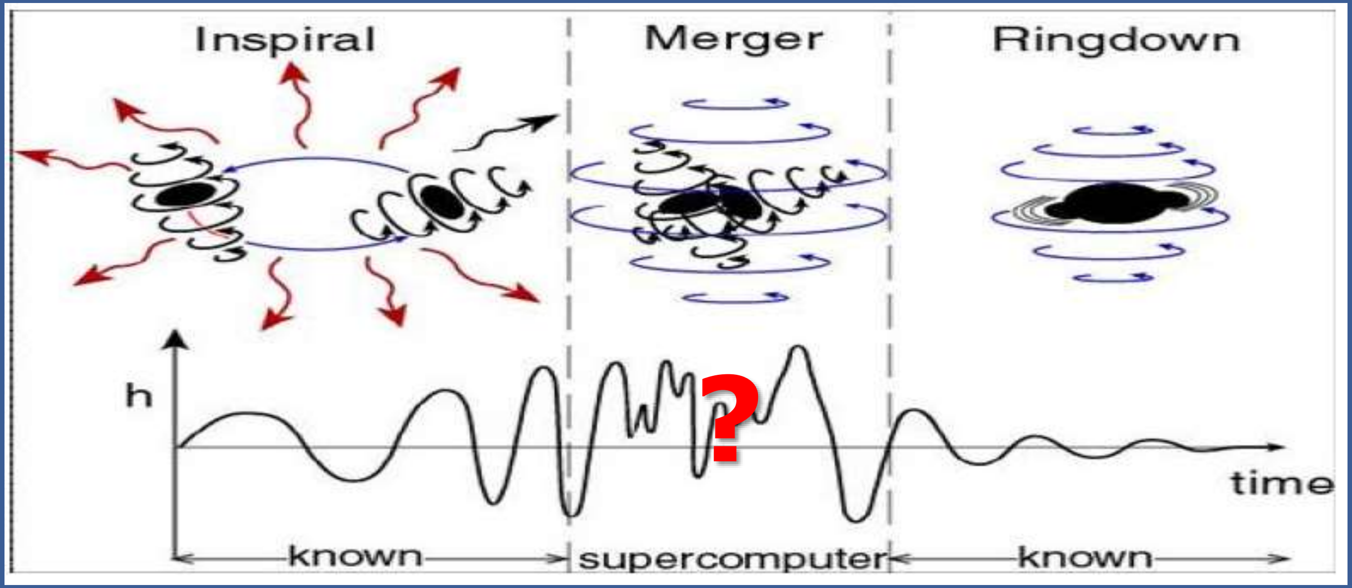


Illustration by Kip Thorne

- Most of the SNR in the best event to date, GW150914, is precisely in the regime where we do not understand beyond-GR physics; have to "nibble at the edges" of the data at present, and the constraints are unsurprisingly a lot weaker

Investigating the FSC in the inspiral within the parameterized post-Einsteinian (ppE) framework

- Detecting the unknown or unexpected, especially with analysis methods that rely on templates, is a nebulous problem
- The idea behind ppE *(Yunes and FP, 2009)* is more modest : take a class of event – binary compact object inspiral here – where there is good evidence GR is at least providing the correct leading order description and then *deform* the GR inspiral templates in a well-motivated manner to capture deviations from the GR baseline. “Well motivated” could include
  - consistent will all existing tests, yet can produce observable deviations in the dynamical, strong field regime
  - predicted by a specific alternative theory
  - characterizes a *plausible* strong-field correction, e.g. more rapid late time inspiral due to excitation of a new degree of freedom (scalar waves, different polarizations, etc)
  - that something like this can practically be applied to BBH mergers is exactly because of the FSC : if didn’t hold, measurement of a ppE deformation from a GR template would not allow one to distinguish from unmodelled “new” BH solutions vs. beyond GR physics (or an anomalous environment)

The minimal ppE inspiral template

$$\tilde{h}(f) = \tilde{h}_I^{GR}(f) \cdot \left(1 + \alpha u^a\right) e^{i\beta u^b}$$

- $h_I^{GR}(f)$ is *some* model of the GR inspiral component, e.g. to leading order

$$\tilde{h}_I^{GR}(f) \propto f^{-7/6} e^{i 2\pi f t_0}$$

  - $u = \pi M f$, with M the chirp mass
  - a, b, α, β are ppE parameters
    - GW observations are most sensitive to the phase parameters *(b, β)*
  - Note : the GR baseline does not need to be the templates used for detection

GR: $\alpha=0$, $\beta=0$

Brans-Dicke: $\alpha=0$, $b=-7/3$

Massive graviton: $\alpha=0$, $b=-1$

Chern-Simons like parity-violation: $a=1$, $\beta=0$

Dynamical Chern-Simons gravity: $a=3$, $b=4/3$

varying G: $a=-8/3$, $b=-13/3$

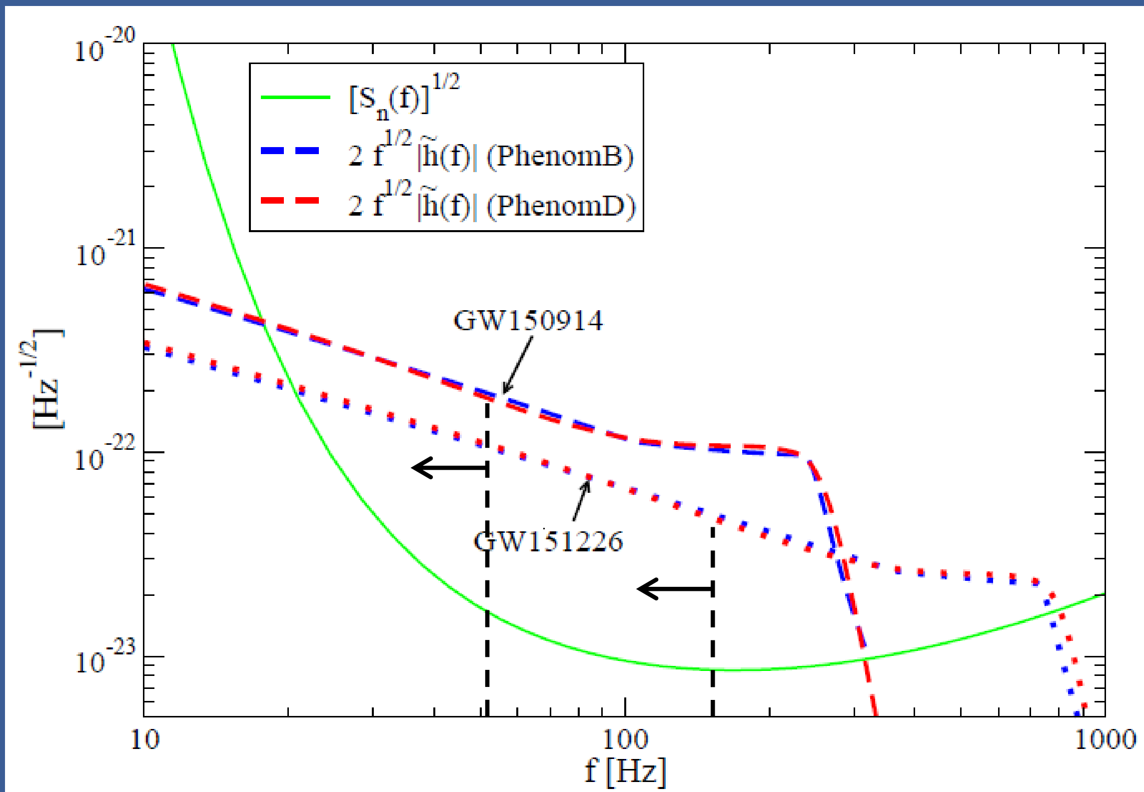
certain extra dimensions: $\alpha=0$, $b=-13/3$

quadratic curvature: $\alpha=0$, $b=-1/3$

modified PN: $a=0$, $b\neq0$, $b=(k-5)/3$, $k \in I$

Inspiral constraints from GW150914/GW151226

Work with Nico Yunes and Kent Yagi, PRD 94 (2016)

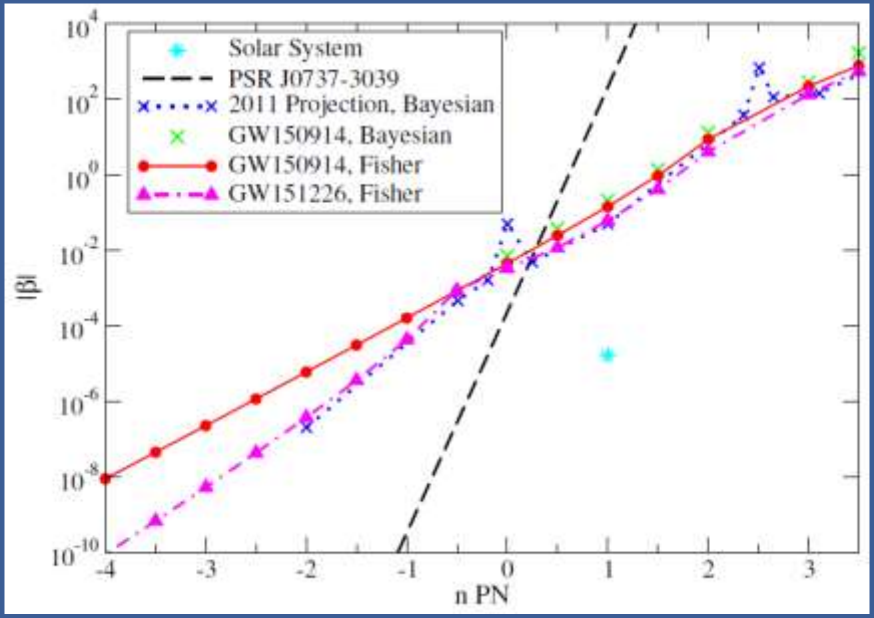


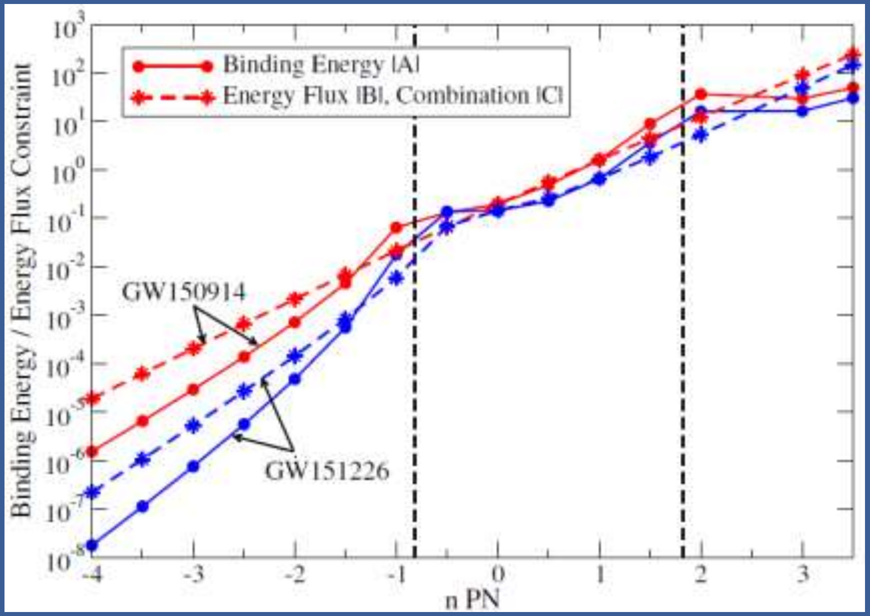
Event	GW150914	GW151226
Signal-to-noise ratio ρ	23.7	13.0
False alarm rate FAR/yr^{-1}	$< 6.0 \times 10^{-7}$	$< 6.0 \times 10^{-7}$
Primary mass $m_1^{\mathrm{source}}/\mathrm{M}_\odot$	$36.2^{+5.2}_{-3.8}$	$14.2^{+8.3}_{-3.7}$
Secondary mass $m_2^{\mathrm{source}}/\mathrm{M}_\odot$	$29.1^{+3.7}_{-4.4}$	$7.5^{+2.3}_{-2.3}$
Chirp mass $\mathscr{M}^{\mathrm{source}}/\mathrm{M}_\odot$	$28.1^{+1.8}_{-1.5}$	$8.9^{+0.3}_{-0.3}$
Total mass $M^{\mathrm{source}}/\mathrm{M}_\odot$	$65.3^{+4.1}_{-3.4}$	$21.8^{+5.9}_{-1.7}$
Effective inspiral spin χ_{eff}	$-0.06^{+0.14}_{-0.14}$	$0.21^{+0.20}_{-0.10}$
Final mass $M_{\mathrm{f}}^{\mathrm{source}}/\mathrm{M}_\odot$	$62.3^{+3.7}_{-3.1}$	$20.8^{+6.1}_{-1.7}$
Final spin a_{f}	$0.68^{+0.05}_{-0.06}$	$0.74^{+0.06}_{-0.06}$
Luminosity distance $D_{\mathrm{L}}/\mathrm{Mpc}$	420^{+150}_{-180}	440^{+180}_{-190}

Using the "IMRPhenom" model of LIGO (P. Ajith et al) excluding spin for the ppE baseline, truncated above 154 hz (52hz) for GW150914 (GW151226), and an analytic approximation to the aLIGO noise curve

LIGO/Virgo, arXiv:1606.04856

Inspiral constraints from GW150914/GW151226





- Upper bound on β vs. PN order *n (n=b+5)*

  Note: Solar system, binary pulsar, and BBH GW tests should really ***NOT*** be displayed together on this kind of plot : apples vs. oranges comparison, constraining different "sectors", and only within GR can they be mapped onto the same *(β,n)* plane. View this as the relative strength of GW vs. Binary Pulsar vs. solar system constraints in their respective "sectors"

- Sample of mapping of constraints on β to physical properties of the binary, here constraints to relative deviations in the binding energy and GW flux to those of the GR inspiral model, defined via

$$E_b = E_{b,\mathrm{GR}}\left(1 + A v^{2p}\right)$$

$$\dot{E} = \dot{E}_{\mathrm{GR}}\left(1 + B v^{2q}\right)$$

where the velocity $v=(m\pi f)^{1/3}$, and $p=p(n)$, $q=q(n)$

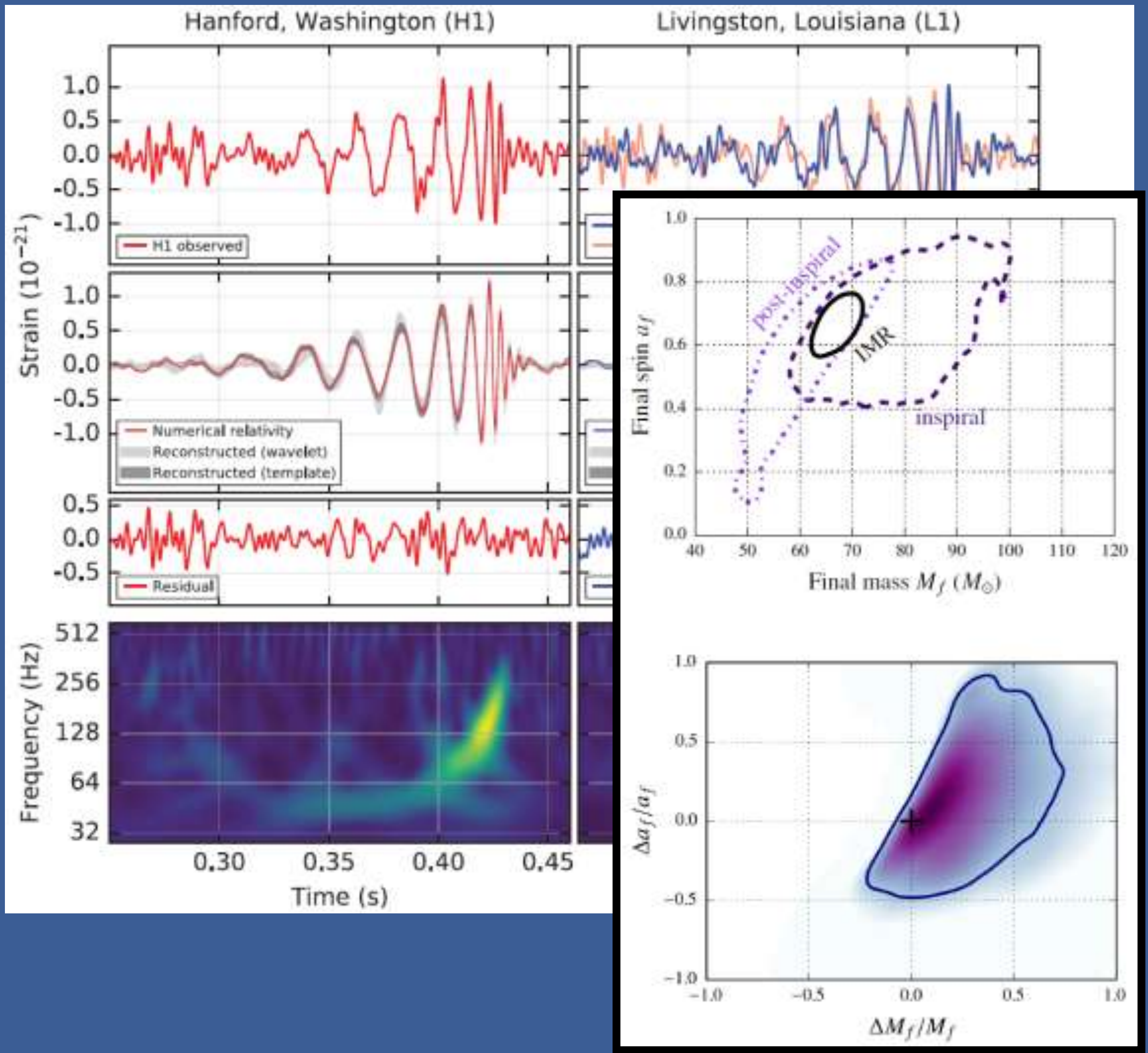
GW150914: Testing the FSC via independent estimates of the properties of the remnant Kerr black hole

- **The two-body inspiral :**

given the parameters of the initial binary, GR uniquely predicts the mass *(M)* and spin *(a)* of the remnant

- **The ringdown of the remnant to Kerr :**

the properties of the quasi-normal ringdown modes again uniquely identify the remnant black hole



arxiv:1602.03841, *LIGO & Virgo Collaboration*

Adding dimensions to the *(Δa, Δm)* space

- Can further over-constrain the mass and spin of the remnant by going after higher order quasi-normal modes (QNM) in the ring-down phase
  - every spheroidal harmonic (l,m) and overtone (n) has a different characteristic frequency/decay constant, but are uniquely determined by (a,m) of the remnant
  - moreover, due to the FSC, the initial amplitude and phase of each mode excited in a merger is uniquely determined by the parameters of the progenitor binary
    - "Initial" is arbitrary and more an artifact of trying to simplify the analysis by only using knowledge of the linear perturbation spectrum of Kerr
    - the FSC does not care about linearity, and in fact for comparable mass mergers the non-linear nature of the initial "perturbation" of the remnant will need to be taken into account

Subleading Quasi-normal Modes

- The promise of higher-order QNMs is with larger (l,m) smaller spatio-temporal scales about the horizon can be probed

- The problem with these modes is that they are excited with much lower amplitude than the $(2,2)$ mode in comparable mass mergers, and they decay more rapidly
  - expect an SNR ~200 event will be needed to detect one or more of the higher order QNM modes from a single event

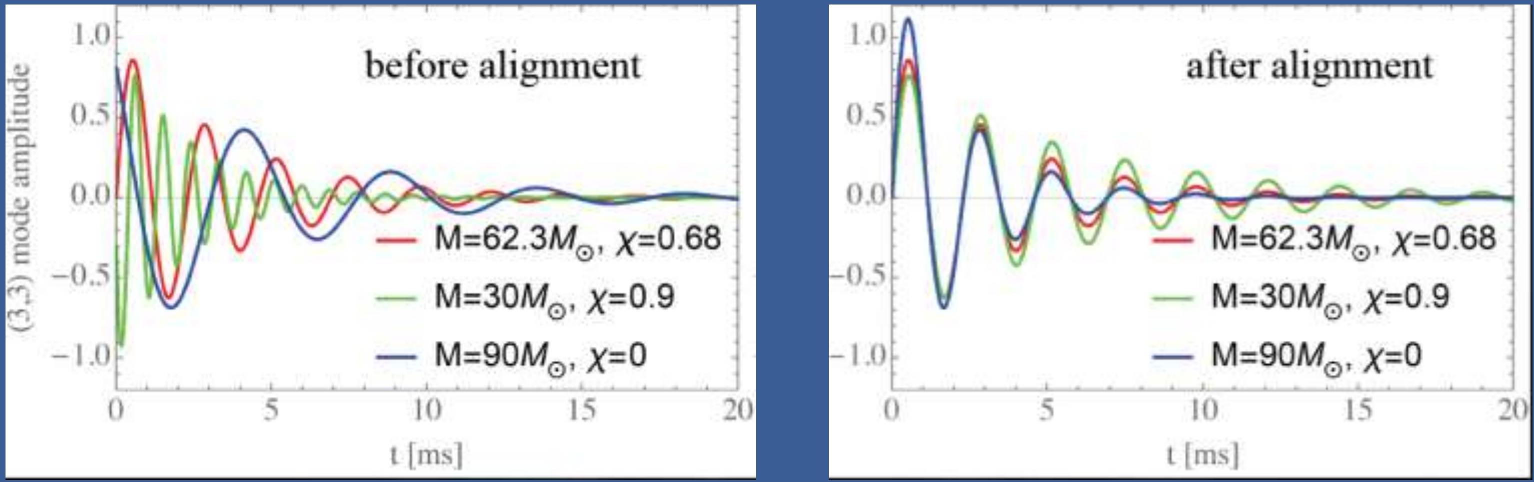
Stacking Data from Multiple Events

Work with H. Yang, K. Yagi, L. Lehner, V. Paschalidis, N. Yunes and J. Blackman, PRL 118 (2017)

- Enhance the effective sensitivity of gravitational wave data analysis to features common to a population of events
  - expect to have O(10's-1000's) of binary black hole (BBH) and binary neutron star (BNS) events by the end of advanced LIGO's operation
- Two approaches suggested to do this
  - *power stacking* : add excess power in select time/frequency bins; or similarly multiply Bayes factors of some common parameter post-detection
  - *coherent stacking* : directly add detector signals, appropriately scaling/aligning them so that the desired feature adds coherently before analysis, and assuming detector noise does not
  - if phase information is available, generically expect coherent to outperform power stacking, in particular for measuring a faint signal component not detectable in any individually event

Stacking to find Subleading QNMs

- Because each event with have a different spectrum of QNMs, cannot simply "add" all the signals
- Instead, target a *single* mode within each event : we can then scale/shift each signal by appropriate constants to phase and frequency align the target mode amongst all events

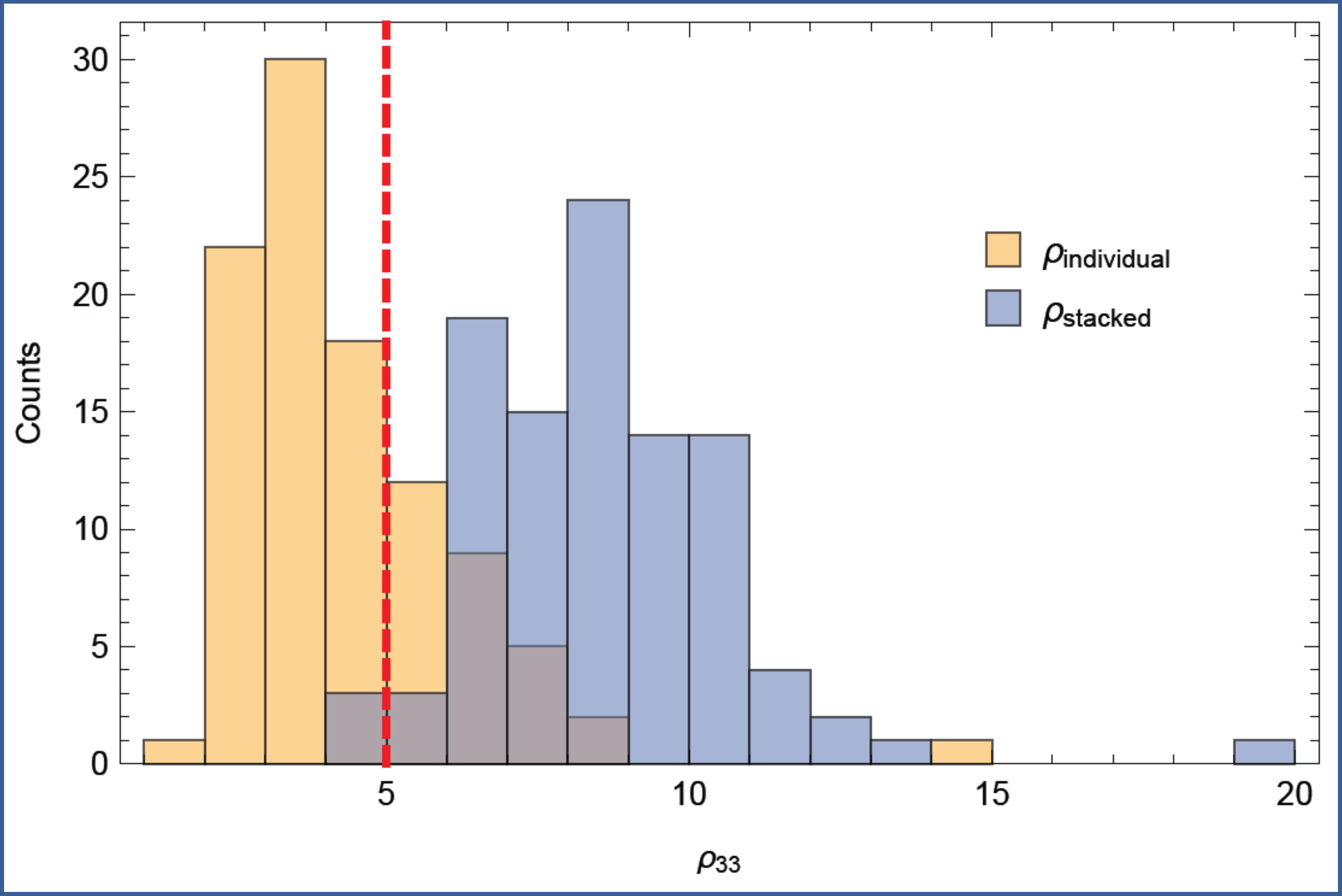


(3,3) mode in equal mass mergers; Image credit K. Yagi

Coherent mode stacking

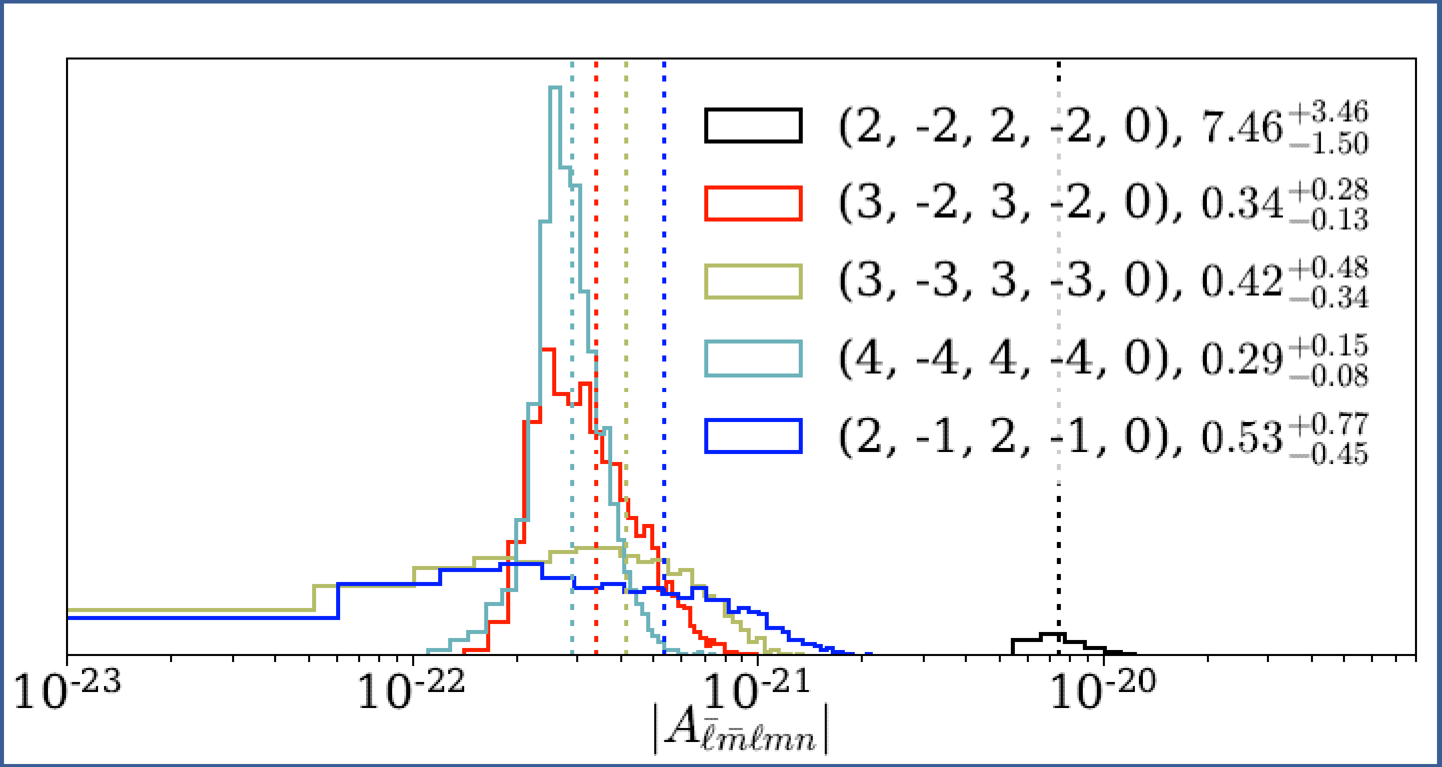
- This introduces a few additional complications, most notably
  - The amplitude/phase of each mode is calculated from measured properties of the inspiral; this introduces an additional parameter estimation uncertainty "noise"
  - We are adding *scaled* detector noise in the stacking
  - How to properly weight the different events in the sum as the population will not be homogeneous, in particular in SNR
- For this first "proof of principle" result for aLIGO, we do the following
  - Restrict to initially non-spinning black holes
  - Assume a uniform distribution of black hole masses from 10-50 $M_\odot$, and the optimistic end of the merger rate of 40/Gpc^3/yr
  - Only select events where the (2,2) mode by itself is detectable with SNR > 8 (in our 100 Monte Carlo runs there were 40-65 such events per year); and for now only stacking the 15 loudest
  - Assume a parameter estimation noise that scales like 1/SNR, calibrated (for all) by that of GW150914
  - Use the "downhill simplex optimization" method to choose stacking weights to maximize the SNR

Result : "Proof of principle" Targeting the (3,3) mode



- Counts from 100 Monte Carlo simulations of 1 year of detections at AdLIGO design sensitivity : 30% chance for detection of *(3,3)* mode from single loudest event, 97% chance from stacked signals

Can repeat the analysis for any desired target mode



- Image taken from *L. London, arxiv 1801.08208 (2018),* illustrating the SNR for the dominant and 4-subleading modes from a GW150914 like event

Stacking Binary Neutron Star Events

Work with H. Yang, K. Yagi, L. Lehner,
V. Paschalidis and N. Yunes , PRD 024049 (2018)

- NSs do not share the uniqueness properties of BHs, and consequently BNS merger events are not ideal candidates for stacking

- However, the post-merger signal is not easily within reach of aLIGO, yet a tremendous amount could be learned by observing this part of the event in GWs

  – prompt vs delayed collapse to a black hole, or even a stable remnant

  – if a long lived remnant, matter dynamics will produce GWs that encode information about the structure of the NS, and the equation of state of hot nuclear matter

    - expect aLIGO to be able to measure the post-merger signal in this case for events within ~10 Mpc; if GW170817 is indicative, this will only happen around one/decade

- Thus, should at least *try* to go after some common signal

Stacking Binary Neutron Star Events

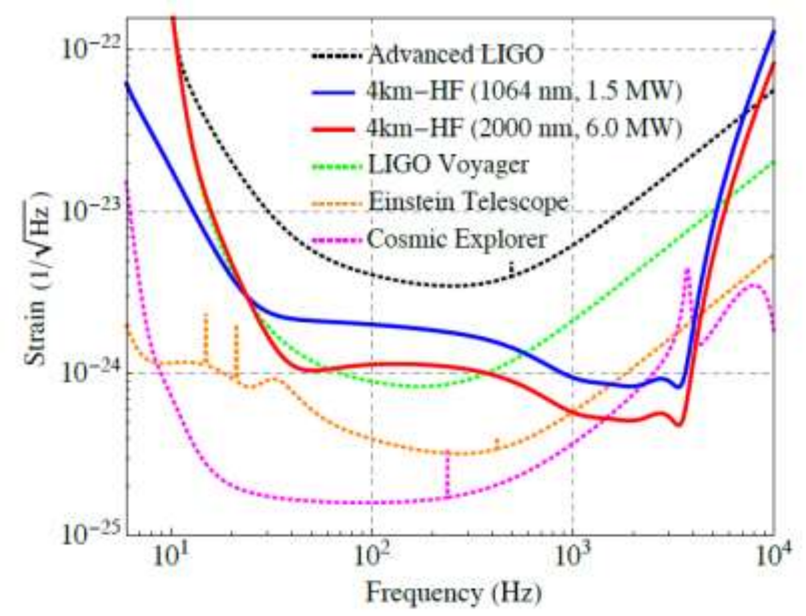
- The f-modes of perturbed NS's are natural targets here
  - estimate the mass and spin of the remnant from the inspiral chirp
  - choose an EOS; based on this can cut events expected to promptly collapse to BHs, for the rest, estimate the dominant GW emitting modes of the remnant
    - typically the (2,2) f-mode; could be a (2,1) mode if the remnant exhibits the "one-arm" instability
  - If simulations are sufficiently advanced by the time we have enough events to stack, can estimate the phases of the modes and coherently stack; otherwise power stack

“Proof of principle” study targeting the remnant (2,2) f-mode in BNS mergers

- Using several model EOS, a BNS merger rate of 1.54 Mpc^{-3} Myr^{-1} (LIGO)
- Prospects for detection are not good with aLIGO; focusing instead on the planned next generation detectors : Cosmic Explorer (CE) and Einstein Telescope (ET)

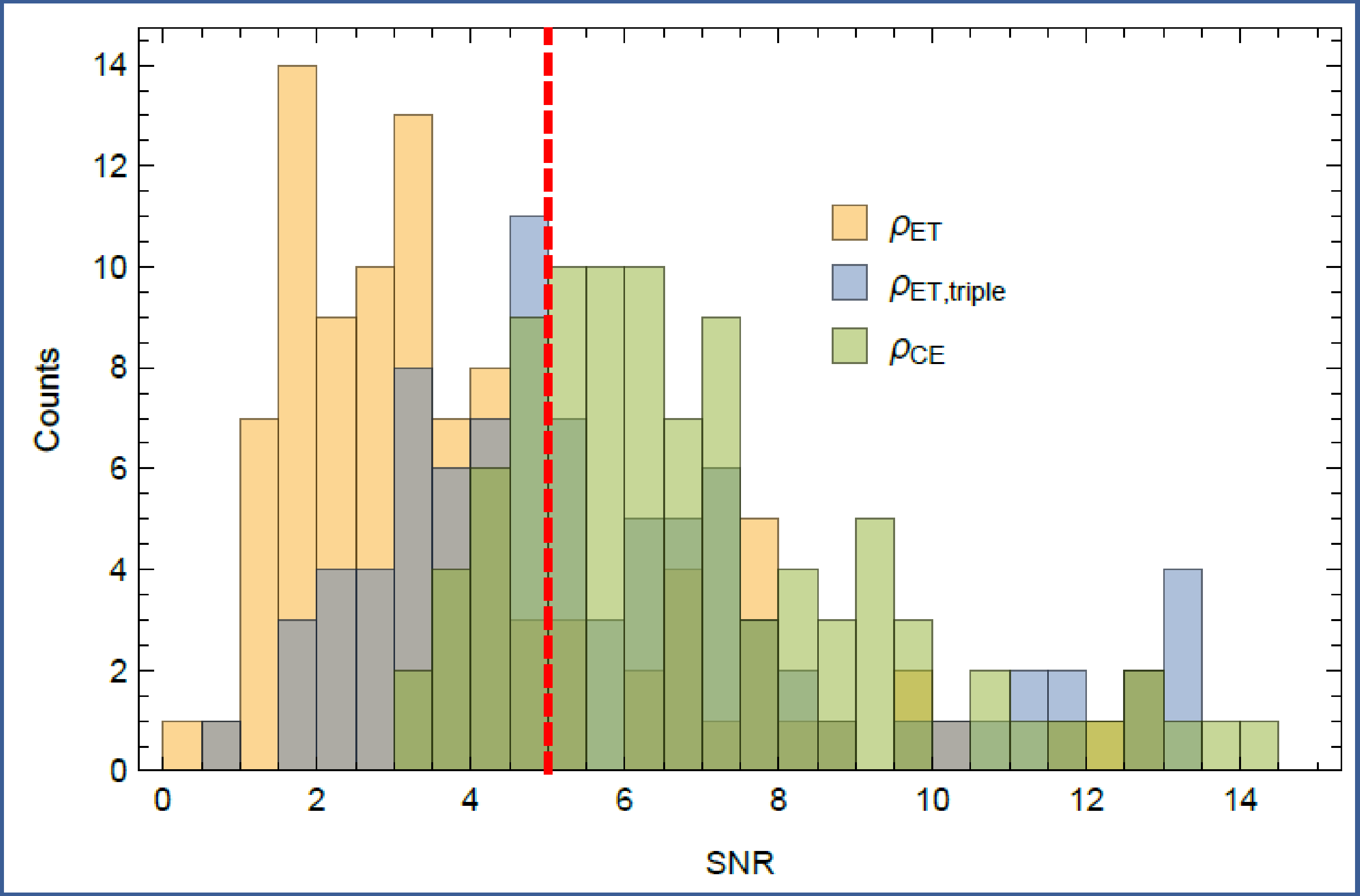
TABLE I. Parameters for different EOS

EOS	$R_{1.6M_\odot}$	$f_{\rm peak}(\rm kHz)\frac{M_\odot}{m_1+m_2}$	$\frac{A'(50\rm Mpc)}{10^{-22}}$	Q	$\frac{M_{\rm thres}}{M_\odot}$
SFHo	11.77	1.21	2.7	25.7	2.95
LS220	12.5	1.09	4.3	25.7	3.05
DD2	13.26	0.98	2.8	12.7	3.35
Shen	14.42	0.84	5.0	23.3	3.45
TM1	14.36	0.85	2.5	34.2	3.1



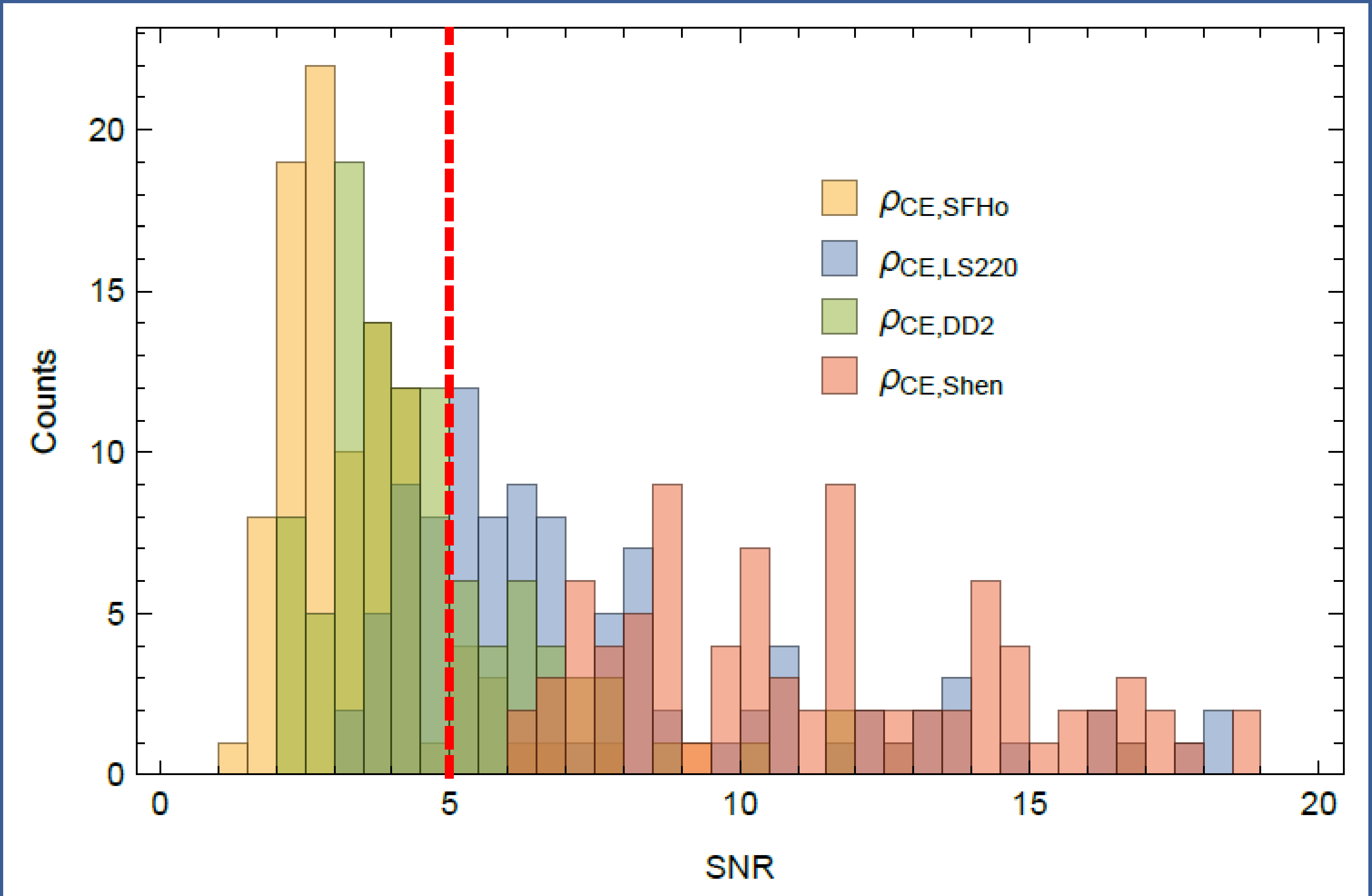
From Miao et al, arxiv:1712.07345

TM1 EOS, different next-generation detector designs



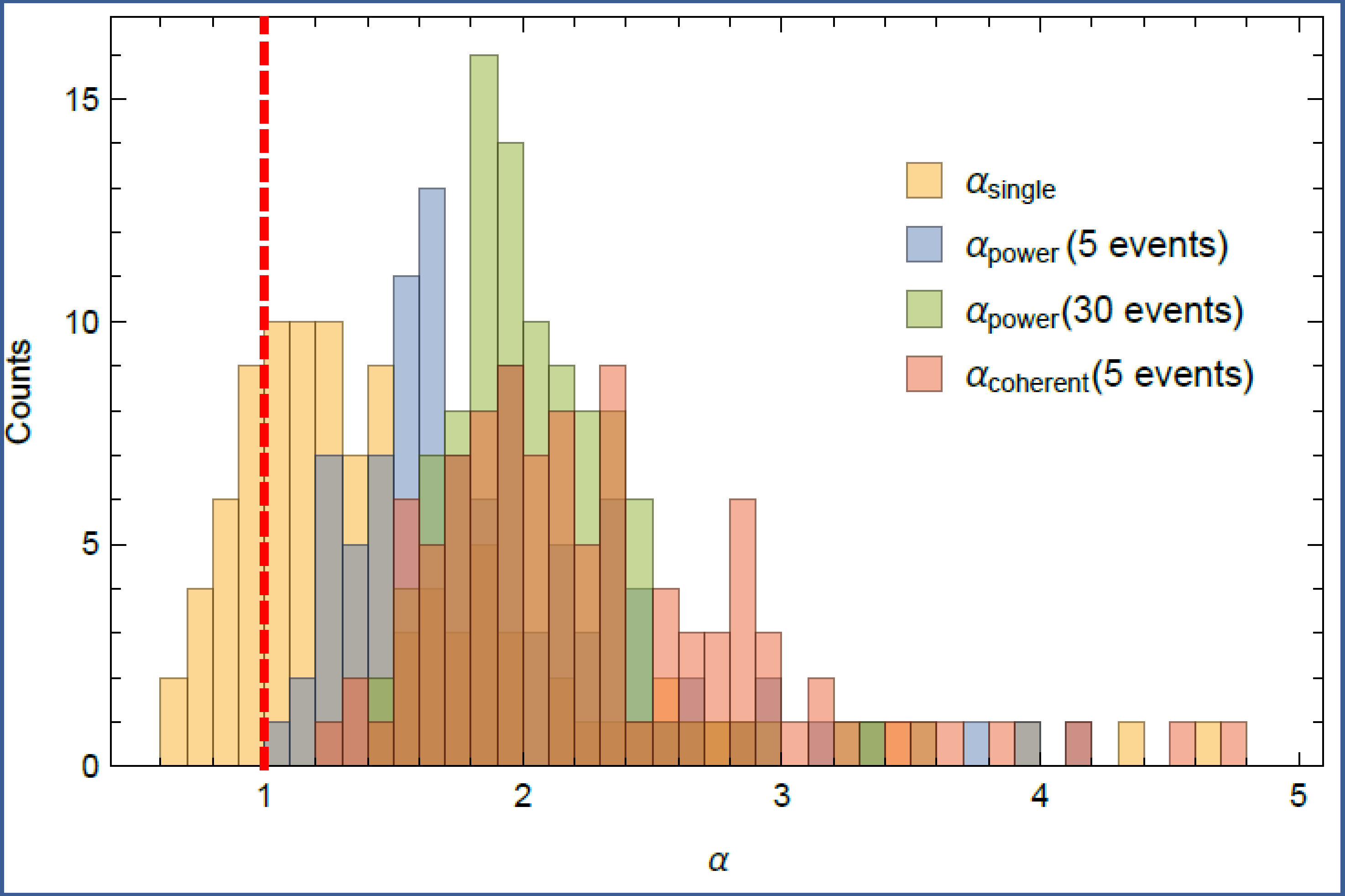
- Single loudest (2,2) f-mode SNR over 100 MC realizations

Different EOSs, Cosmic Explorer



- Single loudest (2,2) f-mode SNR over 100 MC realizations

Stacking, TM1 EOS1, Cosmic Explorer



- Stacked (2,2) f-mode SNR-proxy α (1 is equivalent to SNR 5 for single event) over 100 MC realizations

Conclusions

- Many possible features of GW events to go after combining data from multiple detections
- For binary black hole systems
  - the Final State Conjectures makes BBH mergers ideal probes of physics beyond GR, or of an unexpected circumbinary environment
  - include the non-linear phase of the ringdown into the analysis
  - stack scaled inspirals : measured PN parameters, constrain/discover beyond GR ppE parameters, etc.
- For binary neutron star systems
  - must deal with the lack of uniqueness in NS structure, in addition to what is likely extreme sensitivity of the detailed properties of a NS remnant to small variation in parameters of the progenitor binary
  - could, as with BBHs, stack BNS inspirals : measure PN parameters describing tidal deformability, parameters that try to capture poorly understood conjectured properties including crust cracking and excitation of resonant modes in the star, and constrain/measure non-GR phenomena (dynamical scalarization, etc).