

Pure gravitational wave estimation of Hubble’s constant using neutron star–black hole mergers

Leo W. H. Fung^{1,2}*, Tom Broadhurst^{3,4,5} and George F. Smoot^{1,4,6,7,8}†

¹Department of Physics and Jockey Club Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

²Institute for Computational Cosmology & Centre for Extragalactic Astronomy, Durham University, Stockton Road, Durham DH1 3LE, UK

³Department of Theoretical Physics, University of the Basque Country UPV/EHU, E-48080 Bilbao, Spain

⁴DIPC, Basque Country UPV/EHU, E-48080 San Sebastian, Spain

⁵Ikerbasque, Basque Foundation for Science, E-48011 Bilbao, Spain

⁶Energetic Cosmos Laboratory, Nazarbayev University, Qabanbay Batyr Ave 53, Astana 010000, Kazakhstan

⁷Paris Centre for Cosmological Physics, APC – AstroParticule et Cosmologie, Université de Paris, CNRS/IN2P3 CEA/Irfu, 10 Rue Alice Domon et Leonie Duquet, F-75205 Paris Cedex 13, France

⁸Physics Department & LBNL, University of California at Berkeley, Berkeley, CA 94720, USA

Accepted 2025 February 6. Received 2025 January 14; in original form 2023 September 30

ABSTRACT

It has been predicted before the first detection of such events by LIGO (Laser Interferometer Gravitational-Wave Observatory) that the gravitational wave (GW) emission from compact binary coalescence can be used to constrain the expansion rate of the Universe. Here we show how H_0 can be derived purely from the GW of neutron star–black hole (NSBH) mergers using a Bayesian hierarchical inference framework. This method provides an estimate of H_0 spanning the redshift range $z < 0.25$ with current GW sensitivity and without the need for any afterglow detection. We utilize the inherent distribution of neutron star masses together with the NSBH waveform amplitude and frequency to estimate distance and redshift, respectively, thereby measuring H_0 for the NSBH events up to systematics and observational uncertainties. With the most updated observations, our first estimate is $H_0 = 96_{-50}^{+44}$ km s^{−1} Mpc^{−1} for the likely NSBH events GW190426 and GW200115 using the proof-of-concept analysis outlined in this work. Taking into account of realistic detectors response and sensitivity, we forecast that soon, with 10 more such NSBH events, we can reach competitive precision of $\delta H_0/H_0 \lesssim 20$ per cent. It is valuable to have a completely independent, both in terms of observations and relevant physics, method of determining H_0 as is emphasized by the current controversy on H_0 and attempts to explain it by invoking new physics.

Key words: gravitational waves – methods: observational – cosmological parameters – distance scale.

1 INTRODUCTION

Knowledge of the expansion rate of the Universe is essential for deriving the physical characteristics of extragalactic objects, including sizes and masses of galaxies and clusters and the age and luminosity of their stellar contents. Increasingly precise standard candles and rulers have revealed that Hubble expansion has actually accelerated over the past few Gyr, indicating a constant energy density with negative pressure governs the expansion rate today. As H_0 measurements have improved in precision, consistent claims that smaller values of H_0 have been inferred from distant cosmological measurements compared with local estimates of H_0 from the stellar calibrated standard candles (Scolnic et al. 2018; Verde, Treu & Riess 2019; Knox & Millea 2020; Riess et al. 2022).

A careful recalibration of local Cepheids is now claimed to slightly relieve this H_0 tension when supernovae are classified according to

metallicity, or from a recalibration of evolved stars at the tip of red giant branch (Freedman et al. 2019, 2020; and see Freedman 2021 for a review).

Alternatively, theoretical proposals to explain the Hubble tension include a reinterpretation of the cosmic microwave background (CMB) data, invoking the modifications of recombination physics (for a review, see Valentino et al. 2021), may allow the ‘standard’ sound horizon scale length to be slightly smaller. This is achieved by shifting the surface of the last scattering to a little earlier in time, corresponding to a higher redshift allowing agreement with the local H_0 estimates. It has been shown that this may be physically implemented by modifying recombination physics (Ivanov, Ali-Haïmoud & Lesgourgues 2020; Fung et al. 2021, 2023; Jedamzik, Pogosian & Zhao 2021; Luu 2023), allowing room for reinterpreting the CMB data with a somewhat higher H_0 . There are also studies that attempted to break down the Λ cold dark matter (Λ CDM) model and analyse the origin of such tension (Krishnan et al. 2021a, b).

The seemingly irreconcilable differences among H_0 measurements obtained from different ‘standardizable’ calibration (candles and rulers) strengthen the need for measuring H_0 with independent

* E-mail: leowhfung@gmail.com

† Emeritus.

methods, where the calibrations are driven by unrelated physical mechanisms (Cervantes-Cota, Galindo-Uribarri & Smoot 2023). Gravitational wave (GW) observations provide exciting new opportunities for constraining such cosmological expansion. Compared to conventional electromagnetic (EM) observations where flux falls as the square of the luminosity distance, the GW signal strength scales only linearly inverse with luminosity distance – as the detected strain amplitude depends on the square root of energy density per unit time. As a result, the scaling of GW signal strength can be used to distinguish models that involve modification of photon trajectories [consequently modifying the cosmic distance duality relation (Etherington 1933; Bassett & Kunz 2004)], or other proposals that introduce strong redshift-dependent physics (e.g. Camarena & Marra 2021). Furthermore, GW observables are independent of the nuclear processes that drive the catastrophic supernova explosion, and are dependent mostly on gravitational physics. Consequently, the GW channel provides unique measurements of H_0 , allowing for comparison with those derived from supernovae measurements.

Aside from the unique scaling of signal strength, there are also some applications of GW that depend on physical assumptions lying outside GW data, and consequently, such methods do not fully benefit from the new GW-based window. It has been proposed that spatial cross-correlations between GW sources and galaxy surveys can provide a ‘siren’-based angular scale that does not rely on individual (unknown) GW host galaxy identification (Oguri 2016; Mukherjee et al. 2021; Cigarrán Díaz & Mukherjee 2022). This rests on a statistical comparison of the ensemble of binary black hole (BBH) sky locations with the Dark Energy Survey (DES) – an optical galaxy survey – as no unique host galaxy identification (Abbott et al. 2023a) is yet possible. Hence, this type of method that relies on the establishment of a standard ruler in general shares almost the same set of systematic problems as traditional standard ruler approaches, in particular the baryonic acoustic oscillations. As a result, the GW data do not represent a truly independent means for constraining H_0 . For the consideration of Hubble tension, this would be unable to offer an independent check of the systematics for the late time measurement of H_0 . Alternatively, a galaxy catalogue can also be used to assign (a distribution of) probable redshifts to each of the BBH events, therefore, allowing the determination of H_0 using the population of BBH events. This is implemented in Finke et al. (2021), Abbott et al. (2021a), and Gray et al. (2023), and the LIGO–Virgo–Kagra Collaboration (LVK) analysis yields $H_0 = 68^{+8}_{-6} \text{ km s}^{-1} \text{ Mpc}^{-1}$ (Abbott et al. 2023a).

To overcome the possible contamination from the galaxy catalogue where a specific cosmological model is implied during data reduction, we propose to rely on pure GW observation, as described below.

The long-anticipated GW-based method to constrain H_0 using binary neutron stars (BNSs), pioneered by Schutz (1986), falls in the ‘standard candle’ class and requires a prompt detection of EM nova emission generated during the merger, to define the redshift. This method has finally been realized with the first GW detection of a BNS event GW170817, for which there is an optical redshift (Abbott et al. 2017a, b). Fortunately, *Fermi* satellite detection of prompt gamma-ray emission from associated kilonova emission has localized the host galaxy, providing an optical spectroscopic redshift for the host. The recession velocity of the host galaxy is only 3000 km s^{-1} , thus requiring a sizeable peculiar motion correction, at an estimated level of $\simeq 800 \text{ km s}^{-1}$, resulting in $H_0 = 70^{+12}_{-8.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$. The uncertainty on this estimate brackets the disputed range of H_0 , and implies approximately 10 more such BNS events are required to address independently the current H_0 tension between

Type Ia supernovae (SNe Ia) and the CMB/large-scale structure based estimates (Nissanke et al. 2013). However, such coincident, multimessenger observations are not anticipated to be available frequently, unless the angular resolution of GW observatories can be drastically improved (by ~ 3 orders of magnitude) to localize the position of the event, and indeed subsequent events classified as BNS were not detected at other wavelengths in prompt follow-up searches.

These GW-based methods for H_0 rely on an independent EM-based redshift estimation, either individually in the case of BNS mergers or statistically with redshift surveys described above. Alternatively, if primordial black holes (BHs) exist and their mass function are known, their event rate evolution across redshift also encoded the information about H_0 (Ding 2024).

A pure GW-based method for estimating H_0 has been proposed for the future, requiring much-improved sensitivity, when the tidal deformation of the neutron star (NS) by the BH during a neutron star–black hole (NSBH) merger is detectable. This can provide in principle a GW-based estimate of H_0 from the additional distance dependence on this physical deformation (Chatterjee et al. 2021; Ghosh, Biswas & Bose 2022). Here we propose another more practical ‘standard’ strain method without relying on any detectable EM counterparts, which have apparently been overlooked, provided by the mildly scattered NS mass function established by radio, X-ray, and optical observations in the Milky Way and Local Group, including BNS pulsars.

Below, we will highlight a few aspects in which NSBH events are more advantageous than BNS events, for the purpose of deriving a reliable estimation of H_0 . This is partially owing to the larger detection horizon of NSBH events over BNS events, so that the fractional error in redshift estimates caused by local peculiar motions is smaller. We apply this method to the currently secure NSBH event GW200115 and other possible and proposed candidate NSBH events variously reported. We demonstrate that the existing instruments (Buikema et al. 2020) already suffice for constraining H_0 within $\delta H_0/H_0 \lesssim 20$ per cent, simply by gathering 10 yr of measurements, even in the absence of significant sensitivity upgrades that may be anticipated. We will emphasize throughout that NSBH events and to a lesser extent BNS events can in combination provide competitive precision on H_0 that is purely derived from the GW radiation alone, without the need for a spectroscopic redshift. Such additional independent information would of course provide further improvement, so that this combined general NS-based estimate of H_0 we can anticipate will prove useful in practice. A prototype of this method has been proposed long ago (Chernoff & Finn 1993; Marković 1993), with more detailed predictions before the first successful detection of GW event (Taylor & Gair 2012; Taylor, Gair & Mandel 2012) that focused on using BNS instead of NSBH as proposed here. More recently after the successful detection of compact binary coalescence events, there are also studies suggesting the features of the BH mass function can also be used for a similar purpose (Farr et al. 2019; Mastrogiovanni et al. 2021; Ezquiaga & Holz 2022). This work is inspired by the now-understood practicalities of GW detection, in particular the NSBH event detections for which the massive BH member provides a much larger detection horizon ($z < 0.25$) compared to BNS events where detected events are currently so local that peculiar motions dominate the distance estimate. We present a simplified, comparative analysis of using BNS or NSBH for the purpose of measuring H_0 , with realistic uncertainties from current observations that allow us to assess the reliability and the overall consistency with the existing data, highlighting the practical merits of using NSBH events.

This paper is organized as follows. In Section 2, we sketch out the foundation of the method, and then we describe our choice of data sets for GW events and local X-ray binaries in Section 3. We present a rigorous Bayesian hierarchical analysis applied to the data in Section 4 and present our analysis of the first reported GW events compatible with being NSBH binaries, focusing in particular on the secure NSBH detection claimed to date, GW200115 (Abbott et al. 2021c), for a first estimate of H_0 . We then make a forecast for the precision attainable for H_0 and the evaluation of associated systematics in Section 5. We finally conclude in Section 6.

2 INTUITIVE OVERVIEW OF THE METHOD

The widely applied standard candle measurement of H_0 compares independently determined luminosity distances, d_L , with spectroscopic redshifts, z , for SNe Ia where the intrinsic luminosity is calibrated with local Cepheids. Here, we emphasize that a pure GW-based measure of H_0 is now feasible without the need for a follow-up spectroscopic redshift. Owing to stellar stability constraints, the mass range of NSs is fairly tight about the Chandrasekhar mass. Hence the lower mass binary member (NS) modulation of the NSBH waveform is predictable, so the degeneracy between luminosity distance and redshift of NSBH is relatively small. Consequently, GW detections alone provide useful measurements of both d_L and z and hence a statistical measure of H_0 , as we outline below. This method we apply to GW waveforms identified as NSBH binaries, where the mass ratio can be defined from the waveform modulation and the presence of the relatively massive BH extends the detection horizon to $z \lesssim 0.25$, large enough for statistical measure of H_0 .

2.1 Observables

For the purpose of providing insight into our method, in this section, we would associate the spectral features of the GW waveform to physical quantities in an order-by-order manner from the post-Newtonian (PN) expansion of the waveform. We emphasize that such order-by-order association of spectral features to physical quantities is not exact; physical quantities are best determined by fitting to the full waveform numerically, and this is indeed our approach in the full analysis as outlined in Section 4. Therefore, throughout the below discussion, we use the notation \sim to refer to such order-by-order associations. The observables are highlighted by the subscript $_{\text{obs}}$. None the less, this analysis also provides an intuitive interpretation of the uncertainty budget involved in the proposed H_0 measurement method.

The amplitude of the strain h_{obs} according to the PN calculations (Poisson & Will 1995; Ajith et al. 2008) for compact binary sources is given by the following combination of physical parameters:

$$h_{\text{obs}} \propto \frac{((1+z)\mathcal{M}_c)^{5/6}}{d_L}. \quad (1)$$

The $(1+z)$ factor can be understood as follows: as h_{obs} scales with $\sqrt{E_{\text{energy}}}$, the rate of radiation energy received by the observer is time dilated by the cosmological redshift. Also, the inverse square law $1/d_L^2$ diffusion of energy (flux) implies an inverse dependence of h_{obs} on luminosity distance d_L . Hence, to determine d_L , we therefore need another observable, ideally independent to h_{obs} , to pin down $(1+z)\mathcal{M}_c$. If we take the Fourier transform of the time series of strain over different time windows (e.g. Q -transform), the time evolution

of the corresponding frequency spectrum can also be determined.¹ The first quantity is the redshifted chirp mass $(1+z)\mathcal{M}_c$. The chirp mass is defined as

$$\mathcal{M}_c \equiv \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} = \frac{1/q^{3/5}}{(1+1/q)^{1/5}} m_2, \quad (2)$$

where we have defined the mass ratio $q = m_2/m_1$, and implicitly by definition $m_2 < m_1$ for sorting the masses in the binary system. This quantity, upon redshifted, is indeed imprinted on the spectral moment of the 0.5PN chirping spectrum $f_{\text{obs}}(t)$: (Note, however, that solving the spectral evolution must be done by solving the full differential equation, and the order-by-order approach as outlined below only makes sense under some strict conditions.)²

$$f_{\text{obs}}^{-11/5} \left(\frac{df_{\text{obs}}}{dt} \right)^{3/5} \sim (1+z)\mathcal{M}_c. \quad (3)$$

The redshift effect is easy to understand: the frequency of GW in the source frame is redshifted to $(1+z)^{-1}f_{\text{obs}}$ and thus the time derivative $(1+z)^{-1}t_{\text{obs}}$. If we can pin down m_2 and q , it would then be possible to estimate z . Indeed, considering the NSBH system, the NS mass m_2 can be pinned down by Milky Way observations as $m_2 = 1.4 \pm 0.1 M_\odot$ (see Section 3 for the evidence of this number).

In an order-by-order manner, it is also possible to determine the mass ratio q by another combination of spectral evolution:

$$f_{\text{obs}}^{-13/3} \frac{df_{\text{obs}}}{dt} \sim \frac{1/q}{(1+1/q)^2}, \quad (4)$$

where all the redshift dependencies are cancelled by the multiplicative scaling term that we have neglected here (Poisson & Will 1995).³

Therefore, this observable provides a measurement on q .

We summarize the relations between the observables and physical source quantities in Fig. 1. Clearly, the observed quantities provide largely independent measurements of z and d_L , respectively. The $(z_i, d_{L,i})$ pairs from each observed NSBH event i can therefore be used to determine the Hubble constant H_0 . In the big bang cosmology, we have

$$\frac{d_L}{(1+z)} = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')}. \quad (5)$$

In the subsequent numerical analysis, we will assume $\Omega_m = 0.3$, so that $E(z) = \sqrt{0.3(1+z)^3 + 0.7}$, as the current data quality does not allow strong constrain on Ω_m together with H_0 .

¹The situation is more complicated when the signal is contaminated by noise, where parametric models of the GW waveform are used as matched filters for deriving the waveform evolution. We will ignore this point in this section for illustration.

²Strictly speaking, the spectral moments cannot be separated in this way. However, by using multiscale perturbation methods (e.g. Poincaré-Lindstedt), the order-by-order separation can be a good approximation in a narrow time window, where the window time-scale is defined by the leading coefficients of each spectral moment.

³Precisely, the spectral evolution df_{obs}/dt couples with the polynomial in f_{obs} . As a result, the combination of moments of the spectrum is not orthogonal to each other. The full PN calculation for relating the observables to the chirp spectrum can be found in the pioneering work (Poisson & Will 1995, in particular equation 1.3 therein) and for accurate estimation of the spectrum must be predicted numerically using ordinary differential equation solvers.

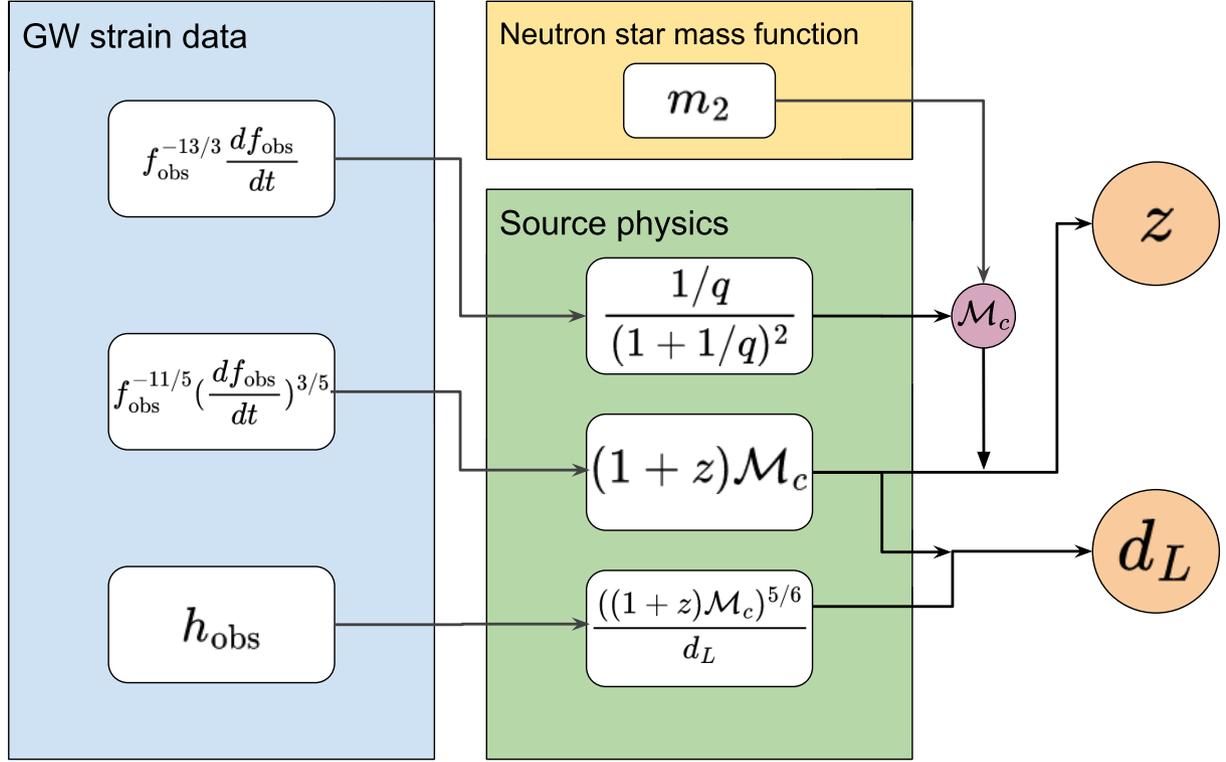


Figure 1. Summary of the relations between observables and the physical quantities of the GW source, and the associated relation to quantities of cosmological interest. In the box on the left, we show the observables that can be derived from the time series of GW strain, and how these observables are related to some physical parameter combinations in the middle lower box. These physical parameter combinations can be used to solve for z and d_L statistically when an NS mass function is supplied, therefore, allowing one to determine H_0 .

2.2 Comparison with binary neutron star events

Here we compare and contrast the ability of BNS and NSBH events for estimating H_0 by the method proposed here, including BNS events that do not have an independent spectroscopic redshift estimate because neither host galaxy nor EM emission is identified. The absence of observed EM counterparts likely comprises the majority of such NS-related GW events given the follow-up success to date, with EM emission found for just one very local BNS event, at only 40 Mpc, and the absence of EM detections for the other BNS and NSBH candidates.

One may argue BNS systems are more suitable for the purpose of establishing an independent estimate on z and thus H_0 . This is seemingly evident from equation (2): if both $m_{s,1}$ and $m_{s,2}$ are pinned down, \mathcal{M}_c are known a priori.

There are though two advantages for NSBH events compared to BNS in terms of realizing our H_0 method. First, the detection horizon for NSBH observations is much larger than is BNS observations. As shown in Ajith et al. (2008), the signal-to-noise ratio (SNR) of GW events scale as $\rho(m_{s,1}, m_{s,2}, z) \sim ((1+z)\mathcal{M}_c)^{5/6}/(d_L(z))$. For BNS with mass ratio $q \approx 1$, $\mathcal{M}_c \approx 0.87m_2$, whereas the BH component serves as an efficient amplifier, boosting the detectable SNR of the companion NS: on average, the BH mass can be as high as $m_{s,1} \approx 20 M_\odot$, corresponding to an amplification factor of $1/q \approx 13$. Thus $\mathcal{M}_c \approx 2.75m_2$, thereby extending the detection horizon for NSBH events to $d_L^{(\text{NSBH})}/d_L^{(\text{BNS})} \approx 4$, four times larger than BNS events.

As shown in Appendix C, the enlarged horizon offered by NSBH – which leads to a larger absolute error in determining the event distance – does not translate into deteriorated precision in determining H_0 .

The enlarged detection horizon is, however, beneficial in terms of the precision in pinning down the cosmological redshift. In practice, the measured redshift is contributed by both the cosmological expansion and the local peculiar motion. When the fractional redshift error is assessed, a smaller detection horizon would be translated to a larger fractional redshift uncertainty. This effect can be mitigated by making the peculiar motion-induced redshift contribute only a tiny portion of the total redshift, effectively by using the GW radiated at a farther distance.

The second advantage for NSBH over BNS events comes from the absolute uncertainty in determining the mass ratio $q \equiv m_{s,2}/m_{s,1}$, which is smaller with $1/q \gg 1$. In our method, the source frame chirp mass is (implicitly) inferred by combining an NS mass function prior that supplies the value of $m_{s,2}$, with the observed constraint on q to constrain the source frame chirp mass \mathcal{M}_c . The significant reduction in the observational uncertainty δq can thus be translated into a tighter constraint in the source frame chirp mass \mathcal{M}_c , which improves the constraint in both z and d_L . For BNS systems $q \approx 1$ and thus the mass ratio is relatively less well constrained by the data, leading to significant H_0 uncertainty, whereas the tight prior restriction from the NS mass function instead provides a desirable tight prior on the source frame chirp mass \mathcal{M}_c for BNS events. This is verified via the simulated observation outlined below in Section 5.1 and Fig. 6, along with other detailed examinations to contrast between BNS and NSBH.

As a closing remark, we would like to highlight a caveat in the NS mass function that is shared by both BNS- and NSBH-based methods. In this work, we assumed implicitly that the population of NS that participates in binary coalescence shares the same characteristics as those observed in the Milky Way. Specifically, we assumed that

there is no preferred mass ratio for binary coalescence to take place, which may not be the astrophysical. If such a preferred mass ratio exists, we would expect that BNS-based method would be biased the most. In our analysis, we tried to minimize the bias effect by using an uninformative mass ratio prior when determining the event redshift. As such, the mass ratio q of an NSBH event is determined mostly from the observed GW waveform – and similarly for the redshifted chirp mass $(1+z)\mathcal{M}_c$. Note importantly that NSBH events inherently generate stronger signals than BNS events, so that the observed waveform is more informative in pinning down q and $(1+z)\mathcal{M}_c$. As such, the possible influence of mass ratio bias is mitigated by this data-driven approach. With the foreseen improvement in sensitivity of GW observatories, the subtle effect of preferred mass ratio would definitely become relevant; at the same time, more observations of binary systems involving NS would provide more data-driven constraints towards a more suitable treatment to the uncertain mass ratio prior.

3 DATA

3.1 NSBH mass function in the Milky Way

The initial birth masses of NSs are predicted to lie in the range $M_{\text{birth}} \sim 1.08\text{--}1.57 M_\odot$, depending on the modelling of the underlying hydrodynamical processes and thermodynamics with subsequent evolution. In the case of close binaries, NSs accrete material from companions, leading to a dependence of the NS mass on the physical properties of the secondary, companion star. The upper limiting mass of an NS star is also understood to sensitively depend on uncertain nuclear reactions and also rotation rate affecting the equation of state (Kiziltan et al. 2013), under physically extreme nuclear conditions, allowing unique tests of the standard model in conditions as yet unrealized in the laboratory.

Fortunately for our purposes, the empirically determined NS mass function is well defined, with a mild spread of masses. Decades of observations have established that the majority of NSs have masses consistent with a single mass scale of $\langle M_{\text{NS}} \rangle = 1.4 M_\odot$, coinciding with the classically predicted for a universal iron core (Lattimer 2012), including over ~ 50 radio pulsars (Thorsett & Chakrabarty 1999). Hence, we first adopt this empirically well-defined prior mass function, approximately described simply by a Gaussian centred at $\langle M_{\text{NS}} \rangle = 1.4 M_\odot$, and consider an intrinsic spread of NS mass ranging over $\sigma_{\text{NS}} \in [0.05, 0.2] M_\odot$ in our subsequent analysis. A possible bimodal mass distribution for Galactic NSs has been claimed in recent analyses for a classification by the type of binary companion (Kiziltan et al. 2013), and fitted to two Gaussians with differing mean mass and variance. For binary NS systems, a mean NS mass of $\langle M_{\text{NS}} \rangle = 1.35 M_\odot$ is claimed to be significantly smaller than the mean NS mass of NS–white dwarf binaries, estimated to be $\langle M_{\text{NS}} \rangle = 1.5 M_\odot$ and with a smaller spread in mass. Other classifications are also proposed (Kiziltan et al. 2013) that do not separate companion types (Valentim, Rangel & Horvath 2011; see Horvath & Valentim 2017 for a review of recent progress of constraining NS mass function), may also suggest two separated mass peaks with masses similar to Kiziltan et al. (2013) and Zhang et al. (2011). We replot the EM-observed NS mass data from Kiziltan et al. (2013) in Fig. 2, along with the three fiducial models of the NS mass function that we will use throughout this analysis. We emphasize that our framework of analysis can be easily adopted to any NS mass function and in the more generic case in which a binary mass distribution is considered. We find that the H_0 posterior inferred from such a simple Gaussian model does not significantly

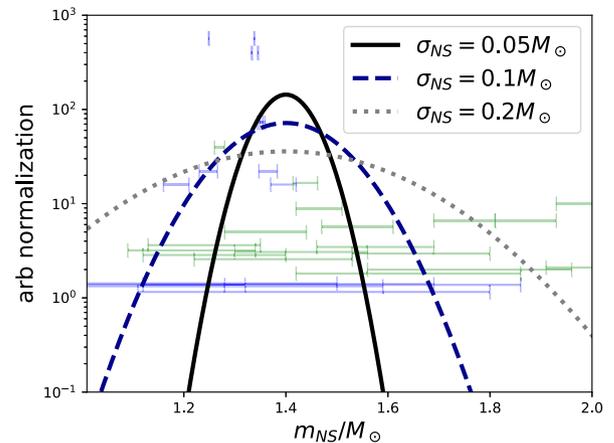


Figure 2. The EM-observed masses of NSs in double NS systems and NS–white dwarf systems, compiled by Kiziltan et al. (2013) and they are replotted here. For clarity, we distribute these data points along the y-axis, whereby their positions on the y-axis scale with the inverse of mass measurement uncertainty. We also plot the three fiducial models of the underlying NS mass function – as a Gaussian with a peak at $\langle m_{\text{NS}} \rangle = 1.4 M_\odot$ and with various variances as shown in the legend.

differ from adopting a more complicated NS mass function, so that our H_0 forecast is insensitive to the details of the NS mass function given the GW data quality anticipated. This argument would be verified in Section 5.

Lastly, there is also evidence from the gamma-ray that there can be some NS-like compact objects with relatively high masses (Romani et al. 2012), with a massive tail up to $M_{\text{NS}} \lesssim 2.3 M_\odot$ established empirically recently for the well-studied ‘black widow’ pulsar PSR B1957+20 (Fruchter, Stinebring & Taylor 1988; and a more recent analysis by Kandel & Romani 2022). The abundance of these relatively massive MS pulsars is much lower than the ‘classical’ population at $\langle M_{\text{NS}} \rangle \pm \sigma_{\text{NS}}$ and these massive examples appear separated from the ‘classical’ population tightly clustered around $1.4 M_\odot$, allowing the possibility of excluding this uncertain, massive population from our analysis, so we may restrict ourselves to the predominant NS population both in calibrating the NS mass distribution and in comparison to the NSBH events observed from the GW channel. We anticipate that as long as we consistently exclude conspicuously massive NS-related events, this prior selection will not induce a significant bias in our subsequent analysis.

In essence, the exclusion of high-mass NS events is implemented in two different steps of our analysis. First of all, we choose an NS mass function prior which ignores the uncertain details for $m_{\text{NS}} > 2 M_\odot$. For the varieties of NS mass function prior we chose, they typically peak at $m_{\text{NS}} = 1.4 \pm 0.1 M_\odot$, thus providing minimal support to $m_{\text{NS}} > 2 M_\odot$ with a probability mass at $\gtrsim 5\sigma$ Gaussian tail. Secondly, when we choose to include a specific NSBH event in our analysis, we pick particularly those where the posterior average NS mass is lower than $2 M_\odot$. Equivalently, this is implicitly applying an additional zero-valued prior to adjust the likelihood of high-mass NS events.

3.2 Gravitational waves from NSBH events

The LVK completed the third observational run in 2020 March and released the GWTC-3 catalogue (Abbott et al. 2023b), containing 90 observed binary coalescence events. The majority are classified as BBH mergers, with initially two events classified as NSBH mergers (Abbott et al. 2021c). Subsequently only one of these

Table 1. List of candidate NSBH events. We list the reported observed frame quantities here, rather than the more model-dependent source frame quantities as shown in the table compiled by LVK (Abbott et al. 2023b). The uncertainty ranges are the 68 per cent highest posterior density intervals. Remarkably, the uncertainty is correlated and clearly visible in terms of $q \equiv m_2/m_1$, where the error bar on this mass ratio is much smaller than $m_1(1+z)$ and $m_2(1+z)$. Values shown are only for quick reference; we use the full posterior samples in deriving these summaries, and similarly in all the subsequent analyses only the events in **bold font** are regarded as reliable for the analysis in this paper, in reference to the reported reliability of the classification of these events, as described in Section 3.

Event ID	$m_1(1+z)/M_\odot$	$m_2(1+z)/M_\odot$	$\mathcal{M}(1+z)/M_\odot$	q	d_L/Mpc	Ref.
GW190426.152155	$6.45^{+1.07}_{-1.78}$	$1.61^{+0.24}_{-0.34}$	$2.60^{+0.007}_{-0.006}$	$0.28^{+0.05}_{-0.15}$	394^{+105}_{-109}	Abbott et al. (2024), Li et al. (2020), Aubin et al. (2021)
GW190917.114630	$11.2^{+2.08}_{-2.46}$	$2.45^{+0.24}_{-0.50}$	$4.21^{+0.014}_{-0.018}$	$0.24^{+0.03}_{-0.12}$	740^{+173}_{-195}	Abbott et al. (2024)
GW191219.163120	$33.2^{+2.44}_{-3.61}$	$1.35^{+0.07}_{-0.08}$	$4.81^{+0.021}_{-0.045}$	$0.04^{+0.005}_{-0.007}$	682^{+187}_{-168}	Abbott et al. (2023b)
GW200105.162426	$9.26^{+1.07}_{-1.64}$	$2.11^{+0.19}_{-0.28}$	$3.62^{+0.005}_{-0.007}$	$0.24^{+0.04}_{-0.08}$	$283^{+79.6}_{-76.7}$	Abbott et al. (2021c, 2023b) ⁴
GW200115.042309	$6.82^{+1.13}_{-1.26}$	$1.48^{+0.13}_{-0.25}$	$2.58^{+0.004}_{-0.004}$	$0.23^{+0.04}_{-0.09}$	$283^{+62.1}_{-82.7}$	Abbott et al. (2021c, 2023b)

⁴ $p_{\text{astro}} = 0.36$.

events GW200115 has survived scrutiny and is now reported in the combined final catalogue (Abbott et al. 2023b). An earlier event initially classed as BBH may now conform better to the NSBH definition, as we tentatively argue below, namely GW190426.

The identification of NSBH events is predominantly based on the observed secondary mass $m_{o,2} = (1+z)m_{s,2}$, unlike BNS events. In principle, NSBH can produce a more complicated waveform than is BBH due to the tidal deformation of the NS. However, the current data quality is not yet able to constrain such higher order tidal effects (Chatterjee et al. 2021). Therefore, the claimed NSBH event(s) mostly rests on a low secondary mass estimated to lie below a maximum of $m_{o,2} \lesssim 2.5 M_\odot$. In Table 1, we list all the reported candidates events, selected by their posterior mean primary masses $m_{o,1} > 5 M_\odot$ and secondary masses $m_{o,2} < 2.5 M_\odot$, comprising a total of five events satisfying this criterion. These events are in general of relatively low SNR (as the SNR scales almost linearly with chirp mass), but are compensated to some extent by the better instrumental sensitivity at higher frequencies. GW events with $m_{s,2} = 2.6 M_\odot$ (GW190814) and $m_{s,2} = 2.8 M_\odot$ (GW200210) have been reported in GWTC-3. Some of us have discussed these separately in the context of possible lensing of NSBH events that allows detection at higher redshift where the chirp mass is enhanced significantly by $1+z$ (Broadhurst, Diego & Smoot 2020). Furthermore, as will be shown in Section 5, the robustness of our method is not strongly influenced by the inclusion or the exclusion of such high-mass events.

In Table 1, we indicate the status of the five possible NSBH candidate events reported to date.⁵ Three of these are reported as marginal, including GW190426 with a high false alarm rate 32 yr^{-1} in the MBTA (Multi-Band Template Analysis) pipeline, first reported in the GWTC-2.1 marginal catalogue (Aubin et al. 2021; Abbott et al. 2024). As commented by LVK (Abbott et al. 2024), such a high false alarm rate is mainly attributed to the initially low NSBH event rate adopted by prior expectations based on the absence of NSBH events, which may be regarded as a somewhat circular argument resulting in the probability of an astrophysical origin assigned to this event of $p_{\text{astro}} \leq 0.12$ across all the analysis pipelines. Hence, this ‘prior’ may now be revised upward in retrospect with the subsequent confident identification of GW200115 as an NSBH event. The event GW191219 is also a marginal detection with a high false alarm rate of 4.0 yr^{-1} , and the secondary mass in the source frame is a relatively low-mass $m_{s,2} < 1.2 M_\odot$ at 68 per cent credibility. This marginal

event is also with by far the most extreme mass ratio $q = 0.04^{+0.005}_{-0.007}$ among the observed events, raising further doubt about its viability as a real astronomical event and this is why it was flagged as by the LVK as marginal (Abbott et al. 2023b, table II).

A third marginal NSBH candidate event is GW200105 of low SNR, partially due to LIGO Hanford being offline at the time of detection. The only detector that reached the desired SNR is the LIGO Livingston (Abbott et al. 2021c), where a chirping signal is clearly observed when the data stream is plotted on a Q -transform spectrogram (Abbott et al. 2021c). Despite the relatively strong SNR = 13.9 reported in Abbott et al. (2023b), the event has a $p_{\text{astro}} = 0.36$. For this reason, this event is classified as marginal by the LVK in Abbott et al. (2023b).

This leaves two NSBH candidates, including GW200115, which has been consistently classed by the LVK as a secure NSBH event (Abbott et al. 2021c), whereas the earlier revised detection GW190426 which although not classed as NSBH or claimed as such does seem to satisfy the NSBH requirement in terms of the security of its detection and the reported component masses. So we focus on these two events here, and in particular, GW200115 given the secure status of this NSBH event as claimed by the LVT (Abbott et al. 2021c).

4 STATISTICAL FRAMEWORK

Throughout this paper, we use the notation $\pi(\cdot)$ to denote the prior distribution, in contrast to the probability distribution $p(\cdot)$ to describe relations among the model parameters and the data. We provide a quick reference to the forms of the priors used throughout the analysis in Table 2.

4.1 H_0 posterior

We aim here to find the posterior for H_0 given the event waveforms, $\{D_i\} \equiv \bigcup_i D_i$, namely $p(H_0|\{D_i\})$. For each event with observed waveform D_i , the LVK collaboration has fitted a cosmology-free event-level posterior $p_i(d_L, z, \theta|D_i)$, with θ are the remaining event level parameters (excluding d_L and z). We wish to ‘sum’ all the events properly to infer the (population level) value of H_0 .

To avoid ab initio fitting, we use the following reweighing scheme as proposed originally for galaxy morphology data set in Hogg, Myers & Bovy (2010), and employed subsequently in GW analyses (Mandel, Farr & Gair 2019). The posterior can be rewritten using Bayes theorem to flip H_0 as the condition, and expand the

⁵We refer to the GW events in the table by the first half of the event name, ignoring the ID after the underscore.

Table 2. A summary of all the priors that we use throughout this paper.

Physical meaning	Notation	Form	Remark
BH mass function	$\pi(m_{s,1})$	Power law + peak	Using the best-fitting parameter of the power law + peak model in Abbott et al. (2021b) and equations (B5)–(B7) therein
NS mass function	$\pi(m_{s,2})$	Gaussian + variation	Various distributional parameters are used as detailed in Section. 3.1
Redshift prior	$\pi(z)$	Uniform	Avoid any informative prior that implicitly assumed cosmology
‘Irrelevant’ NSBH merger parameters	$\pi(\theta)$	Uniform	Contribute to an overall constant in the likelihood
Hubble constant prior	$\pi(H_0)$	Uniform (20, 180) km s ⁻¹ Mpc ⁻¹	Uninformative

marginalized event level parameters as

$$\begin{aligned}
 p(H_0|\{D_i\}) &\propto \pi(H_0) \int p(\{D_i\}|z, d_L, \theta) p(z, d_L, \theta|H_0) dd_L dz d\theta \\
 &= \pi(H_0) \prod_i^{N_{\text{obs}}} \int p(D_i|z, d_L, \theta) p(z, d_L, \theta|H_0) dd_L dz d\theta. \quad (6)
 \end{aligned}$$

The next step is to invoke the Bayes theorem to rewrite $p(D_i|d_L, z, \theta)$:

$$p(D_i|d_L, z, \theta) = \frac{p(z, d_L, \theta|D_i, \text{LVK})p(D_i|\text{LVK})}{p(z, d_L, \theta|\pi_{\text{LVK}})}, \quad (7)$$

with $p(z, d_L, \theta|\pi_{\text{LVK}})$ being an uninformative parameter prior adopted by LVK. Following a similar notation, the term $p(z, d_L, \theta|D_i, \text{LVK})$ represents the parameter fitting results when the LVK uninformative prior is used to fit the data. For a similar reason, the term $p(D_i|\text{LVK})$ is an uninformative data prior, which is mathematically equal to a constant, and could be absorbed as an overall normalization constant. In general, these terms can have a non-trivial form to account for the selection bias as well, and we will elaborate on this in Section 4.3.

Note the K_i sample points of the posterior $p(z, d_L, \theta|D_i, \text{LVK})$ are available in the LVK data releases. Using these samples, one can rewrite equation (6) as

$$\begin{aligned}
 p(H_0|\{D_i\}) \\
 \propto \pi(H_0) \prod_i^{N_{\text{obs}}} \frac{1}{K_i} \left[\sum_{\sim \theta, d_L, z}^{K_i} \frac{p(z, d_L, \theta|H_0)}{p(z, d_L, \theta|\pi_{\text{LVK}})} p(D_i|\text{LVK}) \right], \quad (8)
 \end{aligned}$$

where the integration $\int d\theta dd_L dz$ is replaced by the Monte Carlo sum $\frac{1}{K_i} \sum^{K_i}$. As argued above, every term except $p(z, d_L, \theta|H_0)$ in the square bracket can be modelled as an overall multiplicative constant. The data fitting results come into the above likelihood function via the Monte Carlo summation \sum , while the goodness of fit to the model is assessed in the term inside the Monte Carlo sum. In the subsequent subsection, we will explain how the goodness of fit term can be formulated.

4.2 Redshift–distance likelihood

There is a non-trivial point about the Markov chain Monte Carlo (MCMC) samples released by LVK. Despite values of redshift z being reported in the MCMC chain, those reported values of z should be interpreted with extra care – indeed, their face values are not applicable to our application at all.

The LVK fitting pipeline does not fit for the $(1+z)$ factor directly as it is completely degenerate with the rest-frame chirp mass equation (3) in terms of the frequency response behaviour.

Instead, the LVK pipeline utilized the *Planck* cosmology (Planck Collaboration VI 2020) with $H_0 = 67.9$ km s⁻¹ Mpc⁻¹, so that the luminosity distance d_L that parametrized the GW waveform is directly converted to the redshift, and these directly converted

values of z are reported in the MCMC chain. Therefore, the MCMC samples from LVK collaboration are samples of $\{d_L, m_{o,2}, \theta\}$ instead of directly containing z . For completeness, we provided an in-depth discussion about this point in Appendix A.

For this reason, all the probabilities containing z in equation (8) must be replaced by $m_{s,2}$, which indirectly constrains z with the use of the Milky Way NS mass function.

Our prior here is the independent knowledge of the distribution of $m_{s,2}$ from the long-standing Milky Way NS mass distribution, which we can factorize and write H_0 likelihood (the term inside the bracket in equation 8) as

$$\begin{aligned}
 p(m_{o,2}, d_L, \theta|H_0) &= \int dz dm_{s,2} p(m_{o,2}|z, m_{s,2}) p(d_L|z, H_0) \\
 &\quad \times \pi(z) \pi(m_{s,2}) \pi(\theta). \quad (9)
 \end{aligned}$$

It is clear how the information regarding z indirectly comes into the expression: if $m_{s,2}$ are known precisely, $z = m_{o,2}/m_{s,2} - 1$ would inherit the data uncertainty in $m_{o,2}$. The above equation simply takes the intrinsic scatter of $m_{s,2}$ into account, promoting $m_{s,2}$ to a probability $\pi(m_{s,2})$.

Note the conditionals $p(m_{o,2}|z, m_{s,2})$ and $p(d_L|z, H_0)$ (i.e. the first two terms before the multiplication symbol) are deterministic, and are thus possible to be expressed in terms of Dirac delta. This simplifies the expression to

$$\begin{aligned}
 p(m_{o,2}, d_L, \theta|H_0) &= \pi(m_{s,2} = m_{o,2}/(1+z(d_L; H_0))) \pi(\theta) \\
 &\quad \times \pi(z = z(d_L; H_0)). \quad (10)
 \end{aligned}$$

While the inverse function $z(d_L; H_0)$ does not admit an analytical form, a computationally inexpensive numerical solution can be easily obtained because $d_L(z; H_0)$ is a monotonic function of z .

When d_L is known and a fixed value of H_0 is chosen in an MCMC sampling step, the corresponding value of $z(d_L, H_0)$ is uniquely determined. And the likelihood of generating this specific $z(d_L, H_0)$ is given by the conditional $\pi(z = z(d_L; H_0))$.

As the actual redshift prior $\pi(z)$ can be a complicated function that models both the unknown evolution of the intrinsic NSBH event rate and also the luminosity-limited selection effect, we will simply assume an uninformative prior. In this way, the above model likelihood function would depend on the NS mass function $\pi(m_{s,2} = m_{o,2}/(1+z(d_L; H_0)))$ only (up to unimportant normalization constants). As we will show below, the effect of selection bias would further complicate this likelihood.

4.3 Selection bias

In a flux-limited survey, the observed luminosity distribution is always biased to the brighter end, so that even if the intrinsic NS mass distribution is symmetric, this selection effect skews the detectable mass distribution.

However, as we will show below, this selection effect has only a negligible effect on our results considering the uncertainty of

existing measurements limited by the event-level parameters recovery. Furthermore, there is selection bias involved when using EM-constrained NS mass function compiled by other studies (as discussed in Section. 3.1) as the prior. This effect is beyond the scope of our exploratory work focusing on GW-related effects and we here (implicitly) absorbed them into the effective NS mass function priors by choosing various possible σ_{NS} .

Correction for selection bias in GW surveys is difficult, despite the existence of several attempts focused on the BBH systems (Messenger & Veitch 2013; Mandel et al. 2019; Veske et al. 2021). In particular, the full treatment outlined in Mandel et al. (2019) also incorporates the Poissonian count noise contributed by finite number of observed events. In our analysis, we ignore the Poissonian noise in order to manage the computational cost. The selection effect for systems involving NS can be more subtle with the extra requirement of quantifying the tidal evolution of the NS during the NSBH merger. In the treatment quantifying selection bias below we also ignore all the extra contributions arisen from the tidal effect. In other words, we treat NSs as essentially low-mass BHs without any internal structure and deformability.

Denote the subset of data that passes through the selection criteria as $\mathcal{S} = \{D_i \in \text{detected}\}$. As all observed data are sampled from \mathcal{S} instead of all possible classes of waveforms, the size of the sampling space shrinks, similarly for the likelihood of observing the event-level parameters ($d_L, m_{o,2}, \theta$) as shown in equation (10). To account for this, we would like to model the likelihood $p(d_L, m_{o,2}, \theta|H_0, \mathcal{S})$, which is conditioned on detectability.

Clearly, the BH mass $m_{s,1}$ is one of the dominant factors that change the SNR. As a result, we must reverse the action of marginalization to obtain an expression that depends on $m_{s,1}$. As $m_{o,1}$ is also measured in the event level MCMC chains, the likelihood would thus explicitly depend on $m_{o,1}$. Applying the Bayesian theorem again, we can express the selection-biased model likelihood as

$$p(d_L, m_{o,1}, m_{o,2}, \theta|H_0, \mathcal{S}) = \frac{1}{\mathcal{Z}(H_0, \mathcal{S})} \left[p(\mathcal{S}|m_{o,1}, m_{o,2}, d_L; H_0) \times p(m_{o,1}, m_{o,2}, d_L|H_0) \right]. \quad (11)$$

As usual, the factor $\mathcal{Z}(H_0, \mathcal{S})$ is the normalization of the likelihood, which is also known as the Bayesian evidence. In this case, the calculation of $\mathcal{Z}(H_0, \mathcal{S})$ involves a non-trivial triple integral to be evaluated numerically on regular parameter grids. The calculation of the model term $p(m_{o,1}, m_{o,2}, d_L|H_0)$ has been outlined above, particularly in equation (10). The additional dependence of $m_{o,1}$ can be handled using the same approach for $m_{o,2}$ except for replacing the NS mass function $\pi(m_{s,2})$ prior with the BH mass function prior $\pi(m_{s,1})$.

The computation of $p(\mathcal{S}|m_{o,1}, m_{o,2}, d_L; H_0)$ requires expensive numerical simulation to project the simulated signal onto the interferometer. Fortunately, the readily available software library GWDET (Gerosa 2017) provides an efficient calculation of $p(\mathcal{S}|m_{s,1}, m_{s,2}, d_L; H_0 = H_0^{(P15)})$ (i.e. in terms of source frame quantities) based on a library of pre-computed waveforms. We describe how we approximate $p(\mathcal{S}|m_{o,1}, m_{o,2}, d_L; H_0)$ from GWDET in Appendix. B.

After correcting for selection bias as briefly outlined above, we first examine the strength of its effect relative to the absence of any selection bias. In Fig. 3, we compare the expected NS mass function with and without correcting for selection bias. The effect of selection bias can be seen to increase with distance as a higher NS mass is necessary to produce sufficient SNR for event detection. Among the overlapped curves, the curve peaked at the leftmost shows

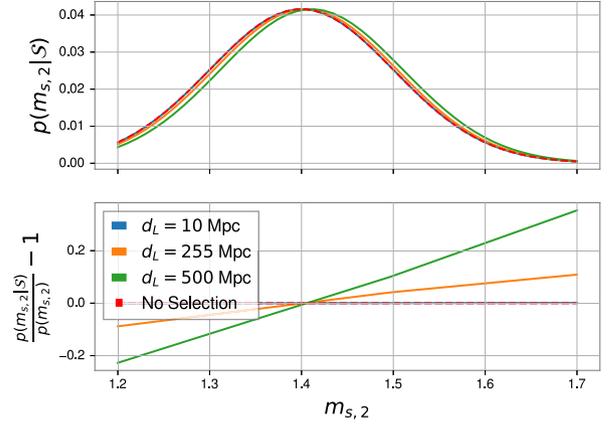


Figure 3. Demonstration of the effect of selection bias. In the upper panel, we show the likelihood of finding mass $m_{s,2}$ at various observed distance d_L . In the lower panel, we show the fractional difference between the mass function with selection bias to the one without the bias (i.e. the mass function evaluated at $d_L = 0$).

the intrinsic NS mass function assumed in this comparison. It is clear that for events in the nearby Universe the observed NS mass function is not noticeably biased, and this is still the case even if we consider events out to the detection limits of approximately $d_L \gtrsim 500$ Mpc, despite a slight potential skew towards the higher masses. Quantitatively, the peak NS mass is shifted only by $0.02 M_\odot$, which corresponds to $\sim 0.2-0.4\sigma_{\text{NS}}$.

While this effect is not important for a few NSBH events as the current detection horizon of NSBH events is approximately $d_L \approx 800$ Mpc, note importantly that in the H_0 likelihood (equation 8), the selection bias multiplies together, and, therefore, the correction of the bias scales with the number of detected events. Therefore, we decided to include this effect anyway in all the subsequent analyses for completeness. However, the uncertainties from the event-level parameter recovery still dominate the total error budget, so the correction from selection bias is considered subdominant.

4.4 Results: GW190426 and GW200115

With the current limited number of observed NSBH events, inference of H_0 suffers from significant sampling effects, as the NS mass of these events may not fairly represent the underlying NS mass function.

The purpose of this current analysis should be understood as a proof of concept, rather than a rigorous inference of H_0 from the existing GW data. In particular, we focus on answering whether, with the current SNR, it is yet possible to deliver any meaningful constraint on H_0 . The main uncertainty currently, as will become clear below, is the distance estimate, d_L , derived from the event amplitude, rather than the observed mass $\mathcal{M}_c(1+z)$ for which the precision is relatively good for the current NSBH detections. In Fig. 5, we show the d_L versus $(1+z)m_2$ posterior for the NSBH events observed so far, with the fiducial model assuming different value of H_0 overlaid on the posteriors. As can be seen, the H_0 constraint is substantially worsened by the intrinsic scatter of the NS mass function, thus the one-to-one correspondence between z and d_L at a fixed H_0 are mapped into a band covering a range of possible $m_{s,2}(1+z)$ values. Focusing on the shaded bands, the bands for different H_0 are increasingly separated at larger distances, implying that models of different H_0 are increasingly distinguishable. Considering *both* the fact that NSBH events can be detected at larger

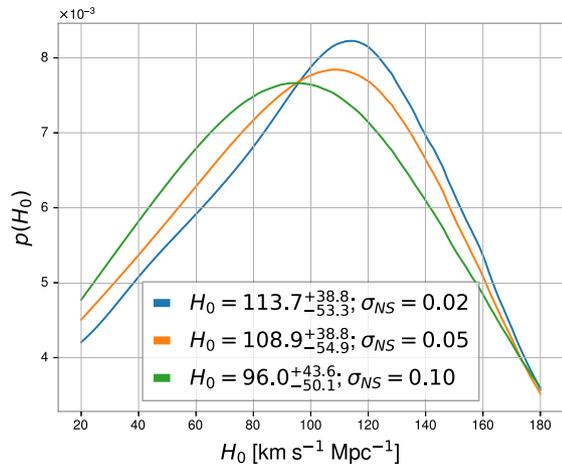


Figure 4. H_0 constraint from the combined constraint of GW190426 and GW200115. H_0 values shown in the legend are in the unit of $\text{km s}^{-1} \text{Mpc}^{-1}$, with the quoted uncertainty as the 68 per cent highest density interval. Note the posterior maxima shift as a function of σ_{NS} and such effect comes from selection bias correction.

distances and $m_{0,2}$ can be measured with higher precision due to the more extreme mass ratio than BNS, NSBH is more preferred over BNS system. In the plot, we also show the posterior of GW190917 in a dotted contour, which has an unusually high $m_{s,2}(1+z)$ and is by now the most distant event with the highest mass ratio observed so far.

Using the more secure events as argued in Section 3, namely GW190426.152155 and GW200115.042309, we perform the fitting procedure as outlined above to determine the recent constraint in H_0 . As the scatter in the NS mass function is still in debate, we repeat the analysis assuming three characteristic values of variance. A fairly broad uniform prior of $H_0 \in [20, 180] \text{ km s}^{-1} \text{Mpc}^{-1}$ is applied. The result is shown in Fig. 4. As the plot indicates, with the very few NSBH events and the poor event-level parameters recovery precision, the constraint on H_0 is not very strong, but a mild exclusion of $H_0 \gtrsim 140 \text{ km s}^{-1} \text{Mpc}^{-1}$ and $H_0 \lesssim 50 \text{ km s}^{-1} \text{Mpc}^{-1}$ can be noticed. This is in line with the intuition from Fig. 5. where the event posterior contours spread different models assuming different H_0 values. None the less, this demonstrates the applicability of our method. It is promising to notice the method can still deliver some exclusion to H_0 models even with the scarcity of data, which is expected to gradually improve in the near future. A forecast of the improvement in precision and a study of the systematics will follow in Section 5.

5 FORECASTING NEAR FUTURE PRECISION ON H_0

As the observed number of NSBH events is still limited, we now present a forecast to predict the achievable precision on H_0 that can be anticipated in the near future. This is estimated by a dedicated waveform injection pipeline. We perform simulations of NSBH events drawn from an assumed event distribution, and examine how the waveform fitting procedure induces uncertainty in addition to the intrinsic scatter on the NS mass distribution adopted in our prior.

The required modelling for the population of simulated events comprises three distributions, namely: the BH mass function $\pi(m_{s,1})$, the NS mass function $\pi(m_{s,2})$, and the event rate $\pi(z)$. We randomly draw samples from these distributions, and subsequently shifted the

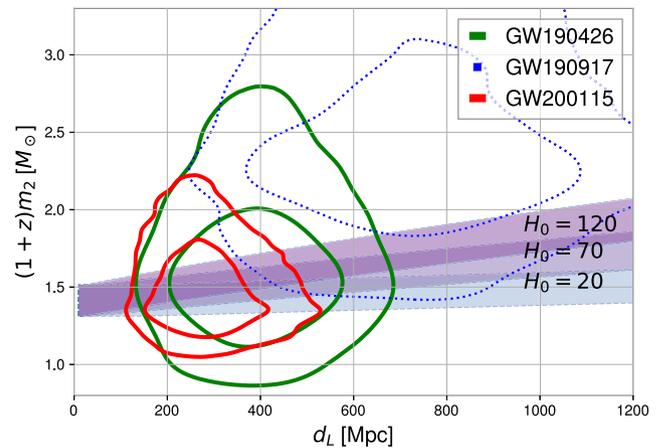


Figure 5. Confidence contour the NSBH events listed in Table 1, overplotted with models that assume different values of H_0 (shaded bands). The y-position of the bands is determined by pinning down the source frame NS mass $m_{s,2} = 1.4 \pm 0.1 M_\odot$, with the boundaries adopting $1.4 + 0.1 M_\odot$ and $1.4 - 0.1 M_\odot$, respectively.

source frame quantity $m_{s,1}$, $m_{s,2}$ to the observed frame quantity $m_{o,1}$ and $m_{o,2}$. The $\{m_{o,1}, m_{o,2}, z\}$ samples are then filtered to rule out events that generate weak signals undetectable before any waveform evaluation to cut down the computational cost. Afterward, we calculate the waveform parametrized only by *observed frame quantities* $\{m_{o,1}, m_{o,2}, d_L\}$ using the publicly available package BIBLY (Ashton et al. 2019; Romero-Shaw et al. 2020; Ashton & Talbot 2021), which is the default analysis package for LVK (Abbott et al. 2021c) to analysis NSBH event. In particular, we use the approximant *IMRPhenomXPHM* to better handle the extreme mass ratio for NSBH events. We then project the waveform to the detectors LIGO L1, H1, and Virgo, and inject noise realizations to the projected waveform from their corresponding power spectral density. These procedures create simulated observations, namely time series of strain variations among each of the detectors.

These time series are then passed to our fitting pipeline. Our goal is to obtain the H_0 posterior using the simulated time series via equation (6). However, as will be explained in the next subsection, we would leave flexibility in choosing the presumed population parameters characterizing $\pi(m_{s,2})$. Therefore, the time series would first be fitted assuming a uniform prior in $m_{s,2}$ and z via the MCMC sampler DYNesty (Speagle 2020; Ashton & Talbot 2021). We have the freedom to enforce different mass priors $\pi(m_{s,2})$ when the MCMC samples are subsequently used to calculate H_0 posterior, by reweighing the MCMC samples as explained in equation (8). This approach effectively reduces the undesirable repetitions in MCMC sampling when we switch between priors, and thus cut the computational cost down to an affordable amount. In the fitting procedure, we use the same likelihood function (Pankow et al. 2015) as adopted in LVK analysis (Abbott et al. 2023b). It is numerically implemented in BIBLY, with the phase parameter being analytically marginalized over as demonstrated in Pankow et al. (2015).

5.1 Event level uncertainty

While the event level parameter recovery and the corresponding uncertainty estimates from noisy signal have been extensively demonstrated by LVK (Abbott et al. 2023b), most of the existing data are on BBH signals. For our application on NSBH, we would

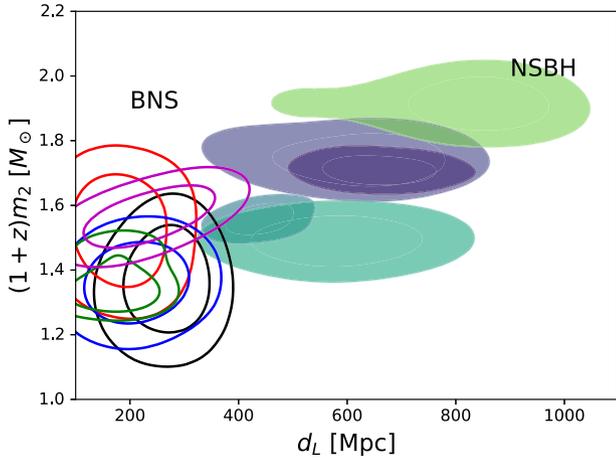


Figure 6. The posterior constraint on five simulated BNS (unfilled contours) and NSBH events (filled contours), respectively. Clearly, despite the uncertainty in d_L is slight larger for NSBH, the uncertainty in $(1+z)m_2$ is substantially reduced when compared to BNS event. Therefore, a better precision on event redshift z is archived by using NSBH event.

like to focus more on the parameter degeneracy related to z and d_L , and how the correlations vary as switching between different injected event parameters, thus elucidating the limit on H_0 precision due to parameter degeneracy.

5.1.1 Parameter degeneracy

In Section 2, our estimation shows the constraint on z and d_L are inferred from the two (almost independent) observables, namely the chirping spectrum (constrain q and thus $(1+z)m_2$) and the signal amplitude, respectively. It is therefore natural to investigate the shape of the posterior given the waveform data: whether the z and d_L are minimally correlated. In Fig. 6, we plotted the confidence contours on the d_L and $(1+z)m_2$ plane. If the two parameters are inferred from independent observables, the semimajor and semiminor axis of the contours should be parallel to the d_L -axis and $(1+z)m_2$ -axis, respectively.⁶ Indeed, our intuition is valid. Because of the absence of a significant correlation between $(1+z)m_2$ and d_L , we can determine z once m_2 is constrained in the NS mass function defined in the prior. The independent measurement of z and d_L allows one to determine H_0 via the classic Hubble diagram.

5.1.2 BNS versus NSBH

In particular, we would like to understand the advantage of using NSBH over the BNS event in a deeper detail. As argued previously in Section 2 due to the efficient amplification powered by the BH companion, NSBH provides the ≈ 4 enlargement of the detection horizon, thus providing improved constraints on H_0 .

Recall the precision on constraining H_0 also depends on the estimation uncertainty in z . Apart from the extension of the detection horizon, NSBH also provides a tighter estimate on z . In particular, mergers with fairly unequal mass ratio generate stronger signals,

⁶If the quantities are inferred from correlated observables, the quantities would consequently be linear combinations of all observables. Thus, the covariance matrix would possess off-diagonal terms that tilt the contours at some angle to the d_L -axis and $(1+z)m_2$ -axis.

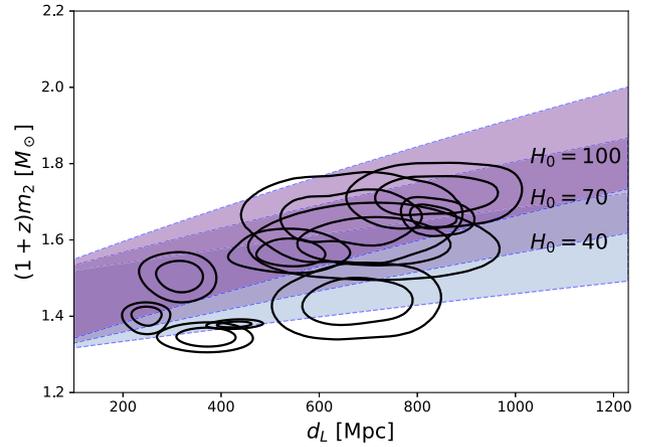


Figure 7. The simulated Hubble diagrams for the 10 injected events, assuming $\sigma_{\text{NS}} = 0.1 M_\odot$.

thus improving the constraints on q . To test this intuition, we inject a population of 10 BNS events with both the component masses sampled from the same NS mass function $\pi(m_{s,2})$, and assign the heavier component as m_1 . Similar to what we did for the NSBH events, we fit the injected waveforms using a flat mass prior for both the masses, but adopting the approximant *IMRPhenomPv2-NRTidal* for BNS. The information of the NS mass function is used only when inferring the H_0 posterior (equation 9). Such a choice would allow a fair comparison among the two types of sources.

The recovered posterior on d_L and $(1+z)m_2$ for each event is plotted in Fig. 6. As can be seen, the detectable BNS mostly fall in the more nearby Universe, and the associated estimation of $(1+z)m_2$ is more uncertain. On the contrary, most NSBH injected are originated at a larger distance as the sampling volume increases. However, the estimation uncertainty in d_L does not worsen significantly. Furthermore, there is a notable improvement in constraining $(1+z)m_2$ and thus z . This demonstrates the advantages of using NSBH over BNS.

5.2 Population level uncertainty

Another limit on precision determination of H_0 comes from (1) the intrinsic scatter of the source population, and (2) the finite sampling from the source distribution.

5.2.1 Sensitivity to source distribution

To examine effect (1), we generate events from $\pi(m_{s,2}; \sigma_{\text{NS}})$ assuming different levels of intrinsic scatter σ_{NS} , and feed the simulated data to our fitting pipeline to obtain the posterior $p(H_0|D)$ following equation (8). Note that we do not enforce to take the same intrinsic scatter σ_{NS} in the H_0 fitting procedure as in the simulation procedure, hence allowing tests of the sensitivity against the incorrect characterization of σ_{NS} .

We injected a population of NSBH with $\langle m_{\text{NS}} \rangle = 1.4 M_\odot$ and $\sigma_{\text{NS}}^{(\text{sim})} = 0.1 M_\odot$; the large intrinsic scatter of $\sigma_{\text{NS}}^{(\text{sim})} = 0.1 M_\odot$ is probably a conservative estimation. We sampled 10 events from the simulation, and determine the H_0 posterior assuming $\sigma_{\text{NS}}^{(\text{fit})} \in \{0.02, 0.05, 0.1, 0.2\} M_\odot$. The Hubble diagram of these simulated events is shown in Fig. 7.

It is clear from Fig. 8 that an underestimated $\sigma_{\text{N}}^{(\text{fit})} < \sigma_{\text{N}}^{(\text{sim})}$ would not only lead to the incorrect shrunk of error, but can in principle, substantially shift the mean $\langle H_0 \rangle$ to an incorrect value, depending

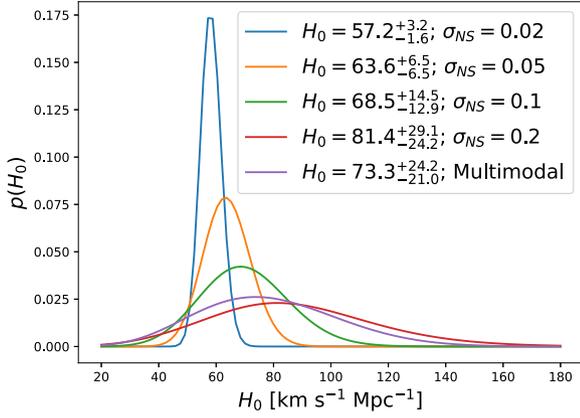


Figure 8. The constraint in H_0 when 10 simulated NSBH events are observed, assuming $\sigma_N^{(\text{sim})} = 0.10 M_\odot$. The multimodal NS mass function (marked as NSMF on the plot) assumes the two peaks at $m_2 = 1.3$ and $1.5 M_\odot$, respectively, with an equal dispersion on each peak $\sigma = 0.1 M_\odot$ and equal relative abundance.

on the sampled mean NS mass $\langle m_{\text{NS}} \rangle^{(\text{fit})}$ in contrast to the simulated mean NS mass $\langle m_{\text{NS}} \rangle^{(\text{sim})}$. An overestimate of $\sigma_{\text{NS}}^{(\text{fit})} > \sigma_{\text{NS}}^{(\text{sim})}$ would exaggerate the uncertainty in H_0 .

5.2.2 Source distribution: multimodal

It has been reported that the NS mass distribution is more complicated than a simple Gaussian: instead of a peak around the Chandrasekhar limit $m_{s,2} \sim 1.4 M_\odot$, the NS mass function has (at least) two neighbouring peaks at $m_{s,2} \approx 1.37$ and $1.57 M_\odot$ (Zhang et al. 2011) despite the exact location differs in different literatures, depending on the selection of NS samples (see a summary in Section 3, or from Horvath & Valentim 2017 and reference therein).

As can be seen in Fig. 8 (particularly the skewed, second-lowest peaked curve), the presence of multiple peaks in the NS mass function would in general skew the distribution towards the direction that corresponds to the less dominant peak. Having said that, the mean value of H_0 can still be recovered. If the peaks in the intrinsic NS mass function are not well separated due to intrinsic scatter, the corresponding H_0 posterior would also hide the peaks due to the propagated uncertainty in H_0 .

The reasons for the striking similarity between the multimodal mass function and the simple Gaussian distribution can be understood as follows. The mean of a two Gaussians mixture $\Gamma \mathcal{G}(\mu_1, \sigma_1) + (1 - \Gamma) \mathcal{G}(\mu_2, \sigma_2)$, with a relative abundance of $\Gamma \in [0, 1]$, is $\bar{\mu} = \Gamma \mu_1 + (1 - \Gamma) \mu_2$. Similarly, the associated variance is $\bar{\sigma}^2 = \Gamma(\sigma_1^2 + \mu_1^2) + (1 - \Gamma)(\sigma_2^2 + \mu_2^2) - \bar{\mu}^2$. We can statistically assess whether it is possible to significantly distinguish a Gaussian mixture with an effective Gaussian $\mathcal{G}(\bar{\mu}, \bar{\sigma})$ via a two-sample Kolmogorov–Smirnov (KS) test. In particular, we focus on the limit that the number of samples from those distributions is finite, so as to mimic the fact that we would observe limited number of NSBH events. As a visual comparison, we plot a few multimodal mass functions with benchmark parameters chosen to mimic some of the claimed forms in Horvath & Valentim (2017). The results are shown in Fig. 9. While the appearance of the effective Gaussian description may look very different in terms of the probability distribution function, they look fairly similar in terms of the cumulative distribution function (CDF) – based on which the KS test is built. In particular, the difference is the most noticeable when $\Delta \equiv (\mu_2 - \mu_1) / \min(\sigma_1, \sigma_2) \gg 1$, so that the two peaks can be ‘resolved’.

For realistic constraints from the Milky Way observation of NS binaries, $\Delta < 1$, and thus the two descriptions are statistically indistinguishable. This conservative assessment did not take into account the substantial GW measurement uncertainties in the redshifted NS masses $m_{s,2}$. One would therefore expect the measurement uncertainties would further diminish the difference between the multimodal distribution and the effective Gaussian. This result is indeed expected: in the effective Gaussian formalism, we force the corresponding mean and variance (the two leading statistical moments) to be consistent with the multimodal distribution. Thus, any distinguishable feature would only contribute from the third-order moment onwards, where the amplitude is suppressed at least by $[(x - \bar{\mu}) / \bar{\sigma}]^3 \ll 1^3$.

This demonstrates our method is insensitive to the fine detail of the underlying NS mass function, and is therefore robust to the selection of the samples from the electromagnetically observed NSs. More precisely, the fine details would not be possible to significantly bias the analysis – with the consideration of the limited number of observed NSBH events foreseen in the near future. In the more distant future with both upgraded sensitivity of GW observatories and improved EM-based measurement of the NS mass function, the fine details would eventually become relevant for the precision determination of H_0 .

5.2.3 Dependence on number of events

Intuitively when the events are completely independent, the uncertainty in H_0 is reduced with increased numbers of events: $\delta^{(N)} H_0 = (1/\sqrt{N}) \delta^{(1)} H_0$. However, we emphasize that the events are not independent despite the fitting procedures that do not rely on the knowledge of other events. The data generation process links the $m_{s,2}$ of each event together via $\pi(m_{s,2})$: the presence of a subset with $m_{s,2} < \langle m_{\text{NS}} \rangle$ implies the likely presence of another subset with $m_{s,2} > \langle m_{\text{NS}} \rangle$, so that the population mean $\langle m_{\text{NS}} \rangle$ is probabilistically restored. The presence of this correlation makes the simple $1/\sqrt{N}$ scaling too optimistic.

To study the impact of those possible hidden correlations, we reuse the simulated events and evaluate the posterior equation (8) by supplying randomly sampled subsets of simulated events, with each subset containing a different number of events. The results are shown in Fig. 10. It is clear that the improvement is slightly slower than $1/\sqrt{N}$, yet the improvement is still significant, so that, for example, the accumulation of 10 events successfully shrinks the error to 1/2 of that from a single event, meanwhile driving the posterior mean $\langle H_0 \rangle$ to match the injected value.

5.3 Discussion

As we have shown above, shrinking down the H_0 uncertainty innate to this method requires more detected NSBH events, which would be achievable in the near future. Apart from waiting for more observed NSBH GW events, other efforts on EM-based observation of the Milky Way NS systems could also help improve the H_0 constraint. One dominant uncertainty is inherited from the estimation of z from the GW data, which requires the external, ‘calibration’ input from the NSs observed in the Milky Way. As explained in Section 2 (and the associated schematic Fig. 1), once the NS mass $m_{s,2}$ is determined, it could be combined with the spectral evolution of the GW events to determine the mass ratio q and thus the source frame chirp mass \mathcal{M}_c . Such source frame chirp mass \mathcal{M}_c can then be compared with the redshifted chirp mass $\mathcal{M}_c(1+z)$ to obtain an estimate of z . On the

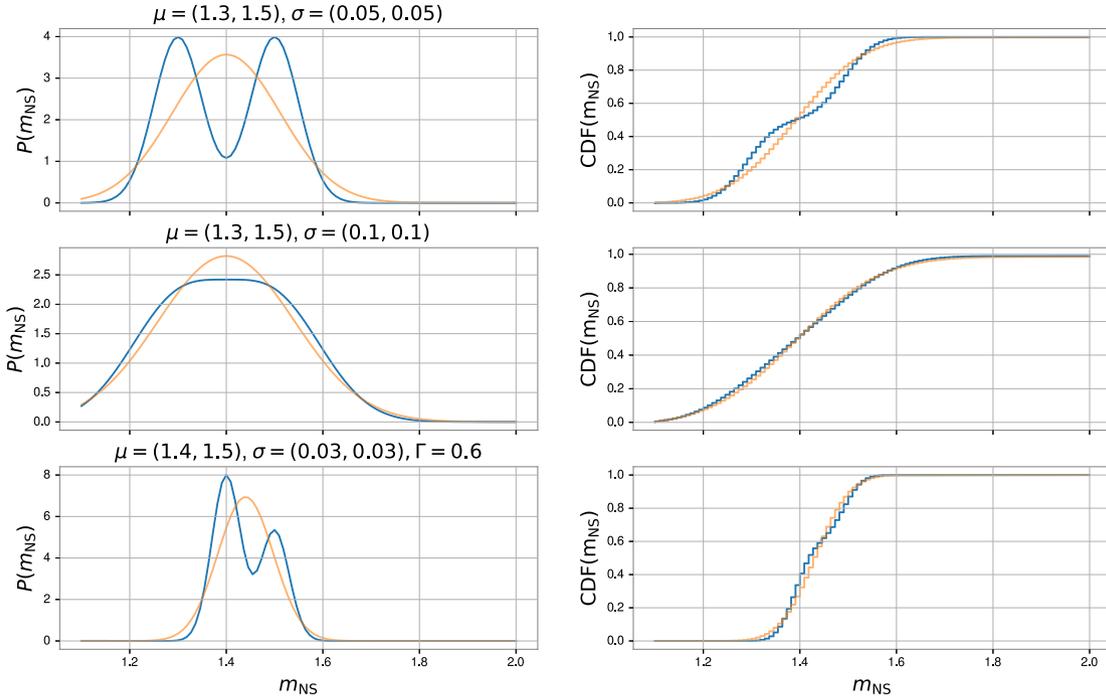


Figure 9. We show the significance of the possible multimodal feature as reported from some EM observation, and the approximation error for describing the multimodal distribution by a single effective Gaussian. Three sets of multimodal parameters are shown, with the corresponding effective Gaussian plotted as the pale, bell-shaped curve. The multimodal distribution parameters are described in the plot title. On the right column, we show the same comparison, but in terms of the CDF, which are traditionally used to define the distinguishability between distributions (via a KS test). As can be seen, in terms of the CDF, the effective Gaussian description is highly indifferent from the full, multimodal description as long as the multimodal peaks are not largely separated.

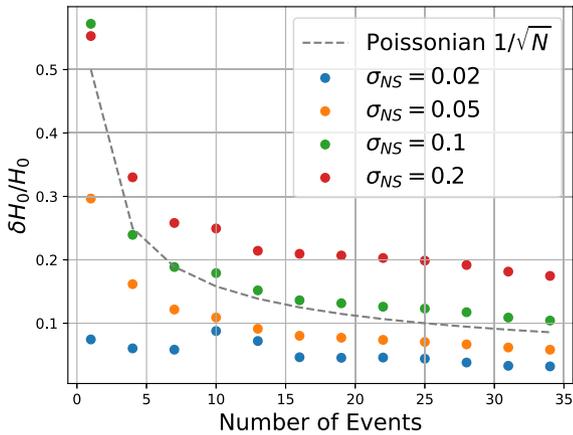


Figure 10. The shrinkage forecast in fractional error in H_0 as a function of a number of observed events. It is clear that the shrink in uncertainty improves slightly slower than the Poissonian decay, as the events are not totally independent: they share the calibration parameters (i.e. intrinsic properties of the mass function) in our fitting process.

other hand, there is an independent method to construct a data-driven external prior on the mass ratio q as well. With an external estimate of q , the uncertainty in the GW spectral evolution constraint on q would be significantly reduced.

The EM measurement of the mass ratio of the binary system involving an NS is not generally possible, unless at least one star of the binary system is a pulsar (or other time transients). This could be understood as follows: astrometry measurements are able to measure the Keplerian parameters, which constrain the component masses up

to an unknown orbit inclination. Such parameter degeneracy can be removed by a companion pulsar – exploiting the relativistic time delay when the pulsar signal is deflected by the gravitational field (see e.g. Lattimer 2012 for a review). Indeed, in the NS mass function, we picked as demonstration (Valentim et al. 2011; Kiziltan et al. 2013; Horvath & Valentim 2017), the NS mass is obtained from system that hosts at least a pulsar. The measurements of q for each system allow the construction of the distribution of q . Such external constraint distribution can then be input as a prior to our analysis pipeline.

Another subtle aspect of the prior on q is to handle its potential correlation with the mass scale m_{NS} of the corresponding binary system. This is theoretically expected as the formation mechanism of NS with viable mass can be different – possibly including tidal interaction within the binary system where q would be an important control parameter. Owing to the limited amount of available of EM-detected NS systems, such an effect is still unclear; it is also for this reason we stay conservative in this analysis and do not model such effect. However, this should be of practical interest in the near future with the consortium of EM observation and GW observation to lift up such degeneracy, thus helping tighten the constraint of H_0 .

A final remark on the application of the BNS system versus the NSBH system as discussed in this work. As the NS mass function has been fixed in a priori, in principle the expected chirp mass \mathcal{M}_c can also be fixed by convolving the NS mass function with itself. Such a process naturally leads to a prior directly on the chirp mass \mathcal{M}_c for the BNS system. This is different from the NSBH approach we discussed here – where the primary mass $m_{s,1}$ (BH mass) and equivalently the mass ratio q has to be constrained by the data.

It should be now clear that the BNS approach would be strongly influenced by the presumption on the q - m_{NS} correlation, so that the expected chirp mass function is no longer a simple convolution of the NS mass function with itself. That is also one reason why NSBH is favoured at the current stage, where the dependence on q is solely addressed by the GW waveform, without exploiting any unclear presumption. Furthermore, the measurement precision for q in the NSBH merger is better than is in the BNS merger, thanks to the long-lasting waveform during the NSBH merger event. However, the synergy of BNS and NSBH to deliver a joint constraint on H_0 is still foreseen in the future upon the rapid accumulation of both EM and GW data.

6 CONCLUSIONS

We have demonstrated here that NSBH events provide an independent and potentially competitive means of estimating H_0 by relying on the characteristic Chandrasekhar mass of NSs – using only the GW data, and without the need for a supporting EM redshift. Thanks to the population of NS masses measured to lie in a confined mass range $m_{\text{NS}} \approx 1.4 \pm 0.1 M_{\odot}$, provides a simple reasonably well-understood ‘standard siren’ for interpreting the GW waveforms of the GW events of this class. At least one NSBH event is now undisputed, namely GW200115 with a redshifted NS mass of $1.48 M_{\odot}$ determined in the observer frame from the LVK waveform analysis and with distances estimated to be $283^{+62.1}_{-82.7}$ Mpc using the measured strain amplitude.

Reconsideration of other GW events recorded prior to GW200115 may now be redefined with an NSBH classification, namely event GW190426 and possibly also GW190917 (for which a less conservative ‘prior’ on the rate of NSBH merging is forced by the clear detection of GW200115).

None the less, we have shown that it is possible to obtain a consistent constraint in H_0 from the existing NSBH events, albeit with larger distance uncertainty due to relatively low SNR in the detections.

We may look forward to upgraded sensitivity that would allow NSBH events with higher SNR, i.e. better-defined waveforms for which the degeneracy with orbital/spin parameters can be much improved allowing significantly more precise distance estimation and hence better defined H_0 that will be defined from a joint analysis of a larger sample of future NSBH events. We have emphasized that even with the current sensitivity, with only 10 more useful NSBH events, we can anticipate $\delta H_0/H_0 \simeq 20$ per cent. In particular, this uncertainty is inherently different from the precisely determined H_0 from supernovae surveys: GW-based methods are based on a vastly different set of physical processes, thus allowing for comparison and validation of the assumptions made in supernovae-based measurements.

We have emphasized that this NSBH-based method for determining H_0 is unlike that of multimessengers BNS events where an external redshift is required for a precision measurement of H_0 . In fact, only one BNS case has been detected this way (GW170817) rather fortuitously via gamma-ray flare time coincidence and the unusually close proximity of this BNS event. For NSBH we have the advantage of a much larger detection horizon thanks to the presence of the associated massive BH that enhances the chirp mass and hence the detectable SNR. Of course for NSBH events with associated EM emission whereby an independent redshift can be established, the precision on H_0 is much improved. Thus it should be regarded as a priority to be prepared for prompt follow-up of NSBH candidate GW events, but nevertheless, this is not a requirement of our method,

for which GW alone is sufficient for competitive precision on H_0 at current sensitivities.

ACKNOWLEDGEMENTS

We thank Stephen Taylor for the useful discussion of the approaches based on BNSs and the historical remarks about the topic. We also thank the editor and the anonymous reviewer of the paper, who helped clarify the presentation of the work and provided detailed insightful feedback.

This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gwosc.org), a service of the LIGO Scientific Collaboration, the Virgo Collaboration, and Kagra. This paper is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation, as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN), and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, and Spain. KAGRA is supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation of Korea (NRF) and Ministry of Science and ICT (MSIT) in Korea; and Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan.

DATA AVAILABILITY

The GW data used in this work are provided by the LIGO–Virgo–Kagra Collaboration (Abbott et al. 2023b) on the web portal (gwosc.org). Additional data derived in this work are largely based on the aforementioned data and are available upon request.

REFERENCES

- Abbott B. P. et al., 2017a, *Phys. Rev. Lett.*, 119, 161101
 Abbott B. P. et al., 2017b, *Nature*, 551, 85
 Abbott B. P. et al., 2021a, *ApJ*, 909, 218
 Abbott R. et al., 2021b, *ApJ*, 913, L7
 Abbott R. et al., 2021c, *ApJ*, 915, L5
 Abbott R. et al., 2023a, *ApJ*, 949, 76
 Abbott R. et al., 2023b, *Phys. Rev. X*, 13, 041039
 Abbott R. et al., 2024, *Phys. Rev. D*, 109, 022001
 Ajith P. et al., 2008, *Phys. Rev. D*, 77, 104017
 Ashton G., Talbot C., 2021, *MNRAS*, 507, 2037
 Ashton G. et al., 2019, *ApJS*, 241, 27
 Aubin F. et al., 2021, *Class. Quantum Gravity*, 38, 095004
 Bassett B. A., Kunz M., 2004, *Phys. Rev. D*, 69, 101305
 Broadhurst T., Diego J. M., Smoot G. F., 2020, preprint ([arXiv:2006.13219](https://arxiv.org/abs/2006.13219))
 Buikema A. et al., 2020, *Phys. Rev. D*, 102, 062003
 Camarena D., Marra V., 2021, *MNRAS*, 504, 5164
 Cervantes-Cota J. L., Galindo-Uribarri S., Smoot G. F., 2023, *Universe*, 9, 501
 Chatterjee D., Hegade K. R. A., Holder G., Holz D. E., Perkins S., Yagi K., Yunes N., 2021, *Phys. Rev. D*, 104, 083528
 Chernoff D. F., Finn L. S., 1993, *ApJ*, 411, L5
 Cigarrán Díaz C., Mukherjee S., 2022, *MNRAS*, 511, 2782

- Ding Q., 2024, *Phys. Rev. D*, 110, 063542
- Dominik M. et al., 2015, *ApJ*, 806, 263
- Etherington I., 1933, *London Edinburgh Dublin Philos. Mag. & J. Sci.*, 15, 761
- Ezquiaga J. M., Holz D. E., 2022, *Phys. Rev. Lett.*, 129, 061102
- Farr W. M., Fishbach M., Ye J., Holz D. E., 2019, *ApJ*, 883, L42
- Finke A., Foffa S., Iacovelli F., Maggiore M., Mancarella M., 2021, *J. Cosmol. Astropart. Phys.*, 08, 026
- Freedman W. L., 2021, *ApJ*, 919, 16
- Freedman W. L. et al., 2019, *ApJ*, 882, 34
- Freedman W. L. et al., 2020, *ApJ*, 891, 57
- Fruchter A. S., Stinebring D. R., Taylor J. H., 1988, *Nature*, 333, 237
- Fung L. W., Li L., Liu T., Luu H. N., Qiu Y.-C., Tye S.-H. H., 2021, *J. Cosmol. Astropart. Phys.*, 08, 057
- Fung L. W. H., Li L., Liu T., Luu H. N., Qiu Y.-C., Tye S. H. H., 2023, *Phys. Rev. Res.*, 5, L022059
- Gerosa D., 2017, *dgerosa/gwdet: v0.1*. Zenodo. Available at: <https://doi.org/10.5281/zenodo.889966>
- Ghosh T., Biswas B., Bose S., 2022, *Phys. Rev. D*, 106, 123529
- Gray R. et al., 2023, *J. Cosmol. Astropart. Phys.*, 12, 023
- Hogg D. W., Myers A. D., Bovy J., 2010, *ApJ*, 725, 2166
- Horvath J. E., Valentim R., 2017, in Alsabti A. W., Murdin P., eds, *Handbook of Supernovae*. Springer, Cham, Switzerland, p. 1317
- Ivanov M. M., Ali-Haïmoud Y., Lesgourgues J., 2020, *Phys. Rev. D*, 102, 063515
- Jedamzik K., Pogosian L., Zhao G.-B., 2021, *Commun. Phys.*, 4, 123
- Kandel D., Romani R. W., 2022, *ApJ*, 942, 6
- Kiziltan B., Kottas A., De Yoreo M. D., Thorsett S. E., 2013, *ApJ*, 778, 66
- Knox L., Millea M., 2020, *Phys. Rev. D*, 101, 043533
- Krishnan C., Mohayaee R., Colgáin E. Ó., Sheikh-Jabbari M. M., Yin L., 2021a, *Class. Quantum Gravity*, 38, 184001
- Krishnan C., Colgáin E. Ó., Sheikh-Jabbari M., Yang T., 2021b, *Phys. Rev. D*, 103, 103509
- Lattimer J. M., 2012, *Annu. Rev. Nucl. Part. Sci.*, 62, 485
- Li Y.-J., Han M.-Z., Tang S.-P., Wang Y.-Z., Hu Y.-M., Yuan Q., Fan Y.-Z., Wei D.-M., 2020, preprint ([arXiv:2012.04978](https://arxiv.org/abs/2012.04978))
- Luu H. N., 2023, *Phys. Rev. D*, 107, 023513
- Mandel I., Farr W. M., Gair J. R., 2019, *MNRAS*, 486, 1086
- Marković D., 1993, *Phys. Rev. D*, 48, 4738
- Mastrogiovanni S. et al., 2021, *Phys. Rev. D*, 104, 062009
- Messenger C., Veitch J., 2013, *New J. Phys.*, 15, 053027
- Mukherjee S., Wandelt B. D., Nissanke S. M., Silvestri A., 2021, *Phys. Rev. D*, 103, 043520
- Nissanke S., Holz D. E., Dalal N., Hughes S. A., Sievers J. L., Hirata C. M., 2013, preprint ([arXiv:1307.2638](https://arxiv.org/abs/1307.2638))
- Oguri M., 2016, *Phys. Rev. D*, 93, 083511
- Pankow C., Brady P., Ochsner E., O’Shaughnessy R., 2015, *Phys. Rev. D*, 92, 023002
- Planck Collaboration VI, 2020, *A&A*, 641, A6
- Poisson E., Will C. M., 1995, *Phys. Rev. D*, 52, 848
- Riess A. G. et al., 2022, *ApJ*, 934, L7
- Romani R. W., Filippenko A. V., Silverman J. M., Cenko S. B., Greiner J., Rau A., Elliott J., Plutsch H. J., 2012, *ApJ*, 760, L36
- Romero-Shaw I. M. et al., 2020, *MNRAS*, 499, 3295
- Schutz B. F., 1986, *Nature*, 323, 310
- Scolnic D. M. et al., 2018, *ApJ*, 859, 101
- Speagle J. S., 2020, *MNRAS*, 493, 3132
- Taylor S. R., Gair J. R., 2012, *Phys. Rev. D*, 86, 023502
- Taylor S. R., Gair J. R., Mandel I., 2012, *Phys. Rev. D*, 85, 023535
- Thorsett S. E., Chakrabarty D., 1999, *ApJ*, 512, 288
- Thrane E., Talbot C., 2019, *Publ. Astron. Soc. Aust.*, 36, e010
- Valentim R., Rangel E., Horvath J. E., 2011, *MNRAS*, 414, 1427
- Valentino E. D. et al., 2021, *Class. Quantum Gravity*, 38, 153001
- Verde L., Treu T., Riess A. G., 2019, *Nat. Astron.*, 3, 891
- Veske D., Bartos I., Márka Z., Márka S., 2021, *ApJ*, 922, 258
- Zhang C. M. et al., 2011, *A&A*, 527, A83

APPENDIX A: REMARKS ON COSMOLOGY DEPENDENCE OF LVK MCMC CHAINS

We note that the PYCBC and BILBY GW parameters inference package used the accelerated method outlined in Thrane & Talbot (2019) and Ashton & Talbot (2021) to compare between model templates and the data, and thus evaluating the likelihood. As an illustrative example (Pankow et al. 2015), this class of methods relies on the analytical marginalization of extrinsic parameters (in particular, d_L) while only numerically samples the intrinsic parameters (\mathcal{M}_c, θ). As the effect of extrinsic parameters is shiftings and scalings of the waveform – in contrast to more complicated dependence involved for intrinsic parameters – analytical marginalization over extrinsic parameters is allowed. These marginalized likelihoods are then used to reconstruct the extrinsic parameters, particularly for d_L following the procedure outlined in appendix C of Thrane & Talbot (2019).

Consequently, in the LVK pipeline, a model of the GW waveform radiated by compact binary coalescence is first created using the intrinsic parameters, including \mathcal{M}_c . The extrinsic effect of d_L is reconstructed afterwards, and such d_L information is converted into redshift z by assuming the nominal value of *Planck 2018* cosmology (Planck Collaboration VI 2020) with $H_0 = 67.9 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Using this redshift derived from d_L , the source frame chirp mass \mathcal{M}_c posterior samples are reinterpreted by rescaling it with $1/(1+z)$. As a result, the source frame masses provided in the MCMC chains provided in the LVK data releases have all assumed the *Planck 2018* cosmology implicitly, and this therefore *forbids the use of the source frame masses and redshifts at their face values in the MCMC chains (for the purpose of cosmography)*. We have verified that the d_L samples in the LVK chains that we used in this paper have a perfect correlation with the z samples provided in these chains with zero scatter (up to floating point precision), and the conversion rules for converting d_L samples into z samples are indeed the distance–redshift relation implied by *Planck 2018* cosmology.

For this reason, in our analysis demonstrated in Section 4, we avoid the use of any source frame quantities provided in the MCMC chains and stick with the observer frame quantities that are cosmology independent.

APPENDIX B: TRANSFORMATION OF SELECTION LIKELIHOOD

A crucial step for our selection-corrected model likelihood is to evaluate the selection bias term, namely $p(S|m_{0,1}, m_{0,2}, d_L; H_0)$. To do this, we use an approximation method built upon the software GWDET (Gerosa 2017). Before proceeding, we quickly review the method implemented in GWDET. In GWDET, the detection completeness – as a function of source frame component masses and distance – is evaluated over millions of parameter grids. The completeness is evaluated via the method outlined in Dominik et al. (2015) as follows. Given a parameter combination, the optimal SNR of the event ρ_{opt} – optimizing the inclination and alignment with the GW interferometer (i.e. extrinsic parameters) – is computed via numerical relativity. Following this, a ratio $w \equiv 8/\rho_{\text{opt}}$ is computed, where 8 is the SNR threshold of detection. Using analytical modelling of the sensitivity variation among different combinations of extrinsic parameters, upon marginalization the selection completeness can be modelled as a function of the form $p(> w)$, which is monotonically decreasing. As w is a function of $(m_{s,1}, m_{s,2}, d_L)$, we can identify this with the selection probability:

$$p(> w = 8/\rho_{\text{opt}}) = p(S|m_{s,1}, m_{s,2}, d_L; H_0). \quad (\text{B1})$$

In particular, the calculation of ρ_{opt} pre-computed in GWDET assumes *Planck 2015* cosmology. In our application, we would like to implement two extra steps: (1) rewriting the functional dependence from source frame to observer frame; and (2) generalizing this selection probability to arbitrary H_0 by suitable approximation – as a compromise of calculation speed and precision.

The first step – rewriting it in the observer frame – is simple. The selection probability can be rewritten as

$$p(\mathcal{S}|m_{o,1}, m_{o,2}, d_L; H_0) = \int dm_{s,1} dm_{s,2} p(\mathcal{S}|m_{s,1}, m_{s,2}, d_L; H_0) \times p(m_{s,1}, m_{s,2}|m_{o,1}, m_{o,2}, d_L; H_0) \times p(d_L|H_0). \quad (\text{B2})$$

Such integral is efficient to evaluate as $p(m_{s,1}, m_{s,2}|m_{o,1}, m_{o,2}, d_L; H_0)$ is a delta function, relating the frames via $m_{s,i} = m_{o,i}/(1 + z(d_L; H_0))$.

The second step to relate between different cosmology is however more tricky, and here we use an approximate approach to the first order. First of all, we evaluate $p(> w = 8/\rho_{\text{opt}}^{(\text{P15})})$ using GWDET. As $p(> w)$ is monotonic, we can inverse the function to calculate $\rho_{\text{opt}}^{(\text{P15})}$ as a function of $(m_{o,1}, m_{o,2}, d_L)$ efficiently with the output of GWDET. Note that to the first order, ρ_{opt} has the following cosmology-dependent scaling (Ajith et al. 2008):

$$\rho_{\text{opt}} \propto \frac{(1+z)^{5/6}}{d_L}. \quad (\text{B3})$$

Therefore, to transform from *Planck 2015* cosmology to arbitrary cosmology, we can scale ρ_{opt} approximately as

$$\rho_{\text{opt}}(H_0) = \rho_{\text{opt}}^{(\text{P15})} \frac{(1+z(d_L; H_0))^{5/6}}{(1+z(d_L; \text{P15}))^{5/6}}. \quad (\text{B4})$$

The argument of the GWDET selection probability $p(> w = 8/\rho_{\text{opt}})$ is scaled accordingly to get the cosmology-adapted selection probability. Following the above procedure, we yield a numerically highly efficient method to calculate the selection probability, therefore, allowing us to quantify the selection bias involved in the inference of H_0 .

APPENDIX C: FISHER ANALYSIS OF NSBH CONSTRAINTS ON H_0

We would like to estimate if the enlarged detection horizon offered by NSBH events can significantly improve the constraints on H_0 . By significantly, we mean an improvement in the error terms at the leading order, derived using Fisher analysis.

In the following, we present a bottom-up approach. The first step is to evaluate the distance measurement error δd of NSBH events and BNS events. We are particularly interested in studying the marginally detectable events, defined by having their SNR $\rho^* = 8$. The NSBH horizon is larger, $d_{\text{NSBH}} > d_{\text{BNS}}$, and they both give rise to the same SNR ($\rho_{\text{NSBH}} = \rho_{\text{BNS}} = \rho^* = 8$). It is well known that $\rho \propto \hat{d}_L^{-1}$ (Ajith et al. 2008). For later convenience, we denote the proportionality constants for both NSBH and BNS events as α s:

$$\rho^* = \frac{\alpha_{\text{NSBH}}}{d_{\text{NSBH}}} = \frac{\alpha_{\text{BNS}}}{d_{\text{BNS}}} = 8. \quad (\text{C1})$$

The distance is then determined from GW data by a matched-filter likelihood of the form

$$p(d_L) \propto \exp(-(\rho - \langle \rho \rangle)^2/2), \quad (\text{C2})$$

where $\langle \rho \rangle$ is the ideal SNR achievable by the specific GW event, particularly parametrized by d_L , and ρ is the observed SNR that is subject to random fluctuation.

The Fisher matrix \mathcal{F}_{dd} can then be derived, which corresponds to the distance measurement uncertainty, this is

$$\mathcal{F}_{\text{dd}} = \left\langle \frac{\partial \log p^2}{\partial d_L} \right\rangle_{\sim \rho} \propto \frac{\alpha^2}{\hat{d}_L^4} \langle (\rho - \langle \rho \rangle)^2 \rangle_{\sim \rho} = \frac{\alpha^2}{\hat{d}_L^4}. \quad (\text{C3})$$

In the second line, we use the fact that ρ is normally distributed with variance equal to unity. Invoking the well-known Cramer–Rao bound from elementary statistics, we can infer that the distance measurement error is

$$\delta d \geq \mathcal{F}_{\text{dd}}^{-1/2} = \hat{d}_L^2/\alpha = \hat{d}_L/\langle \rho \rangle. \quad (\text{C4})$$

Thus, the fractional measurement error in the distance ($\delta d/d_L$) scales as $\langle \rho \rangle$. Note, however, that for the marginally detectable events – regardless of their BNS/NSBH identification and their luminosity distance – we have $\langle \rho \rangle = \rho^* = 8$. As a result, both the furthest detectable NSBH event and the BNS event share the same fractional distance measurement uncertainty to the leading order.

With this result, we can now proceed to derive a Fisher estimate of the error budget on H_0 . Here, we would like to assume a Gaussian error model for measuring the distance d_L of a GW emission event (regardless of BNS/NSBH origin). This is

$$p(d_L|\hat{d}_L(z; H_0)) \propto \exp\left(-\frac{(d_L - \hat{d}_L(z; H_0))^2}{2\sigma_d^2}\right). \quad (\text{C5})$$

As the above analysis suggested, the error term σ_d scales with \hat{d}_L , which we can express as $\sigma_d = \sigma_0 \hat{d}_L$. Repeating the same Fisher analysis – now evaluating the H_0 element of the Fisher matrix \mathcal{F}_{HH} :

$$\begin{aligned} \mathcal{F}_{\text{HH}} &= \left\langle \frac{\partial \log p^2}{\partial H_0} \right\rangle_{\sim d_L} = \int dd_L \left(\frac{\partial \log p}{\partial d_L} \frac{\partial \hat{d}_L}{\partial H_0} \right)^2 p(d_L|\hat{d}_L(z; H_0)) \\ &= \left(\frac{\hat{d}_L}{H_0} \right)^2 \frac{1}{\sigma_0^4} \left\langle \frac{d_L^4}{\hat{d}_L^6} - \frac{d_L^2}{\hat{d}_L^4} \right\rangle_{\sim d_L} \\ &= \frac{(\sigma_0^{-2} + 3)}{H_0^2}. \end{aligned} \quad (\text{C6})$$

The Cramer–Rao bound asserts that the uncertainty in determining H_0 with noisy measurements of d_L is lower bounded as

$$\frac{\delta H_0}{H_0} \geq \frac{\mathcal{F}_{\text{HH}}^{-1/2}}{H_0} = (\sigma_0^{-2} + 3)^{-1/2}. \quad (\text{C7})$$

It should be clear from the above expression that the fractional measurement uncertainty in H_0 does not depend on the NSBH/BNS classification of the GW event. Here, only the uncertainty in measuring d_L is taken into account. In fact, as argued in the main text (Section 2), it is possible to measure q to a better precision for NSBH events. Therefore, NSBH events are more powerful in determining the value of H_0 .

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.