Cosmology with the lights off: Standard sirens in the Einstein Telescope era

Stephen R. Taylor[*] and Jonathan R. Gair[†]

Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK
(Received 1 May 2012; published 3 July 2012)

We explore the prospects for constraining cosmology using gravitational-wave (GW) observations of neutron-star binaries by the proposed Einstein Telescope (ET), exploiting the narrowness of the neutron-star mass function. This builds on our previous work in the context of advanced-era GW detectors. Double neutron-star (DNS) binaries are expected to be one of the first sources detected after ''first-light'' of Advanced LIGO. DNS systems are expected to be detected at a rate of a few tens per year in the advanced era, but the proposed ET could catalog tens, if not hundreds, of thousands per year. Combining the measured source redshift distributions with GW-network distance determinations will permit not only the precision measurement of background cosmological parameters, but will provide an insight into the astrophysical properties of these DNS systems. Of particular interest will be to probe the distribution of delay times between DNS-binary creation and subsequent merger, as well as the evolution of the star-formation rate density within ET's detection horizon. Keeping H_0, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ fixed and investigating the precision with which the dark-energy equation-of-state parameters could be recovered, we found that with 10^5 detected DNS binaries, we could constrain these parameters to an accuracy similar to forecasted constraints from future CMB + BAO + SNIa measurements. Furthermore, modeling the merger delay-time distribution as a power-law ($\propto t^\alpha$) and the star-formation rate density as a parametrized version of the Porciani and Madau SF2 model, we find that the associated astrophysical parameters are constrained to within $\sim 10\%$. All parameter precisions scaled as $1/\sqrt{N}$, where N is the number of cataloged detections. We also investigated how parameter precisions varied with the intrinsic underlying properties of the Universe and with the distance reach of the network (which is affected, for instance, by the low-frequency cutoff of the detector). We also consider various sources of distance-measurement errors in the third-generation era and how these can be folded into the analysis.

I. INTRODUCTION

The era of advanced gravitational-wave (GW) detectors is approaching quickly. The previous decade has seen significant improvements in the sensitivity of GW interferometers, leading to the construction and operation of two Laser Interferometer Gravitational-wave Observatory (LIGO) ([1] detectors in the USA, GEO-600 in Germany [2], Virgo in Italy [3] and TAMA-300 in Japan [4]. The latter detector was designed as a test bed to develop new technologies for the proposed underground, cryogenically cooled KAGRA (formerly LCGT [5]) detector [6]. The LIGO, Virgo and GEO-600 detectors have conducted joint searches since 2007.

The most promising source for the first detection of gravitational waves is the inspiral and merger of a compact-object binary consisting of neutron stars (NSs) and/or black holes [7]. The first joint search for compact-binary coalescence signals during the LIGO S5 science run and the Virgo VSR1 data did not result in direct detections [8], nor did the ''enhanced'' detector search during the LIGO S6 science run and the Virgo VSR2 + 3 data [9]. Furthermore, the upper limits placed on compact-binary

coalescence rates from the latter search remain two to 3 orders of magnitude above existing astrophysically predicted rates. However, the LIGO detectors are currently being upgraded to their ''advanced'' configuration [10], due for completion in ~ 2015, for which the horizon distance for NS-NS inspiral detection will be boosted to ~ 450 Mpc, giving an almost thousandfold gain in volume sensitivity of the detectors. The advanced detectors are expected to detect double NS inspirals at a rate of ~ 40 yr^{-1}, although this may vary by approximately 2 orders of magnitude in either direction [11].

Complementing AdLIGO will be a global network of advanced detectors, including AdVirgo [12], KAGRA [6] and possibly a third LIGO detector in India, LIGO-India [13]. There are currently no prospects for a Southern Hemisphere GW interferometer operating in the advanced era. A global network comprising these detectors will help turn the field from the search for the first detection into a precise astronomical tool.

The GWs emitted by a compact-binary system directly encode the *redshifted* masses and luminosity distance of the system. Double NS (DNS) binary systems are commonly referred to as *self-calibrating standard sirens* because their distance is directly encoded in the waveform, and a means of determining their redshift would mean we could probe the cosmic distance ladder and extract cosmological parameters

---

[*]staylor@ast.cam.ac.uk
[†]jgair@ast.cam.ac.uk

[14–17]. While the phase evolution of the strain signal in a single interferometer can give precise constraints on the redshifted mass of the system, we require a global network of detectors to constrain the sky location, orbital inclination and GW polarization so that we can break the degeneracies of these angular factors with the luminosity distance.

Unfortunately, the redshift and intrinsic mass of the systems enter the waveform only in a combination as the redshifted mass parameter; hence, previous techniques for performing GW cosmology using these standard sirens has relied on the association of the GW source with an electromagnetic counterpart, which can provide an independent redshift measurement [18–21]. In our previous paper, [22], we studied a technique for probing the Hubble constant and NS mass-distribution parameters using only the GWs detected by an advanced-era network. This work relied on the assumption, supported by observations, that the NS mass distribution is sufficiently narrow, which means that we already have a good idea of the intrinsic masses in the system, and the measured redshifted mass parameter then provides a narrow distribution of possible redshifts for each observed source. Combining these redshift distributions with the network-measured luminosity distance for a catalog of observed DNS systems provides constraints on the underlying cosmological parameters, as well as the astrophysical distribution of these systems. This technique (first considered by Marković in Ref. [23] and extended in Ref. [24,25]) relies on measurements of the *redshifted chirp mass* (expected to be the best-determined parameter) and luminosity distance for a catalog of detected systems.

In our previous analysis, the cosmological parameters which we could constrain were restricted by the sub-Gpc reach of an advanced-era network. We now extend this technique to a third-generation network, which could have a reach out to tens of Gpcs. Proposed third-generation detectors aim for a broadband factor of 10 sensitivity improvement with respect to advanced detectors and to increase sensitivity in the range $\sim$1–10 Hz, compared to the $\sim$10–20 Hz low-frequency cutoff of advanced detectors. As a prototypical third-generation detector, we use the Einstein Telescope, consisting of three overlapping 10 km arm-length interferometers in a triangular configuration [26–28]. Each interferometer may actually be two detectors: a cryogenically cooled, underground detector with good low-frequency sensitivity, and a high laser-power detector with good high-frequency sensitivity [29]. Keeping H_0, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ fixed, we find that the sensitivity provided by such a network will be large enough to constrain the dark-energy equation-of-state parameters and NS mass-distribution parameters, as well as the astrophysical distribution of the systems. The latter will inform us about the average time delay between the formation of these compact-binary systems and their merger, as well as the shape of the underlying star-formation-rate density.

Third-generation detectors are unlikely to be online before the mid-2020s, but, if realized, the ambitious and novel design for the Einstein Telescope will have far-reaching scientific advantages. Such a detector could detect as many as hundreds of thousands of DNS inspirals per year, which, along with the distance reach of the detectors, will permit precision GW astronomy. In this paper, we will not consider other methods which have been proposed for using GW observations as cosmological probes. In particular, we do not consider association of GW detections with an electromagnetic (EM) counterpart, which has been studied in Refs. [16,17], nor do we consider tidal-coupling corrections to the phase evolution of the strain signal [30]. The latter method is also a GW-only technique with significant potential, in that these phase-evolution corrections break the mass-redshift degeneracy at 5 PN order and, assuming the NS equation-of-state is well known, will permit the distance-redshift relation to be probed. It may also be possible to apply the method used by Ref. [31], which was investigated in the context of future space-based detectors, to third-generation ground-based detectors. Their method relies on the measurement of cosmologically induced shifts in the GW phase at 4 PN order.

This paper is laid out as follows. Section II describes the characteristics of the Einstein Telescope, as well as possible third-generation networks and detection criteria. In Sec. III, we discuss aspects of DNS systems, including the mass distribution of the constituent NSs, and the redshift distribution of DNS mergers. Section IV describes the effect of the dark-energy equation-of-state parameter on cataloged luminosity distances, while Sec. V describes how we construct and analyze catalogs of detected DNS-system inspirals. Results are shown in Sec. VI, followed by our conclusions in Section VII.

II. DETECTOR CHARACTERISTICS AND NETWORKS

A. The Einstein Telescope

The Einstein Telescope (ET) is a proposed third-generation ground-based interferometer. A recent design study has been carried out within the European Commission's FP7 framework [32] to evaluate the science case for such a detector and to consider the technological advances required for the science goals to be achieved. Through this three-year design study, some favored designs and configurations have emerged.

The aim for third-generation ground-based detector designs is to achieve a broadband factor of 10 sensitivity improvement with respect to advanced detectors and to push the sensitivity down into the $\sim$1–10 Hz range. Early designs examined the prospects for pushing conventional techniques used in advanced detectors to their limits to construct a third-generation interferometer [33]. This gave the ET-B noise curve in Fig. 1. Beyond the extension of the arm length to 10 km, several techniques were proposed to suppress high- and low-frequency noise, including siting the detector underground.
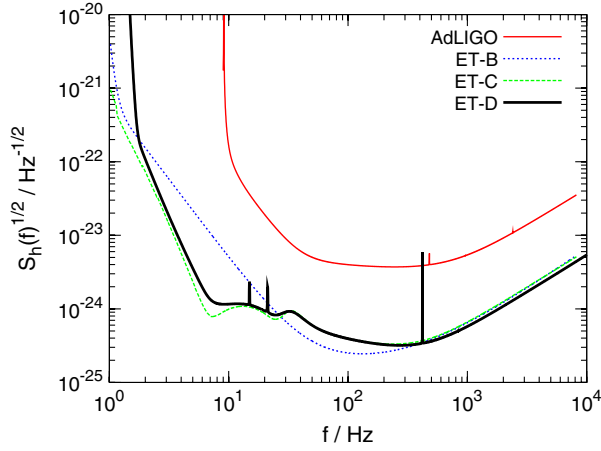
FIG. 1 (color online).   Comparison of the different noise curves for AdLIGO (high-power zero-detuning) [40], the initial Einstein Telescope noise curve based on conventional techniques, ET-B [33], the initial xylophone noise curve, ET-C [34], and the improved, more realistic xylophone noise curve, ET-D [29]. These noise curves are for one 10 km right-angled interferometer.

Crucially, the techniques used to suppress high-frequency noise are not necessarily compatible with the suppression of low-frequency noise. Increasing the laser power will reduce the photon shot noise which dominates the high-frequency range, but this worsens the thermal noise which dominates at low frequencies. The "xylophone" design was proposed to address this issue. Instead of having a single broadband instrument, the xylophone design comprises a high-power, high-frequency interferometer (ET-HF) and a cryogenic low-power, low-frequency interferometer (ET-LF) [34]. ET-LF would be an underground instrument, and limited at low frequencies by gravity-gradient noise, while ET-HF would be colocated and co-oriented with ET-LF but surface-sited. ET-HF would employ high-power lasers to suppress high-frequency photon shot noise.

The initial xylophone design gave the ET-C curve [34] in Fig. 1, which was refined to give the ET-D xylophone design [29]. We will use the ET-D noise curve in the ensuing analysis.

Current and advanced-era ground-based interferometers are right-angled interferometers, since an arm-opening angle of 90° maximizes their sensitivity. However, if both GW polarization states are to be measured at a single site, then two or more colocated nonaligned interferometers are required. Furthermore, at least three colocated interferometers are required to construct a null stream, i.e., a sum of individual interferometer responses which is insensitive to GWs and can be used to identify noise transients in the data stream.

Taking these goals into account, the design requiring the shortest total length of tunnels is a triangular configuration with three identical interferometers positioned at each vertex of the triangle, an arm-opening angle of 60° and rotated

relative to each other by 120° [26–28]. A triangular configuration also provides redundancy, since polarization constraints are still possible with only two vertices operational.

In the following, we consider three ET-D interferometers in the triangular configuration, which we denote as a "single ET." More than one ET would be very optimistic, so we complement our single ET with a network of individual third-generation right-angled interferometers (also with ET-D sensitivity) to permit source distance determination. While different locations are being mooted, we choose the Virgo location as the reference ET site [35].

B. Signal-to-noise ratio

The optimal matched filtering signal-to-noise ratio of a GW detection is given by

$$\rho_{\text{opt}} = 2\left[\int_0^\infty df \frac{|\tilde{h}(f)|^2}{S_h(f)}\right]^{1/2}, \tag{1}$$

where $S_h(f)$ denotes the detector's noise power spectral density. In the quadrupolar approximation, the Fourier transform of the signal amplitude of GWs from an inspiraling binary system takes the following form [36,37]:

$$|\tilde{h}(f)| = \frac{2c}{D_L}\left(\frac{5G\mu}{96c^3}\right)^{1/2}\left(\frac{GM}{\pi^2 c^3}\right)^{1/3}\left(\frac{\Theta}{4}\right)f^{-7/6}. \tag{2}$$

The function Θ is defined by

$$\Theta \equiv 2[F_+^2(1 + \cos^2\iota)^2 + 4F_\times^2\cos^2\iota]^{1/2}, \tag{3}$$

where $0 < \Theta < 4$, and

$$\begin{aligned} F_+ &\equiv \tfrac{1}{2}(1 + \cos^2\theta)\cos 2\phi\cos 2\psi - \cos\theta\sin 2\phi\sin 2\psi, \\ F_\times &\equiv \tfrac{1}{2}(1 + \cos^2\theta)\cos 2\phi\sin 2\psi + \cos\theta\sin 2\phi\cos 2\psi \end{aligned} \tag{4}$$

are the interferometer's strain responses to the different GW polarizations.

Following Ref. [25], we can write the matched filtering signal-to-noise ratio in a single detector as

$$\rho = 8\Theta\frac{r_0}{D_L}\left(\frac{\mathcal{M}_z}{1.2\,M_\odot}\right)^{5/6}\sqrt{\zeta(f_{\max})}, \tag{5}$$

where $\mathcal{M}_z = (1 + z)\mathcal{M}$ is the redshifted chirp mass,

$$r_0^2 \equiv \frac{5}{192\pi}\left(\frac{3G}{20}\right)^{5/3}x_{7/3}\frac{M_\odot^2}{c^3},$$

$$x_{7/3} \equiv \int_0^\infty \frac{df(\pi M_\odot)^2}{(\pi f M_\odot)^{7/3}S_h(f)},$$

$$\zeta(f_{\max}) \equiv \frac{1}{x_{7/3}}\int_0^{2f_{\max}} \frac{df(\pi M_\odot)^2}{(\pi f M_\odot)^{7/3}S_h(f)}, \tag{6}$$

and $2f_{\max}$ is the wave frequency at which the inspiral detection template ends [38]. The intrinsic chirp mass, $\mathcal{M}$, is given in terms of the component masses by,

$$\mathcal{M} = \left(\frac{m_1 m_2}{(m_1 + m_2)^2}\right)^{3/5}(m_1 + m_2). \tag{7}$$

The phase evolution of the strain signal in a single interferometer can constrain $\mathcal{M}_z$ to subpercent precision [39]. In order to measure the luminosity distance, D_L, we require

TABLE I.   The characteristic distance sensitivities [as given by evaluating Eq. (6)] of some advanced-detector configurations and various design studies for the third-generation Einstein Telescope.

Detector	r_0/Mpc
AdLIGO[a]	80
AdLIGO[b]	110
AdLIGO[c]	119
AdVirgo[d]	85
ET[e]	1527
ET-B[f]	1587
ET-C[g]	1918
ET-D[h]	1591

[a]Low-power zero detuning [40]
[b]High-power zero detuning [40]
[c]NS-NS optimized [40]
[d]Ref. [41]
[e]Polynomial noise-curve approximation [42]
[f]Conventional technology [32]
[g]3rd-generation technology, xylophone configuration [32]
[h]3rd-generation technology, xylophone configuration (updated and more realistic) [32]

a network of separated detectors to break the waveform degeneracy between D_L and Θ [see Eq. (2)]. Distance-measurement errors in a third-generation network will be discussed in Sec. VI D.

The signal-to-noise ratio (SNR) of a detected system will vary between the individual network sites, as a result of the different $S_h(f)$'s and angular dependencies. However, following Refs. [25,38], we assume the network SNR of a detected system is given by the quadrature summation of the individual interferometer SNRs,

$$\rho_{\text{net}} = \sqrt{\sum_k \rho_k^2} = \frac{8}{D_L}\left(\frac{\mathcal{M}_z}{1.2M_\odot}\right)^{5/6}\sqrt{\sum_k (r_{0,k}\Theta_k\zeta_k(f_{\text{max}}))^2},$$

(8)

where $r_{0,k}$, $\zeta_k(f_{\text{max}})$ and Θ_k encode the distance, frequency and angular sensitivities of the kth detector. A comparison of the characteristic distance sensitivities of some second- and third-generation detectors is shown in Table I.

C. Network antenna patterns

The angular dependence of the SNR is encapsulated by the variable Θ. The sky location and binary orientation can be deduced from the network analysis; however, in our analysis, we will use only D_L and $\mathcal{M}_z$ measurements. We calculate the probability density function for Θ [25] numerically using Eq. (3) by choosing $\cos\theta$, ϕ/π, $\cos\iota$ and ψ/π to be uncorrelated and distributed uniformly over the range $[-1, 1]$.

It is unlikely that more than one ET will be constructed. A more likely network configuration will involve a single ET with single third-generation right-angled detectors at other sites. In the interest of verisimilitude, we take into account possible detector locations for such a third-generation

network. Table II contains the locations and arm-bisector orientations of various current and planned detectors.

To write down the antenna pattern function as a function of the detector position,[1] we use the notation and formalism of Ref. [43].

For a GW source at coordinates (θ, ϕ) on the sky, with polarization angle ψ and a detector with opening angle η at latitude β and longitude λ and such that the bisector of its arms points at an angle χ, counterclockwise from East, the antenna pattern functions are

$$\begin{pmatrix} F_+ \\ F_\times \end{pmatrix} = \sin\eta \begin{pmatrix} \cos(2\psi) & \sin(2\psi) \\ -\sin(2\psi) & \cos(2\psi) \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix},$$

(9)

where,

$$a = \frac{1}{16}\sin(2\chi)[3 - \cos(2\beta)][3 - \cos(2\theta)]\cos[2(\phi + \lambda)]$$
$$+ \frac{1}{4}\cos(2\chi)\sin\beta[3 - \cos(2\theta)]\sin[2(\phi + \lambda)]$$
$$+ \frac{1}{4}\sin(2\chi)\sin(2\beta)\sin(2\theta)\cos(\phi + \lambda)$$
$$+ \frac{1}{2}\cos(2\chi)\cos\beta\sin(2\theta)\sin(\phi + \lambda)$$
$$+ \frac{3}{4}\sin(2\chi)\cos^2\beta\sin^2\theta,$$

(10)

$$b = \cos(2\chi)\sin\beta\cos\theta\cos[2(\phi + \lambda)]$$
$$- \frac{1}{4}\sin(2\chi)[3 - \cos(2\beta)]\cos\theta\sin[2(\phi + \lambda)]$$
$$+ \cos(2\chi)\cos\beta\sin\theta\cos(\phi + \lambda)$$
$$- \frac{1}{2}\sin(2\chi)\sin(2\beta)\sin\theta\sin(\phi + \lambda).$$

As a reference, we use a network comprising three 60° ET-D sensitivity interferometers at the Virgo location (a single ET), plus right-angled interferometers at the LIGO-Livingston and LIGO-India locations. The characteristic distance reach of all of the interferometers in the network is taken as 1591 Mpc, corresponding to ET-D sensitivity [29]. This is the sensitivity of a 10 km right-angle interferometer. We account for the different detector arm-opening angles in the antenna pattern functions.

The network SNR given by Eq. (8) also depends on $\zeta(f_{\text{max}})$, which describes the overlap of the signal power with the detector bandwidth [25]. The frequency at the end of the inspiral (taken to correspond to the innermost stable circular orbit) is at

$$f_{\text{max}} = \frac{785 \text{ Hz}}{1 + z}\left(\frac{2.8 \ M_\odot}{M}\right),$$

(11)

where M is the total mass of the binary system [36]. The maximum binary-system mass could conceivably be $\sim 4.2 \ M_\odot$.[2] The ET horizon distance for a system with a

---

[1]We do not consider modulation of the antenna patterns due to the rotation of the Earth. We justify this in Sec. VI F.

[2]Both neutron stars in the binary system would need to have masses 2σ above the distribution mean at the maximum considered μ and σ, where $\mu_{\text{NS}} \in [1.0, 1.5] \ M_\odot$, $\sigma_{\text{NS}} \in [0, 0.3] \ M_\odot$.

TABLE II.   A reproduction of the GW-interferometer geographical locations and arm-bisector orientations from Schutz [43]. We include updated IndIGO information [44].

Detector	Label	Longitude	Latitude	Orientation
LIGO-Livingston, LA, USA	L	$90°46'27.3''$ W	$30°33'46.4''$ N	$208.0°$(WSW)
LIGO-Hanford, WA, USA	H	$119°24'27.6''$ W	$46°27'18.5''$ N	$279.0°$(NW)
Virgo, Italy	V	$10°30'16''$ E	$43°37'53''$ N	$333.5°$(NNW)
KAGRA (formerly LCGT), Japan	J	$137°10'48''$ E	$36°15'00''$ N	$20.0°$(WNW)
LIGO-India, India	I	$76°26'$ E	$14°14'$ N	$45.0°$(NE)

total mass of $\sim 4 \, M_\odot$ is ~ 25 Gpc [16]. In the ΛCDM cosmology, this corresponds to a redshift of ~ 2.9, and from Eq. (11), this gives $f_{max} \sim 134$ Hz. Given the ET-D noise-curve [29], $\sqrt{\zeta(f_{max} = 134\text{Hz})} \gtrsim 0.98$. Extending the *redshift* reach out to $z \sim 5$ still gives $\sqrt{\zeta(f_{max} = 87 \text{ Hz})} \gtrsim 0.96$. Thus, we feel justified in adopting $\zeta(f_{max}) \simeq 1$ for all interferometers in the ensuing analysis.

Using these expressions, we were able to numerically estimate the probability distribution for the *effective* Θ,

$$\Theta_{eff} = \sqrt{\sum_k \Theta_k^2}, \qquad (12)$$

where the sum is over all detectors in the network. We use this Θ_{eff} distribution to choose SNRs for each source in the catalog via Eq. (5) and then impose a detection criterion. As a reference, we adopt the detection criterion that the network SNR must be greater than 8.

III. DNS SYSTEMS

A. Neutron-star mass distribution

For a full discussion of our assumptions and modeling details of the NS mass distribution in DNS systems, see our previous work (Ref. [22], and references therein). We provide here a brief summary of the main assumptions pertinent to the present study.

To lowest order, the GW signal depends on the two neutron-star masses through the chirp mass, $\mathcal{M}$. We assume that the distribution of individual neutron-star masses is normal, as suggested by analysis of Galactic DNS systems [45,46] and population synthesis studies (see, e.g., Refs. [37,47,48]). For $\sigma_{NS} \ll \mu_{NS}$, this should also lead to an approximately normal distribution for the chirp mass.

We use a simple ansatz for the relationship between the chirp-mass distribution parameters and the underlying neutron-star mass distribution. The chirp mass distribution is modeled as normal,

$$\mathcal{M} \sim N(\mu_c, \sigma_c^2),$$

with mean and standard deviation

$$\mu_c \approx 2(0.25)^{3/5} \, \mu_{NS}, \qquad \sigma_c \approx \sqrt{2}(0.25)^{3/5} \, \sigma_{NS}, \quad (13)$$

where μ_{NS} and σ_{NS} are the mean and standard deviation of the underlying neutron-star mass distribution.

A recent study by Ozel *et al.* [49] has found that DNS data are consistent with both pulsar and companion having been drawn from the same underlying distribution of masses. The literature indicates an underlying neutron-star mass distribution in DNS systems with $\sigma_{NS} \lesssim 0.15 \, M_\odot$ [45,46,49].[3] Hence, we anticipate that Eq. (13) will be appropriate for generating data sets, and we use this in the ensuing analysis. The assumption throughout is that for the volume of the Universe probed by our global network, the neutron-star mass distribution does not change.

Further population synthesis and observational studies in the following decade will help to shed further light on the nature of the NS mass distribution. The assumption of a unimodal (for DNS systems) Gaussian distribution is an approximation, and if future studies show this to be inappropriate, then a more suitable ansatz could be readily incorporated within the framework described in this paper.

B. Merger-rate density

In this analysis, we aim to probe not only the background cosmology and NS mass-distribution parameters, but the astrophysical properties of the binary NS population. To this end, we now consider the factors contributing to the coalescence of a binary NS system.

Following several population synthesis studies (e.g., Refs. [37,50]) and Ref. [51], we define the merger rate per comoving volume as

$$\dot{n}(t) = \int_{t_*}^{t} \lambda \frac{dP_m}{dt}(t - t_b) \frac{d\rho_*}{dt}(t_b) dt_b, \qquad (14)$$

where λ is a *mass efficiency*, defined as the number of coalescing DNS binaries per unit star-forming mass [37]. dP_m/dt is the probability density distribution of a DNS binary merging at a time $(t - t_b)$ after formation, and $d\rho_*/dt$ is the star-formation rate (SFR) density at cosmological time t_b.

Star formation, and the efficiency of double compact-object formation from the progenitor system, may be sensitive to the host-galaxy morphology and environment (e.g., metallicity). Furthermore, the distribution of delay

---

[3]Indeed, Ozel *et al.* [49] indicates that the DNS mass distribution is peculiar, since it cannot be explained via electron-capture or core-collapse supernovae mechanisms; rather, its narrow dispersion may be a result of the evolutionary path of these systems.

times between star formation and the corresponding DNS-system coalescence may have contributions from several different evolutionary paths. However, we are interested here only in a third-generation GW-interferometer network's ability to constrain various astrophysical and cosmological parameters [22]. As such, we consider a single component star-formation distribution, delay-time distribution and mass efficiency, deferring considerations of the other dependencies to a future study. We now discuss the various terms in Eq. (14) in more detail.

1. Mass efficiency, λ

We use values for λ obtained from the population synthesis calculations of Ref. [50]. Smoothed histograms of the mass efficiency are shown in Fig. 4 of that paper, with modes at $\sim 10^{-5} \, M_\odot^{-1}$ for DNS systems formed in both elliptical and spiral conditions. However the distribution of λ ranges over several orders of magnitude, with $10^{-7} \, M_\odot^{-1} \lesssim \lambda \lesssim 10^{-3} \, M_\odot^{-1}$. We adopt $\lambda = 10^{-5} \, M_\odot^{-1}$ as the reference value for our analysis.

2. Merger-delay distribution, dP_m/dt

Massive stars in high-mass binary systems evolve into DNS systems on much shorter time scales than typical galaxy evolution or Hubble time scales, such that dP_m/dt is essentially determined by the initial orbital separation, a, of the DNS system [52]. The evolutionary time delay between the progenitor formation and the formation of the corresponding DNS system is typically $\lesssim 50$ Myr [53] and is therefore negligible compared to the gravitational-wave inspiral time scale, which scales as $\tau_{\text{gr}} \propto a^4$. Assuming the number of binaries, N, born with separation a scales as $dN/da \propto a^\gamma$ [52], we obtain

$$\frac{dP_m}{dt} \propto \frac{dN}{d\tau_{\text{gr}}} = \frac{dN}{da} \frac{da}{d\tau_{\text{gr}}} \propto t^{(\gamma-3)/4} = t^\alpha. \quad (15)$$

If DNS systems have the same orbital separation distribution as normal-abundance main-sequence stars [54,55], then $\gamma = \alpha = -1$. However, this scaling is not well-constrained, and this is discussed in more detail in Appendix A. Instead, we adopt the approach of allowing α to be a free parameter which we attempt to fit from our observations and ask with what precision this can be determined. We use the value $\alpha = -1$ for our reference model, which is justified by current (albeit sparse) analysis of Galactic DNS systems [56–58], and various population synthesis calculations [50,53,59–62]. For normalization purposes, we assume a minimum delay time of 50 Myr and a maximum delay time equal to the cosmology-dependent age of the Universe; these choices are discussed in Appendix A.

3. Star-formation rate density, $d\rho_*/dt$

The star-formation rate density is also rather uncertain. The SF2 model of Porciani and Madau [63] attempts to

factor in the uncertainties in the incompleteness of data sets and the amount of dust extinction at early epochs. The SF2 model has the form

$$\frac{d\rho_*}{dt}(z) \approx 0.16 \times \left(\frac{\exp(\beta_1 z)}{\exp(\beta_2 z) + 22} \right)$$
$$\times \frac{E(z)}{(1+z)^{3/2}} \, M_\odot \text{Mpc}^{-3} \, \text{yr}^{-1}, \quad (16)$$

where

$$E(z) = \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_\Lambda(z)} \quad (17)$$

and $\beta_1 = \beta_2 = 3.4$. In this model, the SFR density remains roughly constant at $z \gtrsim 2$, which may be incompatible with recent observations [64,65]. This is discussed in more detail in Appendix B. To allow for some uncertainty, we treat β_1 and β_2 as free parameters and explore how precisely we can measure them. While this simple ansatz does not cover all possible forms for the SFR density, using it will provide an indication of what GW observations could tell us. The framework is easily adaptable to more complex SFR models.

We must also specify t_*, the lower integration bound in Eq. (14), which represents the time of the earliest period of star formation. The highest redshift objects observed are a long gamma-ray burst (GRB) with a photometric redshift of ~ 9.4 [66] and a candidate $z \sim 10$ galaxy [67]. We therefore use $z = 10$ as the earliest time of star formation. Future observations, for instance with the *James Webb Space Telescope*, may be able to probe back to the first phases of galaxy formation at $z \sim 15$, and if objects are found at that epoch, this assumption should be revised. However, our results are fairly insensitive to this choice.

4. Calculating $\dot{n}(z)$

Equation (14) can be rewritten as a distribution in redshift using $dt/dz = -1/((1+z)E(z)H_0)$

$$\dot{n}(z) = \int_{10}^{z} \lambda \frac{dP_m}{dt}(t - t_b) \frac{d\rho_*}{dt}(t_b) \frac{dt_b}{dz_b} dz_b$$
$$= \frac{0.16\lambda}{H_0} \int_z^{10} \frac{dP_m}{dt}(t(z) - t_b(z_b)) \left(\frac{\exp(\beta_1 z_b)}{\exp(\beta_2 z_b) + 22} \right)$$
$$\times \frac{dz_b}{(1+z_b)^{5/2}}. \quad (18)$$

Evaluating this expression requires an expensive double integral which created a bottleneck in our analysis. However, because the priors on the cosmological parameters are narrow (see Sec. V D), there is little variation in the merger-rate density across this range, as shown in Fig. 2. We therefore fixed the cosmological parameters at their reference values for the cosmological time calculation, which made the merger-rate density calculation considerably faster. Although this throws away some of the
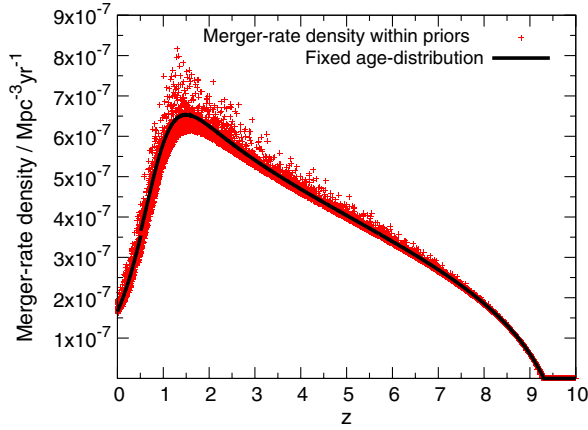
FIG. 2 (color online).   Merger-rate densities computed for the reference cosmology (solid line) and for cosmological parameters chosen randomly from within the prior range (red crosses).

cosmological information, it did not significantly affect the results and made the analysis more tractable.

IV. COSMOLOGICAL MODELS

In our previous analysis [22], we considered only a flat cosmology, but here we allow for curvature and an evolving equation of state (EOS) for the dark energy. From the cosmological field equations, we have

$$\dot{\rho} + 3\left(\frac{\dot{a}}{a}\right)\left(\rho + \frac{p}{c^2}\right) = 0, \tag{19}$$

where ρ and p are the density and pressure of a cosmological fluid, respectively, while a is the scale factor of the Universe. Derivatives are with respect to physical time. For a perfect fluid ($p = w\rho c^2$, where w is the EOS parameter), this reduces to

$$\left(\frac{\dot{\rho}}{\rho}\right) = -3(1 + w)\left(\frac{\dot{a}}{a}\right). \tag{20}$$

Hence,

$$\rho(a) = \rho(a_{\text{today}} = 1) \times e^{-3\int_1^a (1+w)(da'/a')}. \tag{21}$$

The last decade has seen many proposals for w, with different physical motivation. One approach attempts to explain dark energy as a minimally coupled scalar field ("quintessence") slowly rolling down its potential such that it can exert negative pressure. While it is possible to have a constant EOS in this formalism, this requirement places severe constraints on the potential, and so it is natural to expect a time-varying EOS [68].

The simplest approximation is to assume a linear model ($w(z) = w_0 + w_1 z$), but this is only appropriate for local studies due to the divergence at high redshift. The Shafieloo-Sahni-Starobinsky ansatz [69] models the EOS evolution as a "tanh" form which ensures $w = -1$ at early times and $w \to 0$ at low z. This ansatz prevents the crossing of the "phantom divide" at $w = -1$, desirable since

phantom fluids cannot be explained by a minimally coupled scalar field [68]. The ansatz we adopt in this work is the Chevallier-Polarski-Linder ansatz [68,70]

$$w(a) = w_0 + w_a(1 - a), \qquad w(z) = w_0 + w_a\left(\frac{z}{1 + z}\right). \tag{22}$$

This ansatz was adopted by the Dark Energy Task Force [71] and has several desirable features. It depends on only two free parameters, it reduces to the linear model at low z, and it is well-behaved at high redshift, tending to $w_0 + w_a$. Using this EOS,

$$\Omega_\Lambda(z) = \Omega_{\Lambda,0} \times (1 + z)^{3(1+w_0+w_a)} \times e^{-3w_a(z/1+z)}. \tag{23}$$

For different global geometries of the Universe, the luminosity distance, D_L, is given by

$$D_L(z|\mathcal{C}) = (1 + z) \times \mathcal{F}(z|\mathcal{C}),$$

where

$$\mathcal{F}(z|\mathcal{C}) = \begin{cases} \frac{D_H}{\sqrt{\Omega_{k,0}}} \sinh\left(\sqrt{\Omega_{k,0}}\frac{D_c(z|\mathcal{C})}{D_H}\right), & \Omega_{k,0} > 0, \\ D_c(z|\mathcal{C}), & \Omega_{k,0} = 0, \\ \frac{D_H}{\sqrt{|\Omega_{k,0}|}} \sin\left(\sqrt{|\Omega_{k,0}|}\frac{D_c(z|\mathcal{C})}{D_H}\right), & \Omega_{k,0} < 0, \end{cases} \tag{24}$$

in which D_H is the Hubble scale (c/H_0) and $\mathcal{C} = \{H_0, \Omega_{m,0}, \Omega_{\Lambda,0}, \Omega_{k,0}, w_0, w_a\}$ is the set of cosmological parameters describing the large-scale characteristics of the Universe.

The comoving radial distance, $D_c(z)$, is given by

$$D_c(z) = D_H \int_0^z \frac{dz'}{E(z')}, \tag{25}$$

where $E(z)$ is given by Eq. (17). The redshift derivative of the comoving volume is given generally by

$$\frac{dV_c}{dz} = 4\pi D_H \frac{D_L(z)^2}{(1 + z)^2 E(z)}. \tag{26}$$

V. MAKING AND ANALYZING DNS CATALOGUES

We refer the reader to our previous study [22] for full details of our calculation, but we summarize the main details here.

A. Distribution of detectable DNS systems

The two system properties we will use in our analysis are the redshifted chirp mass, $\mathcal{M}_z$, and the luminosity distance, D_L. We assume that only systems with an SNR greater than a given threshold will be detected. We can write down the distribution of the number of events per unit time in the observer's frame with $\mathcal{M}$, z and effective Θ [25,37],

$$\frac{d^4N}{dt d\Theta dz d\mathcal{M}} = \frac{dV_c}{dz} \frac{\dot{n}(z)}{(1+z)} \mathcal{P}(\mathcal{M}) \mathcal{P}_\Theta(\Theta). \qquad (27)$$

The $1/(1+z)$ factor accounts for the redshifting of the merger rate [37].

Converting this to a distribution in $\mathcal{M}_z$, D_L and ρ and integrating over ρ to find the distribution of *detectable* systems (i.e., systems above SNR threshold) gives

$$\frac{d^3N}{dt dD_L d\mathcal{M}_z}\bigg|_{\rho > \rho_0} = \frac{dz}{dD_L} \frac{dV_c}{dz} \frac{\dot{n}(z)}{(1+z)^2} \times \mathcal{P}\left(\frac{\mathcal{M}_z}{1+z}\bigg| D_L\right)$$
$$\times C_\Theta\left[\frac{\rho_0}{8} \frac{D_L}{r_0}\left(\frac{1.2 \, M_\odot}{\mathcal{M}_z}\right)^{5/6}\right], \qquad (28)$$

where the form of (dz/dD_L) will depend on the curvature of the Universe [see Eq. (24)].

To calculate the number of detected systems (given a set of model parameters, $\vec{\mu}$), we integrate over this distribution, which is equivalent to integrating the distribution over redshift and chirp mass, i.e. $N_\mu = T \times \int_0^\infty \int_0^\infty (\frac{d^3N}{dt dz d\mathcal{M}}) dz d\mathcal{M}$, where T is the duration of the observation run.[4]

B. Creating mock catalogs of DNS-binary inspiraling systems

The model-parameter space we investigate is the 7-dimensional space of $[w_0, w_a, \mu_{NS}, \sigma_{NS}, \alpha, \beta_1, \beta_2]$. To generate a catalog of events, we choose a set of reference parameters, motivated by previous analysis in the literature. For our reference cosmology, we adopt $H_0 = 70.4 \, \text{km s}^{-1} \text{Mpc}^{-1}$, $\Omega_{m,0} = 0.2726$, $\Omega_{\Lambda,0} = 0.728$, $w_0 = -1.0$ and $w_a = 0.0$ [72]. The reference parameters of the neutron-star mass distribution are $\mu_{NS} = 1.35 \, M_\odot$ and $\sigma_{NS} = 0.06 \, M_\odot$ [45]. We have previously discussed the delay-time distribution and SFR density in Sec. III B. We adopt a power-law merger-delay distribution with reference power-law index $\alpha = -1.0$, and we take the SFR density to be given by the SF2 ansatz [63], with $\beta_1 = 3.4$ and $\beta_2 = 3.4$.

These reference parameters are used to calculate an expected number of events,[5] and the number of observed events, N_o, is drawn from a Poisson distribution (assuming each binary system is independent of all others) with that mean. Monte Carlo acceptance/rejection sampling is used to draw random redshifts and chirp masses from the distribution in Eq. (27) for each event. The D_L and $\mathcal{M}_z$ are then computed from the sampled $\mathcal{M}$ and z.

For these reference parameters, which give a local merger-rate density of $\sim 2 \times 10^{-7} \, \text{Mpc}^{-3} \text{yr}^{-1}$, and assuming detected systems must have a network SNR greater than 8, we find that the expected number of detections in 1 yr is $\sim 10^5$.[6]

C. Likelihood statistic

In the measurement-parameter space of redshifted chirp mass and luminosity distance, the measured number of detections in a given bin will be a Poisson random variate with a model-dependent mean. If we take the continuum limit of this, such that bin sizes are infinitesimally small and contain either 0 or 1 events, then we can formulate the likelihood of a catalog of discrete events,

$$\mathcal{L}(\vec{\vec{\Lambda}}|\vec{\mu}, \mathcal{H}) = e^{-N_\mu} \prod_{i=1}^{N_o} r(\vec{\lambda}_i|\vec{\mu}), \qquad (29)$$

where $\vec{\vec{\Lambda}} = \{\vec{\lambda}_1, \vec{\lambda}_2, \ldots, \vec{\lambda}_{N_o}\}$ is the vector of measured system properties, with $\vec{\lambda}_i = (\mathcal{M}_z, D_L)_i$ for system i. N_o is the *actual* number of detected systems, while N_μ is the number of DNS inspiral detections *predicted* by the model with parameters $\vec{\mu}$. Finally, $r(\vec{\lambda}_i|\vec{\mu})$ is the rate of events with properties $\mathcal{M}_z$ and D_L, evaluated for the ith detection under model parameters $\vec{\mu}$, which is given by Eq. (28). The trial cosmological parameters are used to calculate a model-dependent redshift from the cataloged luminosity distance, and, in turn, this redshift is used to infer a model-dependent intrinsic chirp mass from the cataloged value of $\mathcal{M}_z$. These values of $\mathcal{M}_z$, D_L, as well as the model-dependent values of z and $\mathcal{M}$ are inserted into Eq. (28) to compute the likelihood.

In our previous study, we employed a modified likelihood statistic which marginalized over the local merger-rate density of DNS systems. This approach reflects our current lack of knowledge of this quantity, estimates of which vary over several orders of magnitude. We adopt the same approach in this analysis, to eliminate the dependence on poorly known scaling factors. This includes the mass efficiency parameter, λ, which is a quantity derived from population-synthesis studies and can vary over several orders of magnitude.

The modified likelihood statistic is

$$\tilde{\mathcal{L}}(\vec{\vec{\Lambda}}|\vec{\mu}, \mathcal{H}) \propto N_\mu^{-(N_o+1)} \prod_{i=1}^{N_o} r(\vec{\lambda}_i|\vec{\mu}). \qquad (30)$$

---

[4]We found that, for the purposes of the calculation, assuming the NS mass distribution was a δ-function, centered at the mean given by the trial parameters, allowed at least a tenfold speedup in the calculation. See Appendix B for further details.

[5]The observation time, T, is assumed to be 1 yr, and the mass efficiency is assumed to be $\lambda = 10^{-5} \, M_\odot^{-1}$ (as mentioned earlier).

[6]This reference local merger-rate density is roughly 5 times smaller than the realistic value quoted in Abadie *et al.* [11], but 20 times larger than the pessimistic estimate. Whilst we could scale our merger-rate density calculations to match the realistic value of $10^{-6} \, \text{Mpc}^{-3} \text{yr}^{-1}$, our modified likelihood statistic makes our analysis insensitive to such scalings. A rescaling to the realistic local merger-rate density of Ref. [11] would lead to an expected detection rate of approximately half a million sources.

We note that we have not included a prior on the scaling factors in the above, which is equivalent to using an improper flat prior over the range $[0, \infty]$. This reflects our current lack of knowledge of the mass efficiency.

D. Calculating the posterior probability

We use a weakly informative prior on the model parameters, so that it does not prejudice our analysis. As a prior on μ_{NS}, we take a normal distribution with parameters $\mu = 1.35 \, M_{\odot}$, $\sigma = 0.13 \, M_{\odot}$. This is motivated by the posterior predictive density estimate for a neutron star in a DNS binary system given in Ref. [45].

Given that ET will most likely not be operational until the mid-2020s, we must consider what constraints conventional observational cosmology can put on cosmological parameters. In the recent study by Zhao *et al.* [17], the authors investigated how the dark-energy EOS could be probed by ET observations of DNS systems to provide distance measurements, complemented by electromagnetic measurements of the associated short gamma-ray burst (sGRB) to provide the redshift. They estimated that a combination of the Planck cosmic microwave background (CMB) prior, JDEM BAO results, and future Type Ia supernova observations could provide cosmological constraints by the ET era of

$$\Delta \Omega_{m,0} = 3.46 \times 10^{-3}, \quad \Delta \Omega_{k,0} = 5.91 \times 10^{-4},$$
$$\Delta H_0 = 0.336, \quad \Delta w_0 = 0.048, \quad \Delta w_a = 0.184. \quad (31)$$

Hence, we assume that H_0, $\Omega_{m,0}$ and $\Omega_{k,0}$ are precisely known quantities, with values of $70.4 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, 0.2726 and -0.0006, respectively. As a prior on w_0 and w_a, we adopt the constraint that $w(z) < -1/3$ at all redshifts. Hence, we use uniform priors on the EOS parameters with $w_0 < -1/3$ *and* $w_0 + w_a < -1/3$ and lower bounds set low enough so as not to affect the posterior probability distribution. We also adopt uniform priors for all other parameters under investigation.

E. Bayesian analysis and an adaptive Markov chain Monte Carlo technique

Bayes' theorem states that the *posterior* probability distribution of the parameters $\vec{\mu}$ describing a hypothesis model $\mathcal{H}$ and given data D is given by

$$p(\vec{\mu}|D, \mathcal{H}) = \frac{\mathcal{L}(D|\vec{\mu}, \mathcal{H})\pi(\vec{\mu}|\mathcal{H})}{p(D|\mathcal{H})}, \quad (32)$$

where $\mathcal{L}(D|\vec{\mu}, \mathcal{H})$ is the *likelihood* (the probability of measuring the data, given the model $\mathcal{H}$ with parameters $\vec{\mu}$), $\pi(\vec{\mu}|\mathcal{H})$ is the *prior* (any constraints already existing on the model parameters) and finally $p(D|\mathcal{H})$ is the *evidence* (this is important in model selection, but in the subsequent analysis in this paper can be ignored as a normalization constant).

Markov chain Monte Carlo (MCMC) techniques provide an efficient way to explore the model-parameter space. An initial point, $\vec{x}_0$, is drawn from the prior distribution and then at each subsequent iteration, i, a new point, $\vec{y}$, is drawn from a *proposal distribution*, $q(\vec{y}|\vec{x})$, and the Metropolis-Hastings ratio evaluated,

$$R = \frac{\pi(\vec{y})\mathcal{L}(D|\vec{y}, \mathcal{H})q(\vec{x}_i|\vec{y})}{\pi(\vec{x}_i)\mathcal{L}(D|\vec{x}_i, \mathcal{H})q(\vec{y}|\vec{x}_i)}. \quad (33)$$

A random sample is drawn from a uniform distribution, $u \in U[0, 1]$, and if $u < R$ the move to the new point is accepted, and we set $\vec{x}_{i+1} = \vec{y}$. If $u > R$, the move is rejected, and we set $\vec{x}_{i+1} = \vec{x}_i$.

The MCMC samples can be used to carry out integrals over the posterior

$$\int f(\vec{x})p(\vec{x}|D, \mathcal{H})d\vec{x} \approx \frac{1}{N}\sum_{i=1}^{N} f(\vec{x}_i). \quad (34)$$

The 1D marginalized posterior probability distributions in individual model parameter follows by binning the chain samples in that parameter.

The trick to using this technique efficiently is to choose an appropriate proposal distribution. In our analysis, we employ an adaptive MCMC procedure, which utilizes an "in-flight" estimation of the sampled chain's covariance matrix to construct an updating proposal distribution. This covariance matrix is updated at each iteration, with a certain chain memory [73–75]. We use several of the procedures outlined in Ref. [75].

For the first 100 points in the chain, simple Gaussian proposal distributions for each individual parameter are used. These first 100 points are merely used to provide a starting point for the covariance matrix evaluation, and so the exact proposal distribution used in this stage is not important. After the first hundred points are sampled, we begin generating points via the adaptive procedure. For a D-dimensional target posterior distribution, we suppose that at the ith iteration, we have sampled at least H points, where the fixed integer H is the *memory* parameter. We then generate a D-dimensional vector of trial parameters, $\vec{y}$, via a linear mapping of an H-dimensional vector of unit-variance Gaussian random scalars, $\vec{\xi}$,

$$\vec{y} = \mathbf{C}^{1/2}\vec{\xi}, \quad (35)$$

where $\mathbf{C}^{1/2}$ is the positive-definite square root of the $D \times D$ covariance matrix evaluated using the previous H points. The covariance matrix may be calculated by collecting the previous H points in the chain into an $H \times D$ matrix $\mathbf{K}$, with each row representing one sampled point. Then,

$$\mathbf{C} = \frac{1}{H-1}\tilde{\mathbf{K}}^T\tilde{\mathbf{K}}, \quad (36)$$

where the centered matrix, $\tilde{\mathbf{K}}$, is constructed by centering each column of $\mathbf{K}$ around the means of the respective parameters, calculated from the H samples.
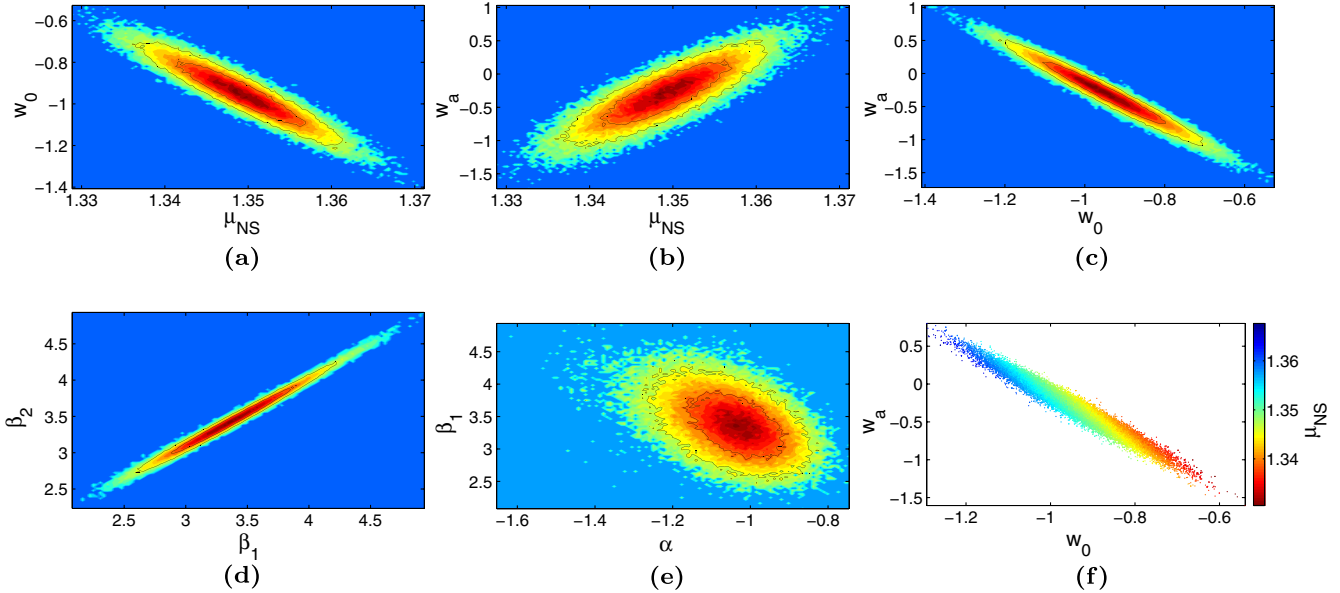
FIG. 3 (color online). Marginalized 2D posterior distributions for the reference catalog of 4500 detections. Only those 2D distributions showing correlations between parameters are shown. The reference parameters are $\mu_{NS} = 1.35 \, M_\odot$, $\sigma_{NS} = 0.06 \, M_\odot$, $w_0 = -1$, $w_a = 0$, $\alpha = -1$ and $\beta_1 = \beta_2 = 3.4$.

We then generate the trial parameter vector $\vec{y}$ via

$$\vec{y} \sim \mathcal{N}(\vec{x}, c_d^2 \mathbf{C}) \sim \vec{x} + \frac{c_d}{\sqrt{H-1}} \tilde{\mathbf{K}}^T \vec{\xi}, \qquad (37)$$

where c_d is a variable which depends only on the dimensionality of the target distribution. This variable is used to optimize the efficiency of the sampling process, and we use the value of $\approx 2.4/\sqrt{D}$ [75,76].

With a memory parameter which is less than the total past history of the chain, this is denoted as the adaptive proposal algorithm [73]. Since the proposal distribution is updated constantly and relies on previous chain information, this procedure is not Markovian, and does not have the correct ergodicity properties for a MCMC algorithm [73]. In principle, this can bias the reconstruction of the target posterior; however, this bias is ignorable in many practical applications and for well-behaved target posterior distributions [73,74]. If the entire previous chain is used to update the covariance matrix, then this algorithm is known as the adaptive Metropolis (AM) algorithm [74]. The AM algorithm does not suffer from the biases which can occur in the adaptive proposal algorithm, and ergodicity is retained.[7] We use the AM algorithm in our work.

---

[7]In the AM algorithm, the covariance of the proposal distribution is actually taken to be $\mathbf{C} + \epsilon \mathbf{I}_D$, where $\mathbf{I}_D$ is the D-dimensional identity matrix. Choosing $\epsilon > 0$ allows for the correct ergodicity properties of a MCMC algorithm to be retained, and in practice, is useful if the covariance of the chain has a tendency to degenerate. However, this parameter can be set very small with respect to the size of the target space, and in practice, can be set to zero.

VI. RESULTS

A. Posterior recovery

An analysis using the full 10^5 observations expected in a year of ET data is computationally prohibitive, so we use a working reference sample of $\sim$4500 detections (corresponding to a shorter observation time or a lower merger-rate density) and extrapolate to the expected number of detections, as discussed in Sec. VI C. For each analysis, we ran 120 independent adaptive MCMC chains of 5000 points on the same data catalog. We then used the last point from each chain to initialize a follow-up run of another 5000 iterations. The first 2000 points from each chain of the follow-up run were discarded as burn-in. This procedure therefore generated 360 000 points, with an average acceptance rate of $\sim$30%. The analysis of the 4500-event reference catalog took $\sim$3.5 hrs in total. Our sampled points were analyzed using the COSMOLOGUI package [77].

B. Marginalized posterior distributions

In Fig. 3, we show the recovered marginalized 2D posterior distributions (with 68% and 95% confidence contours) for the reference catalog. In Fig. 3(c), we observe a correlation between the recovered dark-energy parameters. This is easily explained by the fact that a given cataloged luminosity distance may be consistent[8] with a set of w_0 and w_a combinations, which will depend on the redshift of the source. Since the majority of detected systems will be

---

[8]Here, by "consistent," we mean within $\pm 1\%$ of the reference value.
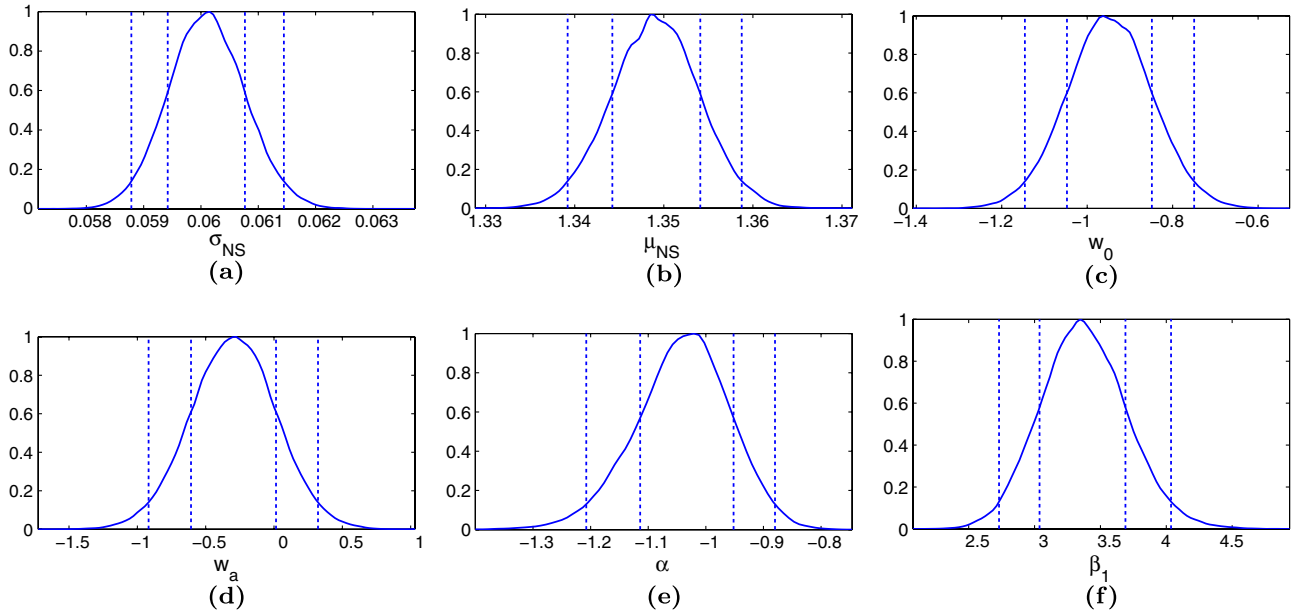
FIG. 4 (color online).   Marginalized 1D posterior distributions for the reference catalog of 4500 detections. Dotted lines indicate the boundaries of the 95% and 68% confidence intervals. The reference parameters are $\mu_{NS} = 1.35\ M_\odot$, $\sigma_{NS} = 0.06\ M_\odot$, $w_0 = -1$, $w_a = 0$, $\alpha = -1$ and $\beta_1 = 3.4$.

centered around $z \sim 1$, the $w_0 - w_a$ correlation will be dominated by these sources.[9] In Fig. 3(a), a negative correlation is observed between the recovered values for w_0 and μ_{NS}. For a given cataloged luminosity distance and fixed w_a, a low value of w_0 will imply a low redshift in that model. When this redshift is used to compute $\mathcal{M}$ from $\mathcal{M}_z$, we obtain a large value of the chirp mass, which is consistent with a chirp-mass distribution (and hence a NS-mass distribution) centered at larger values. Figure 3(b) merely shows the combined information of Figs. 3(a) and 3(c) (where the recovered w_a values are negatively correlated with the w_0 values); therefore, the correlation observed in Fig. 3(b) is positive.

A strong positive correlation is observed between the SFR-density SF2 ansatz parameters, β_1 and β_2, as seen in Fig. 3(d), while Fig. 3(e) shows a weak negative correlation between α and β_1. These correlations correspond to keeping the merger-rate density approximately constant. We calculated which combinations of α, β_1 and β_2 were consistent with a given merger-rate density, at a variety of redshifts. We found that there was a strong positive correlation in these points between β_1 and β_2, but the correlation between α and β_1 changed sign as the redshift increased. The greatest change occurred as the redshift was increased from 0 to 1, where the correlation then reversed;

however, at $z = 4$, the magnitude of the correlation was still not as large as it was at $z = 0.1$. This leads us to believe that although the D_L distribution of detected sources is peaked around $\sim$6 Gpc, with a long tail to $\sim$45 Gpc, the lower-distance sources dominate the $\alpha - \beta_1$ correlation, giving an overall negative correlation.

In Fig. 4, we show the marginalized 1D posterior distributions for the model parameters. The dotted lines in the plots indicate the 68% and 95% confidence regions of the marginalized distributions.[10]

C. Precision scaling with number of detections

We performed similar analyses on catalogs containing various numbers of detections, culminating in a run with 10^5 detections. We can characterize the precision with which we can measure the various model parameters by the 95% confidence intervals. Recording these intervals for all parameters for varying catalog sizes and dividing by the reference sample intervals gave the results shown in Fig. 5. This clearly shows that the precisions scale as $1/\sqrt{N_o}$ as we would expect. Parameter measurement

---

[9]The correlation between the two dark-energy EOS parameters can be reduced by rebinning the MCMC samples using the Wang parametrization [78]. This simply involves a transformation from the (w_0, w_a) parametrization to $(w_0, w_{0.5})$, where $w_{0.5} = w_0 + (w_a/3)$.

[10]While these results were computed using the fast merger-rate approximation, we also analyzed a catalog using the full merger-rate density. The 95% confidence intervals of the marginalized posterior distributions were consistent with our approximate analysis, justifying the use of the approximation to compute the rest of our results. No correlations between the merger-rate density parameters and the dark-energy EOS parameters were found, which supports our earlier statement that the dependence of the merger-rate density on the underlying cosmological parameters is weak within the applied priors.
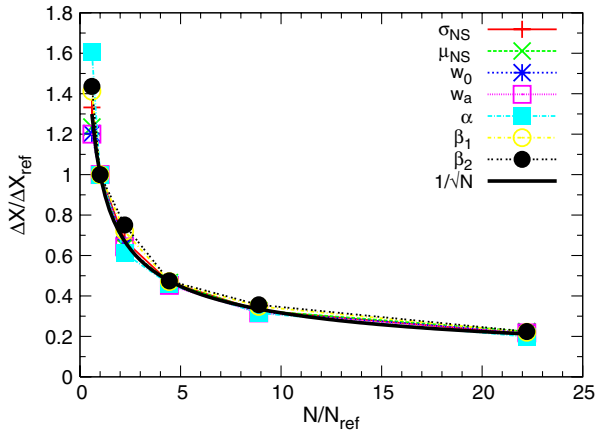
FIG. 5 (color online). 95% confidence intervals of the 1D marginalized distributions relative to those of the 4500-event reference catalog, shown as a function of the number of cataloged events. The same intrinsic parameters of the underlying distributions are used to create the mock catalogs. The expected $\sim 1/\sqrt{N}$ relationship is overlaid on the plot.

accuracies for the 10^5-event catalog are shown in Table III. We see that the measurement precisions of the dark-energy EOS parameters are the same order of magnitude as those forecast for CMB + BAO + SNIa [17], as discussed in Sec. V D.

D. Including and accounting for errors

Distance measurements from a third-generation GW-interferometer network will not be error-free. While a network consisting of a single ET plus one other right-angle interferometer can place constraints on a source's sky location and luminosity distance, the precisions of these properties are improved to almost the 3-ET network level by the inclusion of a second additional right-angle interferometer [35]. The redshifted chirp mass is expected to be very well-constrained ($\lesssim 0.5\%$ error [39]), and so we ignore measurement errors in this parameter. We assume the error in the luminosity distance arising from instrumental-noise scales as $\sim 1/\rho$ and include the ef-

TABLE III. 95% confidence intervals obtained from a catalog of 10^5 detections, with reference parameters used to generate the data. ΔX gives the width of the 95% confidence interval.

Parameter	Reference value	95% confidence interval	ΔX
$\sigma_{NS}/M_\odot$	0.06	[0.059688, 0.060254]	0.000566
$\mu_{NS}/M_\odot$	1.35	[1.347408, 1.351789]	0.00438
w_0	-1.0	[$-1.036403, -0.949623$]	0.0869
w_a	0.0	[$-0.195630, 0.073602$]	0.269
α	-1.0	[$-1.026691, -0.961659$]	0.0650
β_1	3.4	[3.318136, 3.605810]	0.288
β_2	3.4	[3.310287, 3.582895]	0.273

fects of weak lensing as a further source of error. The weak-lensing error on luminosity distance measurements at $z \sim 1$ is approximately 5%, and we linearly extrapolate this to all other redshifts [16,17,79,80]. While several techniques have been proposed to reduce this weak-lensing error [e.g., Refs. [81,82] and references therein], we assume no correction has been done, corresponding to a worst-case scenario.

Errors on the distance-redshift relation from binary-system peculiar velocities are much smaller than instrumental and weak-lensing errors at all but the lowest redshifts, becoming comparable with these at $z \sim 0.1$ where the error is $\lesssim 1\%$ and decreasing sharply at higher redshifts [Ref. [83] and references therein]. The lowest redshift in our reference catalog is ~ 0.05, where the peculiar velocity errors will dominate, but only lead to an error of $\lesssim 2\%$. The sensitivity of the luminosity distance to the dark-energy EOS parameters is very weak in this redshift regime; hence, peculiar velocities are unlikely to introduce significant parameter bias/inaccuracy, and we ignore them here.

We also ignore the effect of detector calibration errors, which, unlike statistical measurement uncertainties, would not be mitigated by boosting the detection rate. Such systematic biases have recently been studied in the case of advanced-era detectors [84] and found to induce a systematic shift in the estimated system parameters which is a small fraction of the statistical measurement errors. We ignore waveform-modelling errors in our analysis, since current post-Newtonian models will only break down close to the onset of the merger phase, and for the neutron-star binaries considered in this analysis, this is at frequencies where the instrumental noise is high and which therefore do not contribute much to the overall signal-to-noise ratio of the system. Furthermore, luminosity-distance determination comes primarily from the network triangulation which will not be significantly affected by modelling uncertainties, and so the distance error will be dominated by instrumental noise and weak lensing, as discussed earlier. Similarly, the error in the distribution of possible source redshifts arising from the measured redshifted chirp mass will be dominated by the intrinsic width of the NS mass distribution rather than the small error in the redshifted chirp mass coming from instrumental noise and modelling uncertainties.

We repeat the run of the working reference sample, offsetting the cataloged luminosity distance by an amount drawn from a Gaussian distribution, with mean at the true distance, and standard deviation,

$$\sigma = D_L \times \sqrt{(1/\rho)^2 + (0.05z)^2}. \qquad (38)$$

The data collected for each event will actually be in the form of posterior probability density functions (PDFs) for the parameters, where previously we have assumed these are δ-functions at the true values. If the offset luminosity

TABLE IV. 95% confidence intervals derived from the reference sample (4500 detections), both in the case where distances are measured precisely, and when distance errors are included and accounted for (using the error averaging technique described in the text, with various numbers of points sampled from the distance posterior PDF). ΔX gives the width of the 95% confidence interval.

Parameter	No errors 95% confidence interval	Errors (50 points) 95% confidence interval	$\Delta X/\Delta X_{\text{ref}}$	Errors (100 points) 95% confidence interval	$\Delta X/\Delta X_{\text{ref}}$	Errors (400 points) 95% confidence interval	$\Delta X/\Delta X_{\text{ref}}$
$\sigma_{\text{NS}}/M_\odot$	[0.058785, 0.061447]	[0.066378, 0.071911]	2.07851	[0.063815, 0.069409]	2.10143	[0.059309, 0.064849]	2.08114
$\mu_{\text{NS}}/M_\odot$	[1.339198, 1.358745]	[1.329060, 1.354066]	1.27928	[1.335499, 1.361690]	1.3399	[1.335782, 1.359339]	1.20515
w_0	[−1.145894, −0.749671]	[−0.880092, −0.338642]	1.36653	[−1.146052, −0.537588]	1.53566	[−1.116809, −0.546566]	1.4392
w_a	[−0.917590, 0.321901]	[−2.345082, −0.452114]	1.52722	[−1.651230, 0.463072]	1.70578	[−1.605299, 0.377340]	1.59956
α	[−1.207554, −0.879888]	[−1.215388, −0.863579]	1.07368	[−1.208005, −0.874856]	1.01673	[−1.198613, −0.890136]	0.941437
β_1	[2.730152, 4.036099]	[2.851895, 4.260085]	1.07829	[2.729217, 4.038989]	1.00293	[2.771780, 4.069150]	0.99343
β_2	[2.842474, 4.059874]	[2.954584, 4.274131]	1.08391	[2.863813, 4.080100]	0.999088	[2.887584, 4.100009]	0.995918

distances are assumed to be the true values with a δ-function posterior PDF, the chain does not move away from its starting point. Hence, we must take these errors into account in the likelihood calculation stage.

We can account for these errors in the analysis by modifying the likelihood in Eq. (29) [85] to

$$\mathcal{L}(\vec{\Lambda}|\vec{\mu}, \mathcal{H}) = e^{-N_\mu} \iint \ldots \int \left[p\left(\vec{n} = \vec{s} - \sum_i \vec{h}_i(\vec{\lambda}_i)\right) \right.$$
$$\left. \times \prod_{i=1}^{N_o} r(\vec{\lambda}_i|\vec{\mu}) \right] \mathrm{d}^k\vec{\lambda}_1 \mathrm{d}^k\vec{\lambda}_2 \ldots \mathrm{d}^k\vec{\lambda}_{N_o}, \quad (39)$$

where, in our case, the number of parameters $k = 2$, and $\vec{s}$ is the detector output, which is a combination of N_o signals, $\vec{h}_i$, and noise, $\vec{n}$. Equation (39) is an integral over all possible values of the source parameters which are consistent with the data. The first term inside the square bracket is the computed posterior PDF for the detected population of sources. Concern has been raised that the high event rate of ET detections may lead to a confusion background. However, the *noise-power-spectral-density-weighted* signals are short enough that there is not expected to be significant overlap [16,39]. Hence, these detections should be uncorrelated, with independent parameter estimates, and so this first term reduces to the product of the posterior PDFs for each detection.

If the posterior PDF for a given source has been obtained via MCMC techniques, then the integral in Eq. (39) may be computed by summing over the chain samples. Thus, errors may be accounted for by making the following replacement in Eq. (29):

$$r(\vec{\lambda}_i|\vec{\mu}) \rightarrow \frac{1}{\mathcal{N}_i} \sum_{j=1}^{\mathcal{N}_i} r(\vec{\lambda}_i^{(j)}|\vec{\mu}), \quad (40)$$

where $\mathcal{N}_i$ is the number of points in the chain for the ith source's PDF, and $\lambda_i^{(j)}$ is the jth element of the discrete chain representing this PDF. This technique does not assume a specific form for the PDF and can be used in the case of multimodal distributions.

We represent the D_L posterior PDF for each source by a chain of 50 points, drawn from a normal distribution with standard deviation as in Eq. (38), and a mean equal to the value in the data catalog, which in this analysis, as discussed earlier, includes an error to offset it from the true value. Results are shown in Table IV. We see that a significant bias in the reconstructed model parameters still exists. We suspected that this bias arose from using only 50 points to evaluate the distance posterior PDFs. We therefore repeated the analysis with an increasing number of points sampled from the distance posterior PDF.[11] With 100 points, all biases are corrected expect for that in σ_{NS}, and the ratio of the 95% confidence interval widths to the reference widths is not significantly different from the 50-point case. This suggests that a larger number of points in the error-averaging technique will be necessary to correct all biases, but this is not necessary to estimate parameter-measurement accuracies in the presence of distance errors. Using 400 points sampled from the distance posterior PDF, all biases in the parameter posterior distributions appear to be corrected, in the sense that the reference parameters then lay within the 95% confidence intervals of the 1D marginalized posterior distributions.

Overall, we find that the result of properly accounting for instrumental and weak-lensing errors is that parameter-measurement precisions are, at worst, approximately halved. Thus, instrumental and weak-lensing induced errors should not affect our general conclusions about the science capabilities of a third-generation GW-interferometer network. We carry out the remainder of this study using catalogs which are generated and analyzed without including errors.

---

[11]The posterior distributions obtained via this analysis should be considered estimates of the true distributions, since the long likelihood computation time required by this error analysis means that we did not collect as many samples as when errors are ignored. We performed burn-in runs and then follow-up runs to estimate the posterior distributions as well as feasible.

TABLE V. The expected detection rates for different choices of intrinsic parameters of the underlying distribution are compared to the reference expected detection rate. One parameter is varied at a time, with all other parameters kept fixed at their reference values. The variation of the expected detection rate with the intrinsic value of σ_{NS} is not shown, since this parameter is not used in the rate calculation (see Appendix B).

$\mu_{NS}/M_\odot$	N/N_{ref}	w_0	N/N_{ref}	w_a	N/N_{ref}	α	N/N_{ref}	β_1	N/N_{ref}	β_2	N/N_{ref}	$\beta_1 = \beta_2$	N/N_{ref}
1.31	0.952	-1.50	1.08	-0.50	1.02	-1.10	1.04	2.90	0.405	3.40	1.00	3.00	0.929
1.33	0.976	-1.25	1.05	-0.25	1.01	-1.00	1.00	3.00	0.475	3.60	0.696	3.20	0.966
1.35	1.00	-1.00	1.00	0.00	1.00	-0.90	0.958	3.10	0.561	3.70	0.595	3.40	1.00
1.37	1.02	-0.75	0.935	0.25	0.986	-0.60	0.810	3.20	0.671	3.80	0.514	3.60	1.03
1.39	1.05	-0.50	0.844	0.50	0.970	-0.50	0.757	3.40	1.00	4.00	0.394	3.80	1.06

E. Precision scaling with intrinsic parameters

We now investigate how the ability of ET to constrain the parameters of the underlying distributions is affected by the values of the intrinsic parameters themselves. This is similar to the kind of analysis performed in our previous study with second-generation interferometers [22]. We perform analyses of data catalogs generated with different intrinsic parameter combinations; multiple runs are performed on each parameter combination. We vary one parameter at a time, fixing all others to the reference values.

Varying the intrinsic parameters with fixed SNR threshold will alter the expected detection rate. This is illustrated in Table V, where the model with reference parameters has an expected detection rate of $\sim 10^5$ yr^{-1}. The expected

$\sim 1/\sqrt{N}$ relationship is well established, as shown in Fig. 5. Hence, we remove this number effect by generating catalogs with the same number of events (4500 each in order to compare against the reference catalog). Therefore, we are testing how the cosmological, astrophysical and intrinsic-mass *distributions* of coalescing DNS binaries impact the precision of parameter recovery.

The results of these analyses are shown in Fig. 6. We see that as σ_{NS} is increased, the measurement precision of both the NS mass distribution and dark-energy EOS parameters decreases. We found a similar trend in Ref. [22]. This makes sense, since if we have an intrinsically narrow NS mass distribution, then we have a good idea of what the intrinsic masses of the systems are, and the range of
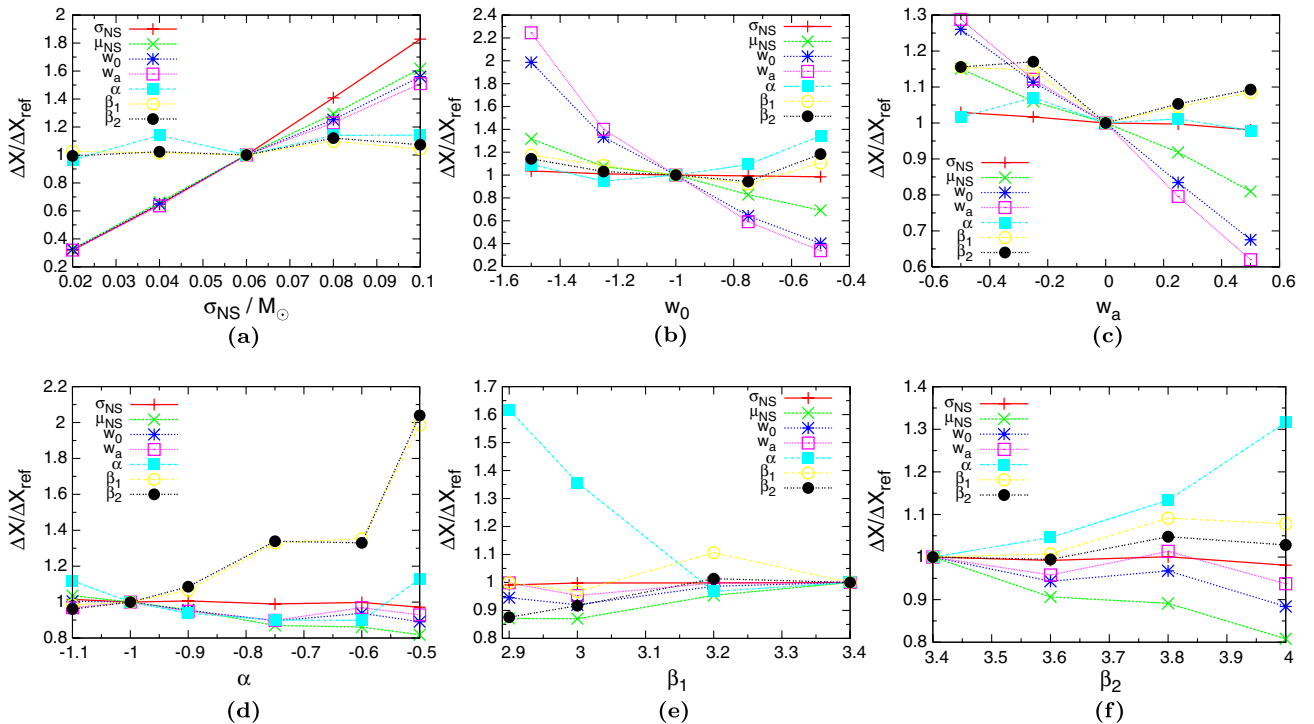

FIG. 6 (color online). The variation of measurement precision with different choices of the intrinsic parameters of the underlying distributions. One parameter is varied at a time, and in the interest of testing how the precision of parameter recovery is affected by the underlying distribution of events, all catalogs are generated with the same number of events (4500 to match the reference catalog). Each point in each panel represents the average 95% confidence interval width of 3 realizations of the catalog.

candidate redshifts produced from a measured redshifted chirp mass will be narrow, improving the precision with which we can recover cosmological parameters.

A variation in the intrinsic μ_{NS} (not shown) produces accuracies comparable to the reference accuracies. Hence, the impact of the intrinsic value of the NS mass-distribution mean on parameter accuracies is predominantly through the change to the expected detection rates, i.e., a larger mean, on average, will lead to larger chirp masses, so that detections can be made from a larger volume [see Eq. (5)].

Increasing the value of the EOS parameter w_0 increases the precision with which we can recover w_0, w_a and μ_{NS}. As w_0 is increased, while the intrinsic w_a is fixed at zero, the recovered posteriors for these parameters are squeezed by the prior restrictions, $w_0 < -1/3$ and $(w_0 + w_a) < -1/3$. A larger intrinsic w_0 increases the horizon distance of detections, which permits greater sensitivity to the dark-energy EOS parameters. Furthermore, a narrowed range of cosmological parameters means that the range of candidate redshifts is also narrowed, such that the precision of the recovered NS mass-distribution mean (deduced from the redshifted chirp mass) improves. We also notice these effects when the intrinsic w_a is increased with the intrinsic w_0 fixed at the reference value. However, the effect is less pronounced in this case, since w_a is a first-order correction to w_0.

As the power-law index, α, is increased the average delay between the formation of the massive progenitor system and the merging of the final compact system increases. This means that more systems formed at higher redshifts survive to merge at lower redshifts, and hence the merger-rate density is boosted to higher values at lower redshifts. In addition, as α increases, the merger-rate density tracks the underlying SFR density to a lesser extent, so it becomes more difficult to extract the details of the SFR density. Hence, the $\beta_{1,2}$ distributions widen to reflect this reduced sensitivity to the underlying SFR density.

When we keep the intrinsic values of β_1 and β_2 equal (not shown), we find that varying these by ± 0.4 around the reference value has a negligible impact on the measurement precision of the parameters. A higher common $\beta_{1,2}$ value leads to a larger expected detection rate, but this is a small effect.

Lowering the intrinsic value of β_1, with β_2 fixed, shifts the distribution of events to lower distances and changes the shape of the underlying merger-rate density. This distribution is consistent with a wider range of α values than the reference distribution, since the sensitivity of the merger-rate density to α is reduced at lower redshifts. This causes the marginalized α-posterior distribution to widen. The same is true when the intrinsic value of β_2 is increased.

F. Varying the SNR threshold

We also generated catalogs for different SNR thresholds defining the detectability of merging systems. Multiple catalogs were analyzed for each SNR threshold, but once
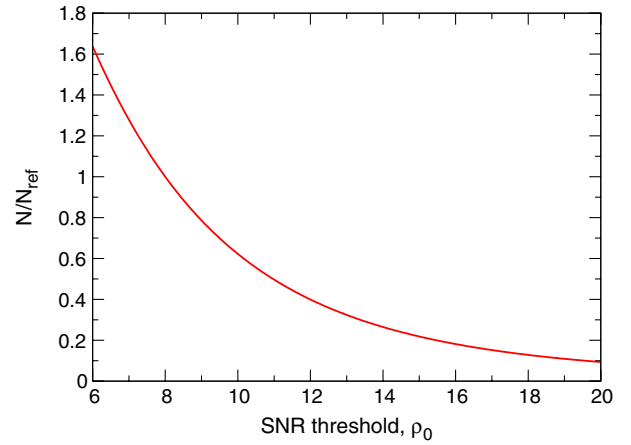

FIG. 7 (color online). We show the variation of the expected detection rate as the SNR threshold, ρ_0, is raised. This can also be interpreted as *lowering* the characteristic distance reach of the network. Since ρ scales as $1/D_L$ [see Eq. (5)], and the difference between the luminosity distance and radial comoving distance becomes smaller at lower redshift, one would expect that at high enough values of ρ_0, the comoving detection volume (and hence detection rate) would scale as $1/\rho_0^3$. This is approximately valid for $\rho_0 \gtrsim 15$.

again, the number of events was fixed at 4500 to match the reference catalog (see Fig. 7). An increase in the SNR threshold is equivalent to the characteristic distance reach of the detectors decreasing. Hence, we would expect the sensitivity of the data to varying dark-energy EOS parameters, which have a greater influence at larger redshifts, to be reduced. A reduced characteristic reach would also result from a larger low-frequency cutoff, f_l, in the detector's noise power spectral density. In the recent mock ET data challenge, Regimbau *et al.* [39] found that confusion between two or more signals rarely affected the analysis performance when $f_l = 25$ Hz. Standard algorithms currently employed for LSC-Virgo analyses cannot handle templates longer than a few minutes; however, multiband filter methods are being developed which will allow f_l to be pushed below 25 Hz. In Fig. 8, one can see that with $f_l = 25$ Hz, the effective SNR threshold is raised from the reference value of 8 (with $f_l = 1$ Hz) to ~ 12.4.

From Fig. 9(a), we see that as the SNR threshold is increased, with the number of cataloged events fixed, the accuracies of β_1 and β_2 degrade sharply. At higher SNR thresholds (or, equivalently, at lower distance reaches), the sensitivity of the merger-rate density to varying $\beta_{1,2}$ is reduced, hence the wider posterior distributions. The measurement precision of α increases slightly as the SNR threshold is increased from 6–12. One might expect α to show the same trend as β_1 and β_2, since an increasing SNR threshold pushes the events to lower redshifts where the sensitivity of the merger-rate density to α is reduced. However, we see in Fig. 10 that the merger-rate density, for various choices of α (but all other parameters fixed), is relatively featureless beyond ~ 2.5. The distribution of the merger-rate density in
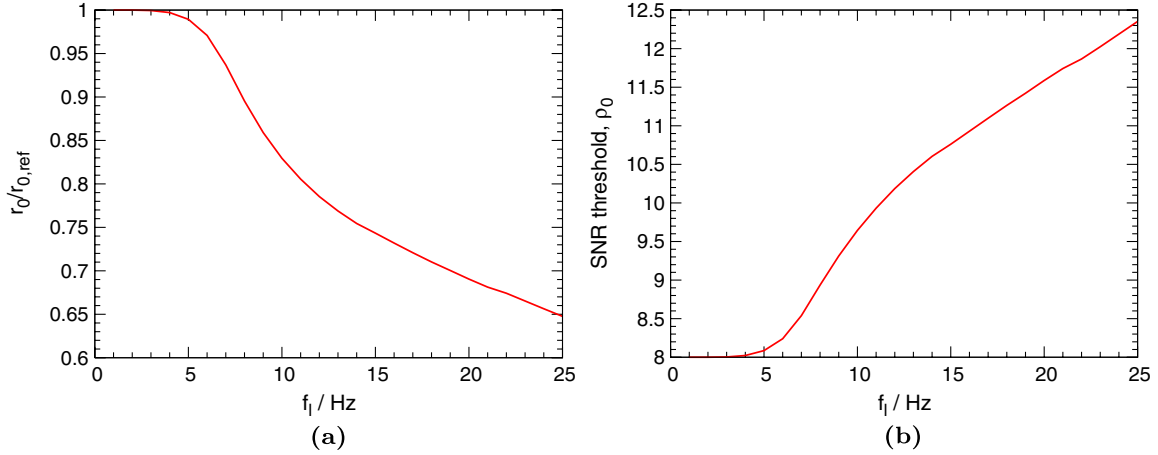
FIG. 8 (color online).   The reduction of the characteristic distance reach associated with raising the low-frequency cutoff, f_l, of 3rd-generation detectors. This can also be interpreted, via Eq. (28), as raising the network SNR threshold. The figures were produced using the ET-D noise curve [32].
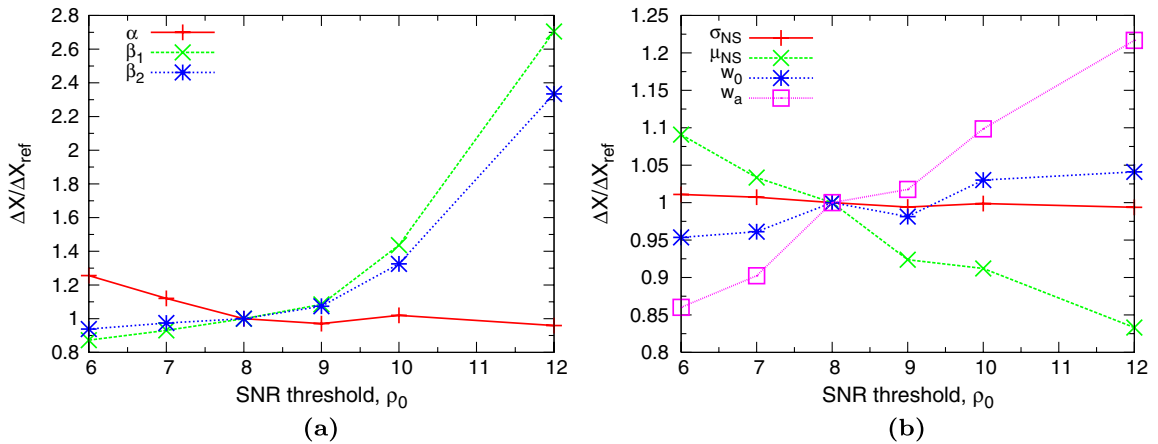

FIG. 9 (color online).   The variation of the parameter measurement precisions with the network SNR threshold. The left panel shows precisions, characterized by the width of the 68% confidence intervals, for the merger-rate density parameters, while the right panel shows precisions, characterized by the 95% confidence intervals used elsewhere, for all other parameters. We use the narrower confidence intervals for the merger-rate parameters to mitigate the effect of poor sampling in the low-α region which was observed in some AM-MCMC chains in this analysis. All catalogs contain the same numbers of events at each threshold value, which, as in the previous subsection, is 4500 to match the reference catalog.

the redshift window of ~ 2.5–7 could be approximately linearly scaled to satisfy a large range of α.[12] Therefore, given that our likelihood statistic is insensitive to linear scalings of the merger-rate density [see Eq. (30)], the significant number of high-redshift detections in a $\rho_0 = 6$ catalog will widen the α posterior distribution, while most α information is found in the redshift window ~ 1–2, where the merger-rate density has more features.

In Fig. 9(b), we see that the measurement accuracy of w_0 and w_a is slightly reduced for higher SNR thresholds; this is a small effect and is expected with a catalog shifted to lower redshifts, where distances are less sensitive to varying dark-energy EOS parameters. The accuracy of w_0 only varies by $\sim \pm 5\%$, since we remain sensitive to detections at tens of Gpcs even with an SNR threshold of 12. However, w_a shows a stronger variation since it is a higher order correction to the EOS parameter, and distances become sensitive to this parameter at higher redshifts than they do to w_0. We also see that the measurement precision of μ_{NS} and σ_{NS} is increased slightly as we move to larger SNR thresholds. This small effect is probably due to the fact that a lower redshift range in the data

---

[12]The same argument did not apply when $\beta_{1,2}$ was varied, since this not only shifted the distribution to lower redshifts but altered the *shape* of the merger-rate density in a way which could not be equated with a linear scaling in any redshift window.
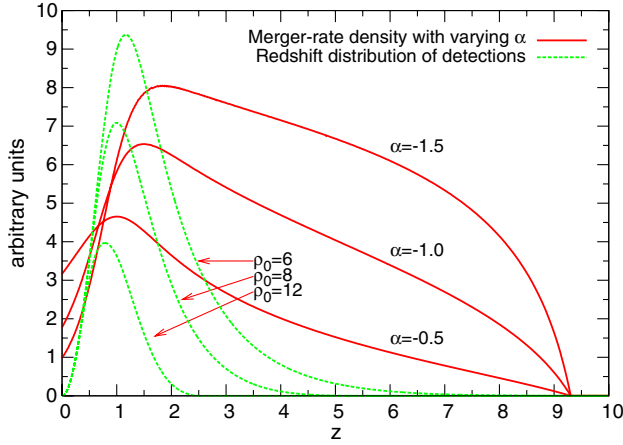
FIG. 10 (color online).   We show the redshift distribution of the DNS merger-rate density for various choices of the power-law index of the delay-time distribution, α (all other parameters are fixed at their reference values). The merger-rate density is relatively featureless beyond $z \sim 2.5$, making it difficult for our analysis (which is insensitive to linear scalings of the merger-rate density) to discriminate between values of α in this range. Overlaid on this figure, we show the redshift distribution of detections for various SNR thresholds.

catalog will mean that the redshifted chirp mass is closer to the intrinsic chirp mass.

Although this suggests that a greater distance reach will improve the precision of cosmological parameter recovery, we have so far ignored distance errors. In fact, instrumental and weak-lensing errors impart an interesting redshift evolution to the w_0 sensitivity, which we approximate as [86]
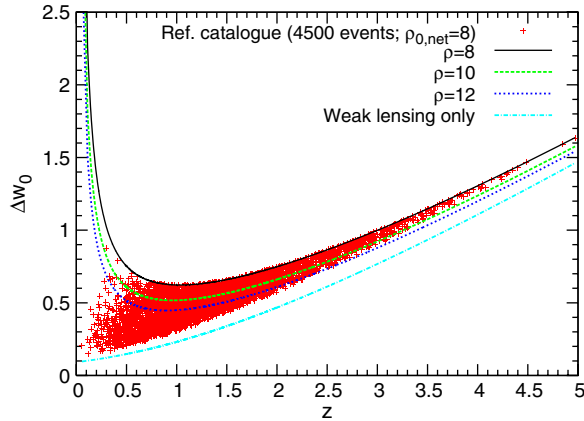

FIG. 11 (color online).   The redshift dependence of the sensitivity of the luminosity distance to w_0. This parameter has a very weak intrinsic impact on D_L at low redshifts, while distance errors from instrumental noise and weak-lensing dominate at higher redshifts. This results in a redshift "sweet spot," where these effects are minimized for lines of constant SNR. We also plot the individual Δw_0 values calculated for the reference catalog events.

$$\Delta w(z) \sim \left| \frac{\partial w_0}{\partial D_L} \right| \times D_L \times \sqrt{(1/\rho)^2 + (0.05z)^2}. \quad (41)$$

In Fig. 11, we see that the sensitivity of the luminosity distance to the cosmological parameter w_0 is greatest at $z \sim 1$, since w_0 has a very weak intrinsic impact on D_L at low redshifts and distance errors dominate at higher redshifts. Increasing a detector's distance reach will raise the fraction of high-redshift cataloged events. We calculate the effective measurement precision of w_0 from our reference catalog by adding the Δw_0 values from each event in quadrature, i.e., $1/\Delta w_{0,\mathrm{eff}} = \sqrt{\sum (1/\Delta w_{0,i})^2}$. This is repeated for various lower and higher SNR threshold values. We perform these calculations for catalogs containing the same number of events (4500 to match the reference catalog) and for catalogs with the number of events scaled by the ratio of the expected detection rate for each SNR threshold to the reference threshold (which in this analysis is 8). The results are shown in Table VI, where we see that for catalogs with the same number of events, lowering the SNR threshold actually worsens the precision of w_0 recovery since the distribution of events is weighted to higher redshifts, where distance errors degrade the precision. Increasing the SNR threshold reduces the number of events at high redshift and hence mitigates the degradation of precision due to distance errors (see Fig. 11). However, this effect slows down with increasing SNR threshold. For catalogs with numbers of events scaled to match the expected detection rate for each SNR threshold, we see that the increased number of events associated with a lower SNR threshold is enough to compensate for degradation of precision from higher redshift events. However, this loss of precision means that lowering the SNR threshold does not lead to the $1/\sqrt{N}$ or $1/\rho_0^{3/2}$ improvement in parameter measurement precision which one would naively expect.

Finally, we address the issue of having assumed that Earth motion does not modulate the antenna patterns of the detectors. The time spent "in band" by an inspiraling event scales as [39]

TABLE VI.   The events from catalogs with different SNR thresholds are used to compute an effective w_0 precision, by adding the Δw_0 values of each event in quadrature. This analysis is performed for catalogs with the same number of events, and for catalogs with the numbers of events scaled to match the expected detection rate for each SNR threshold.

$\rho_{0,\mathrm{net}}$	$f = N/N_{\mathrm{ref}}$	$\Delta w_{0,\mathrm{eff}}/10^{-3}$	
		$(N_o = N_{\mathrm{ref}})$	$(N_o = f \times N_{\mathrm{ref}})$
6	1.64	8.08	6.33
8	1.00	6.82	6.82
12	0.399	5.33	8.35
20	0.0936	4.00	13.1
30	0.0271	3.38	20.6

$$\tau \sim 5.4 \left(\frac{\mathcal{M}_z}{1.22\ M_\odot} \right)^{-5/3} f_l^{-8/3}\ \text{days}. \qquad (42)$$

Hence, a detector with a low-frequency cutoff of 1 Hz (as we have assumed) could have events in band for as long as ~ 5 days. In this case, a correct treatment of the antenna pattern modulation would be needed. However, if we increase f_l to ~ 8 Hz, then the maximum time spent in band is less than 30 minutes, and ignoring the antenna pattern modulation is reasonable. In Fig. 8, we see that a low-frequency cutoff of 8 Hz is equivalent to raising the SNR threshold to ~ 9, and from Fig. 9, we see that the precision of parameter recovery is within $\sim \pm 10\%$ of the reference precisions for an SNR threshold of 9. Therefore, our approximate treatment of the network antenna patterns would seem reasonable.

VII. CONCLUSIONS

We have built on our previous work [22] which explored the capabilities of an advanced (i.e., second-generation) GW-interferometer network to constrain aspects of the NS mass distribution in DNS systems, as well as the Hubble constant. The technique we employed used only information obtained via analysis of the GWs detected in such a network. In this paper, we extended the analysis to a possible third-generation network, consisting of the proposed Einstein Telescope, and complemented by third-generation right-angled interferometers at LIGO Livingston and LIGO-India. The target for the Einstein Telescope is a broadband factor of 10 sensitivity increase with respect to advanced detectors, but to also extend the low-frequency sensitivity of ground-based GW interferometers below 10 Hz. The current design for a single ET consists of three overlapping interferometers, arranged in an equilateral configuration with arm-opening angles of 60° [26–28]. Each interferometer will consist of a cryogenically cooled, underground low-frequency detector, and a high laser-power, high-frequency detector in a xylophone configuration—these two detectors work in tandem to suppress noise over the entire band [29,34]. Current projections for funding and construction of ET place "first light" sometime in the mid-2020s.

The sources of interest in this paper are inspiraling double NS systems, which could be observed at rates of $\sim 40\ \text{yr}^{-1}$ by advanced detectors [11], and rates of $\mathcal{O}(10^5\text{–}10^6)\ \text{yr}^{-1}$ may be achieved by a third-generation network [16,28,39]. These sources are commonly referred to as self-calibrating standard sirens, since their distance from us is directly encoded in the emitted GWs. Combined with a method of redshift determination, these sources can be used to probe the distance-redshift relation and hence extract constraints on background cosmological parameters which are independent of the cosmic distance ladder [14–17].

Our method of cosmography using only GWs relies on the narrowness of the distributions of masses of NSs in these DNS systems. Recent analysis indicates that this mass distribution is indeed narrow, with a Gaussian mean of $\sim 1.35\ M_\odot$ and standard deviation of $0.06\ M_\odot$, which may be a product of a distinct evolutionary path [45,46,49]. Using a measurement of a source's redshifted chirp mass, we can therefore obtain a narrow candidate redshift distribution. A narrower intrinsic NS mass distribution will obviously mean the precision of redshift determination increases. We can combine these with GW-interferometer network determinations of the luminosity distance to constrain cosmological parameters.

We used a Bayesian theoretical framework to assess the capability of a third-generation network to measure cosmological and astrophysical parameters. We performed 7-dimensional adaptive MCMC analysis on the catalogs of detections, using reference parameters $H_0 = 70.4\ \text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$, $\Omega_{m,0} = 0.2726$, $\Omega_{k,0} = -0.0006$, $w_0 = -1$, $w_a = 0$, $\mu_{\text{NS}} = 1.35\ M_\odot$ and $\sigma_{\text{NS}} = 0.06\ M_\odot$. Keeping H_0, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ fixed, we found that the measurement precisions of the dark-energy EOS parameters possible with a 10^5-event catalog were of the same order of magnitude as forecasted constraints from future CMB + BAO + SNIa measurements [17]. Furthermore, the power-law index of the merger delay-time distribution, α, and the parameters of the underlying star-formation-rate density were constrained to within $\sim 10\%$. Accounting for measurement errors degraded precisions by a factor of $\lesssim 2$, while increasing the network SNR threshold required for detection from 8 to 9 (which is equivalent to considering only the ~ 30 minute section of inspiral above 8 Hz) changed the precisions by only $\sim 10\%$.

We also investigated how the precision of parameter recovery scaled with the values of the intrinsic parameters themselves, keeping the number of detected events fixed to factor out pure number-of-event effects. Varying the intrinsic σ_{NS} showed a linear scaling of parameter precision, with narrower intrinsic NS mass distributions favoring tighter parameter constraints. The precisions of the merger-rate density parameters did not appear to be affected in this case. Increasing the intrinsic w_0 and w_a had the effect of increasing their measurement precision, as well as that of μ_{NS}. This was probably due to the fact that larger w_0 and w_a give detections out to greater distances, where the sensitivity to these parameters is higher. Tighter cosmological constraints implies narrower candidate redshift distributions from the cataloged distances, which improves μ_{NS} precision. Increasing the intrinsic value of α meant that the merger-rate density tracked the underlying SFR density to a lesser extent and hence worsened the precision of SFR-density parameter recovery. As we changed the shape of the underlying SFR density to favor closer detections, the measurement precision of α worsened, since the sensitivity of the merger-rate density to α is lower at lower redshifts.

Finally, we varied the criterion for a network detection, which we denoted by a threshold value of the network signal-to-noise ratio. This could also be interpreted as varying the characteristic distance reach of the network, which, in turn, could be caused by varying the detector's low-frequency cutoff. Varying the SNR threshold between 6–12 caused a slight decrease in w_0 and w_a precision, as catalogs with lower distance events will be less sensitive to these cosmological parameters. However, catalogs with, on average, closer events will provide better NS mass-distribution parameter precision, since the redshifted chirp mass will be less offset with respect to the intrinsic chirp mass. Increasing the SNR threshold, and hence decreasing the characteristic distance reach of the network, caused a significant decrease in SFR-density parameter precision, since the merger-rate density is less sensitive to the SFR-density parameters at lower redshift.

While the sensitivity of distances to the dark-energy EOS will obviously be intrinsically weak at low redshifts, distance-measurement errors begin to dominate at higher redshifts. We found that for a fixed number of events in a catalog, lowering the SNR threshold actually worsened the precision of w_0 recovery since events are weighted to higher redshifts, where distance errors degrade the measurement precision. The larger expected detection rate associated with lower SNR thresholds is enough to reverse this effect, but means that lowering the SNR threshold (or increasing the network's distance reach) does not lead to the great improvement in parameter measurement precision which one would naively expect.

We have not considered association of GW detections with an EM counterpart, either through precise sGRB [16,17] or host-galaxy association. The latter technique may only be possible with ∼0.01% of detectable GW events [31]. However, the technique we have used has been shown previously [22] to be well complemented by precision redshift information. In particular, we found that if redshift information (measured to much greater precision than the luminosity distance) is available for ∼10% of a GW-event catalog, then measurement precisions of parameters were more than doubled.

This paper completes our proof-of-principle study of this GW-only cosmographic technique. We have shown the significant potential for a third-generation network including the Einstein Telescope to place interesting constraints on the NS mass distributions in DNS systems, the dark-energy EOS, the average delay between the formation of the DNS-system progenitors and the final merger and the underlying SFR density in the Universe. Over the following decade, tighter constraints will be derived for the NS mass distribution, delay-time distribution of DNS systems, and the SFR density, which can be readily incorporated within this technique. We intend to test this technique in the upcoming ET mock data challenge, as well as study the ability of this technique to discriminate between

NS mass distributions from different metallicity progenitors, different delay-time distributions resulting from different formation paths and possibly multimodal NS mass distributions. Unshackling GW cosmography from its reliance on EM counterparts will be an important step in establishing DNS systems as physical distance indicators and contribute to GW analysis becoming a precision astrophysical tool.

APPENDIX A: DNS MERGER-RATE DENSITY

Here, we provide a more detailed discussion of the astrophysics of DNS mergers and justification for the ansatz we employ for the merger-rate density.

A. Merger-delay distribution, dP_m/dt

Assuming the number, N, of DNS binaries born with separation a follows $dN/da \propto a^\gamma$ [52], the merger-delay distribution is

$$\frac{dP_m}{dt} \propto \frac{dN}{d\tau_{\rm gr}} = \frac{dN}{da}\frac{da}{d\tau_{\rm gr}} \propto t^{\gamma/4}t^{-3/4} = t^\alpha. \quad \text{(A1)}$$

An early suggestion by Piran was to consider newly formed DNS binaries as having the same orbital separation distribution as normal-abundance main-sequence stars [54,55]. For normal-abundance main-sequence binary systems, the distribution of periods has been found to be flat in $\ln(P)$ (where P is the binary period) [55], or lognormal [87,88]. If we follow Ref. [54], and ignoring the progenitor ellipticity distribution, then the initial DNS orbital separation distribution will be flat in $\ln(a)$ (where a is the semi-major axis of the binary orbit). Therefore, $\gamma = -1$ and $dP_m/dt \propto t^{-1}$.[13]

---

[13]The caveats here are that ellipticity can have a significant effect on inspiral time scales, and it is not obvious that DNS systems should have the same orbital separation distribution as main-sequence binaries, since the two supernovae the systems survive would likely modify it. Furthermore, the distribution functions for progenitor evolutionary time scales and merger time scales are not independent. The evolutionary time scale depends on the mass of the progenitor system components, and the gravitational inspiral time scale depends on the chirp mass of the system. Strictly speaking, the joint probability consideration should be considered [53]; however, we ignore this subtlety here.

The catalog of DNS systems in Ref. [56] is used by Refs. [57,58] to estimate the merger-delay distribution of observed systems; it approximately follows $(1/t)$, but there appears to be an excess of systems below an inspiral time of 100 Myr. Selection effects having to do with the difficulty of detecting binary pulsars in close orbit, due to the large and rapidly varying Doppler shift, may significantly affect the reconstructed merger-delay distribution below a few hundred Myr. The authors comment that with such a small sample of systems (~ 6), it is difficult to make predictive conclusions about this distribution, but that it is the best one can presently do with the observations.

Population synthesis calculations in Refs. [50,53] appear to show the cumulative merger-delay distribution being approximately linear in $\ln(t)$ (in which case, the PDF varies as $\sim t^{-1}$), while the studies in Refs. [59–61] show that a $(1/t)$ PDF is an appropriate approximation over several orders of magnitude of the delay-time. Furthermore, the population synthesis calculations of Ref. [62], in their study of the formation rates of short- and long-GRBs, indicate an approximate $(1/t)$ delay-time distribution for NS-NS and NS–black-hole systems.

Population synthesis calculations have also proposed previously unconsidered DNS-formation channels; specifically, Ref. [89] suggests a formation channel via a double common-envelope phase between two low-mass helium stars, such that the subsequently formed NSs would not have had time to accrete matter and be recycled. These DNS systems would be under-represented in Galactic pulsar surveys, since they would be observable as radio pulsars for a much shorter time scale than recycled pulsar systems. Hence, DNS coalescence rate calculations based on the observed Galactic pulsar sample need to take into account any observational biases. Another of these new formation channels involves a stage of hypercritical common-envelope accretion from a low-mass helium giant to the firstborn NS, resulting in a population of tight, short-lived binaries (with merger time scales $\lesssim 1.0$ Myr) which may contribute significantly to the total number of coalescences [90].

In the 2004 study of merging DNS systems as the source of sGRBs by Ando [91], the merger-delay distribution is modeled as a power law ($\sim t^{\alpha}$), and the calculated GRB rate densities are found to be relatively insensitive to the lower cutoff time necessitated by such a parametrization, but considerably sensitive to α. The characteristic upper inspiral time scale is also of interest; several known DNS systems are calculated to have inspiral times exceeding 10 Gyr (Refs. [56,92] and references therein). If these are representatives of a class of DNS systems resulting from a different evolutionary path than the lower time scale systems, then this evolutionary path need not be considered for detectable GW sources.

With the above considerations in mind, in the present study, we adopt a power-law merger-delay distribution for DNS systems, with a reference index of -1. This is supported by the existing (albeit sparse) observational data on the gravitational inspiral times of Galactic DNS systems and the prevalence of power-law delay distributions found in population synthesis studies. For normalization purposes, we adopt a lower delay time of 50 Myr, since the massive progenitor system (containing components with masses between ~ 8–$20\ M_{\odot}$ for NS-NS system formation) may require an *evolutionary* time scale of $\gtrsim 50$ Myr.[14] Taking this as a lower delay time avoids considerations of the (possibly significant) DNS-formation channel with the extra mass-transfer episode (which creates a peak in the delay-time distribution around ~ 20 Myr, and corresponds to a population of tight, short-lived DNS systems [90]). We model only DNS systems formed via the classical formation channel [59,93,94] for which the $\sim t^{-1}$ delay-time distribution is an appropriate approximation over several orders of magnitude of the delay-time. The power-law index will have a greater impact on merger-rate density calculations than the lower cutoff time. We assume an upper inspiral time scale equal to the cosmology-dependent age of the Universe. For the present study, the power-law index in this merger-delay distribution is labeled α.

B. Star-formation rate density, $d\rho_* / dt$

The determination of the low-redshift SFR density has been achieved via a wide variety of techniques, utilizing light at different wavelengths. However, these measurements become more difficult at higher redshifts since many of the techniques successfully employed in the low-z Universe rely on light at wavelengths which cannot be detected beyond $z \sim 4$. This leaves us with only a handful of available techniques to probe the high-redshift star-formation history.[15]

The estimations of star-formation rates at high redshift are obtained from measurements of UV luminosity functions, which tells us how many galaxies emit light in the UV- band in a given epoch. Excepting galaxies with the largest SFRs (which likely suffer from significant dust extinction), UV light has been shown to be a good tracer of the SFR (Refs. [64,95] and references therein). Dust extinction of UV light can be investigated, and hence corrected for, via the measurement of the UV-continuum slope, which has been shown to be well-correlated with dust extinction in the local Universe (Ref. [64] and references therein). A systematic study of the high-redshift SFR density was undertaken in Ref. [64] using *Hubble Space Telescope* data. Correcting their UV luminosity density

---

[14]This evolutionary time scale is an approximate main-sequence lifetime for a 8 $M_{\odot}$ star, burning $\sim 10\%$ of its core hydrogen,
and obtained via the simple scaling relationship, $\tau_{\mathrm{evol}} \sim 10^4 (M/M_{\odot})^{-2.5}$ Myr.

[15]See Ref. [64] for references of low-z techniques for probing the SFR density, as well as the non-UV techniques possible at $z \sim 2$–4.

calculations for dust extinction, and converting this to an estimate of the SFR density, yielded significant evolution of the SFR density between $0 < z \lesssim 6$. The SFR density is shown to rise out to $z \sim 2$–4, followed by a decrease out to $z \sim 6$. This decrease is shown to continue out to $z \sim 8.5$ [65].

Given that only a handful of techniques exist to probe the high-redshift star-formation history, we will have to wait until further studies are carried out, or new techniques are developed, to complement the analyses in Refs. [64,65]. In our present study, we are only interested in a sensible model of the redshift evolution of the SFR density, which we can parametrize for a Bayesian inference analysis. Several of the studies (Refs. [57,58,91]) mentioned in Sec. III B 2, as well as several other studies which attempt to fit GRB densities to delayed SFR-density models (e.g., Ref. [96]), employed the SF2 model of Porciani and Madau [63]. Of the three models considered in the aforementioned paper, the SF2 model attempts to factor in the uncertainties in the incompleteness of data sets and the amount of dust extinction at early epochs. As such, the SFR density remains roughly constant at $z \gtrsim 2$. Its form is,

$$
\frac{d\rho_*}{dt}(z) \approx 0.16 \times \left(\frac{\exp(3.4z)}{\exp(3.4z) + 22} \right)
$$
$$
\times \frac{E(z)}{(1+z)^{3/2}} \, M_\odot \, \mathrm{Mpc}^{-3} \, \mathrm{yr}^{-1}, \quad (A2)
$$

where

$$
E(z) = \sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{k,0}(1+z)^2 + \Omega_\Lambda(z)}. \quad (A3)
$$

Obviously, if studies in the following decade confirm the SFR-density trends found in Refs. [64,65], we would not attempt to fit any ET data with the SF2 model. This model would need to be updated with a more realistic parametrization. But for now, we adopt the SF2 model as a useful ansatz. For the present study, we parametrize the SF2 ansatz by making the factors of 3.4 in the numerator and denominator of Eq. (A2) variables, labeled β_1 and β_2, respectively.

APPENDIX B: A FASTER CALCULATION OF THE EXPECTED DETECTION RATE

The expected detection rate of inspiraling NS-NS binaries is given by

$$
N_D = T \times \int_0^\infty \int_0^\infty \frac{4\pi D_c(z)^2 D_H}{E(z)} \frac{\dot{n}(z)}{(1+z)}
$$
$$
\times \mathcal{P}(\mathcal{M}) C_\Theta \left[\frac{\rho_0}{8} \frac{D_L(z)}{r_0} \left(\frac{1.2 \, M_\odot}{(1+z)\mathcal{M}} \right)^{5/6} \right] dz d\mathcal{M}.
$$
$$
(B1)
$$

In our previous study [22], we found that a simple parametrization of the expected detection rate provided a good approximation to the slower multidimensional integration necessitated by Eq. (B1). However, we now want to extend our model-parameter space to a larger number of dimensions, for which the simple ansatz method becomes cumbersome. We found in our previous analysis that the standard deviation of the NS mass distribution had very little impact on the expected detection rate. Changing σ_{NS} from $0.02 \, M_\odot$ to $0.12 \, M_\odot$ led to a change in the expected detection rate of $\lesssim O(1\%)$.

Although the precision with which we are able to constrain σ_{NS} scales with the number of detections as $1/\sqrt{N}$, it is the distribution of detectable systems rather than their number that provides information on σ_{NS}. If we approximate the Gaussian chirp mass distribution by a δ-function centered on the mean of the chirp mass distribution, we can replace Eq. (B1) by a 1D integral

$$
N_D = T \times \int_0^\infty \frac{4\pi D_c(z)^2 D_H}{E(z)} \frac{\dot{n}(z)}{(1+z)}
$$
$$
\times C_\Theta \left[\frac{\rho_0}{8} \frac{D_L(z)}{r_0} \left(\frac{1.2 \, M_\odot}{(1+z)\mu_\mathcal{M}} \right)^{5/6} \right] dz. \quad (B2)
$$

This integration can be solved at least an order of magnitude faster by standard routines and gives results consistent with the full 2D integration procedure. We also checked this faster method against the ansatz parametrization method, finding that the method used in our previous analysis was sufficiently accurate.

[1] B. P. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. S. Amin, S. B. Anderson, W. G. Anderson, M. A. Arain *et al.*, Rep. Prog. Phys. **72,** 076901 (2009).

[2] H. Grote and the (LIGO Scientific Collaboration), Classical Quantum Gravity **25,** 114043 (2008).

[3] F. Acernese, P. Amico, M. Alshourbagy, F. Antonucci, S. Aoudia, S. Avino, D. Babusci, G. Ballardin *et al.*, Classical Quantum Gravity **23,** S635 (2006).

[4] R. Takahashi (TAMA Collaboration), Classical Quantum Gravity **21,** S403 (2004).

[5] K. Kuroda (LCGT Collaboration), Classical Quantum Gravity **27,** 084004 (2010).

[6] K. Somiya (KAGRA Collaboration), Classical Quantum Gravity **29,** 124007 (2012).

[7] I. Mandel and R. O'Shaughnessy, Classical Quantum Gravity **27,** 114007 (2010).

[8] J. Abadie, B.P. Abbott, R. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari, P. Ajith, B. Allen *et al.*, Phys. Rev. D **82,** 102001 (2010).

[9] J. Abadie, B.P. Abbott, R. Abbott, T.D. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **85**, 082002 (2012).

[10] G.M. Harry and the LIGO Scientific Collaboration, Classical Quantum Gravity **27**, 084006 (2010).

[11] J. Abadie, B.P. Abbott, R. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari, P. Ajith, B. Allen *et al.*, Classical Quantum Gravity **27**, 173001 (2010).

[12] Virgo Technical Report VIR-0027A-09, Virgo (2009).

[13] Indian Initiative in Gravitational-wave Observations, IndIGO, http://www.gw-indigo.org/tiki-index.php?page=Welcome.

[14] B.F. Schutz, Nature (London) **323**, 310 (1986).

[15] S.A. Hughes and D.E. Holz, Classical Quantum Gravity **20**, S65 (2003).

[16] B.S. Sathyaprakash, B.F. Schutz, and C. Van Den Broeck, Classical Quantum Gravity **27**, 215006 (2010).

[17] W. Zhao, C. Van Den Broeck, D. Baskaran, and T.G.F. Li, Phys. Rev. D **83**, 023005 (2011).

[18] D.E. Holz and S.A. Hughes, Astrophys. J. **629**, 15 (2005).

[19] S. Nissanke, D.E. Holz, S.A. Hughes, N. Dalal, and J.L. Sievers, Astrophys. J. **725**, 496 (2010).

[20] C.L. MacLeod and C.J. Hogan, Phys. Rev. D **77**, 043512 (2008).

[21] W. Del Pozzo, arXiv:1108.1317.

[22] S.R. Taylor, J.R. Gair, and I. Mandel, Phys. Rev. D **85**, 023535 (2012).

[23] D. Marković, Phys. Rev. D **48**, 4738 (1993).

[24] D.F. Chernoff and L.S. Finn, Astrophys. J. **411**, L5 (1993).

[25] L.S. Finn, Phys. Rev. D **53**, 2878 (1996).

[26] A. Freise, S. Chelkowski, S. Hild, W. Del Pozzo, A. Perreca, and A. Vecchio, Classical Quantum Gravity **26**, 085012 (2009).

[27] B. Sathyaprakash, M. Abernathy, F. Acernese, P. A.-S. N. Andersson, K. Arun, F. Barone, B. Barr, M. Barsuglia, M. B. N. Beveridge, S. Birindelli *et al.*, Classical Quantum Gravity **29**, 124013 (2012).

[28] M. Punturo, M. Abernathy, F. Acernese, B. Allen, N. Andersson, K. Arun, F. Barone, B. Barr, M. Barsuglia, M. Beker *et al.*, Classical Quantum Gravity **27**, 084007 (2010).

[29] S. Hild, M. Abernathy, F. Acernese, P. Amaro-Seoane, N. Andersson, K. Arun, F. Barone, B. Barr, M. Barsuglia, M. Beker *et al.*, Classical Quantum Gravity **28**, 094013 (2011).

[30] C. Messenger and J. Read, Phys. Rev. Lett. **108**, 091101 (2012).

[31] A. Nishizawa, K. Yagi, A. Taruya, and T. Tanaka, J. Phys. Conf. Ser. **363**, 012052 (2012).

[32] http://www.et-gw.eu/.

[33] S. Hild, S. Chelkowski, and A. Freise, arXiv:0810.0604.

[34] S. Hild, S. Chelkowski, A. Freise, J. Franc, N. Morgado, R. Flaminio, and R. DeSalvo, Classical Quantum Gravity **27**, 015003 (2010).

[35] E.A. Huerta and J.R. Gair, Phys. Rev. D **83**, 044021 (2011).

[36] J. Abadie, B.P. Abbott, R. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari, P. Ajith, B. Allen *et al.*, Phys. Rev. D **82**, 102001 (2010).

[37] R. O'Shaughnessy, V. Kalogera, and K. Belczynski, Astrophys. J. **716**, 615 (2010).

[38] P. Nutzman, V. Kalogera, L. S. Finn, C. Hendrickson, and K. Belczynski, Astrophys. J. **612**, 364 (2004).

[39] T. Regimbau, T. Dent, W. Del Pozzo, S. Giampanis, T.G.F. Li, C. Robinson, C. Van Den Broeck, D. Meacher, C. Rodriguez, B.S. Sathyaprakash *et al.*, arXiv:1201.3563.

[40] Advanced LIGO anticipated sensitivity curves (2010), https://dcc.ligo.org/cgi-bin/DocDB/ShowDocument?docid=2974.

[41] Advanced Virgo Baseline Design (2009), https://pub3.ego-gw.it/itf/tds/file.php?callFile=VIR-0027A-09.pdf.

[42] B.S. Sathyaprakash and B.F. Schutz, Living Rev. Relativity **12**, 2 (2009).

[43] B.F. Schutz, Classical Quantum Gravity **28**, 125023 (2011).

[44] J. Veitch, I. Mandel, B. Aylott, B. Farr, V. Raymond, C. Rodriguez, M. van der Sluys, V. Kalogera, and A. Vecchio, Phys. Rev. D **85**, 104045 (2012).

[45] B. Kiziltan, A. Kottas, and S.E. Thorsett, arXiv:1011.4291.

[46] R. Valentim, E. Rangel, and J.E. Horvath, Mon. Not. R. Astron. Soc. **414**,1427 (2011).

[47] K. Belczynski, V. Kalogera, F.A. Rasio, R.E. Taam, A. Zezas, T. Bulik, T.J. Maccarone, and N. Ivanova, Astrophys. J. Suppl. Ser. **174**, 223 (2008).

[48] I. Mandel and R. O'Shaughnessy, Classical Quantum Gravity **27**, 114007 (2010).

[49] F. Ozel, D. Psaltis, R. Narayan, and A. Santos Villarreal, arXiv:1201.1006.

[50] R. O'Shaughnessy, K. Belczynski, and V. Kalogera, Astrophys. J. **675**, 566 (2008).

[51] J. A. de Freitas Pacheco, Astropart. Phys. **8**, 21 (1997).

[52] T. Totani, Astrophys. J. **486**, L71 (1997).

[53] R. Schneider, V. Ferrari, S. Matarrese, and S. F. Portegies Zwart, Mon. Not. R. Astron. Soc. **324**, 797 (2001).

[54] T. Piran, Astrophys. J. **389**, L45 (1992).

[55] H. A. Abt, Annu. Rev. Astron. Astrophys. **21**, 343 (1983).

[56] D. J. Champion, D. R. Lorimer, M. A. McLaughlin, J. M. Cordes, Z. Arzoumanian, J. M. Weisberg, and J. H. Taylor, Mon. Not. R. Astron. Soc. **350**, L61 (2004).

[57] D. Guetta and T. Piran, Astron. Astrophys. **435**, 421 (2005).

[58] D. Guetta and T. Piran, Astron. Astrophys. **453**, 823 (2006).

[59] K. Belczynski, R. Perna, T. Bulik, V. Kalogera, N. Ivanova, and D.Q. Lamb, Astrophys. J. **648**, 1110 (2006).

[60] K. Belczynski, K.Z. Stanek, and C.L. Fryer, arXiv:0712.3309.

[61] M. Dominik, K. Belczynski, C. Fryer, D. Holz, E. Berti, T. Bulik, I. Mandel, and R. O'Shaughnessy, arXiv:1202.4901.

[62] K. Belczynski, D.E. Holz, C.L. Fryer, E. Berger, D.H. Hartmann, and B. O'Shea, Astrophys. J. **708**, 117 (2010).

[63] C. Porciani and P. Madau, Astrophys. J. **548**, 522 (2001).

[64] R.J. Bouwens, G.D. Illingworth, M. Franx, R.-R. Chary, G.R. Meurer, C.J. Conselice, H. Ford, M. Giavalisco, and P. van Dokkum, Astrophys. J. **705**, 936 (2009).

[65] V. González, I. Labbé, R.J. Bouwens, G. Illingworth, M. Franx, M. Kriek, and G.B. Brammer, Astrophys. J. **713**, 115 (2010).

[66] A. Cucchiara, A. J. Levan, D. B. Fox, N. R. Tanvir, T. N. Ukwatta, E. Berger, T. Krühler, A. Küpcü Yoldaş, X. F. Wu, K. Toma *et al.*, Astrophys. J. **736**, 7 (2011).

[67] R. J. Bouwens, G. D. Illingworth, I. Labbe, P. A. Oesch, M. Trenti, C. M. Carollo, P. G. van Dokkum, M. Franx, M. Stiavelli, V. González *et al.*, Nature (London) **469**, 504 (2011).

[68] M. Chevallier and D. Polarski, Int. J. Mod. Phys. D **10**, 213 (2001).

[69] A. Shafieloo, V. Sahni, and A. A. Starobinsky, Phys. Rev. D **80**, 101301 (2009).

[70] E. V. Linder, Phys. Rev. Lett. **90**, 091301 (2003).

[71] A. Albrecht, L. Amendola, G. Bernstein, D. Clowe, D. Eisenstein, L. Guzzo, C. Hirata, D. Huterer, R. Kirshner, E. Kolb *et al.*, arXiv:0901.0721.

[72] N. Jarosik, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, M. Halpern, R. S. Hill, G. Hinshaw, A. Kogut, E. Komatsu *et al.*, Astrophys. J. Suppl. Ser. **192**, 14 (2011).

[73] H. Haario, E. Saksman, and J. Tamminen, Comput. Stat. **14**, 375 (1999).

[74] H. Haario, E. Saksman, and J. Tamminen, Official Journal for the Bernoulli Society of Mathematical Statistics and Probability **7**, No. 2, 223 (2001).

[75] J. Dunkley, M. Bucher, P. G. Ferreira, K. Moodley, and C. Skordis, Mon. Not. R. Astron. Soc. **356**, 925 (2005).

[76] A. Gelman, G. Roberts, and W. Gilks, in *Bayesian Statistics*, edited by J. M. Bernado *et al.* (Oxford University Press, Oxford, 1996), Vol. 5, p. 599.

[77] COSMOLOGUI package, http://www.sarahbridle.net/cosmologui/.

[78] Y. Wang, Phys. Rev. D **77**, 123525 (2008).

[79] D. E. Holz and E. V. Linder, Astrophys. J. **631**, 678 (2005).

[80] B. Kocsis, Z. Frei, Z. Haiman, and K. Menou, Astrophys. J. **637**, 27 (2006).

[81] C. M. Hirata, D. E. Holz, and C. Cutler, Phys. Rev. D **81**, 124046 (2010).

[82] S. Hilbert, J. R. Gair, and L. J. King, Mon. Not. R. Astron. Soc. **412**, 1023 (2011).

[83] A. Nishizawa, K. Yagi, A. Taruya, and T. Tanaka, Phys. Rev. D **85**, 044047 (2012).

[84] S. Vitale, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, I. Mandel, B. Aylott, and J. Veitch, Phys. Rev. D **85**, 064034 (2012).

[85] J. R. Gair, C. Tang, and M. Volonteri, Phys. Rev. D **81**, 104014 (2010).

[86] J. Wambsganss, R. Cen, G. Xu, and J. P. Ostriker, arXiv: astro-ph/9607084.

[87] A. Duquennoy and M. Mayor, Astron. Astrophys. **248**, 485 (1991).

[88] D. Raghavan, H. A. McAlister, T. J. Henry, D. W. Latham, G. W. Marcy, B. D. Mason, D. R. Gies, R. J. White, and T. A. ten Brummelaar, Astrophys. J. Suppl. Ser. **190**, 1 (2010).

[89] K. Belczyński and V. Kalogera, Astrophys. J. **550**, L183 (2001).

[90] K. Belczynski, T. Bulik, and V. Kalogera, Astrophys. J. **571**, L147 (2002).

[91] S. Ando, J. Cosmol. Astropart. Phys. 6 (2004) 007.

[92] S. Osłowski, T. Bulik, D. Gondek-Rosińska, and K. Belczyński, Mon. Not. R. Astron. Soc. **413**, 461 (2011).

[93] D. Bhattacharya and E. P. J. van den Heuvel, Phys. Rep. **203**, 1 (1991).

[94] K. Belczynski, V. Kalogera, and T. Bulik, Astrophys. J. **572**, 407 (2002).

[95] R. J. Bouwens, G. D. Illingworth, M. Franx, and H. Ford, Astrophys. J. **670**, 928 (2007).

[96] F. J. Virgili, B. Zhang, P. O'Brien, and E. Troja, Astrophys. J. **727**, 109 (2011).