Source confusion from neutron star binaries in ground-based gravitational wave detectors is minimal

Aaron D. Johnson[1,*] Katerina Chatziioannou,[1,2,†] and Will M. Farr[3,4,‡]

[1]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*
[2]*LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA*
[3]*Center for Computational Astrophysics, Flatiron Institute,*
162 5th Avenue, New York, New York 10010, USA
[4]*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*

Upgrades beyond the current second generation of ground-based gravitational wave detectors will allow them to observe tens of thousands neutron star and black hole binaries. Given the typical minute-to-hour duration of neutron star signals in the detector frequency band, a number of them will overlap in the time-frequency plane, resulting in a nonzero cross-correlation. We examine "source confusion" arising from overlapping signals whose time-frequency tracks cross. Adopting the median observed merger rate of 100 Gpc^{-3} yr^{-1}, each neutron star binary signal overlaps with an average of 42(4)[0.5] other signals when observed from 2(5)[10] Hz. The vast majority of overlaps occur at low frequencies where the inspiral evolution is slow: 91% of time-frequency overlaps occur in band below 5 Hz. The combined effect of overlapping signals does not satisfy the central limit theorem and source confusion cannot be treated as stationary, Gaussian noise: on average 0.91(0.17)[0.05] signals are present in a single adaptive time-frequency bin centered at 2(5)[10] Hz. We quantify source confusion under a realistic neutron star binary population and find that parameter uncertainty typically increases by less than 1% unless there are overlapping signals whose detector-frame chirp mass difference is $\lesssim 0.01 M_\odot$ and the overlap frequency is $\gtrsim 40$ Hz. Out of 1×10^6 simulated signals, 0.14% fall within this region of detector-frame chirp mass differences, but their overlap frequencies are typically lower than 40 Hz. Source confusion for ground-based detectors, where events overlap instantaneously, is significantly milder than the equivalent Laser Interferometer Space Antenna problem, where many classes of events overlap for the lifetime of the mission.

I. INTRODUCTION

Planned improvements and upgrades of ground-based gravitational wave (GW) detectors will expand both their detection horizon and their sensitive frequency range [1–3]. The expanded horizon leads to detection of binary neutron stars (BNSs) and binary black holes (BBHs) to larger distances, thus increasing the detection rate by orders of magnitude. The increased bandwidth leads to observation times that reach hours and minutes for BNSs and BBHs, respectively. The combined outcome of these two effects is that multiple signals will overlap in time and frequency in the data streams, leading to source confusion. As discussed in [4,5] and proven analytically in Appendix C, however, the relevant condition is not whether two signals overlap in time or frequency *only*, but rather whether they overlap

simultaneously in both, i.e., if their time-frequency tracks cross. We therefore define "overlapping signals" as those whose time-frequency tracks cross, resulting in a nonzero cross-correlation.[1] "Signal confusion" is then the effect of overlapping signals on inference.

Overlapping signals is not a new problem for GW astronomy. The planned Laser Interferometer Space Antenna (LISA) mission [6] will observe (among other sources) tens of millions of Galactic white dwarf binaries, thousands of which will be individually resolvable with the rest contributing to the unresolvable Gaussian noise [7]. However, the ground-based and LISA overlapping source

---

*aaronj@caltech.edu
†kchatziioannou@caltech.edu
‡wfarr@flatironinstitute.org

[1]The cross-correlation is defined as the noise-weighted inner product between two signals, Eq. (12). In Appendix C we analytically prove under the stationary phase approximation (SPA) that the integral is nonzero if and only if two signals overlap in time and frequency simultaneously. This integral is sometimes also referred to as the "overlap integral" leading to the confusing definition: overlapping signals are those whose overlap is nonzero.

problems are not identical. The vast majority of LISA's white dwarf binaries have negligible frequency evolution during the mission lifetime. As a result, two signals that overlap in frequency at one time will continue doing so practically indefinitely. BNSs as observed by ground-based detectors, on the other hand, are transient sources with strong frequency evolution. Two signals that overlap temporally over a long time will only overlap in both time and frequency *instantaneously*. Moreover, the frequency evolution is faster at higher frequencies, suggesting that most overlaps occur at low frequencies. While one LISA white dwarf binary overlaps with another binary indefinitely, one ground-based BNS overlaps with a large number of BNSs each momentarily and preferentially at lower frequencies. The latter resembles more the case of a single massive BBH that overlaps with multiple white dwarf binaries as it sweeps through the LISA frequency band [8].

This general picture suggests that source confusion is qualitatively different across detectors and astrophysical sources. Through a Fisher formalism, Crowder and Cornish [9] showed that inference accuracy for a single white dwarf binary in LISA deteriorates *exponentially* with the number of overlapping sources. Moving up to the decihertz range, Cutler and Harms [4] showed that BNS source confusion in the Big Bang Observer instead grows as the *square root* of the number of overlapping sources. The scaling difference is exactly due to the fact that white dwarf binaries overlap in the time-frequency domain over a long time [9], while frequency evolution makes BNS overlaps momentary [4]. Each BNS time-frequency intersection happens at a random phase; it is therefore a random walk that adds incoherently. A separate but related question is whether overlapping signals add up to Gaussian noise. Racine and Cutler [10] argued that the answer depends both on the number of sources and on the type of signal we are targeting on top of all other signals. The latter determines how far in the tails of the noise distribution we have to go for detection, i.e., to what σ level the central limit theorem has to be satisfied. In the context of LISA, BBH signals are "sufficiently different" from white dwarf binaries that source confusion can indeed be treated as Gaussian noise [10].

Moving further up to the ground-based detector frequency range, a comparison between the astrophysical rate and observable duration reveals that multiple BNSs will be simultaneously present in the data time series [11–15]. Since signals, however, overlap mostly at low frequencies and "separate" as they approach merger, current detection techniques can identify them [16–20] and measure their coalescence time to $\mathcal{O}(10)$ ms [17]. Ignoring the presence of overlapping signals can lead to parameter biases for two signals that merge sufficiently close [5,13–15,21]. Considering more or louder signals [22] and going beyond masses and aligned spins [5,13–15,21] would likely increase biases. For example, inference of more subtle effects such as spin precession [15], tests of general relativity [18,22,23], and the neutron star (NS) equation of state could be more severely affected by violations of the assumption that the data are consistent with Gaussian noise [24,25]. Quantitative conclusions about source confusion are, however, complicated by the fact that the relevant picture for ground-based detectors is time-frequency overlaps of multiple signals, rather than temporal coincidences between two signals as adopted in [13–15]. To emphasize this distinction, we refer to overlapping signals as those whose time-frequency tracks cross and coinciding signals as those that exist simultaneously in the data stream.

In this study, we revisit overlapping signals in ground-based detectors and quantify source confusion. We restrict ourselves to BNSs that are expected to be the most numerous and long-lasting binary source, thus resulting in more overlapping signals. While BBH signals may suffer from larger source confusion when they overlap, their lower local event rate and short duration in band suggest that this is more rare than BNSs, e.g., [13]. Overall, we expect that source confusion will depend on the astrophysical population properties, astrophysical merger rate, detector sensitivity (affecting the detected rates), and detector low-frequency performance (affecting the signal duration). We therefore consider different networks of proposed detectors, astrophysical populations, and merger rates as described in Sec. II. We address two questions.

(1) How much time-frequency overlap is there? In Sec. III we simulate data with BNSs under different astrophysical rates. We examine time-frequency crossings and confirm the qualitative picture described above. Under the median local rate of 100 Gpc^{-3} yr^{-1} [26], each BNS time-frequency track crosses an average of 42 other BNS tracks from 2 Hz. Because of the slow frequency evolution, the majority of overlaps occurs at low frequencies: 91% in band below 5 Hz and very few above 20 Hz. Splitting data into time-frequency bins that are adapted to the signal morphology, each bin contains on average (at most over 5 days of observation) 0.91(6) signals at 2 Hz, dropping to 0.05(3) signals at 10 Hz. The low occupation number suggests that the central limit theorem is not satisfied and BNS source confusion in ground-based detectors is not another source of Gaussian noise. Though not quantified in this study, we expect this conclusion to hold when further considering BBH and mixed neutron star–black hole (NSBH) events given their shorter duration and lower rates.

(2) What is the impact of overlapping signals on parameter estimation? Again, with simulated data we quantify the impact of overlapping signals on parameter inference of a target BNS of interest. If the overlapping signals were ignored altogether, parameter inference would be subject to systematic biases [5,13–15,21,27]. A "global fit" that simultaneously

analyzes all signals would mitigate such biases. Despite its technical complications, progress in the LISA [28] and ground-based [29–33] contexts suggests that such solutions could be available in the timescale of third-generation ground-based detectors. As such, here we instead focus on the *statistical* uncertainty aspect of source confusion. Following Crowder and Cornish [9], we use the Fisher formalism as described in Sec. IV and compare statistical uncertainty from data with only one signal and data with multiple overlapping signals. In Sec. V we show that source confusion results in a sub-percent increase in parameter uncertainty per signal unless there exist overlapping signals with detector-frame chirp masses $|\Delta\mathcal{M}_z| = |\mathcal{M}_{z2} - \mathcal{M}_{z1}| \lesssim 0.01M_\odot$. Even when events with such similar masses do overlap, parameter uncertainties increase by $\gtrsim 1\%$ only if the frequency of overlap is $\gtrsim 40$ Hz. Out of 1×10^6 simulated signals, 0.14% fall within this chirp mass threshold, but at frequencies lower than 40 Hz, implying that none have significant parameter uncertainty increases. Our results qualitatively agree with those of Ref. [21], generalized over BNS populations and binary parameters.

Overall, we conclude that the confusion problem in third-generation detectors, where signals usually overlap instantaneously, will be a lot more mild than the LISA case, where signals may overlap for the entirety of the mission duration. By exploring the parameter uncertainty increase and comparing to LISA calculations [9], we quantify this comparison and comment on the efficacy of LISA data analysis strategies for third-generation detectors. Global fit analyses that simultaneously model all data components, including instrumental noise, astrophysical/cosmological backgrounds, and transient signals, are likely to be successful for third-generation detector data as well, hopefully without loss of data [34]. We discuss these conclusions and elaborate upon further work in Sec. VI.

II. DETECTOR NETWORK AND ASTROPHYSICAL POPULATIONS

Source confusion depends both on the properties of the detector network and on the astrophysical properties of the signals. In Secs. II A and II B we describe the networks of future detectors and astrophysical populations we consider, respectively.

A. Detector networks

We consider several detectors whose location and orientation are summarized in Table IV. Cosmic Explorer (CE) [2] is envisioned as a 40 km "L" shaped detector. Since the location remains to be determined, we set two CE detectors at the current LIGO sites. We adopt its projected low-frequency cutoff of 5 Hz. While a noise curve tuned to
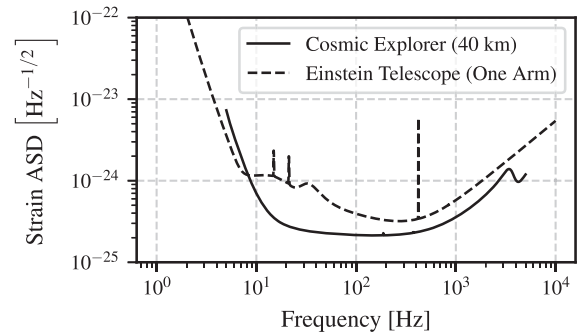

FIG. 1. Projected noise amplitude spectral densities (ASDs) for the two detectors that form the different networks we consider. Despite the nominal ET sensitivity going down to 1 Hz, we adopt a low-frequency cutoff of 2 Hz due to the high ASD values below that frequency. Outside of these frequency ranges, the ASD is assumed to be infinite.

low frequencies exists, we employ the standard noise curve for the CE [35,36], since the projected sensitivity remains the same below 10 Hz where the majority of overlaps occur. The Einstein Telescope (ET) [3] is designed with a triangular shape and 10 km arms; we adopt the possible site of Sardinia [37]. Projected noise curves set the low-frequency sensitivity cutoff at 1 Hz, however, here we adopt a cutoff of 2 Hz, as the noise increases rapidly below this value. Design noise curves for all detectors are shown in Fig. 1.

We combine these detectors to form different networks:
  (i) CE: a single CE detector in the location of LIGO's Hanford detector.
  (ii) ET: the full triangular ET detector.
  (iii) CE + ET: two CE detectors at each of LIGO's sites and an ET detector.

B. Populations of neutron star binaries

We consider populations of quasicircular, spin-aligned BNS inspirals and model the GW signal with the TaylorF2 [38] waveform. Details about the waveform implementation and how we take Earth's rotation into account are given in Appendix D. Binary parameters (other than redshift) are drawn from distributions that are summarized in Table I. Since we assume uniformly distributed masses, we adopt the corresponding local merger rates of 20, 100, and 300 Gpc^{-3} yr^{-1} representing approximately the low, median, and high inferred values [26]. Select results are also presented for the extremely high rate of 1700 Gpc^{-3} yr^{-1} for reference.

We consider different redshift distributions computed as follows. The source-frame merger rate density is

$$\dot{n}(z) \propto \int_{t_d^{\min}}^{t_d^{\max}} \psi(z_f(z, t_d)) P(t_d) dt_d, \quad (1)$$

given a binary formation rate $\psi(z_f)$ and a time delay distribution $P(t_d)$ between formation and merger. The

TABLE I.   Population distributions for BNS parameters. We list the component source-frame masses m_1, m_2; the component spins along the orbital angular momentum χ_1, χ_2; the component dimensionless tidal deformabilities Λ_1, Λ_2 determined through a fixed equation of state SFHo [39] that is consistent with current observational constraints [40]; the declination δ and right ascension α; the polarization angle ψ; the inclination ι; and the time t_c and phase ϕ_c of coalescence.

Parameter	Prior
m_1, m_2	$U[1,2]M_\odot$
χ_1, χ_2	$U[-0.05, 0.05]$
Λ_1, Λ_2	SFHo [39]
$\cos\delta$	$U[-1,1]$
α	$U[0,2\pi]$
ψ	$U[0,\pi]$
$\cos\iota$	$U[-1,1]$
t_c	$U[0,3024000]$ s
ϕ_c	$U[0,2\pi]$

TABLE II.   Simulated populations used throughout this study. We vary the local merger rate between a low, median, high, and very high value inferred in [26] and optionally include a delay between formation and merger. The last two columns give the total number of mergers in a year N_{BNS} and the average time between successive events Δt_c.

Rate [Gpc^{-3} yr^{-1}]	Delay	N_{BNS}	$\langle\Delta t_c\rangle$ (s)
20	Yes	28955	1090
100	Yes	144778	218
300	Yes	434336	73
1700	Yes	2461238	13
20	No	79694	396
100	No	398470	79
300	No	1195412	26
1700	No	6774004	5

In the observer frame, $R_o(z) = R(z)/(1+z)$. The redshift z (equivalently, luminosity distance d_L) distribution is

$$P(z) = \frac{R_o(z)}{\int_0^\infty R_o(z')dz'}. \tag{7}$$

Finally, the total number of BNS events is obtained by integrating over redshift,

$$N_{\mathrm{BNS}} = \int_0^z R_o(z)dz'. \tag{8}$$

Table II lists the total number of events and the average time between them for different choices of the local merger rate. Our results are broadly consistent with equivalent calculations using similar assumptions. Our numbers are similar to those of [12,43–45], but half of those obtained in [18,46] for equivalent local merger rates. All results are highly dependent on the assumed event rates, with higher rates leading to correspondingly higher numbers of overlaps with more severe parameter estimation implications, and vice versa. Unless otherwise noted, all subsequent results incorporate a time delay.

constant of proportionality in Eq. (1) is determined by matching $\dot{n}(z)$ to the local merger rate. We assume that binary formation follows the Madau-Dickinson [41] star formation rate $\psi(z_f) \sim \psi_{\mathrm{SFR}}(z_f)$,

$$\psi_{\mathrm{SFR}}(z_f; \alpha, \beta, z_p) = \frac{(1+z_f)^\alpha}{1 + \left(\frac{1+z_f}{1+z_p}\right)^{\alpha+\beta}}, \tag{2}$$

with $\alpha = 2.7$, $\beta = 2.9$, and $z_p = 1.9$. The delay time between formation and merger is $P(t_d) \propto t_d^{-1}$. We adopt minimum and maximum time delays of $t_d^{\mathrm{min}} = 20$ Myr and a Hubble time $t_d^{\mathrm{max}} = t_H = 14.45$ Gyr, respectively. The mapping between redshift at formation z_f and merger z is obtained by solving

$$t_d - [t_L(z_f) - t_L(z)] = 0, \tag{3}$$

where

$$t_L(z) = \int_0^z \frac{dz'}{(1+z')E(z')} \tag{4}$$

is the look-back time with

$$E(z) = \sqrt{\Omega_\Lambda + \Omega_M(1+z)^3}, \tag{5}$$

and $\Omega_M = 0.3097$ and $\Omega_\Lambda = 0.6903$ [42]. From the rate density of Eq. (1), we obtain the rate in a redshift shell by multiplying by the comoving volume element,

$$R(z) = \dot{n}(z)\frac{dV}{dz}. \tag{6}$$

III. THE PREVALENCE OF OVERLAPPING SIGNALS

At each time, dozens of BNS signals are simultaneously present in the detector data stream [14]. However, as discussed further in Sec. IV and Appendix C, source confusion is not driven by signals overlapping in time (or frequency) alone, but by signals overlapping in time and frequency simultaneously. In this section we study the prevalence of overlapping signals and source confusion through the time-frequency tracks of simulated signals, Eq. (D5).[2] Assuming each value for the local merger rate,

---

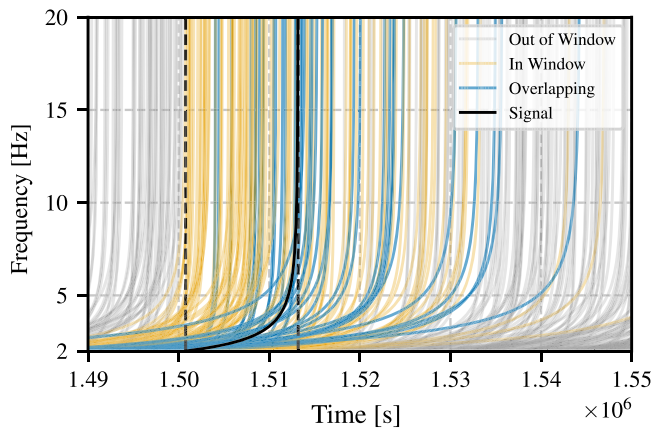[2]A full calculation for the overlap is reserved for Sec. V.

FIG. 2. Approximately 2 days of simulated data for the population with the median merger rate 100 Gpc^{-3} yr^{-1}. Tracks of BNS signals through time-frequency space start with a weak frequency evolution that strengthens until the signals are nearly vertical on this scale. A target signal of interest (black) enters the band and merges within the window denoted by dashed black lines. The signal overlaps in the time-frequency domain with all blue signals. Orange signals exist at any point within the window; they thus coincide with the target signal only temporally. Signals never entering the window are shown in gray. Signals are far more likely to cross at lower frequencies due to the amount of time spent there. Overlapping signals can come into band far before and merge far after the target signal exists.
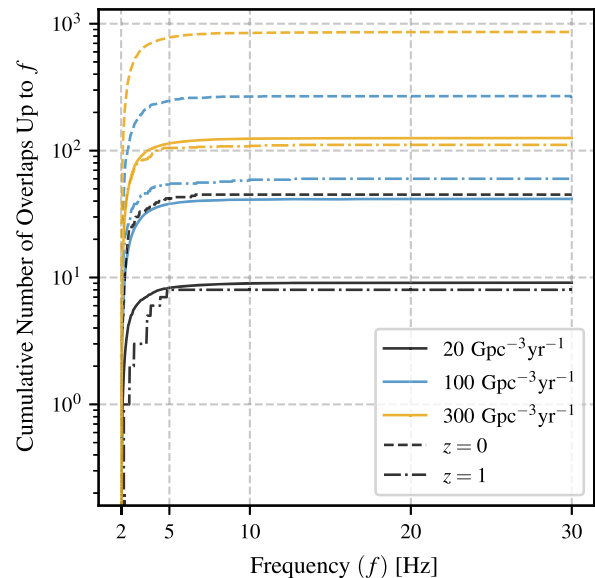

FIG. 3. Cumulative number of time-frequency crossings per signal as a function of frequency. We show results for a binary with the median source-frame chirp mass of $\mathcal{M} = 1.3 M_\odot$ at $z = 0$ (dashed) and at $z = 1$ (dash-dotted). We also show the averaged number of crossings per signal (solid) over the entire population, a result that depends on the assumed mass and redshift distribution. These results are not comparable to calculations of the number of temporally coinciding signals [13–15,21] as we consider time-frequency overlaps.

see Table II, we simulate 35 days of data with BNS signals and use the center 5 days for the analysis. The extra data on either side of the 5-day period reduce edge effects due to events only partially present in that period. Figure 2 shows time-frequency tracks from a representative simulation. Even though a large number of signals (orange) coincide with a target signal (black) temporally, a much smaller subset (blue) overlap in both time and frequency simultaneously. The low, median, and high event rates all have a ratio of signals that overlap in the time-frequency domain to merely coincide temporally with a target signal (number of blue signals divided by orange signals) of ∼0.3.

A. Number of time-frequency crossings per signal

Since source confusion arises from signals overlapping in time and frequency, we begin by studying how often the time-frequency tracks of signals cross. Analytical arguments inspired by Cutler and Harms [4] and repeated in Appendix A show that the rate of crossings is constant along a signal's time-frequency track, only depending on the merger rate, the BNS mass, and redshift distribution. The number of crossings is, instead, higher at low frequencies where frequency evolution is slower.

Turning to numerically simulated BNS populations, Fig. 3 shows cumulative numbers of crossings per signal as a function of frequency for different values of the local

merger rate and BNS masses. Most overlaps occur below 5 Hz. The number of overlaps depends sensitively on the redshift and mass of the signal under consideration, since they affect the detector-frame chirp mass and hence the signal duration. For example, a binary with the median source-frame chirp mass across the population of $\mathcal{M} = 1.3 M_\odot$ undergoes 651(111) overlaps when observed at redshift $z = 0(1)$ in the largest rate examined. When averaged over the entire population, the number of overlaps per signal from 2 Hz is 9, 42, and 145 for the low, median, and high merger rate values, respectively. The analytical calculation of Appendix A reproduces these numerical estimates within a factor of at most 2.

B. Signals per time-frequency bin

The density of signals in time-frequency space controls how large source confusion is and how difficult it is to separate the overlapping signals. Though a full characterization of source confusion hinges on calculating cross-correlations and likelihoods, see Sec. IV, we begin here by obtaining an estimate of the density of signals across time-frequency bins. Qualitatively, if most time-frequency bins contain at least one signal, separating signals (and characterizing the underlying stochastic noise) is more challenging. Moreover, if the number of signals per bin is large, their combined contribution might satisfy the central limit

theorem and amount to Gaussian noise, akin to LISA's white dwarf noise.

The signal density depends on the shape of the bins, namely, their time width δt and frequency height δf that are subject to the uncertainty (Gabor) limit [47],

$$\delta t \, \delta f \geq \frac{1}{2}. \tag{9}$$

In addition to this limit, the shape of the bins must be chosen to be representative of the time-frequency properties of the signals. For example, for a signal with constant frequency (such as a LISA white dwarf binary), increasing δt decreases δf and leads to improved resolution of the signal frequency. In other words, continued observations of the signal offers information about its frequency. However, for a transient signal obtaining more data after the signal terminates or after it has evolved in frequency should not offer additional information. Using a large δt in that case would not be indicative of the frequency resolution δf that is feasible. The bin shapes should, therefore, be adapted to the properties of the signal and specifically the data that are relevant for the frequencies we are trying to resolve.[3]

We adopt bins that saturate the Gabor limit $\delta t \, \delta f = 1/2$ and are adapted to chirps. At low frequencies, signals inspiral slowly, hence longer (large δt, small δf) bins are optimal. As the signal evolves, the frequency evolution speeds up and shorter (large δt, small δf) bins are more appropriate. Formally, we choose bins where the "average" signal enters in the lower left edge and exits in the upper right edge. This condition, together with the Gabor limit, uniquely defines δt and δf at each frequency. An exact derivation of this condition is presented in Appendix B, where we further argue that this process leads to optimal bin sizes for studying the spectral resolution and signal separation that can be achieved.

This procedure requires an average signal whose detector-frame mass is used to determine the bin size. In addition to the astrophysical population properties, the mean/median detector-frame mass depends sensitively on the detector network via its redshift reach. Given the uniform mass distribution, we find that the local merger rate has a minimal impact on the median detector-frame mass, varying by less than 3% between the populations of Table II. Turning to the detector network, we find a median detector-frame chip mass of $2.9 M_\odot$, $2.5 M_\odot$, and $3.2 M_\odot$ for CE, ET, and CE + ET, respectively, for signals with network signal-to-noise ratio (SNR) above 10. Anticipating the eventual necessity of choosing a particular network,

---

[3]This tuning of the bin shapes is only important in order for the back-of-the-envelope calculation we perform here to be indicative of the full cross-correlation. In the context of a full cross-correlation or likelihood calculation, any bin shape should return the same answer. For example, a time-domain and a frequency-domain analysis should be identical if performed consistently [48].
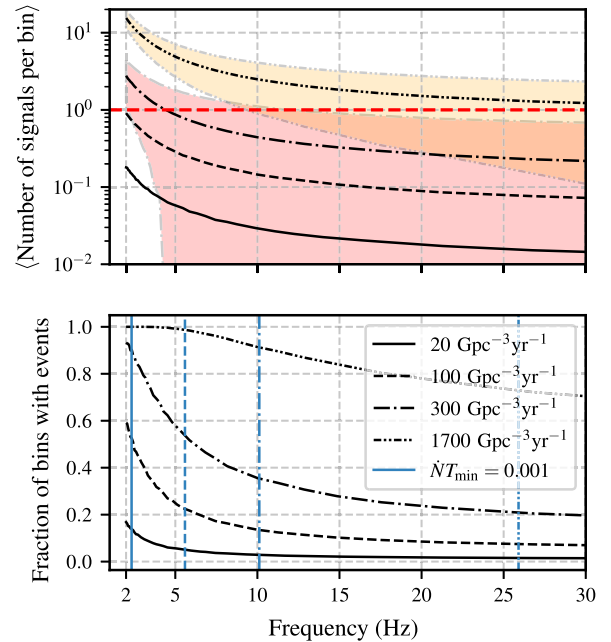

FIG. 4. Average number of signals (top) and occupation fraction (bottom) per time-frequency bin as a function of frequency for populations corresponding to a low, median, high, and extremely high value of the local merger rate. The bin selection follows the procedure laid out in Appendix B. The red horizontal line in the top panel marks the threshold where each bin on average has more than one signal. Shaded regions in the top panel indicate the square root of the variance of the number of signals per bin for the two highest merger rates. We omit shaded regions for the two lowest merger rates as they extend to zero. Vertical blue lines in the bottom panel indicate the frequency at which the occupation fraction has dropped below 1 based on the analytical estimate from Appendix B. Numbers are not comparable to the estimates of [34] who instead use a constant-size time-frequency bin.

in what follows we adopt a detector-frame chirp mass of $2.8 M_\odot$. This choice impacts the following result at the level of a factor of a few. Subsequent results continue including signals with lower SNR, but adopt this median signal for specifying the adaptive grid.

Given the bin selection process, we simulate data, bin the time-frequency plane, and count the number of signals in each bin. We restrict to the signal inspiral phase, as the merger and postmerger last for tens of milliseconds; they therefore do not last long enough to alter the high-frequency bin occupation where the average signal separation is on the order of a minute for the high event rate. The bin occupation fraction is then the percentage of bins that contain at least one signal at a given frequency. Figure 4 shows the average number of signals per bin (top) and the occupation fraction (bottom) as a function of frequency for different values of the local merger rate. These results use all BNSs in the Universe with no cuts on SNR, thus corresponding to the worst-case scenario, and do not depend on the network.

For the low, median, and high populations, the density of signals drops sharply with frequency. The occupation fraction reaches 100% at the lowest frequencies only in the extraordinarily high local rate case of 1700 Gpc^{-3} yr^{-1} (at the top of the two disjoint rate ranges inferred in [26]). For more moderate event rates, the occupation fraction remains below 1 and the average number of signals per bin never exceeds ∼10. Above ∼5 Hz the average number of signals per bin is below 1, suggesting that source confusion is low and the stochastic noise properties (of astrophysical, cosmological, or instrumental origin) can be estimated from the empty bins. For the low rate population, the occupation fraction remains below 20% for all frequencies. Even with the extraordinarily high event rate, at 2 Hz there are on average 15 signals per bin.[4] These results suggest that, even when considering all BNSs in the Universe (i.e., no cuts based on the signal SNR) and unrealistically high rates, the number of signals per bin does not satisfy the conditions for Gaussian noise. Moreover, the density of signals drops sharply with frequency, suggesting that the above-threshold signals can potentially be identified and "traced back" in frequency even in saturated bins. We study these expectations and quantify source confusion in the remaining sections.

IV. EXPLORING SOURCE CONFUSION WITHIN A FISHER FORMALISM

We now turn our attention to source confusion and quantify how overlapping signals affect parameter estimation. Since full parameter estimation via stochastic sampling of the multidimensional posterior of all signals is computationally expensive, we rely on a Fisher formalism to approximate the likelihood. Though such a quadratic likelihood approximation is only valid in the high-SNR limit [49,50] and inadequate for multimodal or complicated likelihood surfaces such as for the sky location parameters, it still corresponds to the Cramér-Rao lower bound on the variance. Moreover, it allows us to explore source confusion under various conditions. We adopt the methodology of Crowder and Cornish [9] as described below.

The data consist of stationary, zero mean, Gaussian noise $n(t)$ and a collection of N signals $H(t; \vec{\Theta})$,

$$s(t) = H(t; \vec{\Theta}) + n(t) = \sum_{n=0}^{N} h_n(t; \vec{\theta}^n) + n(t), \quad (10)$$

where $\vec{\theta}^n$ are the parameters of signal $h_n(t; \vec{\theta}^n)$ and $\vec{\Theta}$ is a concatenation of all parameters, $\vec{\Theta} = (\vec{\theta}^0, \vec{\theta}^1, …, \vec{\theta}^N)$. The likelihood is then

$$p(s|\vec{\Theta}) \propto \exp\left[-\frac{1}{2}(s - H(\vec{\Theta})|s - H(\vec{\Theta}))\right], \quad (11)$$

where the noise-weighted inner product is

$$(h|g) = 4\mathrm{Re} \int_0^\infty \frac{\tilde{h}(f)^* \tilde{g}(f)}{S_n(f)} df, \quad (12)$$

and $S_n(f)$ is the one-sided noise power spectral density (PSD). Tildes denote Fourier transforms and "*" indicate complex conjugation.

The quadratic approximation expands each signal $h(\vec{\theta})$ around its true parameters $\vec{\theta}_t$ to first order

$$h(\vec{\theta}) = h(\vec{\theta}_t) + \partial_i h|_{\vec{\theta}_t} \Delta\theta^i, \quad (13)$$

where $\Delta\theta^i \equiv \theta^i - \theta^i_t$. Equivalently, for the sum of signals,

$$H(\vec{\Theta}) = H(\vec{\Theta}_t) + \partial_i H|_{\vec{\Theta}_t} \Delta\Theta^i. \quad (14)$$

The corresponding quadratic likelihood is

$$p(s|\vec{\Theta}) \propto \exp\left[(n|\partial_i H)\Delta\Theta^i - \frac{1}{2}\Gamma_{ij}\Delta\Theta^i\Delta\Theta^j\right], \quad (15)$$

where $\Gamma_{ij} = (\partial_i H|\partial_j H)$ is the Fisher information matrix. The first term in the likelihood describes the effect of noise realization on the best-fit parameters and can be ignored if the peak of the likelihood coincides with Θ_t, i.e., under a zero-noise realization. The second term describes the measurement uncertainty. The likelihood can be transformed to the posterior if augmented by some prior, $p(\vec{\Theta})$. For the wide and flat priors adopted here, the inverse of Γ_{ij} is the covariance matrix.

Compared to the usual Fisher matrix over a single signal, $\gamma_{ij} = (\partial_i h|\partial_j h)$, here we have the composite Fisher matrix that includes cross terms between signals,

$$\Gamma_{ij} = \left(\frac{\partial H}{\partial\Theta^i}\middle|\frac{\partial H}{\partial\Theta^j}\right), \quad (16)$$

where the indices i, j run over all parameters of all signals. For any signal $h_n(t; \vec{\theta}^n)$, the derivatives $\partial h_n(t; \vec{\theta}^n)/\partial\Theta^i$ are nonzero only if $\Theta^i \in \vec{\theta}^n$, i.e., only if we take derivatives with respect to the parameters of that signal. Moreover, inner products between different signals, such as

$$\left(\frac{\partial h_n(t; \theta^n)}{\partial\Theta^i}\middle|\frac{\partial h_m(t; \theta^m)}{\partial\Theta^j}\right), \quad (17)$$

are nonzero approximately only if the time-frequency tracks of signals h_n and h_m cross, as shown through the SPA in Appendix C.

---

[4]While the projected ET noise curve has a nominal low-frequency cutoff of 1 Hz, the sensitivity is reduced until around 2 Hz where the ASD reaches 1×10^{-22} Hz$^{-1/2}$.

The composite Fisher matrix has a block structure. Along the diagonal, each block consists of each individual signal's Fisher matrix. On the off-diagonal blocks, the matrix contains information about overlaps between signals. Since the covariance is (approximately) the inverse of the Fisher matrix,

$$C^{ij} = (\Gamma^{-1})^{ij}, \qquad (18)$$

any nonzero off-diagonal pieces will affect the covariance. Therefore, time-frequency overlaps will lead to nonzero off-diagonal terms and affect parameter uncertainties,

$$\Delta\Theta^i = (C^{ii})^{1/2}. \qquad (19)$$

These considerations motivated our study of time-frequency crossings in Sec. III.

Evaluation and inversion of the composite Fisher matrix is complicated by the fact that the off-diagonal integrands are highly oscillatory and the matrix might be poorly conditioned. Appendix E describes how we circumvent the first problem by analytically evaluating the off-diagonal terms under the SPA. Matrix condition is described by the condition number c, defined as the ratio of the largest to smallest eigenvalue; matrix inversion loses $\sim \log_{10}(c)$ digits of precision. To circumvent this loss of precision and potentially numerically singular matrices, we use 100 digits of precision, which allows for direct inversion of matrices with condition numbers of $\mathcal{O}(10^{20})$. Before each inversion, we check that the Fisher matrix is positive definite and not numerically singular. Finally, we invert with both a standard inversion method as well as a pseudoinverse and find, in practice, agreement to at least 75 digits, which is far more than necessary. All calculations are performed with *Mathematica*.

V. IMPACT OF SOURCE CONFUSION ON PARAMETER ESTIMATION

In this section, we use the composite Fisher matrix from Sec. IV to explore how parameter uncertainty is impacted by the presence of overlapping sources. For a "target" signal, the parameter uncertainty ratio [9]

$$\varrho = \frac{\Delta\Theta^i}{\Delta\theta^i} = \sqrt{\frac{(\Gamma^{-1})^{ii}}{(\gamma^{-1})^{ii}}} \qquad (20)$$

compares the statistical uncertainty for parameter i when multiple signals are present $\Delta\Theta^i$ with the uncertainty when only the target signal is present $\Delta\theta^i$. The multiple-signal uncertainty $\Delta\Theta^i$ is obtained from the composite Fisher matrix Γ, while the single-signal uncertainty $\Delta\theta^i$ corresponds to only the target signal's block γ.

A. Exploring source confusion with two signals

We begin by studying the dependence of parameter uncertainty ratio on the signal parameters. We consider two overlapping signals and all relevant data: from the time when the first signal enters the band to the time both signals have merged and from 2 to 1024 Hz. The fixed upper frequency cutoff avoids biases from mass-dependent cutoffs [51]. Under the more realistic scenario of multiple overlapping signals, considering all relevant data could result in prohibitively long datasets, cf. Fig. 2. We explore this further in Sec. V B; here we focus on an exploration of the qualitative properties of source confusion.

The Fisher terms that encode source confusion are analytically computed with Eq. (E7) and depend on the following:

(1) the frequency at which the time-frequency tracks overlap f_{ov},[5]
(2) the amplitude $\mathcal{A}(f_{\text{ov}})$ that is a combination of the signal amplitudes and phases,
(3) the phase difference between the signals $\Delta\Phi(f_{\text{ov}})$ when they overlap,
(4) and the second frequency derivative of the phase difference $d^2\Delta\Phi/df^2(f_{\text{ov}})$ which is related to the relative slope between the (tangents of) two signals' time-frequency tracks.

Below we explore the impact of each of the above quantities on the source confusion.

The amplitude $\mathcal{A}(f_{\text{ov}})$ is proportional to the individual signal amplitude and inversely proportional to the noise PSD. It therefore encodes the signals' SNR and the ratio of their SNRs. Changing the SNR ratio while keeping all detector-frame parameters constant does not affect the parameter uncertainty ratio [9]. For two signals, this is proven by considering the inverse of a 2×2 block matrix in Appendix F: the parameter uncertainty of one signal is independent of the SNR of the other signal regardless of whether there are cross terms or not. This conclusion appears to suggest that undetectable signals, as SNR $\to 0$, affect parameter uncertainties as much as loud signals. However, this conclusion is, of course, incorrect [13] and is due to the fact that the Fisher matrix is only valid in the high-SNR limit, thus it does not accurately reflect the low-SNR case [9].

To explore the effect of the overlap frequency f_{ov}, the phase difference at overlap $\Delta\Phi(f_{\text{ov}})$, and the track slope at overlap $d^2\Delta\Phi/df^2(f_{\text{ov}})$ we consider two signals: a target signal with fixed parameters and an overlapping signal

---

[5]As explained in Appendix E, the stationary point of the Fisher off-diagonal terms is not exactly the same as the frequency at which the signals' time-frequency tracks cross. However, the correction is small $\mathcal{O}(10^{-2})$ Hz, especially given that the majority of overlaps occurs at low frequencies. Though Eq. (E7) is computed self-consistently with the correct stationary point throughout, in our discussion we drop this distinction and refer to f_{ov}.
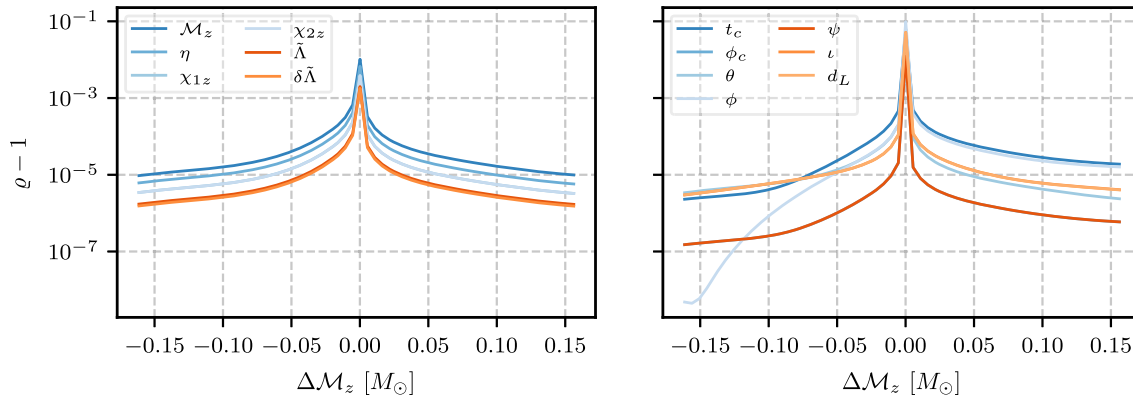
FIG. 5. Uncertainty ratio ϱ (minus one) for the parameters of a target signal when observed alone and when overlapping with a second signal as a function of the difference between the detector-frame chirp masses while fixing the overlap frequency to 2 Hz and the overlap phase difference $\Delta\Phi$ to 0. We consider one arm of the ET network and both intrinsic (left) and extrinsic parameters (right). The uncertainty increase if minimal $\mathcal{O}(10^{-5})$ unless the signals' detector-frame chirp masses differ by less than $0.01M_\odot$.

whose parameters we vary such that the quantities above change. The target signal has SNR $\rho = 100$, no spin, detector-frame chirp mass $\mathcal{M}_z = 1.263M_\odot$, and symmetric mass ratio $\eta = 0.2497$. Unless varied as described below, the overlapping signal has SNR $\rho = 100$, no spin, detector-frame chirp mass $\mathcal{M}_z = 1.2618M_\odot$, and symmetric mass ratio $\eta = 0.2456$. While varying $f_{\rm ov}$ and $\Delta\Psi(f_{\rm ov})$ is straightforward by time or phase shifting the overlapping signal, the slope $d^2\Delta\Psi/df^2(f_{\rm ov})$ is more involved. The slope of the time-frequency track depends (to leading order) on the signal detector-frame chirp mass and the frequency. While the former varies across the BNS population as we consider different masses, the latter is a property of the signal. In other words, increasing $f_{\rm ov}$ at fixed mass will naturally change the slope of the time-frequency track as prescribed by general relativity and this is an effect we wish to retain as $f_{\rm ov}$ changes. We therefore compute the parameter uncertainty ratio for the target signal while varying one of $f_{\rm ov}$, $\Delta\Phi(f_{\rm ov})$, or the mass at a time while keeping the other two fixed.

The impact of the binary masses is explored in Fig. 5, which shows the parameter uncertainty ratio as a function of the difference between the detector-frame chirp mass of the target and the overlapping signal for intrinsic (left) and extrinsic (right) parameters with one arm of the ET network. The ratio is varied by changing the detector-frame chirp mass of the overlapping signal while keeping the overlap frequency (phase) constant at 2 Hz (0). The increase in parameter uncertainty is negligible $\sim 10^{-5}$ unless the binary masses are very similar. Indeed, more unequal masses result in a larger relative slope between the two time-frequency tracks, thus reducing the cross-correlation. More equal masses result in signals that remain closer in the time-frequency domain, thus increasing the cross-correlation. The uncertainty ratio increases to $\sim 10^{-3}$ for chirp masses that differ by less than 1% and spikes to values $\geq 10\%$ across all parameters as the binary masses

become even more equal. Similar chirp masses have also been identified as a necessary condition for large confusion and biases in [21].

The impact of the overlap frequency on the parameter uncertainty ratio is explored in Fig. 6 for intrinsic (left) and extrinsic (right) parameters and for one arm of the ET network (top) and a constant PSD set to the ET noise curve's minimum value (bottom). The overlap frequency is varied by shifting the overlapping signal's time of coalescence, while adjusting its phase to keep the phase difference $\Delta\Phi$ between the two signals constant at 0 at $f_{\rm ov}$. As $f_{\rm ov}$ increases, the difference between the two signals' track slopes, $d^2\Delta\Phi/df^2(f_{\rm ov})$, decreases as the tracks steepen, leading to an increase in the uncertainty ratio for intrinsic parameters. The overall uncertainty ratio depends sensitively on the overlap frequency and the parameter of interest. Both intrinsic and extrinsic parameters display oscillatory behavior for frequencies ≤ 20 Hz, with local minima at a different frequency for each parameter. At higher frequencies, the uncertainty ratio varies more smoothly, reaching $\sim 10\%$ for certain intrinsic parameters. The PSD shape has an effect at the level of a factor of ~ 2.

Finally, the parameter uncertainty ratio as a function of the phase difference between the two signals when they overlap in the time-frequency domain is shown in Fig. 7 for intrinsic (left) and extrinsic (right) parameters using a single arm of the ET network. The phase difference is varied by shifting the overlapping signals' phase of coalescence, while keeping the overlap frequency constant at 2 Hz. This procedure also keeps the difference between the two signals' track slopes $d^2\Delta\Phi/df^2(f_{\rm ov})$ constant. Each parameter's uncertainty ratio shows a similar oscillatory behavior as the phase difference changes, with a period of half the GW one. The uncertainty ratio varies with the phase difference by a factor of ~ 2–3. The oscillations are in phase for all intrinsic parameters and the parameter
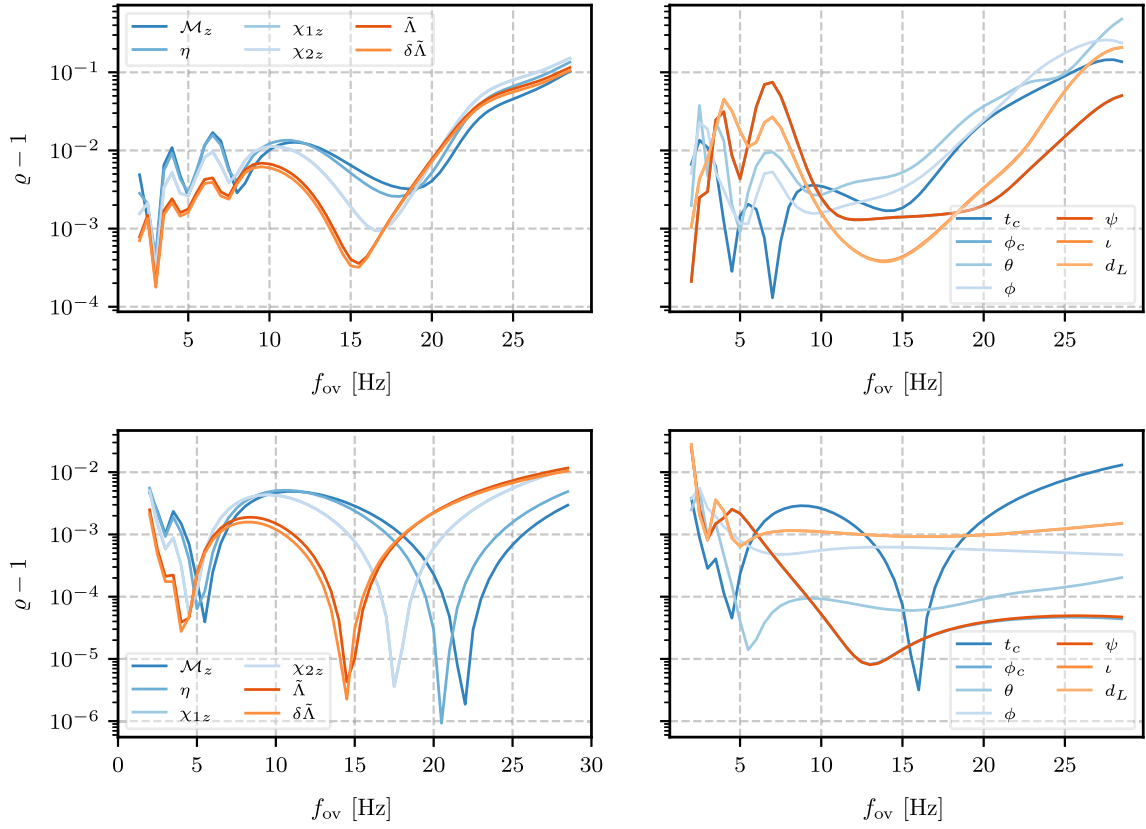
FIG. 6. Uncertainty ratio ϱ (minus one) for the parameters of a target signal when observed alone and when overlapping with a second signal as a function of the overlap frequency $f_{\rm ov}$. We consider one arm of the ET network (top) and a constant PSD (bottom) and both intrinsic (left) and extrinsic parameters (right). The uncertainty ratio is a sensitive function of the overlap frequency and the parameter of interest.
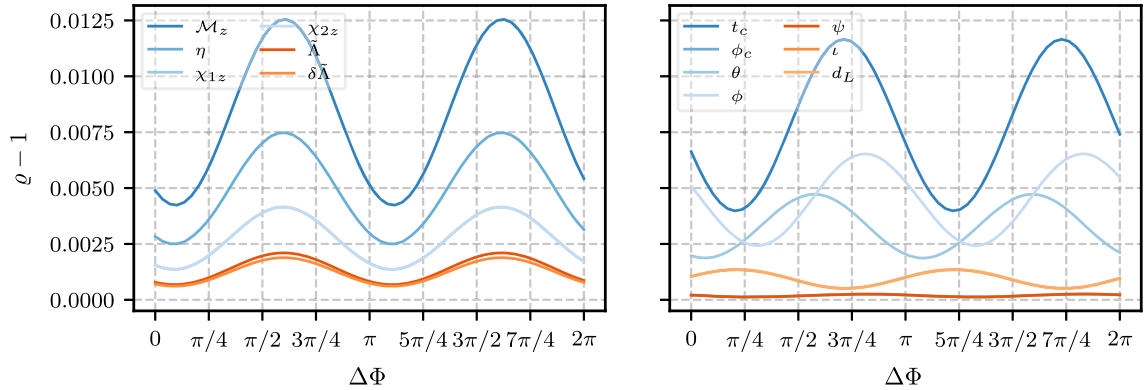

FIG. 7. Uncertainty ratio ϱ (minus one) for the parameters of a target signal when observed alone and when overlapping with a second signal as a function of the phase difference between two signals that overlap at 2 Hz in one arm of the ET network. The uncertainty ratio has a characteristic oscillatory pattern for all parameters.

uncertainty is maximized when the signals are approximately in phase when their tracks cross [recall the factor of $\pi/4$ in Eq. (E7)].

B. Source confusion from BNS populations

Figure 5 shows that, for two overlapping signals, source confusion is negligible unless the binaries have extremely

similar detector-frame chirp masses, which forces their time-frequency tracks to remain close over a range of frequencies. In the realistic case of a BNS population, each binary overlaps with dozens or hundreds of other signals, cf. Fig. 3, and the total source confusion is their combined effect. However, given the sharp drop of source confusion with chirp mass difference, we expect that, even in the full

population case, source confusion will be non-negligible only when there exists an overlapping binary with a similar chirp mass.

Another complication of BNS populations is that the relevant data resemble those of Fig. 2: each signal exists within some finite time window, but it overlaps with signals that enter the detector band and merge over a wide range of times. For example, the target signal of Fig. 2 (black) overlaps in the time-frequency domain with signals (blue) that enter the band (merge) up to 11.5 h (8.5 h) before (after) it. To complicate matters further, a signal (yellow or gray) affects parameter uncertainties for the target signal (black) even if they do not overlap, if they instead both overlap with a *third* signal (one of the blue signals). This is shown in Appendix G and suggests that, in principle, even a signal that enters the band weeks (or years) after the target signal affects its inferred properties. An exact analysis would therefore have to simultaneously analyze years of data, leading to probably a prohibitive computational cost. In practice, however, such multisignal overlaps are sub-dominant to direct two-signal overlaps unless all signals have similar chirp masses.

Even when restricting to two-signal overlaps (the black and the blue signals in Fig. 2), the relevant data extend over long periods of time: the target signal lasts in band for 3.5 h, while the relevant data for any overlapping signal cover a full 20 h. In Sec. V A we presented results analyzing all the relevant data in the case of two overlapping signals. Here we adopt a more moderate setup that resembles current analysis settings [52]: given a target signal we wish to analyze, we consider data from the time they enter the band to the time they merge, i.e., the black-dashed window of Fig. 2.

We considering a simulated population realization with the CE + ET detector network. The target signal has parameters $\mathcal{M}_z = 1.2917 M_\odot$, $\eta = 0.2497$, $\chi_{1z} = \chi_{2z} = 0$, with $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ given by the equation of state; its SNR is $\rho = 320$. We consider the high inferred merger rate of $300 \text{ Gpc}^{-3} \text{ yr}^{-1}$ and all binaries regardless of SNR. Our results are thus an upper limit on source confusion. We simulate 35 days of data and upon injecting the target into the middle of the data, obtain 360 signals that overlap with it. We then compute the composite Fisher matrix and obtain the parameter uncertainty ratio of each of the target signal parameters.

For this simulated population, we find that the parameter uncertainty ratio remains very close to unity, with the largest increase at 0.2% for the sky location parameters. Comparing this result to Fig. 5 suggests that no overlapping signal has a similar chirp mass to the target. Indeed this is the case, as the most similar chirp mass in this realization is $1.5026 M_\odot$. As an extension, we find that the total number of overlapping signals does not affect source confusion, the only condition being whether any of the overlapping signals have a similar detector-frame chirp mass. Given this result, we argue that it is unlikely that BBH and NSBH

signals will cause a significant increase in parameter uncertainty for BNSs, as they never have chirp masses comparable to BNS ones. This conclusion likely applies to subthreshold BNSs as well, which typically have larger redshifts and thus are redshifted outside the typical BNS mass range.

Himemoto *et al.* [21] concluded that both similar detector-frame chirp masses and coalescence time differences $|\Delta t_c| < 0.1$ s are a necessary condition for large biases. Our results qualitatively agree with this conclusion, as the overlapping signals we consider here will also merge close in time if they have similar chirp masses. What is more, our results extend this conclusion to BNS populations and multiple over-lapping signals. Even when the target signal overlaps with hundreds of BNSs, it is only the signals with similar detector-frame chirp masses that result to source confusion. The prevalence of such signals with similar masses depends sensitively on the astrophysical BNS mass distribution and is subject to Poisson uncertainty.

Figure 8 explores the dependence of the parameter uncertainty ratio on the chirp mass difference in more detail. We show the geometric mean of the parameter uncertainty ratio (minus one) over all 13 parameters for different detector-frame chirp mass differences and overlap frequen-cies. For this plot, we simulate two overlapping signals at a phase difference of zero at overlap. As expected, the parameter uncertainty ratio stays generally close to 1. Source confusion becomes more important at higher overlap frequencies; however, at these frequencies we do not expect many signals overlapping to begin with, cf. Fig. 3. Conservatively, we conclude that the parameter uncertainty ratio stays well below 1% unless the detector-frame chirp masses differ by less than $0.01 M_\odot$, see white and blue contours. This estimate is consistent with Ref. [21].

Finally, we quantify how often signals with such similar chirp masses are expected. Instead of simulating a fixed
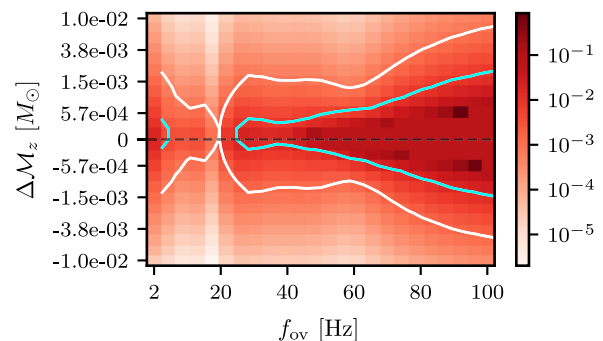

FIG. 8. Geometric mean of the parameter uncertainty ratio (minus one) for all 13 BNS parameters (heat map) as a function of the overlap frequency and difference in detector-frame chirp mass between two overlapping signals. A white (cyan) contour corresponds to a 0.1% (1%) increase in parameter uncertainty. The horizontal dashed line indicates no difference between the two detector-frame chirp masses.

amount of data and thus obtaining results that depend on the assumed merger rate, we present results as a fraction of total signals. Our results still depend on the assumed mass distribution. We simulate a total of 1×10^6 signals and find that 0.14% of these fall within the detector-frame chirp mass difference of $0.01 M_\odot$. However, the overlap frequencies were typically below 40 Hz and thus too low to increase the parameter uncertainties by more than 0.1%. For reference, 1×10^6 total signals would correspond to three years of data at a merger rate of 300 Gpc^{-3} yr^{-1} and 40% of a year of data at 1700 Gpc^{-3} yr^{-1} if all signals were detected. These results are insensitive to the NS mass distribution as we obtain similar numbers for normally distributed source-frame component masses with mean $1.5 M_\odot$ and standard deviation $0.2 M_\odot$. Indeed, the detector-frame chirp mass differences depend on the redshift distribution of the events more than the mass distribution.

VI. CONCLUSIONS

In this work, we examined source confusion from overlapping BNS signals, clarifying the conditions for nonzero cross-correlation in terms of signals that overlap in both time and frequency simultaneously. Overlapping signals appear commonly throughout the data of next-generation ground-based detectors, but with the majority of them occurring at frequencies below 5 Hz. Using nonuniform time-frequency bins that are adapted to BNS signals, and with common assumptions about the high-redshift population and inferred merger rates, bin occupation fraction remains below 100% for all frequencies ≥ 2 Hz. Additionally, the average number of signals per bin is $\lesssim \mathcal{O}(10)$ at 2 Hz and falls off as frequency increases. Even under the highest inferred merger rate of 300 Gpc^{-3} yr^{-1}, the combined contribution of these signals does not satisfy the conditions for Gaussian noise per time-frequency bin. A reasonable fraction of time-frequency bins contain no BNS signal and are therefore suitable for a measurement of underlying stochastic noise. Given that BBH and NSBH signals are rarer and shorter, we do not expect them to qualitatively affect this conclusion.

We then turned to quantifying source confusion, i.e., the impact of overlapping signals on statistical uncertainties for BNS parameters, extending previous literature to higher post-Newtonian (PN) orders, realistic BNS populations, and more parameters. We considered quasicircular non-precessing BNSs described with 13 parameters under the TaylorF2 waveform model, included Earth's rotation, extended to the lowest frequencies that next-generation detectors will be sensitive to, and considered the astro-physically expected number of overlapping signals. Parameter uncertainty typically increases by less than 1% unless the detector-frame chirp masses of the

overlapping signals differ by less than $0.01 M_\odot$ and the overlap frequency is $\gtrsim 40$ Hz. Simulating 1×10^6 signals, we find none of the signals with small detector-frame chirp mass differences have overlaps at high enough frequencies to have significant increases in their param-eter uncertainties. This conclusion is consistent with the analysis of [21], which we have extended to a population of overlapping BNSs with the full TaylorF2 waveform from 2 Hz. Even when there are hundreds of overlapping signals, source confusion is driven only by those that have similar masses and occur with frequency overlaps $\gtrsim 40$ Hz.

The detector-frame chirp mass affects how closely the signals' tracks stay together in the time-frequency domain. As the overlap frequency increases, the signals evolve quickly through frequencies and their tracks stay closer. As an outcome, signals with more different masses can lead to source confusion, cf. Fig. 8, if they overlap at high frequencies. From the population simulations, we find about 0.4% (~ 4000 of 1×10^6) of signals overlap at frequencies higher than 100 Hz at any detector-frame chirp mass. About 50%(25%)[7.5%] of these signals have at least one parameter with an uncertainty increase of more than 1%(10%)[50%]. Of these signals, only a fraction will be detectable. The extent to which this population of signals with small differences in times of coalescence can be disentangled has been considered in previous studies [5,13,14,17,19,31,33].

Our results suggest that source confusion in ground-based detectors is much more mild than in LISA. This is due to two reasons: (i) there are fewer BNSs in the ground-based detector frequency band than white dwarf binaries in LISA's band, and (ii) at ground-based detector frequen-cies, BNSs experience strong frequency evolution, thus most overlaps are instantaneous. The only exception is overlapping signals with similar detector-frame chirp masses whose time-frequency tracks overlap over an extended period of time. Indeed, these are the only signals that can cause significant source confusion. Future exten-sions to NSBH binaries or spin-precessing degrees of freedom [53] are likely to only affect these results at a quantitative level.

From a technical standpoint, such mild source confusion can likely be efficiently addressed with methods similar to LISA's global fit [28] without discarding data. Though parameter uncertainties and biases are low among BNS populations, methods that simultaneously model all data components are still essential. For example, the low number of BNSs suggests that they cannot be treated as Gaussian noise [18,22,54]; thus estimating the actual detector noise in the nonempty bins accurately requires a concurrent treatment of astrophysical signals [55]. Another example concerns reaching the underlying cos-mological stochastic background [32], without resorting to subtracting signals from the data, which is bound

to leave behind residuals that can mask the background [4,8,44,46,56–61]. Finally, a full marginalization over all events is required to fully safeguard against any potential biases [13–15] given the ambitious science goals of next-generation detectors: high-SNR signals, spin-precession, precision constraints on general relativity, and a measurement of the NS equation of state. Our results suggest that source confusion is mild and can thus likely be addressed with global fit techniques developed in the timeline of next-generation detectors.

Software used includes *Mathematica* [62], NumPy [63], SciPy [64], MATPLOTLIB [65], and AstroPy [66].

A software release containing *Mathematica* notebooks for the waveform and PYTHON code to simulate a cosmological population is openly available from the GitHub repository [67].

APPENDIX A: ANALYTICAL ESTIMATE OF THE NUMBER OF TIME-FREQUENCY CROSSINGS

Following the equivalent calculation in Cutler and Harms [4] in the context of the Big Bang Observer, we analytically estimate the number of signals that cross the time-frequency track of a typical BNS, given an observed merger rate of $\dot{N}$. If all signals had the same detector-frame mass and ignoring higher-order modes, then the time-frequency tracks would remain parallel and not cross. So the rate of crossings r_c at a given frequency depends on both the rate of signals and the relative rate of evolution of the time-frequency tracks. That is,

$$r_c = \frac{1}{2}\frac{dN}{df}\frac{d\Delta f}{dt} = \frac{1}{2}\rho(f)\Delta\dot{f}, \tag{A1}$$

where $\rho(f) = dN/df$ is the density of signals in frequency and $\Delta\dot{f}$ is the variation in frequency derivatives, caused by the fact that signals have different masses. The factor of $1/2$ is due to the fact that given two nearby signals there is a 50% chance that they are diverging from each other rather than converging.

The density of signals in frequency $\rho(f)$ can be estimated from the fact that the BNS Universe is stationary, i.e., the signal rate does not change with time,

$$\rho(f)\frac{df}{dt} = \dot{N} = \text{const}, \tag{A2}$$

where df/dt is the GW frequency evolution, Eq. (B1).

For the variation in frequency derivatives between signals $\Delta\dot{f}$, from Eq. (B1) we get to leading PN order

$$\frac{\Delta\dot{f}}{\dot{f}} = \frac{5}{3}\frac{\Delta\mathcal{M}_z}{\mathcal{M}_z}, \tag{A3}$$

where $\Delta\mathcal{M}_z/\mathcal{M}_z$ is the detector-frame chirp mass variation across the BNS population.

Putting everything together,

$$r_c = \frac{1}{2}\rho(f)\dot{f}\frac{5}{3}\frac{\Delta\mathcal{M}_z}{\mathcal{M}_z} = \frac{5}{6}\dot{N}\left[\frac{\Delta\mathcal{M}}{\mathcal{M}} + \frac{\Delta z}{1+z}\right], \tag{A4}$$

where $\mathcal{M}$ is the source-frame mass and z is the redshift. Assuming the fiducial uniform mass distribution of Table I and the redshift distribution of Eq. (7), $\Delta\mathcal{M}/\mathcal{M} \sim 0.14$ and $\Delta z/(1+z) \sim 0.41$, and hence $r_c \sim 0.46\dot{N}$. This rate is independent of the frequency. This is due to a trade-off between the fact that there are more signals at low frequencies, $\rho(f) \sim f^{-11/3}$, but those signals evolve almost "parallel" to each other and do not cross, $\delta\dot{f} \sim f^{11/3}$. Given a constant *rate* of crossings across frequencies, the *number* of crossings is higher at low frequencies where the signals spend the most time. Indeed, in Sec. III we numerically show that over five days of data all but one crossing occurs at or below 20 Hz for any of the event rates considered. Using $\dot{N}$ from Table II, we find that the predicted values from Eq. (A4) agree within a factor of less than 2, as shown in Table III. We computed these values for an average source-frame chirp mass of $\mathcal{M} = 1.3M_\odot$ at redshifts of $z = 0$ and $z = 1$.

TABLE III. Comparison between analytical estimates based on Eq. (A4) and the actual number of overlaps numerically computed in Sec. III A for a BNS with $\mathcal{M} = 1.3M_\odot$ at redshifts $z = 0$ and $z = 1$ for varying local BNS event rates.

Rate (Gpc^{-3} yr^{-1})	z	Analytical	Numerical
20	0	28	49
20	1	9	8
100	0	140	269
100	1	44	60
300	0	417	651
300	1	132	111

APPENDIX B: OPTIMAL TIME-FREQUENCY BINNING AND ANALYTICAL ESTIMATE OF THE NUMBER OF BINS WITH A BNS

Consider BNSs entering the detector band at a rate $\dot{N}$ and merging every T seconds on average. The leading-order observed frequency evolution is

$$\dot{f} = \frac{96}{5}\pi^{8/3}\mathcal{M}_z^{5/3}f^{11/3}, \tag{B1}$$

where f is the observed GW frequency and $\mathcal{M}_z$ is the detector-frame chirp mass. The minimum-time time-frequency bin that contains a signal centered at frequency f has width $T_{\min}$, where

$$\delta f = \frac{1}{2T_{\min}} \simeq \dot{f}T_{\min}. \tag{B2}$$

In the above equation, the first equality is the Gabor condition, while the second (approximate) equality enforces that the signal "exits" the bin at its upper-right edge. Bins with a width larger than $T_{\min}$ would contain data after the signal has evolved in frequency and would thus overestimate how well we can resolve its frequency. Bins with a width smaller than $T_{\min}$ would instead result in the signal also being present in the next time bin and would thus underestimate frequency resolution. We therefore propose this time-frequency bin construction as an optimal way to assess signal spectral resolution and hence how well signals can be separated from each other. This argument is visually presented in Fig. 9.
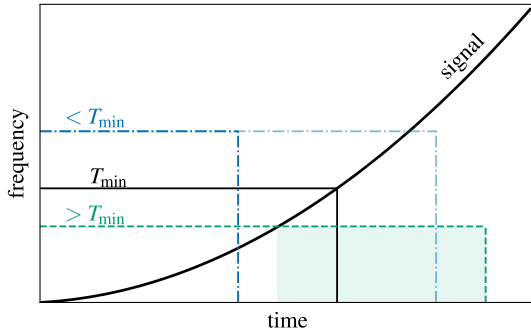

FIG. 9. Visual depiction of the optimal bin selection process. The time-frequency track of the signal is shown in black. The black box represents the optimal bin of width $T_{\min}$, whose lower-left and upper-right edges fall on the signal track. This condition, together with the Gabor limit, defines the bin dimensions. A longer bin (green dashed) would contain data irrelevant to the signal (green shaded region) and thus overestimate spectral resolution. A shorter bin (blue dot-dashed) would leave a portion of the signal to the next bin of the same frequency (pale blue dot-dashed) and thus underestimate spectral resolution. A size $T_{\min}$ therefore defines the most representative bin regarding spectral resolution of a signal at a given frequency.

Equation (B2) implies

$$T_{\min} = \frac{1}{\sqrt{2\dot{f}}} = \sqrt{\frac{5}{192}}\pi^{-4/3}\mathcal{M}_z^{-5/6}f^{-11/6}. \tag{B3}$$

As long as $\dot{N}T_{\min} = T_{\min}/T \ll 1$, there will be many "empty" time-frequency bins at frequency f. This condition is equivalent to an observed rate of signals of

$$\begin{aligned}\dot{N} &\ll \frac{1}{T_{\min}} = \sqrt{\frac{192}{5}}\pi^{4/3}\mathcal{M}_z^{5/6}f^{11/6}\\ &= \frac{5}{\min}\left(\frac{\mathcal{M}_z}{1.2M_\odot}\right)^{5/6}\left(\frac{f}{10\text{ Hz}}\right)^{11/6}.\end{aligned} \tag{B4}$$

The vertical blue lines in the bottom panel of Fig. 4 show the frequency at which $\dot{N}T_{\min} = 0.001$. The analytical estimate is a good approximation for the full numerical results, though it generally overestimates the occupation fraction of bins. We attribute this to the fact that we estimate the bin size with an average signal, while the population contains a range of signal masses.

APPENDIX C: OVERLAP INTEGRAL CALCULATION USING THE STATIONARY PHASE APPROXIMATION

The cross-correlation, or overlap integral, between two signals is nonzero if and only if the signals cross in both time and frequency [4,5,68]. In this appendix, we provide a proof of this through the SPA. Without loss of generality, we restrict ourselves to a toy model with signals with linear frequency evolution.

1. Signal

We consider a sinusoidal signal with constant amplitude A and a liner frequency drift. The signal phase depends on four parameters: the frequency f_0, its derivative $\dot{f}_0$, a constant time t_0, and a constant phase ϕ_0 offset. Overall, $\vec{\theta} = [A, f_0, \dot{f}_0, t_0, \phi_0]$. The frequency evolution is

$$f(t) = f_0 + \dot{f}_0(t - t_0), \tag{C1}$$

and the signal is

$$h(t) = A\cos\left[2\pi\left(f_0(t - t_0) + \frac{1}{2}\dot{f}_0(t - t_0)^2\right) + \phi_0\right]. \tag{C2}$$

The Fourier transform is

$$\tilde{h}(f) = \int h(t)e^{-2\pi i f t}dt = \int A\cos\phi(t)e^{-2\pi i f t}dt, \tag{C3}$$

where

$$\phi(t) = 2\pi\left(f_0(t-t_0) + \frac{1}{2}\dot{f}_0(t-t_0)^2\right) + \phi_0 \quad \text{(C4)}$$

is the time-domain phase. With this definition

$$\tilde{h}(f) = \int A\frac{e^{i\phi(t)} + e^{-i\phi(t)}}{2}e^{-2\pi ift}dt. \quad \text{(C5)}$$

The second term has no stationary point and vanishes under the Riemann-Lebesgue lemma [69]. The first term has a stationary point at

$$2\pi f_{\mathrm{SPA}} = \dot{\phi}(t) = 2\pi[f_0 + \dot{f}_0(t-t_0)]$$
$$\Rightarrow f_{\mathrm{SPA}} = f_0 + \dot{f}_0(t-t_0), \quad \text{(C6)}$$

$$t_{\mathrm{SPA}} = \frac{f-f_0}{\dot{f}_0} + t_0. \quad \text{(C7)}$$

With $\Phi(t) \equiv 2\pi ft - \phi(t)$, the frequency-domain signal is

$$\tilde{h}(f) \approx \frac{A}{2}\int e^{-i\Phi(t)}dt \approx \frac{A}{2}\sqrt{\frac{2\pi}{|\ddot{\Phi}(t_{\mathrm{SPA}})|}}e^{-i\Phi(t_{\mathrm{SPA}})-i\pi/4}. \quad \text{(C8)}$$

Plugging in numbers, we get

$$\tilde{h}(f) = \frac{A}{2\sqrt{\dot{f}_0}}\exp\left[-2\pi ift_0 + i\phi_0 - i\pi\frac{(f-f_0)^2}{\dot{f}_0} - i\frac{\pi}{4}\right]. \quad \text{(C9)}$$

2. Overlapping signals

Given two signals, the cross-correlation is

$$(h_1|h_2) \equiv \Re \int \frac{h_1^*(f)h_2(f)}{S_n(f)}df. \quad \text{(C10)}$$

Substituting Eq. (C9), we have

$$(h_1|h_2) = \frac{A_1 A_2}{4\sqrt{\dot{f}_{0,1}\dot{f}_{0,2}}}\int \frac{\exp\left[2\pi if(t_{0,1}-t_{0,2}) - i(\phi_{0,1}-\phi_{0,2}) + i\pi\frac{(f-f_{0,1})^2}{\dot{f}_{0,1}} - i\pi\frac{(f-f_{0,2})^2}{\dot{f}_{0,2}}\right]}{S_n(f)}df. \quad \text{(C11)}$$

The stationary point is

$$f_{\mathrm{SPA}} = \frac{f_{0,2}\dot{f}_{0,1} - f_{0,1}\dot{f}_{0,2}}{\dot{f}_{0,1} - \dot{f}_{0,2}} + \frac{\dot{f}_{0,1}\dot{f}_{0,2}}{\dot{f}_{0,1} - \dot{f}_{0,2}}(t_{0,1}-t_{0,2}), \quad \text{(C12)}$$

which is reached by each signal at

$$t_1(f_{\mathrm{SPA}}) = t_2(f_{\mathrm{SPA}}) = \frac{f_{0,2} - f_{0,1}}{\dot{f}_{0,1} - \dot{f}_{0,2}} + \frac{t_{0,1}\dot{f}_{0,1} - t_{0,2}\dot{f}_{0,2}}{\dot{f}_{0,1} - \dot{f}_{0,2}}. \quad \text{(C13)}$$

These times coincide, which means that the cross-correlation has a nonzero contribution only at the time when the two signals' time-frequency tracks intersect.

APPENDIX D: WAVEFORM IMPLEMENTATION DETAILS

We model BNS inspirals with the TaylorF2 waveform approximant [38] including spin [70,71] and tidal effects [72]. Parts of our custom implementation were inspired by GWBENCH [73]. A BNS is described in the detector frame by parameters

$$\theta = \{\mathcal{M}_z, \eta, t_c, \phi_c, \chi_1, \chi_2, \Lambda_1, \Lambda_2, \theta, \phi, \psi, \iota, D_L\}, \quad \text{(D1)}$$

consisting of the detector-frame chirp mass, symmetric mass ratio, time of coalescence, phase at coalescence, component spins, component dimensionless tidal deformabilities, declination, right ascension, polarization angle, inclination, and luminosity distance. Additionally, the detector location is described by

$$\lambda = \{\lambda, \varphi, \gamma, \zeta\}, \quad \text{(D2)}$$

including the latitude, longitude, angle of the detector with respect to east, and the angle between the two detector arms. Values for each detector considered here are presented in Table IV. Ignoring higher-order modes, the GW amplitude in the frequency domain and under the SPA is

$$A_0(f;\theta) = \sqrt{\frac{5\pi}{24}}\frac{\mathcal{M}_z^2}{D_L}\eta^{-7/10}u^{-7/2}, \quad \text{(D3)}$$

where $u = (\pi M_z f)^{1/3}$ with the detector-frame total mass $M_z = \mathcal{M}_z\eta^{-3/5}$. The GW phase to 3.5 PN order

$$\Psi_{3.5\mathrm{PN}}(f;\theta,\lambda) = 2\pi ft_c - \phi_c - \frac{\pi}{4} + \frac{3}{128\eta u^5}\sum_{i=0}^{7}B_i u^i, \quad \text{(D4)}$$

where the B_i coefficients (including log terms) exist in a *Mathematica* notebook included in our data release.

TABLE IV. A list of proposed detectors used in future ground-based detector networks, their locations, orientations, arm lengths, and the low-frequency cutoff used in this work.

Detector	Location	Latitude (λ)	Longitude (φ)	Bisector angle (γ)	Arm angle (ζ)	Arm length (km)	f_{low} (Hz)
ET arm 1	Sardinia	40.43°	9.457°	0°	60°	10	2
ET arm 2	Sardinia	40.43°	9.457°	60°	60°	10	2
ET arm 3	Sardinia	40.43°	9.457°	120°	60°	10	2
CE H	Hanford	46.5°	−119.4°	171°	90°	40	5
CE L	Livingston	30.6°	−90.8°	232.7°	90°	40	5

Since BNS signals can last from hours to days when observed from frequencies below 10 Hz, Earth's rotation must be taken into consideration. Our implementation follows GWFAST [74,75]. Overall, the rotation of Earth means that the pattern functions now depend on time [76]. The time to coalescence is

$$t(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = t_c - t_*(f; \boldsymbol{\theta}) + \Delta t_{\text{L}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}), \quad \text{(D5)}$$

where $t_*(f; \boldsymbol{\theta})$ is Eq. (3.8b) of [38], and $\Delta t_{\text{L}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda})$ is the light travel time from the center of Earth to the detector [75],

$$\Delta t_{\text{L}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = -\frac{R_\oplus}{c}[\sin(\theta)\cos(\lambda)\cos(\phi)\cos(2\pi f_\oplus(t_c - t_*(f; \boldsymbol{\theta})) + \varphi)$$
$$+ \sin(\theta)\cos(\lambda)\sin(\phi)\sin(2\pi f_\oplus(t_c - t_*(f; \boldsymbol{\theta})) + \varphi) + \cos(\theta)\sin(\lambda)], \quad \text{(D6)}$$

where $f_\oplus = \text{day}^{-1}$. Earth's rotation causes amplitude and phase modulations through the time dependence of the pattern functions and Doppler shift due to the rotation of Earth. The frequency-domain strain is then

$$\tilde{h}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = A(f; \boldsymbol{\theta}, \boldsymbol{\lambda})e^{i\Phi(f; \boldsymbol{\theta}, \boldsymbol{\lambda})}S(f), \quad \text{(D7)}$$

where

$$A(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = \sqrt{A_+^2(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) + A_\times^2(f; \boldsymbol{\theta}, \boldsymbol{\lambda})}, \quad \text{(D8)}$$

and

$$A_+(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = \frac{1}{2}A_0(f; \boldsymbol{\theta})F_+(f; \boldsymbol{\theta}, \boldsymbol{\lambda})(1 + \cos^2\iota), \quad \text{(D9)}$$

$$A_\times(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = A_0(f; \boldsymbol{\theta})F_\times(f; \boldsymbol{\theta}, \boldsymbol{\lambda})\cos\iota, \quad \text{(D10)}$$

where the beam pattern functions are given in the *Mathematica* notebook. $S(f)$ is a sigmoid function,

$$S(f) = \frac{1}{1 + \exp(f - f_{\text{isco}})}, \quad \text{(D11)}$$

intended to create a gradual falloff of the signal to reduce the bias produced by a hard parameter-dependent frequency cutoff [50,77], where f_{isco} is the frequency at the innermost stable circular orbit. The overall phase including Earth's rotation is then

$$\Phi(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = \Psi_{3.5\text{PN}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) + \phi_{\text{L}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda})$$
$$+ \phi_{\text{P}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}), \quad \text{(D12)}$$

where $\phi_{\text{L}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = 2\pi f \Delta t_{\text{L}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda})$ and

$$\phi_{\text{P}}(f; \boldsymbol{\theta}, \boldsymbol{\lambda}) = -\arctan\left(\frac{A_+(f; \boldsymbol{\theta}, \boldsymbol{\lambda})F_+(f; \boldsymbol{\theta}, \boldsymbol{\lambda})}{A_\times(f; \boldsymbol{\theta}, \boldsymbol{\lambda})F_\times(f; \boldsymbol{\theta}, \boldsymbol{\lambda})}\right). \quad \text{(D13)}$$

APPENDIX E: FISHER CROSS TERMS IN THE SPA LIMIT

Nonzero block cross terms in the composite Fisher matrix involve two signals. In the frequency domain, cross terms correspond to the signals' phases interacting, creating a highly oscillatory integrand that cannot be easily numerically integrated. Block diagonal Fisher terms corresponding to the same signal do not have this issue, as the phase terms are the same and conjugation results in secular integrands. In what follows, we ignore the change in the signal at f_{isco}; however, this should not effect our results since all time-frequency overlaps occur at frequencies far below f_{isco}. The cross terms between two signals $\tilde{h}_1 = A_1 e^{i\Phi_1}$ and $\tilde{h}_2 = A_2 e^{i\Phi_2}$ are

$$I = (\tilde{h}_1'|\tilde{h}_2') = 4\text{Re}\left[\int_{-\infty}^{\infty}\frac{\tilde{h}_1'\tilde{h}_2'^*}{S_n(f)}df\right], \quad \text{(E1)}$$

where a prime denotes a derivative with respect to a single parameter. The integrand is

$$\frac{\tilde{h}_1'\tilde{h}_2'^*}{S_n(f)} = \frac{|\mathcal{A}|e^{i\Delta\Psi}}{S_n(f)}, \quad \text{(E2)}$$

where

$$\mathcal{A} = A_1' A_2' + \Phi_1' \Phi_2' A_1 A_2 + i(A_1 A_2' \Phi_1' - A_1' A_2 \Phi_2'), \quad \text{(E3)}$$

$$\theta = \tan^{-1} \left(\frac{A_1 A_2' \Phi_1' - A_1' A_2 \Phi_2'}{A_1' A_2' + \Phi_1' \Phi_2' A_1 A_2} \right), \quad \text{(E4)}$$

and

$$\Delta \Psi = \Phi_1 - \Phi_2 + \theta = \Delta \Phi + \theta. \quad \text{(E5)}$$

We evaluate these derivatives symbolically. The integrand is highly oscillatory, but it has a stationary point

$$\left. \frac{d\Delta\Psi}{df} \right|_{f_{\text{Fspa}}} = 0. \quad \text{(E6)}$$

We find the stationary point f_{Fspa} numerically. For monotonically increasing frequencies, there is at most a single stationary point. The integral is then

$$I \approx 4\text{Re} \left[\frac{|\mathcal{A}(f_{\text{Fspa}})|}{S_n(f)} \frac{\sqrt{2\pi}}{\sqrt{|d^2\Delta\Psi/df^2(f_{\text{Fspa}})|}} \right.$$
$$\left. \times \exp\left\{ i\Delta\Psi(f_{\text{Fspa}}) + \frac{\pi}{4} i \, \text{sign}[d^2\Delta\Psi/df^2(f_{\text{Fspa}})] \right\} \right]. \quad \text{(E7)}$$

The Fisher stationary point f_{Fspa} is not the same as the frequency at which the time-frequency tracks cross f_{ov},

$$\left. \frac{d\Delta\Phi}{df} \right|_{f_{\text{ov}}} = \left. \frac{d(\Phi_1 - \Phi_2)}{df} \right|_{f_{\text{ov}}} = 0. \quad \text{(E8)}$$

In practice, however, their difference is $\mathcal{O}(10^{-2})$ Hz. Though Eq. (E7) is evaluated self-consistently throughout with f_{Fspa}, in the main text we largely drop the distinction between the two frequencies.

APPENDIX F: INDEPENDENCE OF THE PARAMETER UNCERTAINTY RATIO ON THE SNR FOR TWO SIGNALS

The composite block Fisher matrix for two signals with a set of parameters $\theta = (\mathcal{M}_z, \eta, \ldots, d_L)$ is

$$\Gamma = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}, \quad \text{(F1)}$$

where A is the typical Fisher matrix for the first signal only, D is the typical Fisher matrix for the second signal only,

and B contains the overlap terms for both signals. The inverse of this matrix is [78]

$$\Gamma^{-1} = \begin{pmatrix} M^{-1} & -M^{-1}BD^{-1} \\ -D^{-1}B^T M^{-1} & D^{-1} + D^{-1}B^T M^{-1}BD^{-1} \end{pmatrix}, \quad \text{(F2)}$$

where $M = A - BD^{-1}B^T$ is the Schur complement of the D block. With this inverse, the parameter uncertainty ratio for the first signal is

$$\varrho = \sqrt{\frac{(M^{-1})^{ii}}{(A^{-1})^{ii}}} = \sqrt{\frac{((A - BD^{-1}B^T)^{-1})^{ii}}{(A^{-1})^{ii}}}. \quad \text{(F3)}$$

If we adjust the SNR of the second signal, the Fisher entries are modified,

$$B' = \alpha B \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & & \alpha \end{pmatrix}$$

$$D' = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & \alpha \end{pmatrix} \alpha^2 D \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & \alpha \end{pmatrix},$$

where D' and B' are the adjusted entries with modified SNR, and α is the factor of the SNR decrease, i.e., $\alpha = 5$ for a 5× decrease in the SNR. The new parameter uncertainty ratio is

$$\varrho' = \sqrt{\frac{((A - B'D'^{-1}B'^T)^{-1})^{ii}}{(A^{-1})^{ii}}} \quad \text{(F4)}$$

$$= \sqrt{\frac{((A - BD^{-1}B^T)^{-1})^{ii}}{(A^{-1})^{ii}}} = \varrho, \quad \text{(F5)}$$

thus it does not depend on α and the SNR ration between the two signals.

APPENDIX G: OVERLAP BETWEEN THREE SIGNALS

Correlations propagate between nonoverlapping signals if they both overlap with a third one. Appendix C showed that the overlap between two signals has support only when the signals overlap in time and frequency. Consider the case where signal 1 overlaps with signals 2 and 3 that do not overlap with each other. In this case, the composite Fisher matrix is

$$F = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{12} & F_{22} & 0 \\ F_{13} & 0 & F_{33} \end{pmatrix}, \quad \text{(G1)}$$

where each element is a square matrix with dimensionality equal to the number of signal parameters, subscripts denote the signal, and F_{ij}, $i \neq j$ represent cross terms. Since signals 2 and 3 do not overlap, $F_{23} = 0$. The 33 element of the inverse is

$$F_{33}^{-1} = \frac{F_{12}^2 - F_{11}F_{22}}{F_{13}^2 F_{22} + F_{12}^2 F_{33} - F_{11}F_{22}F_{33}}, \tag{G2}$$

which depends on signal 2, even though signal 3 does not overlap with it. As expected, if signals 1 and 2 did not overlap ($F_{12} = 0$), the signal 2 dependence drops out of signal 3.

[1] R. X. Adhikari *et al.* (LIGO Collaboration), A cryogenic silicon interferometer for gravitational-wave detection, Classical Quantum Gravity **37**, 165003 (2020).

[2] D. Reitze *et al.*, Cosmic Explorer: The U.S. contribution to gravitational-wave astronomy beyond LIGO, Bull. Am. Astron. Soc. **51**, 035 (2019).

[3] S. Hild *et al.*, Sensitivity studies for third-generation gravitational wave observatories, Classical Quantum Gravity **28**, 094013 (2011).

[4] C. Cutler and J. Harms, BBO and the neutron-star-binary subtraction problem, Phys. Rev. D **73**, 042001 (2006).

[5] Z. Wang, D. Liang, J. Zhao, C. Liu, and L. Shao, Anatomy of parameter-estimation biases in overlapping gravitational-wave signals, Classical Quantum Gravity **41**, 055011 (2024).

[6] P. Amaro-Seoane *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[7] D. Hils, P. L. Bender, and R. F. Webbink, Gravitational radiation from the Galaxy, Astrophys. J. **360**, 75 (1990).

[8] T. Robson and N. Cornish, Impact of galactic foreground characterization on a global analysis for the LISA gravitational wave observatory, Classical Quantum Gravity **34**, 244002 (2017).

[9] J. Crowder and N. J. Cornish, LISA source confusion, Phys. Rev. D **70**, 082004 (2004).

[10] E. Racine and C. Cutler, Gaussianity of LISA's confusion backgrounds, Phys. Rev. D **76**, 124033 (2007).

[11] T. Regimbau and S. A. Hughes, Gravitational-wave confusion background from cosmological compact binaries: Implications for future terrestrial detectors, Phys. Rev. D **79**, 062002 (2009).

[12] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170817: Implications for the stochastic gravitational-wave background from compact binary coalescences, Phys. Rev. Lett. **120**, 091101 (2018).

[13] E. Pizzati, S. Sachdev, A. Gupta, and B. Sathyaprakash, Toward inference of overlapping gravitational-wave signals, Phys. Rev. D **105**, 104016 (2022).

[14] A. Samajdar, J. Janquart, C. Van Den Broeck, and T. Dietrich, Biases in parameter estimation from overlapping gravitational-wave signals in the third-generation detector era, Phys. Rev. D **104**, 044003 (2021).

[15] P. Relton and V. Raymond, Parameter estimation bias from overlapping binary black hole events in second generation interferometers, Phys. Rev. D **104**, 084039 (2021).

[16] T. Regimbau *et al.*, A Mock data challenge for the Einstein gravitational-wave telescope, Phys. Rev. D **86**, 122001 (2012).

[17] D. Meacher, K. Cannon, C. Hanna, T. Regimbau, and B. Sathyaprakash, Second Einstein Telescope mock data and science challenge: Low frequency binary neutron star data analysis, Phys. Rev. D **93**, 024018 (2016).

[18] S. Wu and A. H. Nitz, Mock data study for next-generation ground-based detectors: The performance loss of matched filtering due to correlated confusion noise, Phys. Rev. D **107**, 063022 (2023).

[19] P. Relton, A. Virtuoso, S. Bini, V. Raymond, I. Harry, M. Drago, C. Lazzaro, A. Miani, and S. Tiwari, Addressing the challenges of detecting time-overlapping compact binary coalescences, Phys. Rev. D **106**, 104045 (2022).

[20] A. L. Miller, N. Singh, and C. Palomba, Enabling multimessenger astronomy with continuous gravitational waves: Early warning and sky localization of binary neutron stars in Einstein Telescope, Phys. Rev. D **109**, 043021 (2024).

[21] Y. Himemoto, A. Nishizawa, and A. Taruya, Impacts of overlapping gravitational-wave signals on the parameter estimation: Toward the search for cosmological backgrounds, Phys. Rev. D **104**, 044010 (2021).

[22] L. Reali, A. Antonelli, R. Cotesta, S. Borhanian, M. Çalışkan, E. Berti, and B. S. Sathyaprakash, The impact of confusion noise on golden binary neutron-star events in next-generation terrestrial observatories, arXiv:2209.13452.

[23] Y. Dang, Z. Wang, D. Liang, and L. Shao, Impact of overlapping signals on parameterized post-Newtonian coefficients in tests of gravity, arXiv:2311.16184.

[24] J. Y. L. Kwok, R. K. L. Lo, A. J. Weinstein, and T. G. F. Li, Investigation of the effects of non-Gaussian noise transients and their mitigation in parameterized gravitational-wave tests of general relativity, Phys. Rev. D **105**, 024066 (2022).

[25] E. Payne, S. Hourihane, J. Golomb, R. Udall, D. Davis, and K. Chatziioannou, Curious case of GW200129: Interplay between spin-precession inference and data-quality issues, Phys. Rev. D **106**, 104017 (2022).

[26] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaborations), Population of merging compact binaries inferred using gravitational waves through GWTC-3, Phys. Rev. X **13**, 011048 (2023).

[27] A. Antonelli, O. Burke, and J. R. Gair, Noisy neighbours: Inference biases from overlapping gravitational-wave signals, Mon. Not. R. Astron. Soc. **507**, 5069 (2021).

[28] T. B. Littenberg and N. J. Cornish, Prototype global analysis of LISA data with multiple source types, Phys. Rev. D **107,** 063004 (2023).

[29] R. Smith et al., Bayesian inference for gravitational waves from binary neutron star mergers in third generation observatories, Phys. Rev. Lett. **127,** 081102 (2021).

[30] J. Langendorff, A. Kolmus, J. Janquart, and C. Van Den Broeck, Normalizing flows as an avenue to study overlapping gravitational wave signals, Phys. Rev. Lett. **130,** 171402 (2023).

[31] J. Janquart, T. Baka, A. Samajdar, T. Dietrich, and C. Van Den Broeck, Parameter estimation methods for analyzing overlapping gravitational wave signals in the third-generation detector era, arXiv:2211.01304.

[32] S. Biscoveanu, C. Talbot, E. Thrane, and R. Smith, Measuring the primordial gravitational-wave background in the presence of astrophysical foregrounds, Phys. Rev. Lett. **125,** 241101 (2020).

[33] J. Alvey, U. Bhardwaj, S. Nissanke, and C. Weniger, What to do when things get crowded? Scalable joint analysis of overlapping gravitational wave signals, arXiv:2308.06318.

[34] H. Zhong, R. Ormiston, and V. Mandic, Detecting cosmological gravitational waves background after removal of compact binary coalescences in future gravitational wave detectors, Phys. Rev. D **107,** 064048 (2023).

[35] M. Evans et al., A Horizon Study for Cosmic Explorer: Science, observatories, and community, arXiv:2109.09882.

[36] Cosmic explorer noise curves, https://dcc.cosmicexplorer.org/CE-T2000017/public (2022) (accessed September 20, 2023).

[37] M. Di Giovanni et al., Temporal variations of the ambient seismic field at the Sardinia candidate site of the Einstein Telescope, Geophys. J. Int. **234,** 1943 (2023).

[38] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors, Phys. Rev. D **80,** 084043 (2009).

[39] A. W. Steiner, M. Hempel, and T. Fischer, Core-collapse supernova equations of state based on neutron star observations, Astrophys. J. **774,** 17 (2013).

[40] I. Legred, K. Chatziioannou, R. Essick, S. Han, and P. Landry, Impact of the PSR J0740 + 6620 radius constraint on the properties of high-density matter, Phys. Rev. D **104,** 063003 (2021).

[41] P. Madau and M. Dickinson, Cosmic star-formation history, Annu. Rev. Astron. Astrophys. **52,** 415 (2014).

[42] N. Aghanim et al. (Plank Collaboration), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641,** A6 (2020).

[43] T. Regimbau et al., MDC_Generation (2022), https://dcc.ligo.org/LIGO-T1400401/public.

[44] S. Sachdev, T. Regimbau, and B. S. Sathyaprakash, Subtracting compact binary foreground sources to reveal primordial gravitational-wave backgrounds, Phys. Rev. D **102,** 024051 (2020).

[45] S. Borhanian and B. Sathyaprakash, Listening to the Universe with next generation ground-based gravitational-wave detectors, arXiv:2202.11048.

[46] Q. Hu and J. Veitch, Accumulating errors in tests of general relativity with gravitational waves: Overlapping signals and inaccurate waveforms, Astrophys. J. **945,** 103 (2023).

[47] D. Gabor, Theory of communication. Part 1: The analysis of information, J. Inst. Electr. Eng., Part 3 **93,** 429 (1946).

[48] M. Isi and W. M. Farr, Analyzing black-hole ringdowns, arXiv:2107.05609.

[49] M. Vallisneri, Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects, Phys. Rev. D **77,** 042001 (2008).

[50] C. L. Rodriguez, B. Farr, W. M. Farr, and I. Mandel, Inadequacies of the Fisher information matrix in gravitational-wave parameter estimation, Phys. Rev. D **88,** 084013 (2013).

[51] I. Mandel, C. P. L. Berry, F. Ohme, S. Fairhurst, and W. M. Farr, Parameter estimation on compact binary coalescences with abruptly terminating gravitational waveforms, Classical Quantum Gravity **31,** 155005 (2014).

[52] R. Abbott et al. (KAGRA, Virgo, and LIGO Scientific Collaborations), GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, Phys. Rev. X **13,** 041039 (2023).

[53] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries, Phys. Rev. D **49,** 6274 (1994).

[54] L. Reali, A. Maselli, and E. Berti, The impact of compact binary confusion noise on tests of fundamental physics with next-generation gravitational-wave detectors, arXiv:2307.01264.

[55] C. Plunkett, S. Hourihane, and K. Chatziioannou, Concurrent estimation of noise and compact-binary signal parameters in gravitational-wave data, Phys. Rev. D **106,** 104021 (2022).

[56] T. Regimbau, M. Evans, N. Christensen, E. Katsavounidis, B. Sathyaprakash, and S. Vitale, Digging deeper: Observing primordial gravitational waves below the binary black hole produced stochastic background, Phys. Rev. Lett. **118,** 151105 (2017).

[57] B. Zhou, L. Reali, E. Berti, M. Çalışkan, C. Creque-Sarbinowski, M. Kamionkowski, and B. S. Sathyaprakash, Subtracting compact binary foregrounds to search for sub-dominant gravitational-wave backgrounds in next-generation ground-based observatories, Phys. Rev. D **108,** 064040 (2023).

[58] B. Zhou, L. Reali, E. Berti, M. Çalışkan, C. Creque-Sarbinowski, M. Kamionkowski, and B. S. Sathyaprakash, Compact binary foreground subtraction in next-generation ground-based observatories, arXiv:2209.01221.

[59] H. Song, D. Liang, Z. Wang, and L. Shao, Compact binary foreground subtraction for detecting the stochastic gravitational-wave background in ground-based detectors, arXiv:2401.00984.

[60] A. Sharma and J. Harms, Searching for cosmological gravitational-wave backgrounds with third-generation detectors in the presence of an astrophysical foreground, Phys. Rev. D **102,** 063009 (2020).

[61] D. S. Bellie, S. Banagiri, Z. Doctor, and V. Kalogera, The unresolved stochastic background from compact binary mergers detectable by next-generation ground-based gravitational-wave observatories, arXiv:2310.02517.

[62] Mathematica, Version 13.3 (Wolfram, Inc., Champaign, IL, 2023) https://www.wolfram.com/mathematica.

[63] C. R. Harris *et al.*, Array programming with NumPy, Nature (London) **585,** 357 (2020).

[64] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in PYTHON, Nat. Methods **17,** 261 (2020).

[65] J. D. Hunter, MATPLOTLIB: A 2d graphics environment, Comput. Sci. Eng. **9,** 90 (2007).

[66] Astropy Collaboration, The ASTROPY Project: Building an open-science project and status of the v2.0 core package, Astron. J. **156,** 123 (2018).

[67] A. D. Johnson, Source confusion from neutron star binaries in ground-based gravitational wave detectors is minimal software, v0.1.0, Zenodo (2024), 10.5281/zenodo.10850947.

[68] N. Cornish (private communication).

[69] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).

[70] P. Ajith, Addressing the spin question in gravitational-wave searches: Waveform templates for inspiralling compact binaries with nonprecessing spins, Phys. Rev. D **84,** 084037 (2011).

[71] C. K. Mishra, A. Kela, K. Arun, and G. Faye, Ready-to-use post-Newtonian gravitational waveforms for binary black holes with nonprecessing spins: An update, Phys. Rev. D **93,** 084054 (2016).

[72] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, Systematic and statistical errors in a Bayesian approach to the estimation of the neutron-star equation of state using advanced gravitational wave detectors, Phys. Rev. D **89,** 103012 (2014).

[73] S. Borhanian, GWBENCH: A novel Fisher information package for gravitational-wave benchmarking, Classical Quantum Gravity **38,** 175014 (2021).

[74] F. Iacovelli, M. Mancarella, S. Foffa, and M. Maggiore, Forecasting the detection capabilities of third-generation gravitational-wave detectors using GWFAST, Astrophys. J. **941,** 208 (2022).

[75] F. Iacovelli, M. Mancarella, S. Foffa, and M. Maggiore, GWFAST: A Fisher information matrix PYTHON code for third-generation gravitational-wave detectors, Astrophys. J. Suppl. Ser. **263,** 2 (2022).

[76] P. Jaranowski, A. Krolak, and B. F. Schutz, Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection, Phys. Rev. D **58,** 063001 (1998).

[77] I. Mandel, C. P. L. Berry, F. Ohme, S. Fairhurst, and W. M. Farr, Parameter estimation on compact binary coalescences with abruptly terminating gravitational waveforms, Classical Quantum Gravity **31,** 155005 (2014).

[78] T.-T. Lu and S.-H. Shiou, Inverses of 2×2 block matrices, Comput. Math. Appl. **43,** 119 (2002).