

PAPER

Sensitivity of present and future detectors across the black-hole binary gravitational wave spectrum

To cite this article: A R Kaiser and S T McWilliams 2021 *Class. Quantum Grav.* **38** 055009

View the [article online](#) for updates and enhancements.

You may also like

- [The 7th Gravitational Wave Data Analysis Workshop, Kyoto, Japan, 17-19 December 2002](#)

Nobuyuki Kanda, Misao Sasaki and Hideyuki Tagoshi

- [Multimessenger analysis strategy for core-collapse supernova search: gravitational waves and low-energy neutrinos](#)
Odysse Halim, Claudio Casentini, Marco Drago et al.

- [Gravitational Waves, Volume 1: Theory and Experiments](#)
Eric Poisson



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Sensitivity of present and future detectors across the black-hole binary gravitational wave spectrum

A R Kaiser^{1,2,*}  and S T McWilliams^{1,2} 

¹ Department of Physics and Astronomy, West Virginia University, PO Box 6315, Morgantown, WV 26506, United States of America

² Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, United States of America

E-mail: ark0015@mix.wvu.edu

Received 2 October 2020, revised 9 December 2020

Accepted for publication 18 December 2020

Published 11 January 2021



CrossMark

Abstract

Black-holes are known to span at least 9 orders of magnitude in mass: from the stellar-mass objects observed by the Laser Interferometer Gravitational-Wave Observatory Scientific Collaboration and Virgo Collaboration, to supermassive black-holes like the one observed by the Event Horizon Telescope at the heart of M87. Regardless of the mass scale, all of these objects are expected to form binaries and eventually emit observable gravitational radiation, with more massive objects emitting at ever lower gravitational-wave frequencies. We present the tool, *gwent*, for modeling the sensitivities of current and future generations of gravitational wave detectors across the entire gravitational-wave spectrum of coalescing black-hole binaries (BHBs). We provide methods to generate sensitivity curves for pulsar timing arrays (PTAs) using a novel realistic PTA sensitivity curve generator (Hazboun, Romano and Smith 2019 *Phys. Rev. D* **100** 104028), space-based interferometers using adaptive models that can represent a wide range of proposed detector designs (Amaro-Seoane *et al* 2017 arXiv:[1702.00786](https://arxiv.org/abs/1702.00786)), and ground-based interferometers using realistic noise models that can reproduce current (Abbott *et al* 2016 *Phys. Rev. Lett.* **116** 061102), second, and third generation designs (Hild *et al* 2011 *Class. Quantum Grav.* **28** 094013), as well as novel variations of the essential design parameters. To model the signal from BHBs at any mass scale, we use phenomenological waveforms capable of modeling the inspiral, merger, and ringdown for sources with varying mass ratios and spins (Khan *et al* 2016 *Phys. Rev. D* **93** 044007; Husa *et al* 2016 *Phys. Rev. D* **93** 044006). Using this adaptable framework, we produce signal-to-noise ratios (SNR) for the combination of any modeled parameter, associated with either the detector or the source. By allowing

*Author to whom any correspondence should be addressed.

variation across each detector and source parameter, we can pinpoint the most important factors to determining the optimal performance for particular instrument designs. The adaptability of our detector and signal models can easily be extended to new detector designs and other models of gravitational wave signals.

Keywords: gravitational waves, data analysis, gravitational wave detectors, multimessenger astronomy

(Some figures may appear in colour only in the online journal)

1. Introduction

The observation of gravitational waves (GWs) from black-hole binaries (BHBs) allows us to probe the Universe on a wide range of scales. Stellar BHBs up to hundreds of solar masses can be observed from the ground, whereas more massive objects can only be observed from space. Massive binaries that are millions of solar masses require space-based detectors, whereas supermassive binaries that are billions of solar masses require detectors larger than our Solar System, and are currently being pursued with pulsar timing arrays (PTAs).

In September of 2015, the Advanced Laser Interferometer Gravitational Wave Observatory (LIGO) first detected GWs (Abbott *et al* 2016). Now with multiple observing runs under LIGO’s and Virgo’s belt, low mass compact binary mergers are now being studied en masse. More massive sources have not yet been observed, but we expect they will be in the near future using different types of GW observations. Here we focus on three types of GW observatories based on the extension of established technologies: ground-based and space-based interferometers, and PTAs. Future ground-based interferometric detectors, like Cosmic Explorer (Reitze *et al* 2019) or the Einstein Telescope (ET) (Hild *et al* 2011), will probe a wider parameter space of low mass mergers relative to Advanced LIGO/Virgo. Proposed space-based detectors, like the Laser Interferometer Space Antenna (LISA) and the Deci-Hertz Interferometer Gravitational wave Observatory (DECIGO), will probe milli-hertz (Amaro-Seoane *et al* 2017) and deci-hertz frequency GWs (Kawamura *et al* 2006), respectively, providing a bridge between currently covered gravitational wavebands. PTAs are currently attempting to detect low-frequency GWs in the nanohertz regime (Hobbs *et al* 2010; Arzoumanian *et al* 2018).

Because of the disparate ways in which each detector’s sensitivity curve is constructed, it has been difficult to treat them uniformly. In the past, PTA representations have been often overly simplistic when compared to interferometer detection sensitivities. With this paper we wish to provide a concise, broadband comparison of probable observations from a representative sample of detectors in each frequency band, and to provide relevant parameterized expressions so that these examples can be easily extended to alternative designs. Additionally, this uniform treatment we hope to clarify how to accurately and fairly compare the different methods of GW detection and apply their sensitivities to assess each detector’s response to the signal from an inspiraling BHB.

In section 2 we go through the conventions of instrument sensitivity representation, signal parameterization and characterization, and signal-to-noise ratio (SNR) calculation. In sections 3–5 we apply the methodology in section 2 to PTAs, space-based GW detectors, and ground-based GW detectors, respectively. In section 6 we discuss our results and future work.

2. Background

2.1. Sensitivity curves

When comparing GW detectors, it is useful to have a relatively quick estimate of their sensitivities. Typically this sensitivity attempts to include approximations to real analysis techniques to better represent the true measurement capabilities of the instrument (Moore *et al* 2015).

With this in mind, we characterize an instrument's sensitivity using its power spectral density (PSD), or the time-averaged power per unit frequency of the noise, normalized by the instrument's response to a GW. To provide an easy visual comparison of different detectors' sensitivities to various signals, one can convert this normalized PSD to an effective characteristic strain amplitude, which gives the source strain that would be equivalent to the level of instrument noise (Thrane and Romano 2013).

2.1.1. Spectral density. It is first useful to define a few quantities that are frequently used in sensitivity estimation. We refer to the PSD of the detector as $P_n(f)$, and therefore the amplitude spectral density is $\sqrt{P_n(f)}$. The PSD is often defined as a combination of power laws that are set by the different types of noise inherent to the particular detector; for instance, seismic, thermal, and shot noises for (Advanced) LIGO, acceleration and position noises for LISA, and white (or red) noise timing uncertainties for PTAs (Moore *et al* 2015). Because each PSD estimate depends on the specific design, we define the parameters that determine specific detectors' performances in the following sections: we discuss PTAs in section 3, space-based interferometers in section 4, and ground-based interferometers in section 5.

2.1.2. Instrument strain representations. To define a more useful representation of instrument sensitivity, we combine the PSD, $P_n(f)$, and response function, $\mathcal{R}(f)$, of the detector to define the effective noise PSD, $S_n(f)$ as (Wainstein *et al* 1963)

$$S_n(f) = \frac{P_n(f)}{\mathcal{R}(f)}. \quad (1)$$

This quantity is sometimes displayed for sensitivity plots, but a more useful quantity is the noise's effective strain amplitude,

$$h_n(f) = \sqrt{f S_n(f)}, \quad (2)$$

i.e. the PSD is converted from a spectral density to a strain power, and then square-rooted to yield a strain amplitude. As we will see later, one can visually estimate the SNR by simply comparing the height of the effective strain noise amplitude, $h_n(f)$, and the characteristic strain of the source, $h_c(f)$, which we define in section 2.2.

2.1.3. Response functions. The response function, $\mathcal{R}(f)$, characterizes the frequency dependence of the stretching and squeezing experienced by a detector in the presence of a passing GW. Since the detector response also depends on the sky location and polarization of the source, as characterized by the antenna pattern, we are often interested in including this dependence in an averaged sense within $\mathcal{R}(f)$. The instrument can additionally have some correlation between the detector arms (or pulsar pairs in the case of PTAs) (Thrane and Romano 2013). We will summarize the construction of $\mathcal{R}(f)$ in this subsection, but see Maggiore (2000) for further details on response functions, antenna patterns, and inter-instrument correlations.

For a GW propagating in the $\hat{\Omega}$ direction with transverse orthonormal vectors \hat{m}, \hat{n} defined by,

$$\hat{\Omega} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \quad (3)$$

$$\hat{m} = (\sin \phi, -\cos \phi, 0) \quad (4)$$

$$\hat{n} = (\cos \theta \cos \phi, \cos \theta \sin \phi, -\sin \theta), \quad (5)$$

the strain is given by the tensor,

$$\mathbf{h}(t, \mathbf{x}) = h_+ \left(2\pi f(t - \hat{\Omega} \cdot \mathbf{x}); \iota \right) \epsilon^+(\hat{\Omega}, \psi) + h_\times \left(2\pi f(t - \hat{\Omega} \cdot \mathbf{x}); \iota \right) \epsilon^\times(\hat{\Omega}, \psi), \quad (6)$$

made up of the two GW polarizations $+$ and \times , where \mathbf{x} is a vector defining the spatial position of the source, ψ is the polarization angle describing a rotation of \hat{m} and \hat{n} about $\hat{\Omega}$ (i.e. $\hat{m}, \hat{n} \rightarrow \hat{u}, \hat{v}$) such that $\cos \psi = \hat{m} \cdot \hat{u}$, and ι is the inclination between the propagation direction $\hat{\Omega}$ and the orbital angular momentum unit vector \hat{L} , i.e. $\cos \iota = \hat{\Omega} \cdot \hat{L}$. See figures 6 and 7 of Hazboun *et al* (2019) for diagrams of the coordinate system. The polarization tensors

$$\epsilon^+(\hat{\Omega}, \psi) = \mathbf{e}^+(\hat{\Omega}) \cos 2\psi - \mathbf{e}^\times(\hat{\Omega}) \sin 2\psi \quad (7)$$

$$\epsilon^\times(\hat{\Omega}, \psi) = \mathbf{e}^+(\hat{\Omega}) \sin 2\psi + \mathbf{e}^\times(\hat{\Omega}) \cos 2\psi, \quad (8)$$

in equation (6) are made up of the tensor products of the transverse basis vectors:

$$\mathbf{e}^+(\hat{\Omega}) = \hat{m} \otimes \hat{m} - \hat{n} \otimes \hat{n} \quad (9)$$

$$\mathbf{e}^\times(\hat{\Omega}) = \hat{m} \otimes \hat{n} + \hat{n} \otimes \hat{m} \quad (10)$$

(Anholm *et al* 2009). Equivalently, equation (6) can be written in terms of the Fourier transform of $\mathbf{h}(t)$,

$$\mathbf{h}(t, \mathbf{x}) = \sum_{A=+, \times} \int_{-\infty}^{\infty} df e^{i2\pi f(t - \hat{\Omega} \cdot \mathbf{x})} \tilde{h}_A(f; \hat{\Omega}, \iota) \epsilon^A(\hat{\Omega}, \psi) \quad (11)$$

where $\tilde{h}_{+, \times}(f; \iota)$ are the Fourier transforms of $h_{+, \times}(t; \iota)$. Therefore the response in the GW detector is

$$h(t; \hat{\Omega}, \iota, \psi) = \int_{-\infty}^{\infty} df \tilde{h}(f; \hat{\Omega}, \iota, \psi) e^{i2\pi f t} \quad (12)$$

where

$$\tilde{h}(f; \hat{\Omega}, \iota, \psi) = F^+(f, \hat{\Omega}, \psi) \tilde{h}_+(f; \iota) + F^\times(f, \hat{\Omega}, \psi) \tilde{h}_\times(f; \iota), \quad (13)$$

is the source strain made up of the two GW polarizations,

$$\tilde{h}_+(f; \iota) = A(f) \frac{(1 + \cos^2 \iota)}{2} e^{i\Psi_{+,+}(f)} \quad (14)$$

$$\tilde{h}_\times(f; \iota) = iA(f) \cos \iota e^{i\Psi_{\times,+}(f)}, \quad (15)$$

where $A(f)$ is the strain amplitude and $\Psi_{\times,+}(f)$ is the strain phase (Maggiore 2007). $F^{+, \times}(\hat{\Omega}, \psi, f)$ in equation (13) is the complete response function, incorporating the frequency, antenna pattern, and polarization dependencies for a specific instrument. We discuss PTA detector response functions in section 3.2.1 and interferometer response functions in section 4.2.

Since we are interested here in the typical sensitivity to a source, we generally average over the source sky location, inclination, and polarization,

$$\begin{aligned}
\langle \tilde{h}(f; \hat{\Omega}, \iota, \psi) \tilde{h}^*(f; \hat{\Omega}, \iota, \psi) \rangle &= \langle F^+(\hat{\Omega}, \psi, f) F^{+*}(\hat{\Omega}, \psi, f) \rangle \langle \tilde{h}_+(f; \iota) \tilde{h}_+^*(f; \iota) \rangle \\
&\quad + \langle F^\times(\hat{\Omega}, \psi, f) F^{\times*}(\hat{\Omega}, \psi, f) \rangle \langle \tilde{h}_\times(f; \iota) \tilde{h}_\times^*(f; \iota) \rangle \\
&= \langle F^+(\hat{\Omega}, \psi, f) F^{+*}(\hat{\Omega}, \psi, f) \rangle \\
&\quad \times \left(\langle \tilde{h}_+(f; \iota) \tilde{h}_+^*(f; \iota) \rangle + \langle \tilde{h}_\times(f; \iota) \tilde{h}_\times^*(f; \iota) \rangle \right), \quad (16)
\end{aligned}$$

since $\langle F^+(f) F^{+*}(f) \rangle = \langle F^\times(f) F^{\times*}(f) \rangle$ (Maggiore 2007). Thus,

$$\begin{aligned}
\langle \tilde{h}(f; \hat{\Omega}, \iota, \psi) \tilde{h}^*(f; \hat{\Omega}, \iota, \psi) \rangle &= \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \\
&\quad \times \int_0^{2\pi} \frac{d\psi}{2\pi} F^+(\hat{\Omega}, \psi, f) F^{+*}(\hat{\Omega}, \psi, f) \\
&\quad \times \int_{-1}^1 \frac{d(\cos \iota)}{2} \left[|\tilde{h}_+(f; \iota)|^2 + |\tilde{h}_\times(f; \iota)|^2 \right] \\
&= \frac{4}{5} \mathcal{R}(f) |\tilde{h}(f)|^2 \quad (17)
\end{aligned}$$

where the final equality defines $\mathcal{R}(f)$, $|\tilde{h}(f)|^2 = |\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2$ is the Fourier strain amplitude, and the factor of $\frac{4}{5}$ comes from angle averaging of the inclination angle (see e.g. Thorne (1989)). We discuss instrument-specific response functions in the sections that follow, and refer back to this section for general statements.

In figure 1, we display the characteristic strain sensitivity curves for the frequency bands of each detector. The plotted models are representative of current GW detector designs in the literature and all are generated with `gwent`. We discuss the precise make-up of each sensitivity curve in the relevant sections.

2.2. Signal parameterization

The entire panoply of GW signals includes a variety of source types. We choose to focus on the main class of source that is common to all three detector bands: inspiraling BHs.

2.2.1. Inspiring black-hole binaries. BHs are an important source of GW signals detectable with current and future observatories. The loss of energy and angular momentum due to gravitational-wave emission causes the BHB's orbit to decay, eventually driving it to merge (Thorne 1989). The coalescence of two black-holes is generally divided into three epochs: inspiral, merger, and ringdown (Flanagan and Hughes 1998).

The inspiral portion of BHB evolution is the longest of the three stages. During the inspiral, the binary emits gravitational radiation and evolves at an initially slow, well known rate that can be accurately modeled using the post-Newtonian (PN) approximation (Einstein *et al* 1938). The merger portion of coalescence is the final collision of the BHB to form a single black-hole. During this epoch, the separations are small, and the black-holes are moving at relativistic speeds; numerical relativity simulations are needed to extract information from the binary (Boyle *et al* 2019).

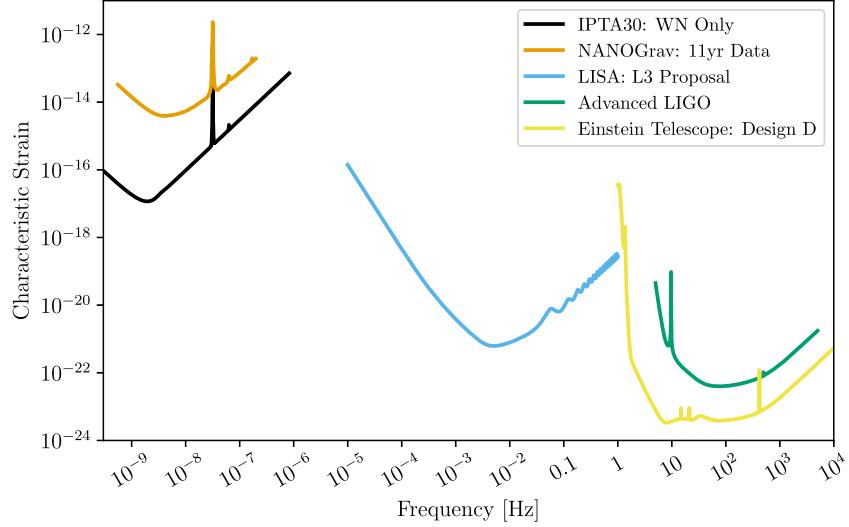


Figure 1. Broadband characteristic strain sensitivity to individual coalescing BH sources. The IPTA30 model uses our white noise only model with $\sigma_{\text{RMS}} = 100$ ns, $T_{\text{obs}} = 20$ yrs, $N_p = 200$, and a one week cadence. The NANOGrav curve uses the 11 yr data modeled with `hasasia`. See table 1 and section 3.3 for more information about each PTA model. The LISA design uses the proposed design from third large mission (L3) of the European Space Agency’s Cosmic Vision Plan (Amaro-Seoane *et al* 2017). See section 4 for more specific details on our LISA model. The aLIGO and ET curves that represent the design sensitivities are taken from the references in section 5.

Post-merger, the resulting Kerr black-hole continues to emit gravitational radiation through oscillations of its quasinormal modes, an epoch known as the ringdown (Teukolsky and Press 1974). During this final phase, the black-hole emits GWs as a superposition of exponentially damped sinusoids.

We use a combination of methods to estimate the signal strength. For binaries that do not evolve over the instrument’s observation time, we simply utilize the leading order quadrupole contribution in the Fourier domain via the stationary phase approximation (Droz *et al* 1999). In the case of evolving, or ‘chirping’ binaries that inspiral (i.e. whose frequency evolves by more than the frequency resolution of the observation) over the detector’s observation period, we use a phenomenological model developed in Husa *et al* (2016) and Khan *et al* (2016).

Chirping sources. As described above, coalescing binaries move through several stages of GW emission. In order to combine the PN approximation of the inspiral with the numerical relativity (NR) results for the merger and ringdown, Husa *et al* (2016) developed a fairly simple, non-precessing model of the GW signal in the frequency domain. Their method uses a hybrid model of inspiral approximates matched with NR simulations to provide an easily adopted, accurate waveform approximate. We implemented their method for the first time in a Python package as opposed to using the LALSuite implementation in C (LIGO Scientific Collaboration 2018).

The output from IMRPhenomD gives the amplitude and phase of the dominant quadrupolar modes of the GW signal,

$$\tilde{h}(f; \Xi, \theta, \phi) = \tilde{h}_+(f; \Xi, \theta, \phi) - i\tilde{h}_\times(f; \Xi, \theta, \phi) \quad (18)$$

$$= \sum_{m=-2,2} \tilde{h}_{2m}(f; \Xi)^{-2} Y_{2m}(\theta, \phi), \quad (19)$$

where $\Xi \in (M, \eta, \chi_1, \chi_2)$ are the total mass, symmetric mass ratio, and dimensionless spin parameters, respectively, (θ, ϕ) are the observer's orientation with respect to the orbital angular momentum of the binary, and Y_{lm} are spin-weighted spherical harmonics. Since $\tilde{h}_{2(-2)}(f) = \tilde{h}_{22}^*(-f)$, we can parameterize the quadrupolar amplitude as

$$\tilde{h}_{22}(f; \Xi) = A(f; \Xi) e^{-i\Psi(f; \Xi)}, \quad (20)$$

where $A(f; \Xi)$ is now the geometrized strain amplitude and $\Psi(f; \Xi)$ is the geometrized strain phase.

The IMRPhenomD signal amplitude has the leading order PN contribution factored out through normalizing $A(f; \Xi)$ by

$$A_0 = \sqrt{\frac{2\eta}{3\pi^{1/3}}} f^{-7/6}, \quad (21)$$

where $\eta = q/(1+q)^2$ is the symmetric mass ratio and $q = m_2/m_1$ is the mass ratio (in our convention $m_1 \leq m_2$) and m_1 and m_2 are the individual black-hole masses (Cutler and Flanagan 1994). Thus we define the output from IMRPhenomD as $\hat{A}(f; \Xi) = A(f; \Xi)/A_0$. To get the non-normalized amplitude in non-geometrized units and extract the raw Fourier strain amplitude ($|\tilde{h}(f)|$), we assume that as $f \rightarrow 0$ (i.e. time until merger $\rightarrow \infty$), $|\tilde{h}(f)|$ goes to the leading order post-Newtonian amplitude,

$$\begin{aligned} |\tilde{h}(f)| &= \sqrt{\frac{5}{16\pi}} \frac{M^2}{D_L(z)} A_0(Mf) \hat{A}(Mf; \Xi) \\ &= \sqrt{\frac{5}{24}} \frac{\mathcal{M}^{5/6}}{\pi^{2/3} D_L(z)} f^{-7/6} \hat{A}(f; \Xi), \end{aligned} \quad (22)$$

where $D_L(z)$ is the source's luminosity distance at a redshift z , $M = m_1 + m_2$ is the total mass of the system, and $\mathcal{M} = \eta^{3/5} M$ is the chirp mass (Cutler and Flanagan 1994). To convert from natural units (i.e. $G = c = 1$), one would multiply each M by GM_\odot/c^3 and divide the luminosity distance by the speed of light ($D_L = D_L/c$). To get the mass in the detector frame, one would redshift the mass to: $M_z = (1+z)M$. See figure 2 for examples of chirping source strain given by the IMRPhenomD formalism.

Monochromatic sources. For a non-evolving binary, a.k.a. one that remains in the same GW frequency bin over many orbital revolutions, the source strain can be accurately approximated using the quadrupole and stationary-phase approximations (Droz *et al* 1999).

To determine whether a source evolves during an observation, either an initial frequency or a time before merger at which the source is observed must be chosen. We select the most sensitive frequency of the detector, f_{opt} . Thus, the change in GW frequency is,

$$\dot{f} = \frac{96}{5\pi\mathcal{M}^2} (\pi\mathcal{M}f)^{11/3}, \quad (23)$$

which is integrated to find the instantaneous GW frequency,

$$f(\tau) = \frac{1}{8\pi\mathcal{M}} \left(\frac{5\mathcal{M}}{\tau} \right)^{3/8}, \quad (24)$$

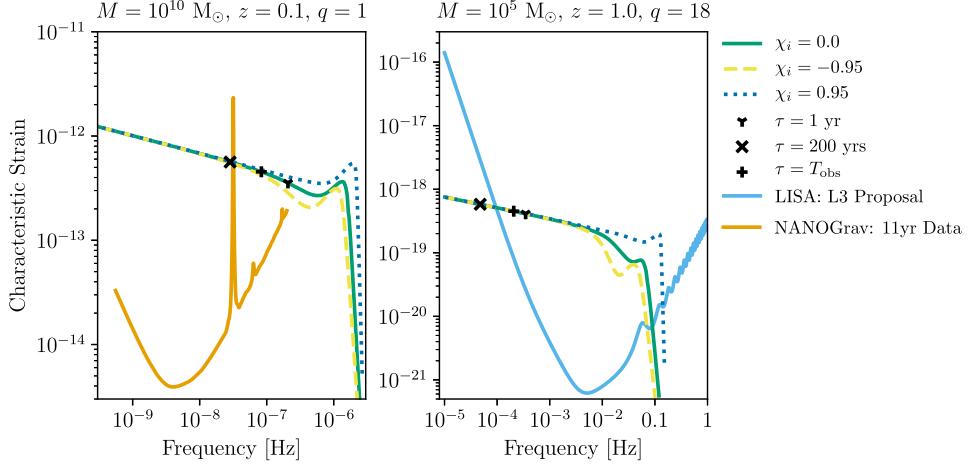


Figure 2. Example GW sources for a PTA source (left) and a LISA source (right) plotted with GW detector sensitivity curves for the NANOGrav 11 yr data and the LISA L3 proposal, respectively. The definitions for the instrument characteristic strain is found in equation (2) and for the source characteristic strain in equation (34). Each chirping GW source, have the same mass ($M = 10^{10}(10^5)M_{\odot}$), redshift ($z = 0.1(1.0)$), and mass ratio ($q = 1(18)$) with three different spin states, no spin ($\chi_i = 0$), aligned spin ($\chi_i = 0.95$), and anti-aligned spin ($\chi_i = -0.95$). The frequency evolution of the chirping sources, we mark three times to merger (in equation (24), $\tau = 1$ yr, T_{obs} yrs, and 200 yrs) for each source.

where $\tau = t_c - t$ and t_c is an integration constant referred to as the time to coalescence. By inverting equation (24), we arrive at the time to merger,

$$\tau = \frac{5\mathcal{M}}{256}(\pi\mathcal{M}f)^{-8/3}. \quad (25)$$

The change in frequency from the initial observed time until some observation time of the detector, using equation (24) and Taylor expanding, is

$$\begin{aligned} \Delta f &= f(\tau) - f(\tau - T_{\text{obs}}) \\ &\simeq \frac{3}{64\pi\mathcal{M}} \left(\frac{5\mathcal{M}}{\tau} \right)^{3/8} \frac{T_{\text{obs}}}{\tau}, \end{aligned} \quad (26)$$

where T_{obs} is the detector's observation time. Thus, if the frequency has evolved by less than the frequency resolution of the detector ($1/T_{\text{obs}}$), i.e.

$$\Delta f < \frac{1}{T_{\text{obs}}} \quad (27)$$

the source is considered monochromatic.

If we assume, without loss of generality, that $t_c = 0$ and $t(f)$ is some time until coalescence, and use the assumed optimal frequency f_{opt} , we can substitute $t(f_{\text{opt}}) = t_{\text{init}} = \tau$ into equation (26), which gives us the desired time until coalescence for a binary initially observed at the detector's optimal frequency. See figure 2 for an example of how the source strain and frequency evolve with time.

The strain from a monochromatic source is given by,

$$h_+(t; \iota, \alpha_0) = h_0 \left(\frac{1 + \cos^2 \iota}{2} \right) \cos(2\pi f_0 t + \alpha_0) \quad (28)$$

$$h_\times(t; \iota, \alpha_0) = h_0 \cos \iota \sin(2\pi f_0 t + \alpha_0), \quad (29)$$

where h_0 is the approximately constant strain amplitude, f_0 is the GW frequency of the binary, ι is the binary's inclination, and the initial phase is α_0 .

The Fourier transform of equation $h_{+,\times}(t; \iota, \alpha_0)$

$$\tilde{h}_+(f; \iota, \alpha_0) = \frac{h_0}{2} \left(\frac{1 + \cos^2 \iota}{2} \right) [e^{i\alpha_0} \delta(f - f_0) + e^{-i\alpha_0} \delta(f + f_0)] \quad (30)$$

$$\tilde{h}_\times(f; \iota, \alpha_0) = \frac{h_0}{2i} \cos \iota [e^{i\alpha_0} \delta(f - f_0) - e^{-i\alpha_0} \delta(f + f_0)], \quad (31)$$

can be used in equation (17) to get the angle averaged detector response.

If we use the methods in Moore *et al* (2015) for an inspiralling source in the stationary-phase approximation, and separate out the inclination terms from equation (28), the Fourier strain amplitude is

$$|\tilde{h}(f)| \simeq \frac{h_0}{\sqrt{2f}}. \quad (32)$$

This relation allows us to use the amplitude of the very early inspiral portion of the phenomenological model as it asymptotes to the stationary-phase approximation to extract the monochromatic strain amplitude h_0 . As there is little to no numerical difference between this approach and the typical leading-order assumptions made for monochromatic sources, which we lay out below, we choose to use the usual approach because of its computationally simpler and faster properties.

Thus we use the constant strain amplitude h_0 , given by

$$h_0 = 4 \frac{\mathcal{M}^{5/3}}{D_L} (\pi f_0)^{2/3}. \quad (33)$$

Note that we use the convention of a factor of 4 in the strain amplitude, because of our definitions of equation (28). In the literature this factor is sometimes represented by a factor of 2, with a corresponding change in the inclination terms that come out to the same result (see e.g. Thorne (1989)). If, for our selected optimal frequency, f_{opt} , a binary obeys the inequality in equation (27), then we assume $f_0 = f_{\text{opt}}$ and that the source emits at f_{opt} for the entire observation. See figure 3 for examples of monochromatic strain amplitudes, h_0 , in the PTA frequency band. In the same vein as equation (22), one must convert the luminosity distance and the chirp mass from natural units.

2.2.2. Stochastic backgrounds. Stochastic backgrounds from various GW sources are present in every frequency band (Thrane and Romano 2013). As a consequence of this, there are many studies on detectability and characterization of backgrounds (Jenet *et al* 2006; Arzoumanian *et al* 2018; Cornish and Robson 2017; Parida *et al* 2016). Because this paper focuses on resolvable, inspiraling BHBs, we treat stochastic backgrounds as an additional noise source. We discuss any specific contributions from stochastic foregrounds in the respective detector sections.

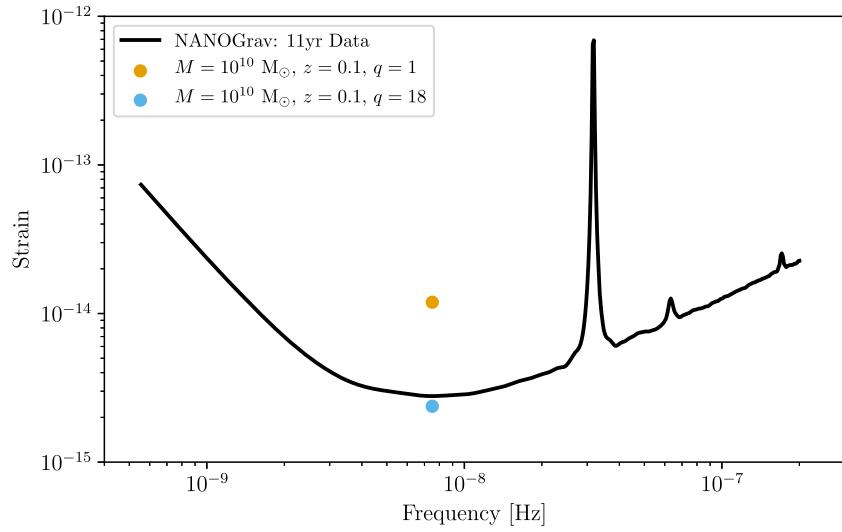


Figure 3. Example monochromatic GW sources for a PTA source plotted with GW detector sensitivity curve for the NANOGrav 11 yr data corresponding to a source strain (h_0) with SNR of one (see equation (39)); note that this is not characteristic strain. Each source, have the same mass ($M = 10^{10} M_\odot$), redshift ($z = 0.1$), and different mass ratios ($q = 1$ and $q = 18$). The sources are assumed to be emitting at a GW frequency corresponding to the instrument’s most sensitive frequency.

2.2.3. Characteristic strain. In order to represent general sources, we choose to absorb the ψ, θ and ϕ terms by using a polarization and orientation averaging (see equations (13) and (17)). This allows us to directly transition between the monochromatic and inspiraling strains. In order to get a more useful representation for comparing the signal strain to the sensitivity curves and for calculating the SNR, we use the characteristic strain of the source,

$$h_c^2(f) = 4f^2|\tilde{h}(f)|^2. \quad (34)$$

As previously mentioned, one can visually estimate the SNR by comparing the ratio of the characteristic strain of the source, $h_c(f)$, to the effective strain noise amplitude, $h_n(f)$, multiplied by the base-10 logarithm of the bandwidth, which is made simpler by the fact that these plots are generally shown in logscale.

In figure 2 we show a comparison of the chirping strain generated by IMRPhenomD for various source configurations. The cases displayed correspond to an array of coalescing BHs with varying spin parameters, and whose masses and redshifts are chosen so that the characteristic strain lies in the frequency bands of PTAs and LISA. For each respective mass and redshift of the BHs in the figure, we display the time to coalescence at different time slices in the inspiral. This allows for some insight into the timescale of inspiral versus the merger of various sources.

2.3. Matched filter and the signal-to-noise ratio

Since the strain emitted from all but the strongest GW sources are below the instrument’s noise fluctuations, we use matched filtering to maximize the sensitivity. Matched filtering can be used for many applications (e.g. parameter estimation, signal significance, etc); in this case, we use it to find local maxima of the match between the signal and the data to maximize the

SNR. This method assumes that the noise in the signal is stationary and Gaussian, and that it is uncorrelated with the signal (Maggiore 2007).

In this work we use the typical optimal filter argument and arrive at,

$$\rho^2 = 4 \int_0^\infty df \frac{|\tilde{h}(f)|^2}{S_n(f)} = (\tilde{h}(f)|\tilde{h}(f)). \quad (35)$$

the most frequently used form of the optimal SNR, ρ . For a full derivation, see section 7.3 of Maggiore (2007). Using equations (2) and (34), we can rewrite the SNR in equation (35) as

$$\rho^2 = \int_{-\infty}^\infty d(\ln f) \left[\frac{h_c(f)}{h_n(f)} \right]^2, \quad (36)$$

which, as mentioned previously, allows for convenient visual estimation of SNR when the characteristic source strain and effective strain sensitivity of detectors are plotted together (Flanagan and Hughes 1998)

2.3.1. Monochromatic sources. For monochromatic sources, the optimal filter is limited to the smallest bandwidth around the frequency of the binary. As mentioned in section 2.2.1, instruments have a finite resolution limited by the observation time, T_{obs} . Since we again use the average over solid angle and polarization, then using equation (35) in tandem with (17), the squared SNR becomes,

$$\rho^2 = h_0^2(f_0) \int_0^\infty df \frac{\delta(f - f_0) \delta(0)}{S_n(f)} \quad (37)$$

where f_0 is the frequency of the binary, h_0 is given in equation (33).

To account for the finite observing time, we must replace the Dirac deltas with finite-time Dirac deltas,

$$\delta(f) = \int_{-\infty}^\infty dt e^{i2\pi ft} \rightarrow \int_{-T_{\text{obs}}/2}^{T_{\text{obs}}/2} dt e^{i2\pi ft}, \quad (38)$$

which is allowed over the assumed window of equation (27) (Maggiore 2007). In particular, from equation (38), $\delta(0) \rightarrow T_{\text{obs}}$.

Thus the squared SNR becomes,

$$\rho^2 = \frac{h_0^2(f_0) T_{\text{obs}}}{S_n(f_0)} = \frac{h_0^2(f_0) N_{\text{cycles}}}{h_n^2(f_0)} = \left(\frac{h_c(f_0)}{h_n(f_0)} \right)^2, \quad (39)$$

where N_{cycles} is the number of waveform cycles that occur over the observation time, and the last equality serves to define h_c for monochromatic binaries (Maggiore 2007). Unlike in the chirping case where the waveform matching builds up over the changing frequency, the signal builds at the frequency of the source over the entire observation time.

In figure 4, we display the match-filtered SNR for the sensitivities of GW detectors across the entire gravitational-wave spectrum of coalescing BHBs. For figure 4 and for all relevant plots we use the flat Λ CDM, nine-year Wilkinson Microwave Anisotropy Prob (WMAP9) cosmology provided by `astropy` (Astropy Collaboration *et al* 2018; Hinshaw *et al* 2013). Using `gwent`, one can easily calculate the SNR for sensitivities of current and future GW detectors for any coalescing BHB parameter. Here, also using this formalism, we display the most relevant multiband overlap in parameter space in the literature, redshift and total source mass.

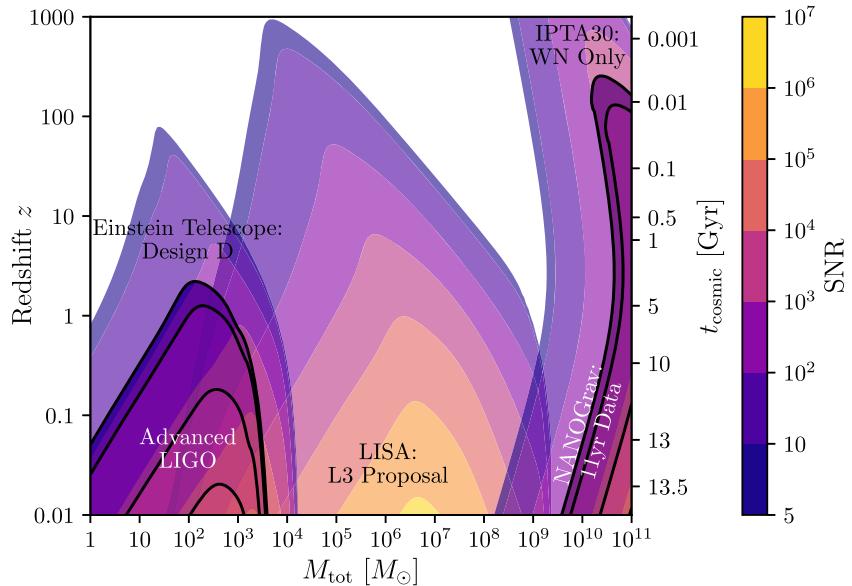


Figure 4. SNRs with respect to variations in source BHBs redshift and total source mass for ET design model D, aLIGO, LISA L3 proposal, IPTA30 WN only, and NANOGrav 11 yr. All curves assume no stochastic foreground, an individual, non-spinning ($\chi_i = 0$), equal mass ($q = 1$) BHB as the source, and WMAP9 cosmological parameters. All detectors have complementary coverage to study the full range of black hole masses at various stages of their evolution across the observable Universe. The particular regions of overlap between different detectors allow for observations of the same source class throughout the BHB's evolution.

Clearly, as instruments become more sensitive, the synergy between detector bands increases dramatically.

3. PTA sensitivity

PTAs search for sources of low-frequency GW signals by observing pulsar times-of-arrival (TOAs) and looking for deviations from models of the intrinsic pulsar TOAs. Most of the analyses take place in residual space, where residuals are the predicted model TOAs subtracted from the observed TOAs. In theory, one could detect GWs using individual pulsars (Sazhin 1978). To attribute potential GW signatures in pulsar residuals to a signal and not some spurious noise source however, one needs to use correlations between pulsars. By using correlations between different pulsars, global variances, like GW signals, can be extracted from the residuals.

Previous PTA sensitivity curve generators have often represented PTAs as a wedge with a sharp cutoff around $f = T_{\text{obs}}^{-1}$ (Thrane and Romano 2013). In fact, PTAs do maintain some sensitivity at frequencies below this cutoff, with the level being limited by pulsar timing model fitting (Hazboun *et al* 2019). One point of confusion with PTA sensitivity curve representation is the various ways in which they are presented. As pointed out in Moore *et al* (2015), one must be careful to not assume a similar response to different GW sources.

Here we focus on searches for GWs from supermassive black-hole binaries (SMBHBs). Often these searches are grouped into two main categories: stochastic signals from a superposition of unresolved SMBHBs, and individually resolvable SMBHBs. This stochastic

signal, referred to as the stochastic GW background (GWB), is common to all pulsars across the sky. The GWB is hypothesized to follow a power law spectrum and induce a quadrupolar signature on the spatial correlations between each pairing of pulsars in the array, known as the Hellings–Downs (HD) correlation (Hellings and Downs 1983).

Signals from individual SMBHBs do not induce an HD correlation between pulsars, but instead cause a correlation that depends on the source’s location. The strength of the correlation depends on the source’s location and its amplitude as given by equation (33). The formalism developed here focuses on these individual SMBHBs, not the GWB.

There are ongoing searches for both predicted types of GW signals, but we choose to focus on individually resolvable SMBHBs as they are directly comparable to BHB mergers across the GW spectrum. For more information about searches for stochastic backgrounds, and the most recent NANOGrav PTA upper-limits, see Arzoumanian *et al* (2018). For the most recent results of the NANOGrav PTA search for individually resolvable SMBHBs, see Aggarwal *et al* (2018).

3.1. PTA sensitivity curves

Since we are interested in presenting sensitivities of PTAs in a similar way to the formalism used in ground and space-based interferometers, we outline PTA noise sources in terms of power spectral densities. In reality, PTAs operate first in the time domain in terms of TOAs and fitting individual pulsar noise properties. This is because unlike interferometers, PTAs do not have noise-only channels. Additionally, each pulsar can have unknown noise properties and sources. Thus, an amount of fitting is needed to characterize the TOAs of each pulsar (Arzoumanian *et al* 2018). Once a fit is generated for each pulsar, the modeled TOAs are subtracted from the observed TOAs, creating residuals that should only contain non-deterministic noise sources and signals (i.e. GWs). Our analysis treats PTAs in terms of these residuals, where the noise in individual pulsars simply contains some white and/or red noise that accounts for any remaining noise in the modeled TOAs.

We use the sensitivity generation results of Hazboun *et al* (2019) (via their code `hasasia`) in constructing the characteristic strain plots for PTAs and in the SNR calculations. We reproduce the relevant equations for easy comparison between PTA sensitivity calculations and interferometers.

3.2. PTA power spectral density

The overall PSD of the residuals for a PTA is made up of the PSD from one or more stochastic backgrounds (SBs), $P_{\text{SB}}(f)$, which is common to all N_p pulsars in the array, and the sum of all individual pulsar noises, $P_i(f)$ (Thrane and Romano 2013; Lam 2018):

$$P_{\text{n}}(f) = P_{\text{SB}}(f) + \sum_i^{N_p} P_i(f). \quad (40)$$

Here we only consider power-law modeled SBs, which can be parameterized as

$$P_{\text{SB}}(f) = \frac{A_{\text{SB}}^2}{12\pi^2} \left(\frac{f}{f_{\text{year}}} \right)^{2\alpha} f^{-3}, \quad (41)$$

where A_{SB} is the strain amplitude at a frequency of 1 yr^{-1} (i.e. f_{year}) and α is the spectral index of the characteristic strain. For a SB dominated by SMBHBs (i.e. the GWB), $\alpha = -2/3$ is assumed (Jenet *et al* 2006). We assume a conservative amplitude $A_{\text{GWB}} = 4 \times 10^{-16}$ found as

the median value of SMBHBs merger rate estimates in Sesana *et al* (2016). If we include the GWB in our framework, it appears as the dominant ‘instrument’ noise at lower frequencies.

Each pulsar’s PSD,

$$P_i(f) = P_{\text{RN},i}(f) + P_{\text{WN},i}, \quad (42)$$

is made up of two components: the pulsar red noise (RN),

$$P_{\text{RN},i}(f) = A_{\text{RN},i} \left(\frac{f}{f_{\text{year}}} \right)^{-\gamma_i}, \quad \gamma_i > 0, \quad (43)$$

where $A_{\text{RN},i}$ is the amplitude of the pulsar red noise and γ_i is the red noise power spectral index, and the pulsar white noise (WN),

$$P_{\text{WN},i} = 2 \frac{\sigma_i^2}{c_i}, \quad (44)$$

which is determined by the observing cadence c_i (c_i is also written as $1/\Delta t_i$ in the literature) and σ_i , the white-noise rms of each pulsar.

To implement these noise PSDs appropriately, one needs to carefully account for the power removed by fitting the timing model. This is represented in the diagonal inverse-noise-weighted transmission function,

$$\mathcal{N}_i^{-1}(f) = \mathcal{T}_i(f)/P_i(f), \quad (45)$$

where $\mathcal{T}_i(f)$ is the pulsar’s transmission function, and $P_i(f)$ is the white noise PSD in equation (44) (Hazboun *et al* 2019). $\mathcal{T}_i(f)$ characterizes the diminished sensitivity to frequencies removed by fitting a timing model, and is treated in detail in Hazboun *et al* (2019), but not replicated here for brevity. For the inclusion of red and white noise in the pulsar (i.e. using (42)), equation (45) becomes approximate and the noise covariance matrix becomes non-diagonal (Hazboun *et al* 2019).

3.2.1. The response function. The pulsar timing residual response function is given by

$$R^{+,\times}(f, \hat{\Omega}, \psi) = \frac{1}{i2\pi f} \left(1 - e^{-i2\pi f D (1 + \hat{\Omega} \cdot \hat{p})/c} \right) \sum_{A=+,\times} F^A(\hat{\Omega}, \psi), \quad (46)$$

where

$$F^A(\hat{\Omega}, \psi) \equiv \epsilon_{ij}^A(\hat{\Omega}, \psi) \frac{\hat{p}^i \hat{p}^j}{2(1 + \hat{\Omega} \cdot \hat{p})}, \quad (47)$$

\hat{p} is the unit vector from the Solar System barycenter to the pulsar, D is the distance to the pulsar, $\epsilon_{ij}^A(\hat{\Omega}, \psi)$ are given in equation (7) and $i, j = x, y, z$ are the spatial indices (Anholm *et al* 2009). Note the factor of $\left(1 - \exp(-i2\pi f D (1 + \hat{\Omega} \cdot \hat{p})/c) \right)$, which comes from the redshifting of signals from a pulsar and is the metric perturbation at the Earth (first term) and the pulsar (second term). We note that the pulsar term is averaged over in Hazboun *et al* (2019) and produces a factor of two increase in sensitivity over simply taking the Earth term only. The factor of $(1/i2\pi f)$ comes from transforming the redshift in pulse arrival to the residuals which are typically used in pulsar timing (Hazboun *et al* 2019). For a more in-depth discussion, see section 2 in Anholm *et al* (2009).

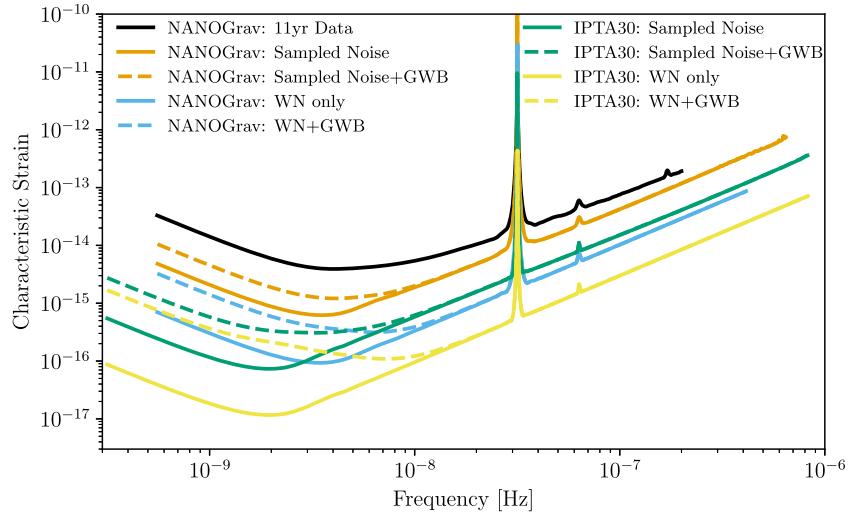


Figure 5. Sensitivities of various PTA designs with noise parameters noted in table 1. Dashed curves use the same values as the solid curves, but with an added GWB of 4×10^{-16} . The NANOGrav 11 yr data sensitivity curve uses the real TOAs and fitting parameters given in Arzoumanian *et al* (2018). The WN only models have identical fixed WN levels for the number of pulsars in the array observed over the total respective timespan. The difference between the IPTA30 and NANOGrav WN only models is that the IPTA30 model includes an increase in the number of pulsars, cadence, and total observation time. The sampled noise cases either assign each pulsar with corresponding values in Arzoumanian *et al* (2018) for the NANOGrav cases, or draw values from uniform distributions given in table 1 for the IPTA30 cases. The specific makeup of the sampled noise cases are more rigorously defined in section 3.3 and in table 1. It is clear that the simplified frameworks employed here tend to underestimate the noise by not including some noise sources, leading to an increase in the sensitivity of the PTA.

The sky and polarization averaged response function in equation (17) for PTAs is then given by

$$\mathcal{R}(f) = \frac{1}{12\pi^2 f^2}, \quad (48)$$

where the detector response functions in equation (17) come from equation (46) to account for the Earth and pulsar terms.

To assemble a full PSD one sums the individual pulsar strain-noise PSD,

$$S_I(f) \equiv \frac{1}{\mathcal{N}_I^{-1}(f)\mathcal{R}(f)}, \quad (49)$$

scaled by the individual pulsar's fractional observation time to get the effective noise PSD,

$$S_{\text{eff}}(f) \equiv \left(\frac{4}{5} \sum_I \frac{T_I}{T_{\text{obs}}} \frac{1}{S_I(f)} \right)^{-1}. \quad (50)$$

The factor of 4/5 again comes from averaging over inclination. We assume that $S_n(f) = S_{\text{eff}}(f)$ when calculating noise curves and SNRs (Hazboun *et al* 2019).

Table 1. PTA component noise values used throughout this work. Brackets indicate the ranges for the relevant noise values. Ranges for the sampled NANOGrav parameters and IPTA30 red noise parameters correspond to the minimum and maximum noise values in the NANOGrav 11 yr dataset (Arzoumanian *et al* 2018). The IPTA30 values correspond to an hypothetical IPTA in the 2030s similar to those in Chen *et al* (2019) and Taylor *et al* (2016). σ_{RMS} is the white-noise rms of each pulsar, the cadence is the number of observations per pulsar per month, T_{obs} is the total observation time. A_{GWB} is the amplitude of the stochastic GW background at $f = 1 \text{ yr}^{-1}$, A_{RN} is the amplitude of the pulsar red noise at $f = 1 \text{ yr}^{-1}$, and γ_{RN} is the red noise power spectral index. We note that we use the overall RMS from Arzoumanian *et al* (2018) as the bounds on the sampled NANOGrav σ_{RMS} . Since the overall RMS may include RN, this may inflate the WN levels, but since we do not include other sources of WN present in real TOAs (e.g. EQUAD, EFAC, and ECORR), we expect this choice to produce little effect on the resulting sensitivity levels.

Model	σ_{RMS} (ns)	cadence (month $^{-1}$)	T_{obs} (yr)	A_{GWB} $N_p \times 10^{-16}$	$\log_{10}(A_{\text{RN}})$	γ_{RN}
NANOGrav: WN only	100	2	11.4	34	—	—
NANOGrav: WN + GWB	100	2	11.4	34	4	—
NANOGrav: sampled noise	[108, 3650]	[1, 4]	[4, 11.4]	34	—	[-19.9, -12.2] [-1.85, 6.24]
NANOGrav: sampled noise + GWB	[108, 3650]	[1, 4]	[4, 11.4]	34	4	[-19.9, -12.2] [-1.85, 6.24]
IPTA30: WN only	100	4	20	200	—	—
IPTA30: WN + GWB	100	4	20	200	4	—
IPTA30: sampled noise	[100, 1000]	[1, 4]	20	200	—	[-19.9, -12.2] [-1.85, 6.24]
IPTA30: sampled noise + GWB	[100, 1000]	[1, 4]	20	200	4	[-19.9, -12.2] [-1.85, 6.24]

3.3. Sampled PTAs

Using the robust methods presented in Hazboun *et al* (2019) in this formalism allows us to compare future detectors to current ones using improved methods to previous PTA SNR calculations. Additionally, using the framework of Hazboun *et al* (2019) we are able to construct sensitivities for PTAs with more realistic noise parameters, and to use real data directly. In figure 5, we show the characteristic strain for five different NANOGrav models and four estimated International PTA (IPTA) models for an IPTA in the 2030s similar to those in Chen *et al* (2019) and Taylor *et al* (2016).

All of the generated models improve upon the sensitivity of the NANOGrav 11 yr data curve in figure 5. The NANOGrav 11 yr data sensitivity curve uses the real TOAs and fitting parameters given in Arzoumanian *et al* (2018) to give a comparison to simulated plots to real data in hasasia.

The WN only models in figure 5 have identical fixed WN levels for the number of pulsars in the array (see table 1 for exact values). The difference between the IPTA30 and NANOGrav WN only models is that the IPTA30 model includes an increase in the number of pulsars, an increase in cadence, and a total observation time of 20 years instead of the NANOGrav model's 11 years. This is reflected by the IPTA30 WN only model's almost order of magnitude lower sensitivity curve in figure 5 when compared to the NANOGrav WN only model. When we take the same two WN only models and add a GWB, the sensitivity of both models is reduced at frequencies below f_{year} . It is expected that adding a GWB as the only source of red noise in all pulsars would cause the low frequency sensitivity to decrease and approach similar values

for PTAs with different WN levels. To fairly and accurately characterize PTA sensitivities, we should be considering the noise properties of the entire array of pulsars instead of the traditional simplistic, WN only models.

To simulate more realistic PTAs, labeled as ‘sampled noise’ in figure 5, we use current WN and RN noise parameters along with pulsar sky locations, observation times, and cadences from Arzoumanian *et al* (2018). Each pulsar that makes up the sampled NANOGrav curves is assigned an RMS and a RN. Thus, for these PTAs, the overall RMS from table 2 in Arzoumanian *et al* (2018) is used as the WN RMS, and the RN levels are set to the maximum likelihood values found in the analysis of Arzoumanian *et al* (2018). For the NANOGrav curves, we use only the 34 pulsars used in Arzoumanian *et al* (2018) with their corresponding locations, and observation and noise parameters. For the IPTA30 sampled noise models, we use the individual NANOGrav 11 yr parameters to create a distribution of values from which to draw each pulsar’s location and RN parameters. The WN parameters for the sampled IPTA30 models use a uniform distribution between the WN-only model’s σ_{RMS} and ten times that value. The per pulsar cadence is likewise drawn from a uniform distribution of observing between once and four times a month. The distribution ranges for all sampled models are given in table 1. Drawing from real pulsar noise data gives us a reasonable representation of sensitivities that a future PTA can obtain without needing to simulate a full dataset.

As can be seen in figure 5, the result of this sampled PTA generation still produces a discrepancy of about a factor of three relative to noise levels present in real TOAs/residuals. To fully simulate a PTA’s sensitivity, one would need to account for additional noise sources, the origins of which are not yet entirely clear (Lam *et al* 2017). In order to utilize gwent’s unique ability to produce SNRs to BHB sources for variations in instrument parameters, we cannot realistically simulate full PTA datasets that attempt to include the unknown sources of noise.

3.4. Methods of sensitivity representation

As mentioned earlier, one must be careful when comparing various PTA sensitivity curves. We again note that one cannot construct a PTA sensitivity curve agnostic of the source. For this reason, one should view PTA curves for deterministic sources as different from curves for stochastic sources. For example, the wedge-shaped sensitivity curves in Thrane and Romano (2013), Hobbs (2011), and Moore *et al* (2015) are often used to represent PTA sensitivities. Since these curves represent PTA’s sensitivities to the GWB, one should not use them in the context of individually resolvable SMBHBs. With that in mind, we briefly examine three typical variations of PTA sensitivities to continuous waves (CWs, i.e. monochromatic GW sources) from SMBHBs and compare them to our methods. We display an example of each of the discussed models in figure 6.

In figure 6, we also present three curves used in this study to highlight the differences between the various methods of assessing PTA sensitivities to CWs. The NANOGrav 11 year data curve and the NANOGrav sampled noise curve use the same parameters in figure 5. For the NANOGrav sampled noise curve that includes a GWB, we use a background amplitude of $A_{\text{GWB}} = 2.0 \times 10^{-15}$. This singular change is more in line with current NANOGrav upper-limits on the GWB, and reduces the difference in sensitivity between the simulated curves and real data to within a factor of around 1.5. Because characteristic strain curves, as seen in figure 5, do not directly relate to the source strain amplitude, we plot the strain of each PTA to a source that produces an SNR of 3 via solving equation (39) for h_0 . This allows us to directly compare to other methods in the literature typically used for assessing detectability of CW sources.

Aggarwal *et al* (2018) used Bayesian inference in addition to frequentist techniques to search for GWs from individual SMBHBs in circular orbits in their 11 yr dataset. For the

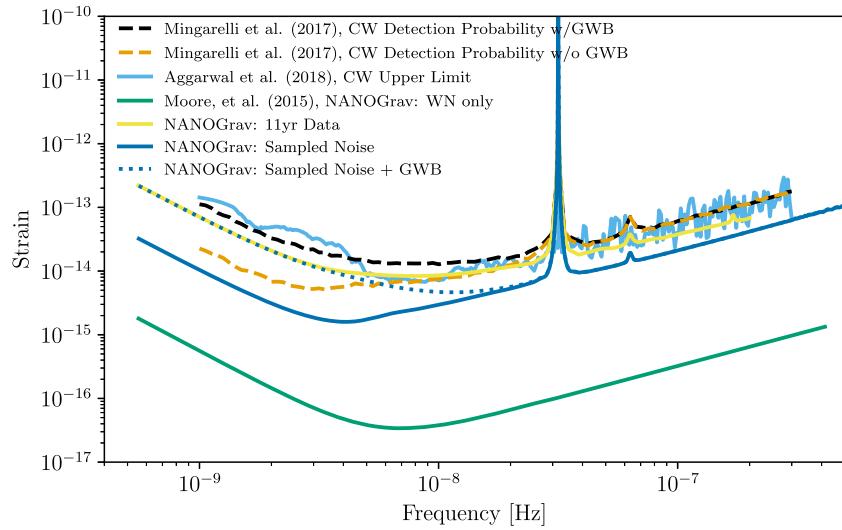


Figure 6. Sensitivities of various PTA designs discussed in section 3.4. The NANOGRAV 11 yr data curve and the NANOGRAV sampled noise curve use the same parameters as in figure 5. For the NANOGRAV sampled noise curve that includes a GWB, we use a background amplitude of $A_{\text{GWB}} = 2.0 \times 10^{-15}$ in order to better match the real data sensitivity. In order to present the curves produced by our formalism with those mentioned in section 3.4, we plot the strain of each PTA to a source that produces an SNR of 3 via solving equation (39) for h_0 . The Mingarelli *et al* (2017) curves use the 95% detection probability with a false alarm probability of 10^{-4} , a 15 yr observation time, and for the GWB case, an unsubtracted GWB amplitude of 4×10^{-16} . The WN only, scaled detection sensitivity curve derived from Moore *et al* (2015) has a similar spectral shape to the other PTAs in this figure, but has a distinctly more sensitive detection strain for the same choice of parameters.

Bayesian GW analyses, the authors used fixed individual pulsar white noise parameters based on previous noise analyses, and sampled the red noise parameter space. To extract the common red noise (i.e. the CW) from the intrinsic pulsar red noise, the authors searched over both simultaneously. For the frequentist analysis, the authors used the \mathcal{F}_p -statistic developed in Ellis *et al* (2012), which maximizes the log of the likelihood ratio of all pulsar-dependent contributions. The \mathcal{F}_p -statistic is the easiest to relate to the analysis techniques presented here because it is directly related to the optimal SNR

$$\langle 2\mathcal{F}_p \rangle = 2M + (\tilde{\mathbf{s}}|\tilde{\mathbf{s}}) = 2M + \rho^2, \quad (51)$$

where $\tilde{\mathbf{s}}$ are the induced pulsar timing residuals, which can be related to the source signal, ρ is the optimal SNR, and $2M$ are the pulsar-dependent amplitude parameters; see Ellis *et al* (2012) for more detailed definitions. To use the \mathcal{F}_p -statistic to produce sensitivity curves similar to the work here, one could calculate the \mathcal{F}_p -statistic corresponding to a minimum detectable SNR in each frequency bin.

To get the SNR to compare directly to the formalism used here using the \mathcal{F}_p -statistic, one would need to either simulate, or use real-pulsar data to populate, a PTA including pulsars with particular noise data over some observation time-length. Then, one would need to inject a signal corresponding to a particular source into the PTA, including its response based on its sky location. Once the signal is injected, one would then compute the \mathcal{F}_p -statistic and recover the SNR from equation (51).

The issue with this method is that one needs to assign a sky-location to the source. This is because the response in the detector is heavily dependent on the arrangement of pulsars in the PTA. A region of the sky with fewer pulsars leads to insensitivities to sources in that particular region (Aggarwal *et al* 2018). This presents a problem as well as an opportunity, because PTA's sky sensitivities evolve as more pulsars are added, especially in underrepresented regions.

Since neither their Bayesian nor their frequentist searches showed significant evidence of a CW, Aggarwal *et al* (2018) computed the sky-averaged 95% upper-limit of the GW strain. Because these upper-limit curves are therefore not directly related to the sensitivity calculations as we have described it, one cannot use upper-limit curves for the SNR calculations we wish to perform. This also means that the use of EPTA upper-limit curves in Rosado *et al* (2016) to produce horizon distance estimates is inconsistent, although one would need to recalculate those results using a realization of the EPTA effective noise PSD to determine how large a difference this would make.

In Mingarelli *et al* (2017), the authors use the Two Micron All Sky Survey (2MASS) (Skrutskie *et al* 2006) in tandem with simulated galaxy merger rates from the Illustris cosmological simulation project (Rodriguez-Gomez *et al* 2015; Genel *et al* 2014) to populate the local Universe (all GW sources are placed at distances < 225 Mpc). After simulating multiple realizations of the Universe, the authors then use the \mathcal{F}_p -statistic to compute detection probability curves at varying false alarm probabilities (FAP), or the probability that a measured \mathcal{F}_p exceeds an \mathcal{F}_p with no signal present (Ellis *et al* 2012). The authors then produce sky-averaged strain sensitivity curves and strain sensitivity sky-maps at the particular locations of the 2MASS galaxies.

Because the curves in Mingarelli *et al* (2017) correspond to particular FAPs, they are detection threshold sensitivity curves. Each curve corresponds to the particular detection confidence threshold corresponding to each FAP. Therefore, one should not use the results of Mingarelli *et al* (2017) directly as a sensitivity curve within our formalism. One could, however, use their results as another comparison of the detection significance and strength of the corresponding signal.

In Moore *et al* (2015), the authors' use both frequentist and Bayesian methods to generate PTA sensitivity curves to both a monochromatic GW source (a CW) and a power-law stochastic background (the GWB). The author's produce scaling laws for the sensitivity to a CW. The frequentist method uses a similar formalism to that of Hazboun *et al* (2019), but assumes only white, Gaussian, and uncorrelated noise between all pulsars in the array. This allows for a great simplification of the PTA's response to a CW. Likewise, they only consider an optimally oriented source ($\iota = \psi = 0$) and average over the sky position angles to get the scaling for the SNR of a monochromatic source. By choosing a detection threshold SNR, one can then invert their result to find the PTA's sensitivity to a CW. This is useful for scaling relations for the sensitivity of a simplified PTA to a CW. Since our model is intended to extend to chirping sources, and we wish to more accurately represent PTAs, we do not use their relations.

3.5. Sensitivity changes to various designs

In figure 7, we show the SNR versus total source mass for variations of a non-spinning, equal-mass GW source at a redshift of $z = 0.1$ (i.e. two of those three parameters are held fixed while the third is varied). We simultaneously present the same changes for three NANOGrav curves in figure 5. Not surprisingly, the full range of sensitive spins, mass ratios, redshifts, and masses decrease as we continue to add noise sources.

In other words, the further away one gets from the WN only model (i.e. the optimal PTA scenario of identical pulsars with only stochastic noise), the more the ranges of detectable source distance and total mass decrease.

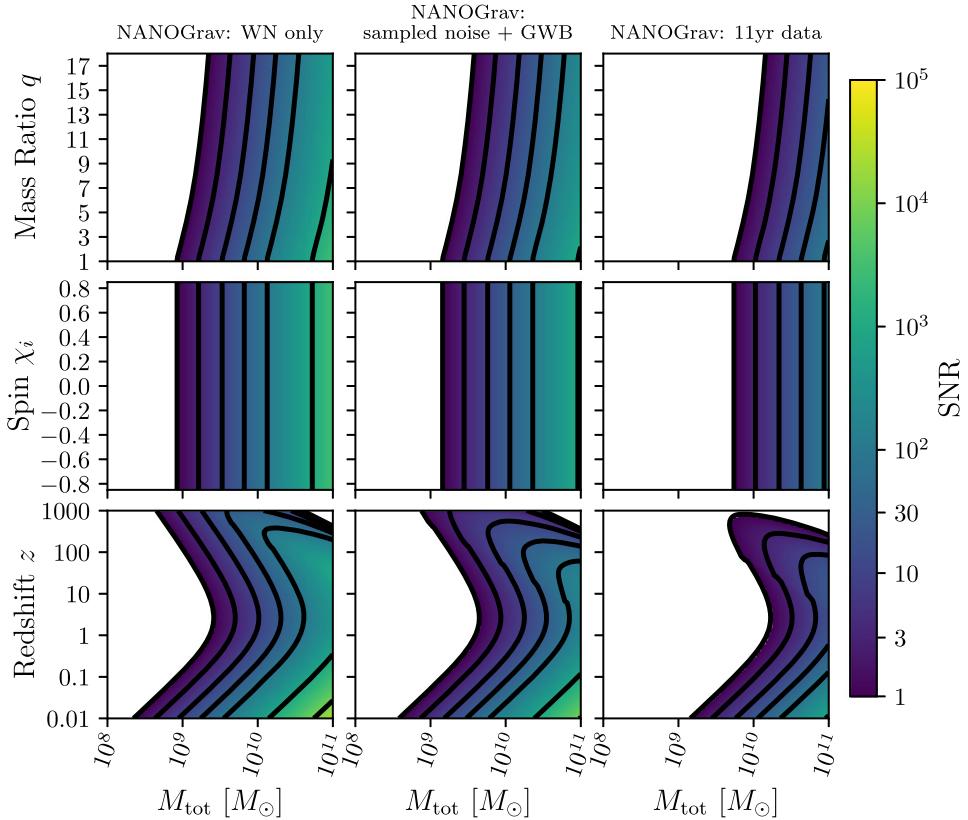


Figure 7. SNRs for changes to a fiducial, non-spinning, equal-mass BHB at a redshift of $z = 0.1$ ($\chi_i = 0$, $q = 1$) with respect to PTAs: NANOGrav: WN only (left column), NANOGrav: sampled noise + GWB (middle column), and NANOGrav: 11 yr data (right column). Solid black contours correspond to the levels marked on the colorbar. Top row: mass ratio vs total source mass, middle row: spins of both BHs $\chi_i = \chi_1 = \chi_2$ vs total source mass, bottom row: redshift vs total source mass.

For higher masses, there can be some change in the waveform close to merger with information about the spins of the individual binaries. However, based on figure 2, PTAs are not expected to be sensitive to spin effects on the waveform amplitude. Even for the most massive and nearby sources that might be considered astrophysically relevant, the difference between a spinning and a non-spinning system is minimal over the frequencies which a PTA could observe the BHB.

As one can see, the effect of varying mass ratio has the same level of impact on the SNR for any time in the BHB's evolution. In all three detectors, there is a significantly larger total mass parameter space observable for equal mass binaries than for binaries with large mass ratios.

We also use this formalism's unique ability to vary instrument parameters, to produce SNRs that show examples of how particular noise parameters affect PTA sensitivities to specific GW sources. In figure 8 we take the fiducial values for the NANOGrav WN only and NANOGrav sampled noise + GWB PTAs (see table 1), then vary the observation time, cadence, WN RMS, and the number of pulsars, respectively. We then calculate the SNR for variations to these parameters versus the total source mass for a non-spinning, equal-mass SMBHB at a redshift of $z = 0.1$ on the other axis. This allows us to explore the change in SNR over the full variable

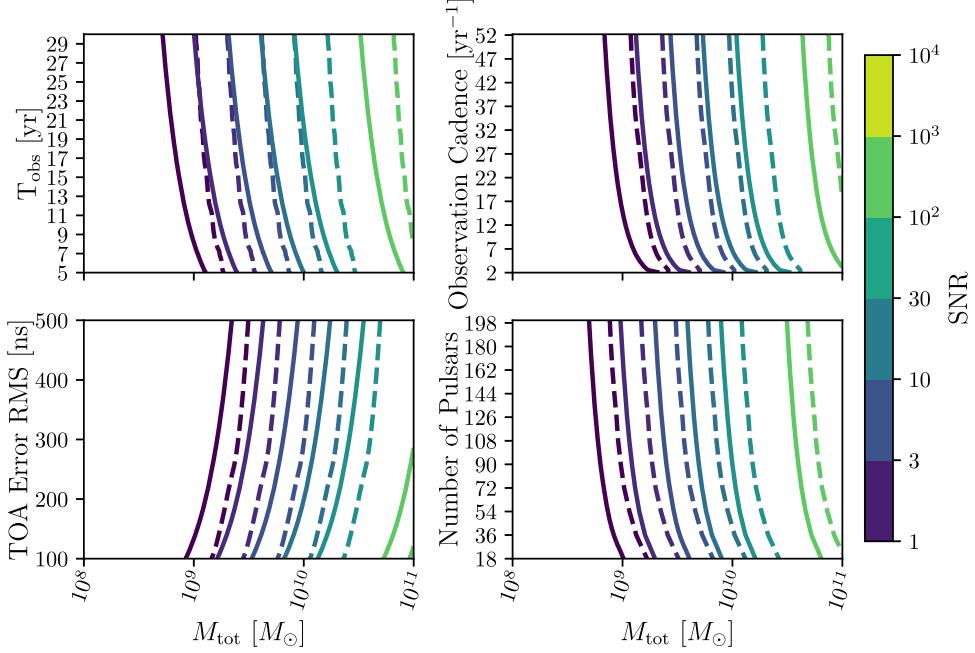


Figure 8. SNRs of non-spinning, equal mass BHB sources at redshift $z = 0.1$ observed by variations on the fiducial NANOGrav: WN Only (solid lines) and NANOGrav: sampled noise + GWB (dashed lines) as a function of total mass. Each parameter is changed in isolation while the other PTA parameters are fixed to their fiducial values (see table 1). Upper left: observing time, upper right: observing cadence, lower left: pulsar timing noise RMS, lower right: number of pulsars.

range for each of the above parameters. We find that the SNRs behave as one would expect: as observation time, cadence, and the number of pulsars increases, and as the timing RMS decreases, the PTA gains sensitivity overall, and consequently a larger range of total mass becomes detectable.

We find that increasing the cadence extends PTA sensitivity to higher frequencies and, more importantly in this context, lowers the overall WN. As the PTA's higher frequency coverage increases, and the optimal frequency does not change. The SNR varies with cadence, c , as \sqrt{c} . The WN is also lowered by decreasing the timing error RMS σ_{RMS} , which is represented by the SNR scaling as $1/\sigma_{\text{RMS}}$. When adding multiple pulsars with similar T_{obs} , σ_{RMS} , and c to a WN only PTA, the SNR scales with the number of pulsars N_p as $\sqrt{N_p}$.

Since the SMBHBs with masses smaller than $\sim 10^{10} M_\odot$ tend to be monochromatic sources in the PTA band, the SNR only builds as $\sqrt{T_{\text{obs}}}$ for those sources. Increasing T_{obs} extends sensitivity to lower frequencies, and also shifts the optimal frequency of the detector to lower frequencies. The shift in optimal frequency adds an additional factor to the typical scaling of $\sqrt{T_{\text{obs}}}$ with SNR for a WN only PTA.

It is clear that, in the case of WN only pulsars, when increasing the observation time, cadence, or number of pulsars, there is a point of diminishing returns because of their effect on SNR. The only parameter that does not follow this trend is the WN timing error. In the case of using sampled noise, the SNRs follow similar trends as the WN only model, but with around an order of magnitude less sensitivity overall. This is to be expected based on a

similar factor in sensitivity decrease between the WN only model and sampled noise models for a NANOGrav-esque PTA seen in figure 5.

The biggest difference in the shape of the sampled noise curves is their smoothness when varying some parameters. When changing parameters in the sampled noise case, the SNR appears to have a non-smooth increase when comparing to the WN only case. We attribute this to the fact that when varying each parameter, the entire array is assigned the new sampled value. We find that simply changing one factor only has a moderate effect on the overall SNR, but we emphasize that this conclusion is highly dependent on our underlying assumption of a homogeneous array. In reality, when increasing the number of pulsars, the SNR of a source in the direction of a newly added pulsar could drastically increase. This is because the non-sky averaged sensitivity is highly dependent on the relative locations of the source and pulsars (Aggarwal *et al* 2018).

Another interesting finding using the capabilities of this formalism is that the SNR dips below the threshold for detection, then comes back into the detectable range at higher redshifts for certain sources. This effect, seen in figure 7, was also noted in Rosado *et al* (2016). Because of the increase in SNR at high redshifts, PTAs could potentially detect sources in the very early Universe, were they to exist. Typically these high mass and high redshift sources have some orbital evolution during the observation time (i.e. they do not obey equation (27)). We emphasize that much of the domain of these plots covers regions of parameter space (e.g. $10^{11} M_{\odot}$ BHs at $z = 1000$) that should not exist within our current understanding of cosmological structure formation; these plots are strictly intended to illustrate the potentiality of a PTA measurement, rather than our expectation based on current theory.

4. Space based detector sensitivity

In this section we discuss space-based detectors in the context of LISA, a detector consisting of three satellites in a triangular constellation connected by two-way laser links (Amaro-Seoane *et al* 2017). The LISA frequency band spans from $\sim 10 \mu\text{Hz}$ to $\sim 1 \text{ Hz}$ (see figure 9). To get to the lower frequency domain, the proposed inter-spacecraft separation of LISA will be on the order of 2.5 million km. The LISA three satellite constellation will trail Earth between 50 and 65 million km in a heliocentric orbit and will fully rotate around two axes over the course of a year (one normal to the ecliptic, the other inclined by 60 degrees), allowing its varying sky-dependent response function to provide additional source position information (Amaro-Seoane *et al* 2017; Cornish and Larson 2001). The different arm orientations of the three legs of the constellation, and their variation due to the orbit, both provide information about the parameters of the emitting bodies. LISA's sweeping orbit and the longer time periods that its sources remain in-band will allow it to gather similar information to the suite of planned ground detectors with a single instrument (Amaro-Seoane *et al* 2017). Variations on the nominal LISA design are of interest both for followup mission concepts, as well as contemporaneous concepts like TianQin (Luo *et al* 2016) and Taiji (Ruan *et al* 2020).

The deployment of space-based interferometric GW detectors requires reducing extraneous non-gravitational accelerations, referred to as acceleration noise, as much as possible, to best approximate a measurement with free-falling test masses. The effect of spurious random forces with respect to the spacecraft's inertial frame limit the instrument's sensitivity at low frequencies, whereas photon shot noise and other optical path noises limit the high-frequency sensitivity (Armano *et al* 2016). The photon shot noise is the same phenomenon that limits ground-based detectors at high frequencies, and the acceleration noise as a limitation requires mitigation in order to reach the required sensitivity.

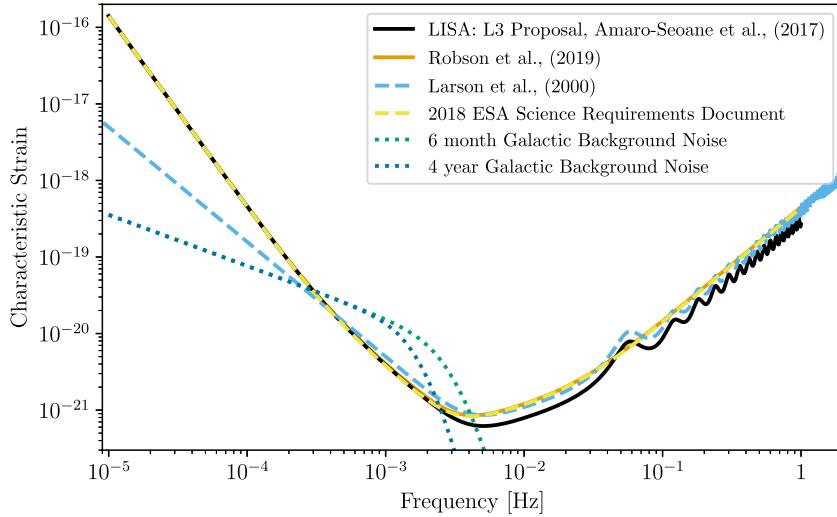


Figure 9. Sensitivities of various proposed LISA designs along with the galactic binary confusion noise. Differences between proposals and models made with `gwent` mostly change with respect to low-frequency spectral indices and high-frequency interferometer noise floors. The models for the 2018 ESA Science Requirements Document and Robson *et al* (2019) use the approximation shown in equation (60).

To filter out non-gravitational forces as much as possible, each LISA spacecraft uses two free-floating test masses as geodesic reference test particles. Each spacecraft is then forced to follow its test mass using micro-Newton thrusters. After applying these corrections to the spacecraft, the distance between test masses along each arm of the interferometer is calculated, with corrections for Doppler shifts and post-processing on Earth (Amaro-Seoane *et al* 2017). The data from the three arms (and six laser links) can be combined in a number of different ways to form sets of three different data channels, although a primary goal of most combinations is to cancel out the effect of the relative orbital motion of the satellites, a technique referred to as time-delay interferometry (TDI). In one such combination, the resulting set of observables is similar to three separate Michelson interferometers, each of which is comparable to a single ground-based detector. For this work, we specifically calculate the SNR of the Michelson X channel of the XYZ TDI combination, but one can straightforwardly modify to calculate other TDI combinations (Tinto and Dhurandhar 2014; Smith and Caldwell 2019), although the network SNR should be the same regardless.

4.1. LISA power spectral density

The noise limitations on LISA in the high frequency regime are determined by laser interferometer noise (IFO), consisting of shot noise and any other optical path noises (e.g. the telescope, optical bench, phase measurement system, laser, clock and TDI processing), and is combined and modeled as

$$P_{\text{IFO}} = (1.0 \times 10^{-12} \text{ m})^2 \left(1 + \left(\frac{2 \text{ mHz}}{f} \right)^4 \right) \text{ Hz}^{-1}, \quad (52)$$

which is the proposed effective total displacement noise in one direction between two of the LISA spacecraft arms from the L3 mission proposal (Amaro-Seoane *et al* 2017). This noise

source can be expressed more generally as

$$P_{\text{IFO}} = (A_{\text{IFO}})^2 \left(1 + \left(\frac{f_{\text{break}}}{f} \right)^4 \right) \text{Hz}^{-1}, \quad (53)$$

where A_{IFO} is the displacement amplitude of a single-link, and f_{break} is the frequency of the turnover for the high frequency noise.

The low frequency regime is dominated by spacecraft acceleration noise (Larson *et al* 2000). This is represented in the L3 proposal by

$$P_{\text{acc}} = (3 \times 10^{-15} \text{ m s}^{-2})^2 \left(1 + \left(\frac{0.4 \text{ mHz}}{f} \right)^2 \right) \left(1 + \left(\frac{f}{8 \text{ mHz}} \right)^4 \right) \text{Hz}^{-1}, \quad (54)$$

which can be expressed more generally as

$$P_{\text{acc}} = (A_{\text{acc}})^2 \left(1 + \left(\frac{f_{\text{break, low}}}{f} \right)^2 \right) \left(1 + \left(\frac{f}{f_{\text{break, high}}} \right)^4 \right) \text{Hz}^{-1}. \quad (55)$$

When looking at LISA's noise curve (i.e. figure 9), the dependencies from equations (52) and (54) determine the entire sensitivity curve (Amaro-Seoane *et al* 2017).

If we assume each detector in the LISA configuration has the same noise spectral density, we can just add additional factors of each noise type. In a LISA-like interferometer there are four contributions to the IFO and sixteen contributions to the acceleration noise. Half of the acceleration noise contributions are scaled by $\cos^2(f/f_*)$, where $f_* = c/(2\pi L)$ is the round-trip light frequency (i.e. the transfer frequency), to account for combining the acceleration noise at two different times (Cornish 2002). To combine the high and low frequency noise into strain, we must convert from displacement to strain by dividing by the square of the round-trip light travel distance, $2L$ (Robson *et al* 2019). Furthermore, to obtain the effective position noise due to spurious accelerations, we must divide the total acceleration noise by the angular frequency of the GW, $2\pi f$, to the fourth power (Cornish 2002). The total noise spectral density is then,

$$P_{\text{n}}(f) = \frac{4P_{\text{IFO}}}{(2L)^2} + 8 \left(1 + \cos^2(f/f_*) \right) \frac{P_{\text{acc}}}{(2\pi f)^4 (2L)^2}. \quad (56)$$

4.2. The transfer function

As mentioned in section 2.1.3, instrument sensitivities depend on the source's properties. Additionally, the response functions of interferometers are dependent on the arm length of the detector in relation to the amplitude of the passing GW. This effect is encapsulated in the transfer function,

$$\mathcal{T}(\mathbf{r} \cdot \hat{\Omega}, f) = \text{sinc} \left[\frac{f}{2f_*} (1 - \mathbf{r} \cdot \hat{\Omega}) \right] e^{i \frac{f}{2f_*} (1 - \mathbf{r} \cdot \hat{\Omega})}. \quad (57)$$

where \mathbf{r} is a unit vector pointing from the first to the second spacecraft (Cornish and Larson 2001). For $f \ll f_*$, the transfer function approaches unity, while at frequencies higher than f_* , the sinc function causes periodic dips at $f = n/2L$ and is proportional to $1/f$ for $f \gg f_*$ (Larson *et al* 2000). This can be seen in figure 9 for the models with numerically calculated transfer functions.

The response function in equation (17) is made up of the instrument antenna functions

$$F^A(\hat{\Omega}, f) = \mathbf{D}(\hat{\Omega}, f) : \mathbf{e}^A(\hat{\Omega}), \quad (58)$$

where the colon denotes the double contraction $\mathbf{a} : \mathbf{b} = a_{ij}b^{ij}$. The antenna pattern is made up of the detector tensor

$$\mathbf{D}(\hat{\Omega}, f) = \frac{1}{2}(\mathbf{r} \otimes \mathbf{r}) \mathcal{T}(\mathbf{r} \cdot \hat{\Omega}, f), \quad (59)$$

and the basis tensors of the GW polarizations $\mathbf{e}^A(\hat{\Omega})$ are given by equation (9).

Because of the complexity of the functions in $\mathcal{R}(f)$ from equation (17), one needs to numerically integrate to get the full LISA response function. We use the method described in Larson *et al* (2000) to produce the LISA response curves in this paper and provide the result in a text file inside of `gwent`.

Alternatively, one can use an analytic fit for the response function

$$\mathcal{R}(f) = \frac{2 \sin^2 \beta}{5} \frac{1}{\left(1 + 0.6(f/f_*)^2\right)}, \quad (60)$$

found in Robson *et al* (2019), where the $2 \sin^2 \beta/5$ prefactor ensures that $\mathcal{R}(f)$ asymptotes to the appropriate low frequency limit, where β is the opening angle between two detectors (Cornish and Larson 2001). The normalization in equation (60) is a factor of 2 larger than what is given in some of the literature, since it accounts for a sum over the two independent low-frequency data channels (Robson *et al* 2019).

4.3. LISA sensitivity curve

We find LISA's effective noise PSD,

$$S_n(f) = \frac{P_n(f)}{\mathcal{R}(f)} \quad (61)$$

$$= \frac{10}{3\mathcal{R}(f)} \left(\frac{P_{\text{IFO}}}{L^2} + 2(1 + \cos^2(f/f_*)) \frac{P_{\text{acc}}}{(2\pi f)^4 L^2} \right) \quad (62)$$

by using LISA's PSD from equation (56) and its response function from either the stored result of the numerical integration in (Larson *et al* 2000) or from equation (60).

4.3.1. Galactic binary background. Current population modeling of interacting white dwarf binaries estimate around 26 million binaries emit GWs in the mHz band (Cornish and Robson 2017). Most of these binaries are unresolved, so that for the purpose of observing BHBs, they constitute an additional noise source and their power should be added to the instrument noise power, in direct analogy to the role of the GWB when trying to observe CWs with PTAs. To represent this confusion noise, we use a modified version of the smooth fit to the co-added signals from the simulated Galactic binary population in Cornish and Robson (2017). We determined that most of the fit parameters employed in Cornish and Robson (2017) were poorly constrained, even with the correction to the fit in Schmitz (2020). Since the parameters had only a very weak effect on the signal; this led to the highly non-monotonic behavior of three of the fit parameters when varying the observation time. Therefore, we digitized the curves in figure 2 of Cornish and Robson (2017), and fit each of them to the simpler function

$$S_c(f) = A f^{-7/3} [1 + \tanh(\gamma(f_k - f))] \text{Hz}^{-1}. \quad (63)$$

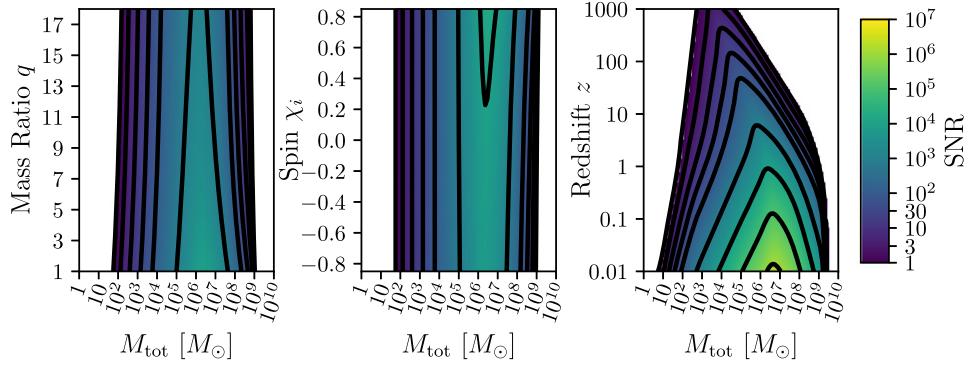


Figure 10. L3 ESA LISA model SNRs for changes to a fiducial, non-spinning, equal-mass BHB at a redshift of one ($\chi_i = 0, z = 1, q = 1$). Left: mass ratio vs total source mass, middle: spins of both BHs $\chi_i = \chi_1 = \chi_2$ vs total source mass, right: redshift vs total source mass. Solid black contours correspond to the levels marked on the colorbar.

We then fit A , γ , and f_k as power laws of T_{obs} , finding $A = 1.4 \times 10^{-44}$, $\gamma = 1100 (T_{\text{obs}}/1 \text{ yr})^{3/10}$, and $f_k = 0.0016 (T_{\text{obs}}/1 \text{ yr})^{-2/9}$ Hz. We note that this fitting function is physically motivated in the asymptotic limits; it is simply the $f^{-7/3}$ spectrum expected for a fixed population of binaries evolving slowly due to gravitational radiation at low frequencies, and is exponentially suppressed as we move to frequencies $f \gg f_k$ since an increasing fraction of binaries in each frequency bin becomes individually resolvable Barack and Cutler (2004). The tanh function provides a simple transition between these two limits.

We show the resulting characteristic strain in figure 9. Summing the PSD of the Galactic WD binaries and that of the instrumental noise allows us to represent the total sensitivity. We note that in the LISA context, the response function is negligible at the low frequencies where Galactic binaries dominate, but for designs with longer arm lengths, the effects may occur in the same band, so it is important to apply the response function to both the galactic foreground and the instrument noise consistently (i.e. $S_{\text{n}}(f) = (P_{\text{n}}(f) + S_{\text{c}}(f)) / \mathcal{R}(f)$).

4.4. Sensitivity changes to various designs

Because of the versatility of our sensitivity model defined in equation (56), using `gwent`'s SNR calculation capabilities, we can explore the wide parameter space of each instrument variable. The differences among the models in figure 9 are due to variations in the low-frequency spectral index, the interferometer noise floor, and exact versus approximate transfer functions.

In figure 10 we show the sensitivity to a fiducial non-spinning, equal mass example BHB at a redshift of one ($\chi_i = 0, z = 1, q = 1$) in the LISA band, to changes in those source parameters versus total mass. We find that for mass ratio, the L3-proposed LISA configuration covers a consistent parameter space in mass no matter the ratio. There is an increase in signal strength for more equal mass BHBs. Similarly, the detectability of this example BHB source spin maintains the same total detectable space in source mass, but has a preference in strength to highly aligned spin. We note that the largest SNR with respect to the marginal effect of changing the spin of each black-hole in the binary happens when the merger occurs over the most sensitive frequencies. For LISA's response to BHBs at different redshifts, we note a similar trend; the region of highest SNR is dominated by sources that merge over the detector's most sensitive

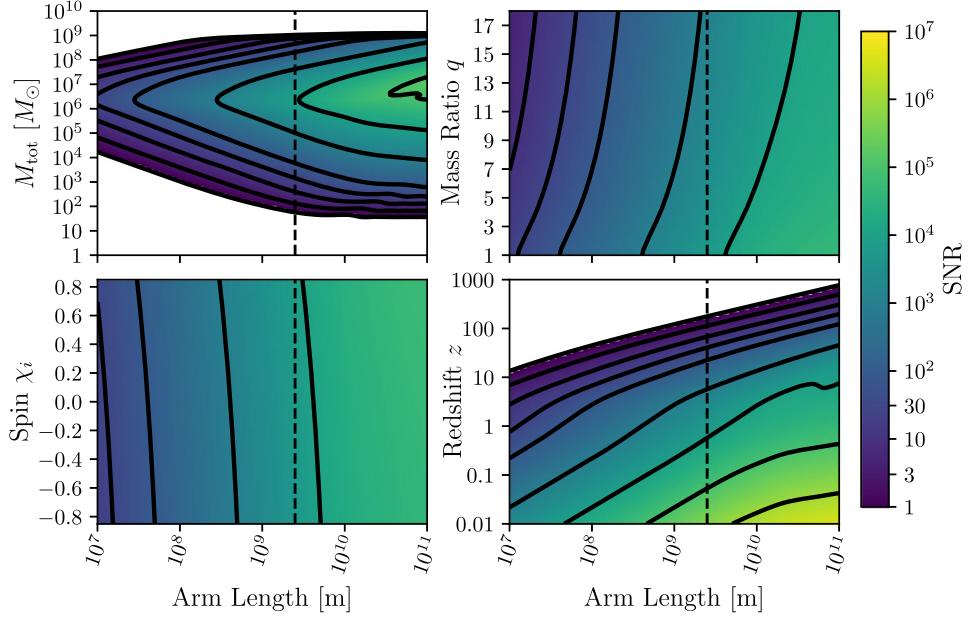


Figure 11. SNRs for LISA arm lengths compared to changes to a fiducial BHB ($M_{\text{tot}} = 10^6 M_{\odot}$, $\chi_i = 0$, $z = 1$, $q = 1$). Dashed lines represent the proposed value for the arm length. Solid black contours correspond to the levels marked on the colorbar. Upper left: total source mass vs arm length, upper right: mass ratio vs arm length lower left: spins of both BHs $\chi_i = \chi_1 = \chi_2$ vs arm length lower right: redshift vs arm length.

frequencies. Much like PTAs, LISA has the potential to see particular BHs coalescing at the earliest times in the Universe, in the unlikely case that they exist.

In figure 11, we look at the same fiducial source BHB as figure 10 ($M_{\text{tot}} = 10^6 M_{\odot}$, $\chi_i = 0$, $z = 1$, $q = 1$) as a function of both the detector's arm length and various source parameters. It is evident that the arm length greatly impacts the sensitivity of the detector. In general, a longer arm length increases the sensitivity across the entire source parameter space.

Increasing the arm length can expand the detectable total BHB mass space to both higher and lower masses until around the arm length planned for LISA. Above this arm length, there appears to be a point of diminishing returns in terms of the observable mass range, although the sensitivity to a given mass continues to increase. As expected, the SNR increases with increasing spin and decreasing (i.e. closer to unity in our convention) mass ratio, with mass ratio being a stronger factor. As the arm length increases, LISA steadily gains sensitivity out to larger redshifts. While this analysis would suggest that a larger constellation is always better, we note that this ignores the increased cost of reaching those orbits, and the potential lack of sufficiently stable orbits until one reaches AU-scale constellations (see e.g. Folkner (2011)).

In addition, sensitivity does not always correlate with parameter measurement accuracy; the larger orbits would modulate a signal more slowly, but would also have a nontrivial frequency dependent response over more of their sensitive band, and these features would compete with one another in terms of affecting measurement accuracy. We therefore emphasize that our calculations can only address the question of observability of given sources, assuming such design parameters can be realized in practice, and does not necessarily reflect what design would yield the most accurate measurements.

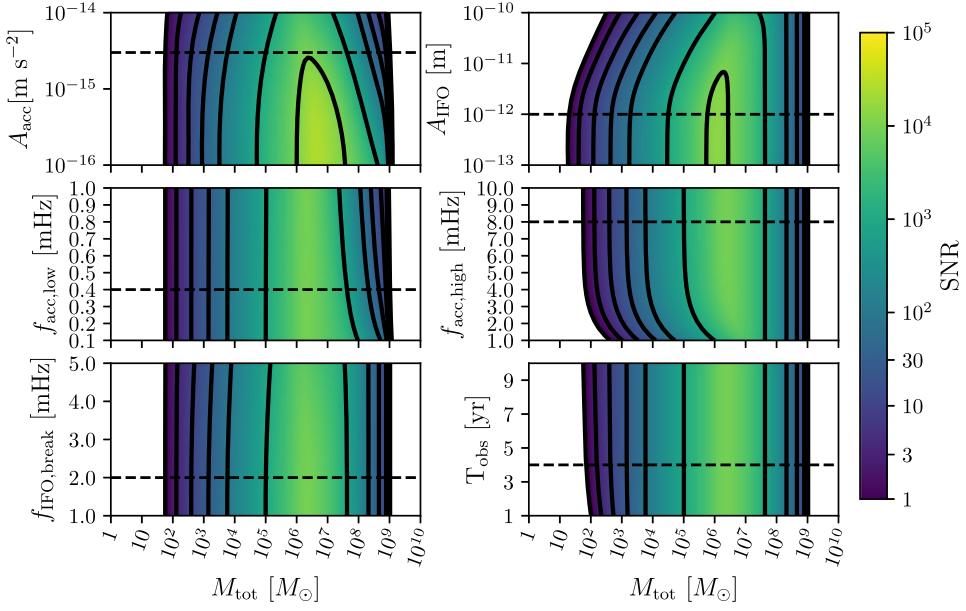


Figure 12. SNRs for a fiducial, non-spinning, equal-mass BHB at a redshift of one ($\chi_i = 0, z = 1, q = 1$) compared to changes to the noise parameters for the L3 ESA LISA model. Top left: acceleration noise amplitude, A_{acc} vs total source mass. Dashed lines represent the fiducial values for each parameter. Solid black contours correspond to the levels marked on the colorbar. Top right: optical metrology noise amplitude, A_{IFO} vs total source mass. Middle left: low frequency acceleration noise break frequency, $f_{\text{acc},\text{low}}$ vs total source mass. Middle right: high frequency acceleration noise break frequency, $f_{\text{acc},\text{high}}$ vs total source mass. Bottom left: optical metrology noise break frequency, $f_{\text{IFO},\text{break}}$ vs total source mass. Bottom right: observation time length T_{obs} vs total source mass.

In figure 12, we look at the dependence of the sensitivity to a non-spinning, equal-mass BHB at $z = 1$, to variations of the other modeled LISA parameters (using the noise values for the L3 ESA LISA proposal given in equations (52) and (54)). In most cases, the detectable total mass parameter space stays constant. In fact, reasonably reducing the noise levels often only marginally increases the SNR. This kind of examination shows that the most impactful noise parameter for LISA is its arm length, as seen in figure 11.

5. Ground based detector sensitivity

In general, the current generation of ground-based GW detectors have a frequency range of ~ 10 Hz to a few kHz. This high frequency band primarily contains compact binaries (some pairing of stellar-mass black-holes and neutron stars) and supernovae sources. Being attached to the Earth, ground-based detectors have significant sources of environmental noise. The lower frequency limit has a steep cutoff due to seismic noise (i.e. vibrations due to the ground motion that cannot be eliminated by the suspension system); design choices like improved seismic isolation and longer arm lengths can somewhat mitigate this noise, but it is a significant challenge to push below the limit due to the Newtonian background (the time-varying mass distribution near the test mass) of a few Hz for terrestrial detectors (Pepper *et al* 2017; Saulson 1994).

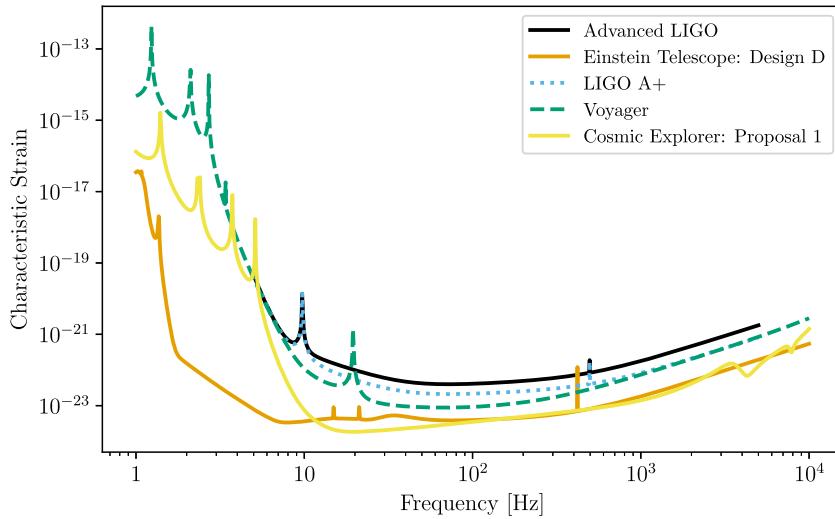


Figure 13. Sensitivities of various ground detector designs.

For the low to mid-range frequencies, there are a combination of noise levels, but the sensitivity is primarily limited by gravity gradient noise and quantum optical noise (photon radiation pressure noise for lower frequencies) (Hild *et al* 2011). At the higher frequencies, the noise is dominated by photon shot noise, which, along with other potential optical path noises, determine the sensitivity for both ground-based interferometers like LIGO/Virgo (Abbott *et al* 2016) and KAGRA (Somiya 2012), and space-based interferometers like LISA (Amaro-Seoane *et al* 2017) at the high end of their respective bands.

5.1. Sensitivity measurements

In principle, we can treat ground-based interferometer sensitivities much like space-based ones. Ground-based detectors also have various noise sources like thermal, seismic, quantum, etc whose power can be summed to generate a sensitivity curve (Hild *et al* 2011). Ground-based instruments also need to take into account the arm length penalty expressed in section 4.2 if their arms are long enough. The transfer frequency for aLIGO is $f_* \approx 10^4$ Hz, thus we can safely assume current ground-based detectors operate in the low frequency limit of a flat response function. Future detectors that are more sensitive at higher frequencies and/or have longer arms will need to include the full response function, as seen in the Cosmic Explorer sensitivity curve in figure 13.

Because of the extensive studies of ground-based sensitivities, there exist tools to model both the micro and macro noise effects in these instruments. Thus one can better characterize noise inside the instrument and better model current and future detectors (Martynov *et al* 2016; Miller *et al* 2015). For all of our SNR calculations and the ground-based sensitivity curves, we use `pygwinc`³ with the lone exception of the sensitivity estimate for ET, for which we use publicly available data⁴.

³ <https://git.ligo.org/gwinc/pygwinc/tree/master>.

⁴ <http://et-gw.eu/index.php/etsensitivities>.

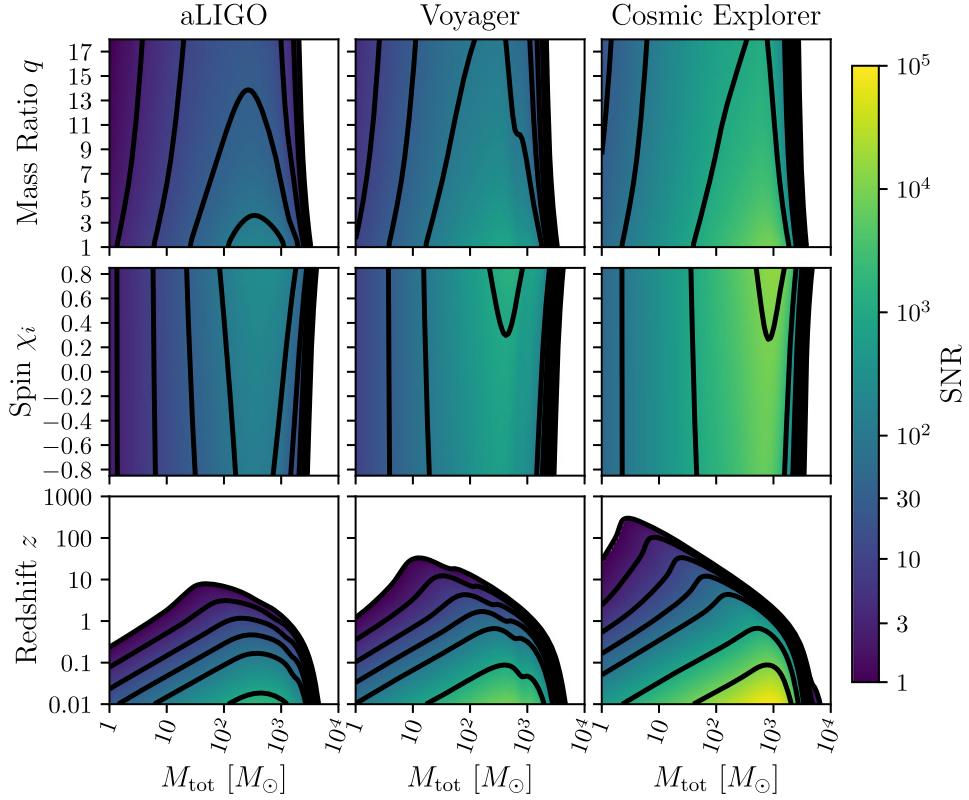


Figure 14. SNRs for changes to a fiducial, non-spinning, equal-mass BHB at a redshift of one ($\chi_i = 0, z = 1, q = 1$) with respect to ground-based detectors: aLIGO (left column), Voyager (middle column), and Cosmic Explorer (right column). Solid black contours correspond to the levels marked on the colorbar. Top row: mass ratio vs total source mass, middle row: spins of both BHs $\chi_i = \chi_1 = \chi_2$ vs total source mass, bottom row: redshift vs total source mass.

`pygwinc` is an analytic sensitivity curve generator for current and future ground-based GW detectors. It combines both public and collaboration-internal noise estimates to provide designs with the most up-to-date information available. For these reasons, we implement `pygwinc` as the primary tool for generating ground-based sensitivity curves. The generated sensitivity curves are integrated into `gwent`'s framework to provide full model adaptability to each variable parameter and ground-based detector design.

In figure 13 we display the default configurations for aLIGO, LIGO A +, Voyager, and Cosmic Explorer using this formalism, along with the sensitivity for ET taken from publicly available data.

5.2. Sensitivity changes to various designs

We examine the detectable GW source parameter space with this formalism using the fiducial detector designs. In figure 14 we show the results of varying the mass ratio, spin, and redshift versus the total mass of a BHB in the ground detector mass regime. It should be noted that all detectable BHBs in the explored parameter space will be evolving in frequency. We find that there is only a moderate bias in SNR strength towards an aligned, highly-spinning BHB in both

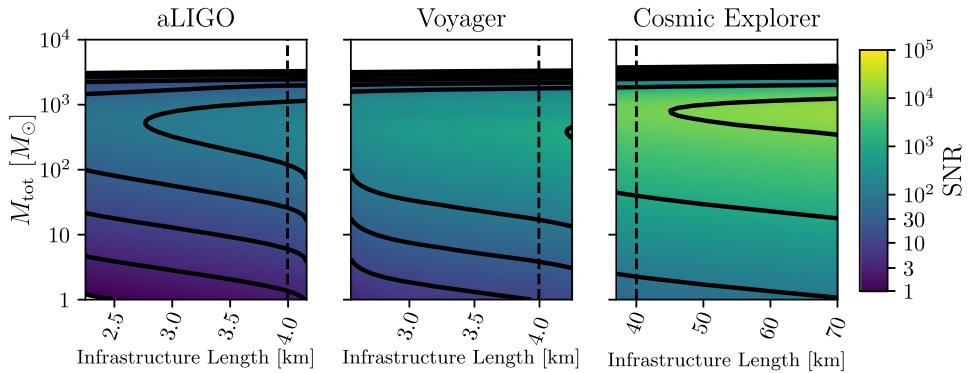


Figure 15. SNRs for the response of variations in arm length to a fiducial, non-spinning, equal-mass BHB at a redshift of one ($\chi_i = 0, z = 1, q = 1$) for aLIGO (left column), Voyager (center column), and CE (right column). Black dashed lines represent the fiducial arm lengths in `pygwinc`. Solid black contours correspond to the levels marked on the colorbar.

the aLIGO and ET models. Both detectors exhibit the consistent behavior we have seen for all GW detectors in this work of higher SNRs for closer to equal-mass mergers.

By extending the arms or making other modifications to the nominal design of current or proposed detectors, we can extend the sensitivity to lower frequencies and/or improve the sensitivity across the band, thereby increasing the SNRs, horizon distances, and sensitive mass ranges. In some cases, the change in sensitivity can be one or more orders of magnitude relative to aLIGO.

We explore varying several of these instrument design parameters around the fiducial instrument values for aLIGO, Voyager, and CE in figures 15 and 16. We find that, for the particular source of a non-spinning, equal-mass BHB at a redshift of $z = 0.1$, the SNR for this example source increases with the arm length for all three detector designs, shown in figure 15. When increasing the arm length (and some other parameters) much outside of the initial design, one needs to simultaneously change other parameters in order to stay in the allowed region of detector parameter space as enforced by `pygwinc`. In this work, we do not simultaneously vary parameters and merely limit the range of variability to the allowed region around the proposed instrument's design value.

Furthermore, for this particular BHB, there is a significant decrease in sensitivity to total mass when increasing the temperature of the mirror substrate. This trend, along with the sensitivity gains leveling out at temperatures below the proposed design levels is common to all of the simulations seen in figure 16. This tendency to reach a point of diminishing returns in sensitivity is common to many parameters because as the individual component's noise floor lowers (in this case by decreasing the substrate temperature), the noise component contributes less to the overall noise as other non-varying sources come to the foreground.

In terms of laser power, we see in figure 16 that aLIGO's sensitivity depends only weakly on the laser power near its design value. In contrast, there appears to be an optimal laser power that maximizes the SNR for particular mass ranges of the example BHB source for Voyager and CE. The optimal value of ~ 200 W for Voyager occurs near its actual planned operating value. For CE, increasing the laser power to the upper bound of the planned operating range yields a steady increase in SNR with increasing laser power for stellar-mass binaries, whereas for a narrow range of intermediate-mass BHs near $10^3 M_\odot$, a lower laser power of ~ 300 W

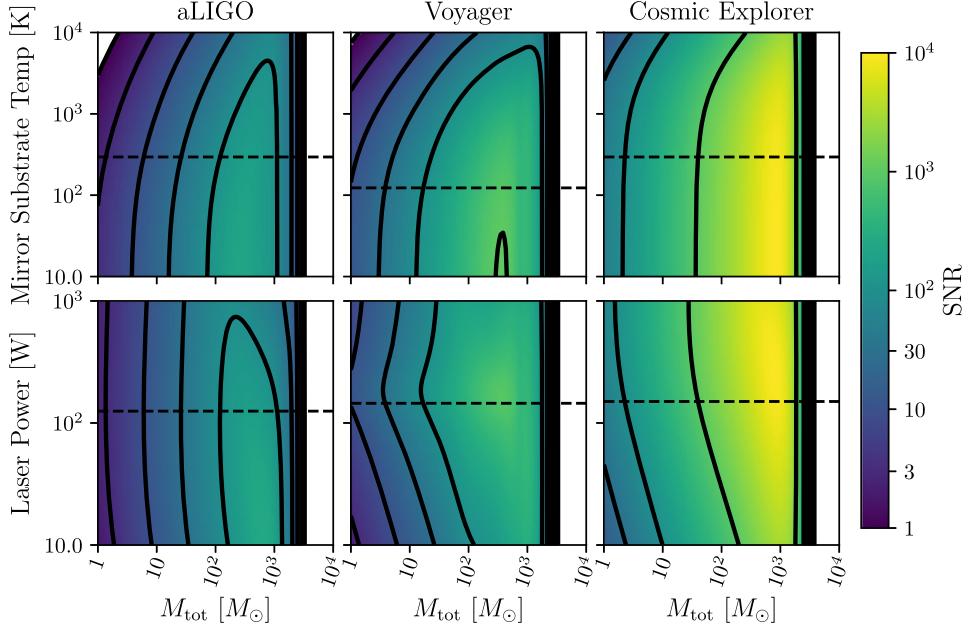


Figure 16. SNRs for the response of variations in instrument parameters to a fiducial, non-spinning, equal-mass BHB at a redshift of one ($\chi_i = 0, z = 1, q = 1$) for aLIGO (left column), Voyager (center column), and CE (right column). Black dashed lines represent the fiducial noise parameters in `pygwinc` for each detector. Solid black contours correspond to the levels marked on the colorbar. Top row: mirror substrate temperature vs total source mass. Bottom row: laser power vs total source mass.

maximizes the SNR. Given the scientific value of such sources as a potential evolutionary missing link between stellar BHBs and SMBHBs, there may be significant benefit in spending some time operating CE closer to this optimal value. Additional investigations for other sources and detector arrangements are left for future work.

As demonstrated in figure 16, investigating the interplay of instrument parameters and BHB source parameters with this formalism can reveal optimal design values. This again highlights the usefulness of the unified formalism implemented here.

6. Discussion

6.1. Application of sensitivities

We have presented the tool, `gwent`⁵, the first open source software for constructing sensitivity curves and performing SNR calculations for current and future GW detectors across the entire GW spectrum of coalescing BHBs. Each detector model contains a sufficiently detailed parameter space to adequately describe the characteristics of a wide variety of detector designs. We present our models for interferometer and PTA GW detectors in the same context to facilitate a fair comparison of capabilities. Additionally, we use a versatile source model to extend the capabilities of our tool to variations in mass, mass ratio, spins, orientation, and redshift.

⁵ Our code, `gwent`, is available at <https://github.com/ark0015/gwent> and on PyPI.

By providing a robust tool for both GW detectors and sources, we are able to clearly show the synergy between PTA, space-based, and ground-based detectors across the relevant parameter space and over the entire GW frequency band.

While examples of similar methods and tools to this study exist (e.g. the work of Moore *et al* (2015)⁶, Katz and Larson (2019)⁷, and Smith and Caldwell (2019)⁸), we believe our treatment is more versatile and extensible as `gwent` is continually adaptable and not limited to a particular instrument.

6.2. Multiband observations of BHBs

Studying the millihertz frequency band with LISA allows us to connect the mostly monochromatic SMBHBs in the PTA band to their merging counterparts, and to study the lower mass mergers at larger redshifts that built up the more massive PTA sources at lower redshifts. Furthermore, LISA is sensitive to the early inspiral of stellar mass BHBs that merge within the ground-based band, providing vital information about their initial evolution and properties; see figure 4 for the overlapping parameter space between instruments.

Including realistic PTA curves using the tools of Hazboun *et al* (2019) alongside other GW detectors allows for a fair comparison of detectors across the GW spectrum. This can elucidate the ability of PTAs to detect individually resolvable GW sources in tandem with space-based detectors, albeit at different times in the BHB's evolution. Furthermore, the modeling ability presented here provides a robust framework for assessing future detector concepts.

6.3. Waterfall plot estimates

The framework we have implemented here allows us to calculate realistic SNRs using multi-parameter GW source models alongside adaptable GW detector designs. Because of the adaptability of both the detector and source models, we can explore changes in detector sensitivity to various sources, as a function of either source or detector parameters. This allows us to assess the optimal instrument parameter choices as a function of either the source population or the detector design.

By combining tools like ours with population synthesis models, one can employ a robust detection probability analysis and source parameter estimation for various GW detectors (Katz *et al* 2020). One can supplement and expand existing methods of PTA detection probability estimation using similar methods (Kelley *et al* 2018).

The adaptability of each of these GW detector models allows us to investigate optimal configurations to maximize sensitivity to particular sources. We plan to investigate these various configurations in future papers. By properly modeling BHB sources and GW detectors across multiple bands and detector classes, we can more accurately assess the true prospects for multi-band GW astronomy.

Acknowledgments

We thank Maura McLaughlin and Jeff Hazboun for helpful feedback on the manuscript and insightful discussion. This work was supported by NSF PFC Grant No. PHY-1430284 and NSF CAREER Grant No. PHY-1945130.

⁶ <https://gwplotter.com/>.

⁷ BOWIE: Binary Observability With Illustrative Exploration.

⁸ <https://doi.org/10.5281/zenodo.3341817>.

Software

Astropy (Astropy Collaboration *et al* 2013, 2018), Hasasia (Hazboun *et al* 2019), Matplotlib (Hunter 2007), NumPy (Oliphant 2006; van der Walt *et al* 2011), Python (Oliphant 2007; Millman and Aivazis 2011), SciPy (Virtanen *et al* 2020).

ORCID iDs

A R Kaiser  <https://orcid.org/0000-0002-3654-980X>
 S T McWilliams  <https://orcid.org/0000-0003-2397-8290>

References

Abbott B P *et al* 2016 *Phys. Rev. Lett.* **116** 061102
 Aggarwal K *et al* 2018 (NANOGrav Collaboration) arXiv:1812.11585
 Amaro-Seoane P *et al* 2017 arXiv:1702.00786
 Anholm M *et al* 2009 *Phys. Rev. D* **79** 084030
 Armano M *et al* 2016 *Phys. Rev. Lett.* **116** 231101
 Arzoumanian Z *et al* 2018a (NANOGrav Collaboration) *Astrophys. J.* **859** 47
 Arzoumanian Z *et al* 2018b (NANOGrav Collaboration) *Astrophys. J. Suppl. Ser.* **235** 37
 Robitaille T P *et al* (Astropy Collaboration) 2013 *Astron. Astrophys.* **558** A33
 Price-Whelan A M *et al* (Astropy Collaboration) 2018 *Astron. J.* **156** 123
 Babak S *et al* 2016 *Mon. Not. R. Astron. Soc.* **455** 1665–79
 Barack L and Cutler C 2004 *Phys. Rev. D* **69** 082005
 Boyle M *et al* 2019 *Class. Quantum Grav.* **36** 195006
 Burke-Spolaor S *et al* 2019 *Astron. Astrophys. Rev.* **27** 5
 Chen S, Sesana A and Conselice C J 2019 *Mon. Not. R. Astron. Soc.* **488** 401
 Cornish N J 2002 *Phys. Rev. D* **65** 022004
 Cornish N J and Larson S L 2001 *Class. Quantum Grav.* **18** 3473–95
 Cornish N and Robson T 2017 Galactic binary science with the new LISA design *J. Phys.: Conf. Ser.* **840** 012024
 Cutler C and Flanagan É 1994 *Phys. Rev. D* **49** 2658–97
 Droz S, Knapp D J, Poisson E and Owen B J 1999 *Phys. Rev. D* **59** 124016
 Einstein A, Infeld L and Hoffmann B 1938 *Ann. Math.* **39** 65
 Ellis J A, Siemens X and Creighton J D E 2012 *Astrophys. J.* **756** 175
 Finn L S 1992 *Phys. Rev. D* **46** 5236
 Flanagan É and Hughes S A 1998 *Phys. Rev. D* **57** 4535–65
 Folkner W M 2011 <https://pcos.gsfc.nasa.gov/studies/rfi/GWRFI-0003-Folkner.pdf>
 Genel S, Vogelsberger M, Springel V, Sijacki D, Nelson D, Snyder G, Rodriguez-Gomez V, Torrey P and Hernquist L 2014 *Mon. Not. R. Astron. Soc.* **445** 175–200
 Hazboun J S, Romano J D and Smith T L 2019 *Phys. Rev. D* **100** 104028
 Hazboun J, Romano J and Smith T 2019 *JOSS* **4** 1775
 Hellings R W and Downs G S 1983 *Astrophys. J.* **265** L39–42
 Hild S *et al* 2011 *Class. Quantum Grav.* **28** 094013
 Hinshaw G *et al* 2013 *Astrophys. J. Suppl. Ser.* **208** 19
 Hobbs G *et al* 2010 *Class. Quantum Grav.* **27** 084013
 Hobbs G 2011 *Astrophys. Space Sci. Proc.* **21** 229
 Hunter J D 2007 *Comput. Sci. Eng.* **9** 90
 Husa S *et al* 2016 *Phys. Rev. D* **93** 044006
 Jenet F A *et al* 2006 *Astrophys. J.* **653** 1571–6
 Katz M L and Larson S L 2019 *Mon. Not. R. Astron. Soc.* **483** 3108
 Katz M L *et al* 2020 *Mon. Not. R. Astron. Soc.* **491** 2301
 Kawamura S *et al* 2006 *Class. Quantum Grav.* **23** S125

Kelley L Z, Blecha L, Hernquist L, Sesana A and Taylor S R 2018 *Mon. Not. R. Astron. Soc.* **477** 964–76

Khan S *et al* 2016 *Phys. Rev. D* **93** 044007

Lam M T *et al* 2017 *Astrophys. J.* **834** 35

Lam M T 2018 *Astrophys. J.* **868** 33

Larson S L, Hiscock W A and Hellings R W 2000 *Phys. Rev. D* **62** 062001

LIGO Scientific Collaboration 2018 LIGO algorithm library—LALSuite free software (GPL) <https://doi.org/10.7935/GT1W-FZ16>

Luo J *et al* 2016 *Class. Quantum Grav.* **33** 035010

Maggiore M 2000 *Phys. Rep.* **331** 283–367

Maggiore M 2007 *Gravitational Waves (Theory and Experiments Oxford Master Series in Physics* vol 1) (Oxford: Oxford University Press)

Martynov D V *et al* 2016 *Phys. Rev. D* **93** 112004

Miller J *et al* 2015 *Phys. Rev. D* **91** 062005

Millman K J and Aivazis M 2011 *Comput. Sci. Eng.* **13** 9

Mingarelli C M F, Lazio T J W, Sesana A, Greene J E, Ellis J A, Ma C-P, Croft S, Burke-Spolaor S and Taylor S R 2017 *Nat. Astron.* **1** 886–92

Moore C J, Cole R H and Berry C P L 2015a *Class. Quantum Grav.* **32** 015014

Moore C J, Taylor S R and Gair J R 2015b *Class. Quantum Grav.* **32** 055004

Oliphant T E 2006 *A Guide to NumPy* (USA: Trelgol Publishing)

Oliphant T E 2007 *Comput. Sci. Eng.* **9** 10

Parida A, Mitra S and Jhingan S 2016 *J. Cosmol. Astropart. Phys.* **JCAP04**(2016)024

Pepper K, Adhikari R and Willems P 2017 https://labcit.ligo.caltech.edu/~cit40m/Docs/T070192_KeenanPepper.pdf

Reitze D *et al* 2019 *Bull. Am. Astron. Soc.* **51** 35

Robson T, Cornish N J and Liu C 2019 *Class. Quantum Grav.* **36** 105011

Rodriguez-Gomez V *et al* 2015 *Mon. Not. R. Astron. Soc.* **449** 49–64

Rosado P A *et al* 2016 *Phys. Rev. Lett.* **116** 101102

Ruan W-H *et al* 2020 *Int. J. Mod. Phys. A* **35** 2050075

Saulson P R 1994 *Fundamentals of Interferometric Gravitational Wave Detectors* ed P Saulson (Singapore: World Scientific)

Sazhin M V 1978 *Sov. Astron.* **22** 36

Schmitz K 2020 arXiv:2005.10789

Sesana A, Shankar F, Bernardi M and Sheth R K 2016 *Mon. Not. R. Astron. Soc. Lett.* **463** L6–L11

Skrutskie M F *et al* 2006 *Astron. J.* **131** 1163–83

Smith T L and Caldwell R R 2019 *Phys. Rev. D* **100** 104055

Somiya K 2012 *Class. Quantum Grav.* **29** 124007

Taylor S R *et al* 2016 *Astrophys. J. Lett.* **819** L6

Teukolsky S A and Press W H 1974 *Astrophys. J.* **193** 443–61

Thorne K S 1989 *Three Hundred Years of Gravitation* ed S W Hawking and W Israel (Cambridge: Cambridge University Press) pp 199–446

Thrane E and Romano J D 2013 *Phys. Rev. D* **88** 124032

Tinto M and Dhurandhar S V 2014 *Living Rev. Relativ.* **17** 6

van der Walt S, Colbert S C and Varoquaux G 2011 *Comput. Sci. Eng.* **13** 22

Virtanen P *et al* 2020 *Nat. Methods* **17** 261

Wainstein L A, Zubakov V D and Mullin A A 1963 *Am. J. Phys.* **31** 68