



Multi-messenger Approaches to Supermassive Black Hole Binary Detection and Parameter Estimation: Implications for Nanohertz Gravitational Wave Searches with Pulsar Timing Arrays

Tingting Liu and Sarah J. Vigeland

Center for Gravitation, Cosmology and Astrophysics, Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA; tingtliu@uwm.edu
 Received 2021 May 17; revised 2021 August 5; accepted 2021 August 13; published 2021 November 12

Abstract

Pulsar timing array (PTA) experiments are becoming increasingly sensitive to gravitational waves (GWs) in the nanohertz frequency range, where the main astrophysical sources are supermassive black hole binaries (SMBHBs), which are expected to form following galaxy mergers. Some of these individual SMBHBs may power active galactic nuclei, and thus their binary parameters could be obtained electromagnetically, which makes it possible to apply electromagnetic (EM) information to aid the search for a GW signal in PTA data. In this work, we investigate the effects of such an EM-informed search on binary detection and parameter estimation by performing mock data analyses on simulated PTA data sets. We find that by applying EM priors, the Bayes factor of some injected signals with originally marginal or sub-threshold detectability (i.e., Bayes factor ~ 1) can increase by a factor of a few to an order of magnitude, and thus an EM-informed targeted search is able to find hints of a signal when an uninformed search fails to find any. Additionally, by combining EM and GW data, one can achieve an overall improvement in parameter estimation, regardless of the source's sky location or GW frequency. We discuss the implications for the multi-messenger studies of SMBHBs with PTAs.

Unified Astronomy Thesaurus concepts: [Gravitational waves \(678\)](#); [Supermassive black holes \(1663\)](#); [Active galaxies \(17\)](#)

1. Introduction

Multi-messenger astrophysics (MMA) began with the observations of neutrinos from the Sun (which dates back to the 1960s) and later from the supernova SN1987A and the 2017 observations of a binary neutron star merger (Abbott et al. 2017) ushered in a new era of the field: for the first time, the messengers were light and gravitational waves (GWs). This burgeoning field of GW MMA is enabling new and unique insight into the source's astrophysics that is otherwise inaccessible with traditional observational astronomy, as well as gravity theories, fundamental physics, and cosmology.

The future prospects of MMA with GW and electromagnetic (EM) observations are tremendously promising. The Laser Interferometer Space Antenna (LISA, Amaro-Seoane et al. 2017) will observe GWs in the mHz regime (10^{-4} – 10^{-2} Hz), whose astrophysical sources include white dwarf binaries, massive black hole binaries (MBHBs, $\gtrsim 10^5 M_\odot$), and extreme mass-ratio inspirals (e.g., Klein et al. 2016; Babak et al. 2017). Pulsar timing array (PTA) experiments (e.g., Hobbs 2013; Kramer & Champion 2013; McLaughlin 2013) probe the nanohertz regime (10^{-9} – 10^{-6} Hz), where the main astrophysical sources are the supermassive black hole binaries (SMBHBs, $\gtrsim 10^9 M_\odot$). Unlike most stellar-mass black hole binaries, MBHB and SMBHB systems are expected to emit EM radiation: the galaxy merger process, which is thought to be the formation channel of MBHBs and SMBHBs, can funnel a large amount of gas to the nuclear region (Barnes & Hernquist 1992), thereby enabling accretion onto the BHs and potentially powering active galactic nucleus (AGN) activity.

As these close SMBHB systems in the GW regime (which are typically at approximately milliparsec separations) are largely beyond the reach of direct imaging and interferometric observations, the indirect observational signatures of these binary systems have been of great interest to a large body of theoretical and numerical work. These EM indicators of a close SMBHB can be roughly classified into one or more of the following categories:

peculiar spectral features (e.g., Gültekin & Miller 2012; Kocsis et al. 2012; Tanaka et al. 2012; McKernan et al. 2013; Roedig et al. 2014; Farris et al. 2015; d'Ascoli et al. 2018), periodic variations or flares as the result of an orbiting binary (e.g., MacFadyen & Milosavljević 2008; Noble et al. 2012, 2021; Shi et al. 2012; D'Orazio et al. 2013; Farris et al. 2014; Gold et al. 2014; Bowen et al. 2018; D'Orazio & Di Stefano 2018; Tang et al. 2018), and other associated transient phenomena in an otherwise quiescent system (such as tidal disruption events with modified dynamics, e.g., Liu et al. 2009; Ricarte et al. 2016). Binary population estimates have further made predictions for the occurrence rates of binaries displaying these signatures (Kelley et al. 2019; Krolik et al. 2019), which are estimated to be $\mathcal{O}(10^2)$ in the volume probed by facilities such as the forthcoming Vera Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 2008). This indicates the possibility of identifying SMBHBs electromagnetically in meaningful numbers in current and future astronomical data sets.

Indeed, the advent of large, sensitive, and modern time-domain surveys in the past decades has opened us to the possibility of searching for those theoretically predicted signatures among a large sample of AGN, making it a promising avenue to identify these intrinsically rare sources. In the past few years alone, more than 100 SMBHB candidates have been proposed in the literature as the results of systematic searches (e.g., Graham et al. 2015a, 2015b; Liu et al. 2015, 2016, 2019; Charisi et al. 2016; Chen et al. 2020) and serendipitous discoveries alike (e.g., Hu et al. 2020; Shu et al. 2020). While observational searches for SMBHBs can be susceptible to false-positive contamination,¹ this can be partially

¹ For instance, false positives can result from time-domain searches when the intrinsic temporal variability of “regular” AGNs is mistaken for periodicity, especially given limited data length and quality (see Vaughan et al. 2016). The high rate of false positives is also suggested by comparing the ensemble GW signal (the gravitational wave background, GWB) implied by the observational sample with the independent PTA constraints on the GWB amplitude (see Sesana et al. 2018).

overcome by, e.g., more sensitive observations and longer temporal baselines provided by e.g., the LSST, and it can be expected that EM observatories will continue to expand their roles in the search for SMBHBs.

In addition to the obvious goal of uncovering the elusive SMBHBs, a wealth of astrophysical information about the binary (or binary candidate) can be extracted from EM observations, either through standard observational techniques developed for “regular” AGNs, e.g., to estimate the total black hole mass; or applying a specific binary model, e.g., to obtain the mass ratio or orbital period. A remarkable example is the well-known (and currently the best) SMBHB candidate OJ 287 (e.g., Sillanpaa et al. 1988; Lehto & Valtonen 1996; Valtonen et al. 2008, 2016), where its long-term light curve allows one to obtain an orbital solution and derive parameters including orbital period, primary and secondary BH masses, eccentricity, and even precession rate and primary BH spin. More generally, for binary candidates discovered in time-domain surveys that display apparent periodic variability, the putative period can be determined with time-series analysis techniques, and constraints on other parameters such as orbital inclination and mass ratio may be made by modeling the source light curve under the assumption of the binary model that best describes the possible periodicity (an example is the interpretation of the apparent optical periodicity of PG 1302-102 as relativistic beaming by D’Orazio et al. 2015).

This has particularly interesting implications for MMA with SMBHBs. In the context of searching for and studying SMBHBs with the PTA, a “continuous wave” (CW) signal from a circular, non-spinning SMBHB can be described by eight independent parameters: sky position (two parameters), chirp mass, GW frequency, inclination, characteristic strain (or luminosity distance), GW polarization angle, and GW phase. With the exception of the last two GW parameters, they could all be determined or constrained by EM (including time-domain) observations. EM information obtained from the source may then be utilized in the search for a CW signal in PTA data and potentially enhance the detectability of the signal, which may otherwise be marginal. Conversely, the PTA GW data estimate the binary parameters in a way that is completely independent from EM observations, and the detection of a second messenger, GWs, from an electromagnetically selected SMBHB candidate and the independent verification of its source parameters would be more robust against false-positive detections and could provide the ultimate confirmation.

The effort to search for GWs in PTA data using the EM information of an SMBHB candidate dates back to the early years of PTA experiments, when Jenet et al. (2004) searched in 7 yr long observations of one pulsar for the timing residuals (that is, difference between the observed and expected pulse arrival times) induced by the proposed SMBHB in the center of the radio galaxy 3C 66B (Sudou et al. 2003), and placed upper limits on its mass and eccentricity from the non-detection of GWs. More recently, Arzoumanian et al. (2020b) searched for GWs from 3C 66B in the timing data of 34 pulsars spanning up to 11 yr and placed a constraint on its mass at a level that is almost competitive with EM observations (Iguchi et al. 2010). This is in part attributed to the boosted sensitivity by using EM information as priors: they demonstrated that they were able to gain a factor of two in upper-limit sensitivity by targeting the CW search at the

sky position of the source, rather than performing an all-sky search; and an order of magnitude improvement in sensitivity by searching for a signal at the GW frequency suggested by 3C 66B’s apparent orbital motion, instead of performing an all-frequency search. This approach is analogous to the “directed search” (when sky location is known) and “targeted search” (when frequency and other parameters are also known) for the continuous GW signals from rotating, non-axisymmetric neutron stars with LIGO (e.g., Aasi et al. 2015; Abbott et al. 2019). These searches similarly reduce the computational cost and enhances the sensitivity compared to an all-sky search for an unknown source.

As the PTA sensitivity to nanohertz CWs continues to improve with longer observations and more pulsars, it is poised to put tighter constraints on EM-selected SMBHB candidates, provide more powerful tests of their binary hypothesis, and even make a detection in the near future. Indeed, binary population simulations suggest that a CW detection is possible in the next decade with an improved or next-generation PTA (Rosado et al. 2015; Kelley et al. 2018), and predictions based on observational samples of possible binary hosts are equally promising (Mingarelli et al. 2017; Xin et al. 2021). For instance, the latter study suggests that, OJ 287, 3C 66B, and a handful of SMBHB candidates from Graham et al. (2015a) (if they are genuine SMBHBs) would be within the detection sensitivity of the next-generation PTA experiment with the Square Kilometre Array (Janssen et al. 2015) in the 2030s.

With the dawn of nanohertz GW and multi-messenger astronomy upon us, so is the urgency to systematically and thoroughly understand the PTA’s capability as a GW observatory, including its ability to estimate binary parameters and its potential to search for and study the source in a multi-messenger approach. This paper is a step toward that goal, by examining the effects of a multi-messenger search on CW detection and source-parameter estimation using a realistic, mock, data analysis approach. To that end, we inject CW signals of modest GW amplitudes in simulated 11 yr long PTA data sets as proxies of the first detectable CW signals in some future, improved PTA, data set (Section 2.1). We attempt to recover the signal parameters by performing mock uninformed and EM-informed searches (Sections 2.2 and 2.3). We quantify and compare the results of those searches (Sections 3.1 and 3.3) and discuss implications for future CW searches and SMBHB studies (Section 4). We conclude in Section 5. Throughout the paper, we assume a nine-year WMAP cosmology (Hinshaw et al. 2013) and adopt geometrized units with $G = c = 1$.

2. Methods

2.1. A Simulated PTA and CW Injections

We first construct a simulated PTA using properties of the pulsars from the NANOGrav 11 yr data set² (Arzoumanian et al. 2018b). Of the 45 pulsars that are timed over an 11 yr period, 34 have time spans of more than three years and were used for NANOGrav’s suite of 11 yr GW analyses (Arzoumanian et al. 2018a, 2020b, 2021; Aggarwal et al. 2019, 2020). We additionally remove PSR B1937+21 from the simulated PTA, whose timing residual is known to be characterized by a

² <https://data.nanograv.org>

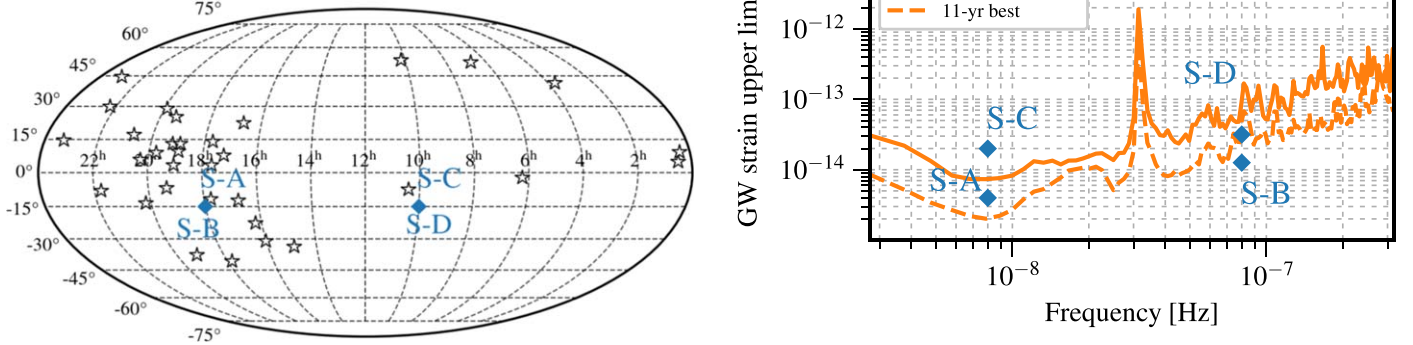


Figure 1. Left panel: sky locations of injected CW sources, S-A–S-D (diamonds). Locations of the pulsars used in the simulated PTA data set are marked with stars. Right panel: for illustration purposes only, we also show the strain amplitudes of the injected signals compared to the actual 11 yr upper-limit sensitivity on CWs averaged over the sky and at the best sky location (solid and dashed lines, respectively; reproduced from Aggarwal et al. 2019). Note that S-A and S-B are located near the best sky location.

large amount of red noise above the white noise level (e.g., Lam et al. 2017; Arzoumanian et al. 2018b). Our simulated PTA thus consists of 33 pulsars that retain the observing cadence and sky locations (Figure 1, left panel) of the 11 yr pulsars.

To simulate their timing residuals, we used the `libstempo` software package³ and modeled them as a combination of white noise, red noise, and residuals induced by a CW signal. The white noise component represents the nominal TOA measurement uncertainties ($\text{EFAC} = 1$), and the red noise component encompasses various effects such as spin noise intrinsic to the pulsar and interstellar medium propagation effects and is modeled as a power law. The power-law model is parameterized by an amplitude A_m and a spectral index γ_m , which have been measured separately for each pulsar.

To inject a CW signal using the functionality in `libstempo`, the following signal model parameters are required: $\{\phi, \theta, d_L, \mathcal{M}, f_{\text{gw}}, i, \Phi_0, \psi\}$.

1. The parameters ϕ and θ are the azimuthal angle and polar angle, respectively, and are conventionally used in PTA GW analyses to represent the CW source’s position on the sky. In observational astronomy, they correspond to R.A. α and decl. δ , respectively: $\phi = \alpha$, $\theta = \pi/2 - \delta$. We inject CW sources at two representative sky locations (Figure 1, left panel): a sensitive sky location where a large number of pulsars are being monitored nearby (R.A. 18^{h} , decl. -15°), and a less sensitive sky location where very few of them are nearby (R.A. 10^{h} , decl. -15°).
2. The parameter d_L is the luminosity distance of the CW source. Assuming a cosmology, this can be obtained observationally by measuring the redshift z of the host galaxy of the SMBHB, or the AGN for an active SMBHB. We injected mock CW sources in the very nearby universe at $z < 0.1$ to mimic the detectable signals that are already present in the respective simulated data sets.
3. The parameter f_{gw} is the GW frequency and is related to the orbital frequency of the SMBHB: $f_{\text{gw}} = 2f_{\text{orb}}$ (assuming a circular orbit). We inject CW signals at two different frequencies (Figure 1; right panel): $f_{\text{gw}} = 8$ nHz,

which is approximately the most sensitive frequency in the 11 yr data set; and $f_{\text{gw}} = 80$ nHz, which is motivated by the observational search for AGN periodicity: in time-series data with a baseline of a few years (which is typical for current time-domain surveys), at least a few cycles are required before claiming a periodic detection, hence making the EM search more sensitive to periods of approximately months. Furthermore, the stochastic GW background (GWB), on which the CW signal would be superimposed, has a canonical power-law spectral shape and thus has a lower amplitude at high GW frequencies, potentially making the search for a CW source emitting at a high GW frequency more fruitful.

4. \mathcal{M} is the (redshifted) “chirp mass” of the SMBHB and is a combination of the component masses: $\mathcal{M} \equiv (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$. Since our injected CW sources are in the very local universe, we make the approximation that the rest-frame chirp mass $\mathcal{M}_r \approx \mathcal{M}$. The amplitude of the GW signal is then related to d_L , f_{gw} , and \mathcal{M} :

$$h_0 = \frac{2\mathcal{M}^{5/3}(\pi f_{\text{gw}})^{2/3}}{d_L}. \quad (1)$$

In our analyses, we parameterize the signal strength in terms of the total mass $M_{\text{tot}} = m_1 + m_2$ and the mass ratio $q \equiv m_2/m_1 \leq 1$ rather than the chirp mass because those parameters are more closely tied to EM observations. Observationally, the total mass of the system can be measured via standard methods such as single-epoch virial black hole mass estimation (assuming the method is still valid for close binary systems), which suffers from an ~ 0.3 dex systematic uncertainty (e.g., Shen 2013 and references therein). In this study, we assign astrophysically reasonable black hole masses to the injected CW sources: $\log M/M_\odot = 9.4\text{--}9.5$. The mass ratio of a binary system cannot always be directly determined from the light curve of a periodically varying SMBHB (see Section 4 for further discussion), and we keep the mass ratio fixed at $q = 0.8$ in all injections for simplicity.

³ <https://github.com/vallis/libstempo/>

Table 1
Injected CW Parameters

Parameter	Simulation				
	0	A	B	C	D
R.A. (hr)	(18)	18	18	10	10
Decl. (deg)	(−15)	−15	−15	−15	−15
f_{gw} (nHz)	(8)	8	80	8	80
$\log \mathcal{M}$ (M_{\odot})	(9.4)	9.4	9.4	9.5	9.5
z	(0.06)	0.06	0.08	0.02	0.05
$\log h_0$...	−14.4	−13.9	−13.7	−13.5

Note. For all simulations: $q = 0.8$, $i = \pi/2$, $\Phi_0 = 2.0$, $\psi = 2.0$. Additionally, S-0 does not contain an injected CW signal, but the search was performed using the parameters of S-A (shown in parentheses).

- The parameter i is the inclination of the binary orbit and could be observationally constrained, including through light-curve fitting. We keep the inclination fixed at $i = \pi/2$ (which corresponds to an edge-on binary) for all injections for simplicity. We note that, while an edge-on or near edge-on orientation is optimal for EM observers in the cases of relativistic beaming and binary self-lensing, it is the least optimal for GW detection because it induces the smallest pulsar timing residuals.
- The parameters Φ_0 and ψ are the initial orbital phase and the GW polarization angle in units of radians, respectively. As they cannot be measured from EM observations, we keep them fixed at $\Phi_0 = 2.0$ and $\psi = 2.0$ in all injections for simplicity.

The parameter values for each injected CW signal are summarized in Table 1. Additionally, we have included the so-called “pulsar term” in the injected signal, which is the signal induced at the pulsar (whereas “Earth term” is the signal induced at the Earth). In contrast to the alternative approach that uses a signal model that only contain the Earth term, our analysis is sensitive to the evolution, or “chirping,” of the signal, and hence the chirp mass, as GWs pass through the pulsar and the Earth.

For each simulation (Simulations A-D), we first calculate the GW strain h_0 from the binary parameters (Figure 1, right panel; Table 1). We have generated two realizations for each simulation, denoted S-X1 and S-X2, in order to study the potential effect of pulsar white and red noise realizations on the CW search. For each realization, we compute the \mathcal{F} -statistic, which is the log-likelihood ratio maximized over the so-called extrinsic parameters of the source and is used as a detection statistic for GW searches (Jaranowski et al. 1998; Cornish & Porter 2007; Babak & Sesana 2012). In the context of this work, we evaluate the pulsar-term \mathcal{F} -statistic (\mathcal{F}_p , Ellis et al. 2012) at the injected GW frequency. We additionally compute the corresponding SNR, since $2\mathcal{F}_p$ follows a χ^2 distribution with $2N$ degrees of freedom, where N is the number of the pulsars: $\text{SNR} = \sqrt{2\mathcal{F}_p} - 2N$. We show the \mathcal{F}_p and SNR of each realization of S-A–S-D in Table 2. The different mock data sets containing the same injected CW signal differ slightly in \mathcal{F}_p and SNR, as expected given their different noise realizations.

As we see in Tables 1 and 2, Simulation A has the highest \mathcal{F}_p statistic or SNR despite having the lowest nominal GW amplitude, which can be expected as it is emitting at a GW frequency favorable to the PTA. Conversely, Simulation D has

Table 2
 \mathcal{F}_p Values and SNRs of different Realizations of S-A–S-D

	\mathcal{F}_p	SNR
S-A1	55	6.6
S-A2	45	4.9
S-B1	40	3.7
S-B2	44	4.7
S-C1	42	4.2
S-C2	43	4.5
S-D1	30	<1
S-D2	32	<1

the largest GW amplitude but has the lowest detection statistic or SNR, due to its GW frequency to which the NANOGrav PTA is less sensitive. Overall, Simulations A-D represent CW signals of a range of parameters and are capable of serving as test cases in this proof-of-concept study to investigate any effects of sky location and GW frequency on CW searches. Their comparable SNRs further permit reasonable comparisons between the mock CW search results. Moreover, their modest SNRs are suitable proxies for the first CW signals in some future PTA data set, which are expected to be weak.

We note that, in real PTA data, any CWs detectable as discrete signals would be superimposed on the stochastic GWB, which is the superposition of GW signals from a cosmological population of SMBHBs, and it is probable that multiple CW signals may be detectable simultaneously above the GWB. However, since the 11 yr NANOGrav analysis does not show evidence for the GWB (Arzoumanian et al. 2018a), which is expected to emerge in a data set spanning at least 15 yr (Taylor et al. 2016; Pol et al. 2021), we do not inject a GWB in our simulated PTA data sets. Additionally, the joint search for one or more CW signals in addition to a GWB (Bécsy & Cornish 2020) requires techniques such as a trans-dimensional Reversible Jump MCMC sampler (e.g., Green 1995) to explore a parameter space whose dimensionality is not fixed, which is beyond the scope of this work. For these reasons, we only inject one CW source in a given simulated PTA data set.

We further note that we have based our simulated PTA on the NANOGrav 11 yr data set, instead of the more recent 12.5 yr data set (Alam et al. 2021). Since a red noise process common among the 12.5 yr pulsars is present in this data set, but whose origin cannot yet be determined (Arzoumanian et al. 2020a), we do not utilize this data set in our analysis. Also, an 11 yr like simulated data set permits reasonable comparisons with NANOGrav results from the latest CW analysis, which was performed with the 11 yr data set (Aggarwal et al. 2019). Furthermore, we argue that the main results from this proof-of-concept work should not strongly depend on the length and size of the particular data set, since the SNRs of the injected CW signals are relative to the sensitivity of the data set. That said, extending this work to a later data set, either actual or projected, would be of interest to future work.

Finally, we generate two additional mock data sets with no injected CW signals (S-0). In Sections 2.2 and 2.3, we will treat S-0 in the same manner as S-A when performing uninformed and targeted searches (Table 1) and will compare and contrast the search results. This is motivated by the high false-positive rate of SMBHB searches (see, e.g., the footnote in Section 1 for a brief discussion), and searching for a CW signal in S-0 thus

Table 3
Priors in Uninformed Searches

Parameter	Prior	Range	Symbol in Table 5
ϕ	Uniform	$[0, 2\pi]$	\times
$\cos \theta$	Uniform	$[-1.0, 1.0]$	\times
$\log d_L$	Uniform	$[-2.0, 4.0]$	\times
$\log M_{\text{tot}}$	Uniform	$[6.0, 12.0]$	\times
$\log f_{\text{gw}}$	Uniform	$[-9.0, -7.0]$	\times
q	Uniform	$[0.0001, 1.0]$	\dots
Φ_0	Uniform	$[0, \pi]$	\dots
ψ	Uniform	$[0, \pi]$	\dots
$\cos i$	Uniform	$[-1.0, 1.0]$	\dots

Note. q , $\cos i$, Φ_0 , and ψ are always searched over their respective priors and are hence omitted in Table 5.

represents the case where a CW search is performed using the information of an EM candidate which turns out to be a false positive.

2.2. Mock Uninformed Searches

For each injection, we first perform a “blind” search, where the MCMC sampler is free to explore the prior ranges of all parameters (Table 3). This procedure thus mimics an unguided and uninformed search for a CW signal in PTA data. We perform these mock searches using the NANOGrav GW detection software package *enterprise*.⁴ As we largely adopt the same data analysis methods as the previous CW analyses (modified for a simulated PTA data set), we refer the reader to Ellis et al. (2012) or Aggarwal et al. (2019) for a detailed description of the CW signal model and search methods. To explore the parameter space, we use *PTMCMCSampler*⁵ (Ellis & van Haasteren 2017), with a geometrically spaced temperature ladder for parallel tempering (e.g., Earl & Deem 2005). Each chain is then sampled for $(1-2) \times 10^6$ iterations and individually inspected for convergence using the Geweke diagnostic (Geweke 1992).

2.3. Mock Targeted Searches

To study the effects of applying EM-informed priors on the detection and parameter estimation of a CW signal, we perform three separate searches for the same injected signal in a progressive scheme:

1. *ra_dec_z*: We search for the signal at its injected sky location and luminosity distance. This mimics the search for a CW signal in the PTA data set which is prompted by the detection (or observation) of its EM counterpart (or candidate). In this scenario, we assume the SMBHB’s location and redshift can be measured exactly, which is a reasonable assumption for EM observations of a source.
2. *ra_dec_z_dmtot*: We search for the signal at its injected sky location, luminosity distance, and in addition, we take into account a 0.3 dex systematic uncertainty on the black hole mass measurement. In practice, this is implemented through adopting a normal prior for $\log M_{\text{tot}}$ which is centered at the injected value. We note

Table 4
Priors in Targeted Searches

Parameter	Prior	Symbol in Table 5
ϕ	Fixed	\checkmark
$\cos \theta$	Fixed	\checkmark
$\log d_L$	Fixed	\checkmark
$\log M_{\text{tot}}$	Normal	0.3 dex
$\log f_{\text{gw}}$	Uniform	$\times 4$

that, in practice, this scenario almost always contains the previous case; in other words, if an EM counterpart can be identified at all, its black hole mass can usually be estimated, and therefore in practice, *ra_dec_z* is never a stand-alone case for a targeted search. However, here we study both cases in order to isolate the effects of knowing the black hole mass on the targeted search.

3. *ra_dec_z_dmtot_dfgw*: This represents the case where the SMBHB is detected electromagnetically as a periodically varying AGN, where the observed period is interpreted as the imprint of the binary orbital period. In addition to the priors above, we take into account the uncertainty of the GW frequency, which is motivated by factors including measurement uncertainties and being agnostic about how the observed variability period translates to the binary orbital period (which may have a binary model and/or mass-ratio dependence, see Section 4 for details). In practice, we apply a uniform prior to $\log f_{\text{gw}}$, while allowing a factor of two uncertainty around the injected value, i.e., a total of a factor of four uncertainty on f_{gw} .

The priors applied in targeted searches are listed in Table 4, and we summarize the uninformed search and all targeted searches in Table 5.

3. Results and Discussion

3.1. Detectability

We first investigate the effects of a targeted search on the detectability of the CW signal. To this aim, we compute the Bayes factor for the presence of a CW signal. We use the Savage–Dickey formula (Dickey 1971), which can be used to compute the Bayes factor \mathcal{B}_{10} for two nested models \mathcal{H}_1 and \mathcal{H}_0 . In this case, \mathcal{H}_0 is a model with pulsar noise only, and \mathcal{H}_1 is a model with noise plus a CW signal, which is equivalent to noise only for $h_0 = 0$. For more details, see Appendix A.

In Figure 2, we show the evolution of the Bayes factor from an uninformed search to targeted searches. In *uninformed*, all Bayes factors $\mathcal{B}_{10} < 1$, indicating that no evidence for the CW signal can be found. This is expected from the low SNRs of the signals. In *ra_dec_z*, \mathcal{B}_{10} has increased slightly, but stays below the nominal threshold of $\mathcal{B}_{10} = 1$, or barely above it. Thus, knowing the source’s sky location does not appear to significantly improve the detectability of the source. However, by additionally knowing the black hole mass (*ra_dec_dmtot*), the increases in \mathcal{B}_{10} are more significant (except for S-C1); in particular, \mathcal{B}_{10} in S-A and S-B have increased by at least a factor of a few. In *ra_dec_dmtot_dfgw*, where the frequency is also known, \mathcal{B}_{10} increases even further in some cases. S-A and S-B have the larger final \mathcal{B}_{10} values ($\sim 5-10$) among the four injections. While a Bayes factor of ~ 10 is shy of being able to claim a confident detection, it is noteworthy that evidence for

⁴ <https://github.com/nanograv/enterprise>

⁵ <https://github.com/jellis18/PTMCMCSampler>

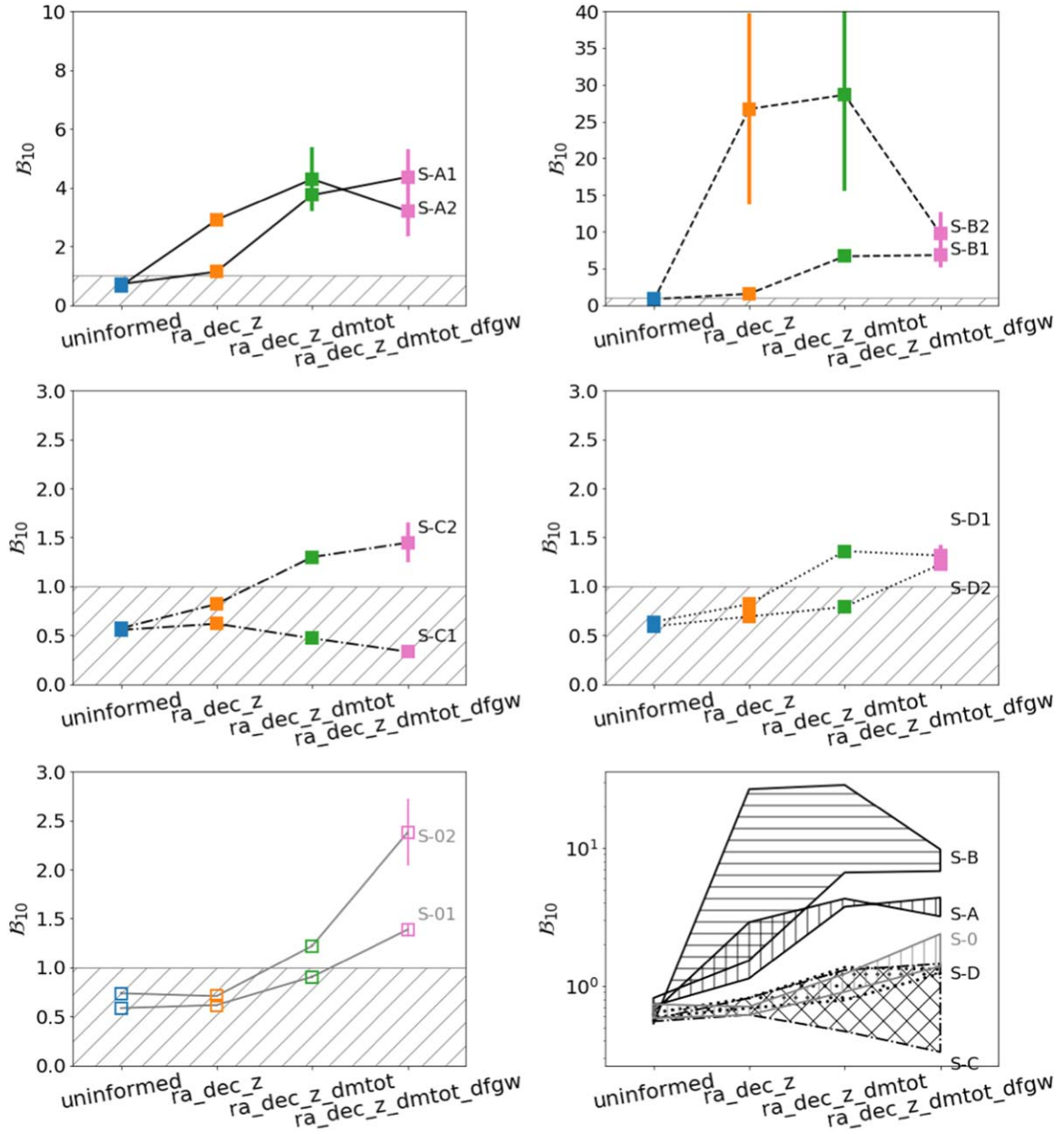


Figure 2. The first five panels show the mean Bayes factors B_{10} for the presence of a CW signal in *uninformed*, *ra_dec_z*, *ra_dec_dmtot*, *ra_dec_dmtot_dfgw* (blue, orange, green, and purple squares, respectively) for the two realizations of S-0–S-D. The error bars represent the uncertainties of B_{10} (see Appendix A). Note that some of the error bars are not visible due to their small size. The gray hashed region below the nominal detection threshold of $B_{10} = 1$ is shown for visual reference. Note the different ranges of the y-axes. The last panel compares the evolution of the B_{10} shown in the first five panels: S-A, vertical hashed; S-B, horizontal hashed; S-C, crossed hashed; S-D, dotted hashed; S-0, gray vertical hashed. The upper and lower bounds are given by the respective mean B_{10} of the two realizations S-X1 and S-X2; uncertainties in individual realizations are omitted for clarity. The y-axis is shown in log-scale for clarity.

Table 5
Names of Uninformed and Targeted Searches in Figures 3–10

Parameter	Name			
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	×	✓	✓	✓
$\cos \theta$	×	✓	✓	✓
$\log d_L$	×	✓	✓	✓
$\log M_{\text{tot}}$	×	×	0.3 dex	0.3 dex
$\log f_{\text{gw}}$	×	×	×	×4

the (weak) CW signal can strengthen by at least a factor of a few via the EM observation of the source.

Additionally, the overall trend described above appears to hold for both realizations, albeit with mild variation between realizations. The variation is likely the result of the MCMC analysis or the noise realization, rather than a simple correlation with SNR of the source. For example, S-A2 contains the same CW signal as S-A1 and has a slighter lower SNR, but has the highest final \mathcal{B}_{10} . Similarly, S-C1 and S-C2 have almost identical SNRs, but an overall increase in \mathcal{B}_{10} is seen in S-C2, while a slight decrease is observed in S-C1.

Despite the individual variations, the evolution trends of S-A and S-B appear to be separated from those of S-C and S-D (last panel in Figure 2). This is likely due to the fact that S-A and S-B are located at a sensitive sky location, which is favorable for CW searches originally (see, e.g., Aggarwal et al. 2019) and benefits from EM information even further. However, we note that the Savage–Dickey approximation starts to break down for stronger detections, and therefore we advise the reader to take the \mathcal{B}_{10} values (and their error bars) of S-B2 with a grain of salt.

Furthermore, comparing the trends of S-A versus S-B, the latter appears to be more detectable than the former: both realizations, S-B1 and S-B2, have higher final \mathcal{B}_{10} than S-A. This is despite the fact that S-B has a quieter signal than S-A. We speculate that this is because S-B is emitting at a higher GW frequency than S-A, so that the CW signal is observed for more cycles over the same data length. While higher frequencies are conventionally considered “less sensitive” frequencies in terms of the GW amplitude, our results suggest that sources emitting at high frequencies could be as detectable as (or even more detectable than) those emitting at “sensitive” frequencies. This is encouraging for the possibility of strong synergies with EM observations (see Section 4 for further discussion). However, we are unable to observe the same dependence on frequency for S-C versus S-D, as their improvements in \mathcal{B}_{10} are much weaker.

We caution that these trends ought to be confirmed with a larger number of realizations and parameters than we have included here, especially given the aforementioned variation between different noise realizations. Unfortunately, it is prohibitively expensive to perform MCMC analysis on a statistically meaningful number of CW searches, for which a more robust and efficient detection pipeline is still needed. Therefore, while our current analyses suggest a dependence of CW detectability on sky location and GW frequency, at present we are not able to fully investigate the effect of different noise realizations on this trend.

3.2. False-positive Detections

We then proceed to investigate whether the above trend applies to S-0, which does not contain an injected CW signal. As we show in the bottom left panel in Figure 2, both realizations have slightly increased Bayes factors as a result of targeted searches. Therefore, applying (incorrect) EM priors appears to also risk increasing the false alarm probability. Additionally, the final Bayes factor in $ra_dec_z_dmtot_dfgw$ is ~ 2 , which is comparable to that of S-C or S-D. Thus in actual searches, it would be difficult to distinguish a false-positive detection from a low SNR source, if the Bayes factor and its increase are interpreted at face value.

However, it is important to note that the final Bayes factor is significantly below that of S-A, whose source parameters were

used in the search for a signal in S-0. Therefore, we argue that the trend of increased \mathcal{B}_{10} of true CW signals (and S-A in particular) cannot simply be the result of possibly increased false alarm probability due to the EM-informed search. Instead, this indicates that a false-positive detection is less likely to be mistaken for a true CW source if it is located at a favorable sky location.

More importantly, this emphasizes that the Bayes factor is not a “one size fits all” number to quantify the evidence for a signal, nor is there an absolute threshold between detection and non-detection. In actual CW searches (and especially targeted searches), the resulting Bayes factor should be interpreted in the appropriate context, for example by performing a false alarm analysis similar to this one. Finally, the “detection” of a CW signal is only part of the story. EM observations carry with them the parameters of the possible binary, which should be used for cross validation with GW searches in order to further reduce possible false alarms. In the next section, we will investigate the ability of the PTA to estimate, and independently verify, binary parameters in targeted searches.

3.3. Parameter Estimation

In this section, we focus on the ability of the PTA to estimate binary parameters, and the effects of an EM-informed targeted search on parameter estimation or limits at various sky locations and GW frequencies. We discuss the results separately for each injected signal (S-A to S-D), which are the posteriors we obtained for the parameters $\{M_{\text{tot}}, q, f_{\text{gw}}, i, \Phi_0, \psi\}$ from the MCMC analysis. The posterior distributions are shown in Figures 3–10. Note that the x -axis plotting ranges and histogram bin sizes for M_{tot} and f_{gw} vary between simulations S-A–S-D for presentation and are kept fixed between realizations (e.g., S-A1 and S-A2) for visual comparison.

To quantify the improvements in parameter estimation, we compute the width of the 90% percent of the posterior distribution (90% confidence interval, CI), which directly measures the precision of the parameter estimation. We also compute the Kullback–Leibler (KL) divergence (Kullback & Leibler 1951), which quantifies the difference between two probability distributions, and is given by

$$D_{\text{KL}} = \sum_x p(x) \ln \left[\frac{p(x)}{q(x)} \right], \quad (2)$$

where here $p(x)$ and $q(x)$ are the posterior and prior probability distributions of the parameter x . Therefore the KL divergence measures the difference between the prior used for a given parameter in a given search and the resultant posterior. A small value of D_{KL} indicates that the posterior is indistinguishable from the prior, whereas a larger value suggests that the parameter estimation is improved relative to the prior (recall that the parameters have different EM priors in the targeted searches). In Tables 6 and 7, we show the 90% CI and D_{KL} for each parameter from each search. A more detailed discussion of the KL divergence can be found in Appendix B. We note that we calculate the KL divergence for each parameter in terms of the marginalized prior and posterior distributions, instead of a single N -dimensional KL divergence for all parameters. We also note that here we compute the KL divergence using the summation form (as opposed to integral form), because the posterior distributions cannot be fit to simple, known functions.

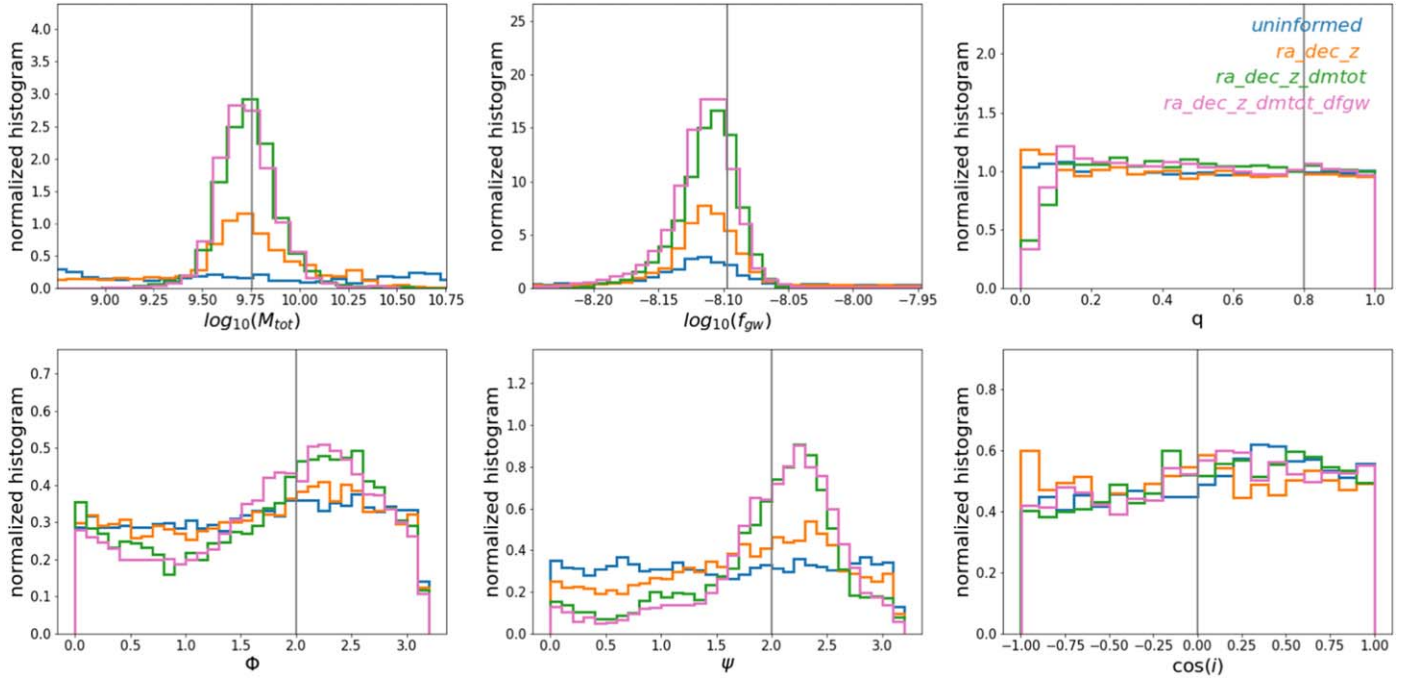


Figure 3. An uninformed search and three types of targeted searches are performed for Simulation A1 in order to obtain the posterior distributions for $\{M_{\text{tot}}, f_{\text{gw}}, q, \Phi, \Psi, i\}$. The blue, orange, green, and purple colors correspond to *uninformed*, *ra_dec_z*, *ra_dec_dmtot*, and *ra_dec_dmtot_dfgw*, respectively (see Table 5 and text for details). The areas under the histograms have been normalized to one. The injected value is marked with a gray solid line. The injected CW signal in S-A1 is located at a favorable sky location and has a f_{gw} of 8 nHz.

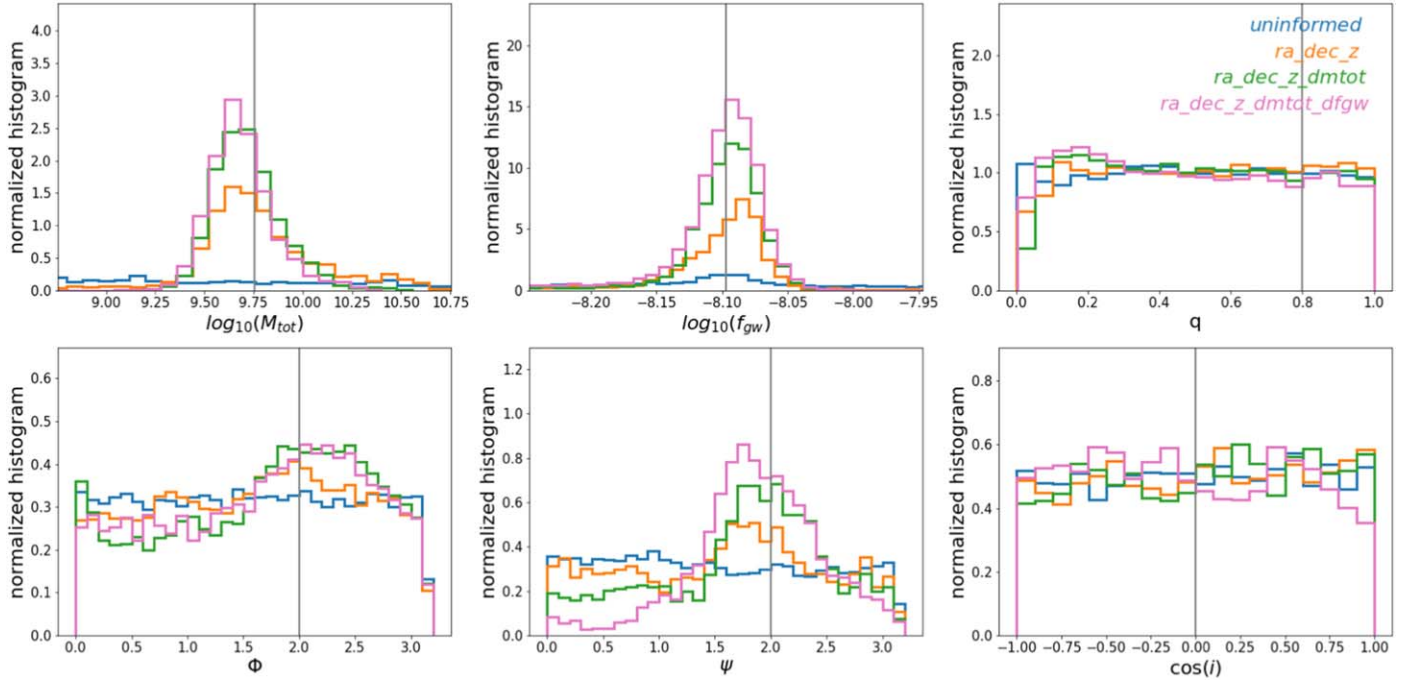


Figure 4. Same as Figure 3, but for S-A2.

Therefore, we choose the discrete form, to which the histograms of the priors and posteriors can be directly applied.

Simulation A (Figures 3 and 4): S-A is located at a sensitive sky location and a sensitive GW frequency. In *uninformed*, evidence for a CW signal at the injected f_{gw} is very weak, with all parameters essentially unconstrained (blue histograms). Searching

for the source at its injected sky location and distance (*ra_dec_z*) assists the algorithm in locating the correct M_{tot} (which manifests as a small peak in the posterior distribution, see the orange histogram in the top-left panel). It has also improved the estimation of f_{gw} (orange histogram in top center), even though both parameters are being searched over their respective full prior

Table 6
90% Confidence Interval and Kullback–Leibler Divergence in Uninformed and Targeted Searches (S-X1)

Simulation A1 (90%CI/ D_{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.524/0.021
$\cos \theta$	1.801/0.014
$\log d_L$	5.253/0.087
$\log M_{\text{tot}}$	4.480/0.185	3.867/0.575	0.504/0.313	0.460/0.366
$\log f_{\text{gw}}$	1.776/0.135	1.666/0.616	0.574/2.162	0.103/1.555
q	0.902/0.002	0.904/0.003	0.855/0.018	0.850/0.028
Φ_0	2.827/0.005	2.819/0.010	2.822/0.047	2.767/0.049
ψ	2.847/0.004	2.741/0.046	2.460/0.279	2.171/0.343
$\cos i$	1.794/0.011	1.817/0.006	1.769/0.014	1.785/0.017
Simulation B1 (90%CI/ D_{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.810/0.025
$\cos \theta$	1.825/0.015
$\log d_L$	5.249/0.038
$\log M_{\text{tot}}$	4.498/0.141	3.810/1.100	0.468/0.549	0.378/0.791
$\log f_{\text{gw}}$	1.822/0.175	1.796/1.352	1.644/2.741	0.014/3.188
q	0.909/0.005	0.887/0.038	0.823/0.056	0.869/0.035
Φ_0	2.856/0.002	2.617/0.098	2.401/0.173	2.158/0.289
ψ	2.803/0.005	2.607/0.180	1.903/0.477	1.168/0.754
$\cos i$	1.801/0.005	1.579/0.126	1.267/0.253	0.980/0.456
Simulation C1 (90%CI/ D_{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.646/0.004
$\cos \theta$	1.791/0.003
$\log d_L$	5.220/0.051
$\log M_{\text{tot}}$	3.801/0.340	3.512/0.393	0.797/0.318	0.741/0.411
$\log f_{\text{gw}}$	1.804/0.012	1.796/0.015	1.794/0.428	0.486/0.225
q	0.907/0.002	0.904/0.002	0.909/0.048	0.910/0.077
Φ_0	2.829/0.001	2.838/0.001	2.797/0.007	2.786/0.007
ψ	2.812/0.003	2.833/0.003	2.829/0.020	2.700/0.107
$\cos i$	1.795/0.005	1.774/0.005	1.652/0.056	1.597/0.106
Simulation D1 (90%CI/ D_{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.821/0.014
$\cos \theta$	1.800/0.008
$\log d_L$	5.384/0.042
$\log M_{\text{tot}}$	4.080/0.267	3.742/0.301	0.841/0.120	0.609/0.317
$\log f_{\text{gw}}$	1.835/0.047	1.815/0.030	1.824/0.355	0.462/0.765
q	0.910/0.002	0.905/0.001	0.905/0.011	0.900/0.006
Φ_0	2.821/0.001	2.817/0.001	2.841/0.001	2.806/0.009
ψ	2.815/0.004	2.820/0.002	2.802/0.011	2.709/0.097
$\cos i$	1.792/0.004	1.784/0.003	1.753/0.015	1.624/0.121

ranges. Applying prior information on M_{tot} in *ra_dec_z_dmtot* slightly improves the actual estimation of the parameter (green histogram in top left), but improves the estimation of f_{gw} more significantly (green histogram in top center); see their respective D_{KL} values in Tables 6 and 7. The improvement on f_{gw} is worth noting, since the parameter is still being searched over the wide prior range at this stage. Including the additional prior on f_{gw} in *ra_dec_z_dmtot_dfgw* does not further improve the estimation of M_{tot} or f_{gw} significantly, as the green and purple histograms in the

top-left and top-center panels, respectively, essentially converge and the D_{KL} value largely plateaus. Note that the estimation of f_{gw} greatly exceed the level of accuracy provided by the EM prior, which is indicated by the large final KL divergence (recall that the prior on f_{gw} is log-uniform with a factor of four uncertainty on f_{gw} on the linear scale). However, the mass ratio q remains unconstrained in all searches (top right).

Additionally, we observe an overall improvement in the estimation of ψ (bottom center), especially in *ra_dec_z_dmtot*

Table 7
Same as Table 6, but for S-X2

Simulation A2 (90%CI/D _{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.559/0.016
$\cos \theta$	1.793/0.006
$\log d_L$	5.287/0.048
$\log M_{\text{tot}}$	4.229/0.227	3.557/1.031	0.565/0.262	0.541/0.320
$\log f_{\text{gw}}$	1.780/0.047	1.348/0.726	0.581/1.656	0.290/1.256
q	0.903/0.002	0.878/0.007	0.866/0.022	0.886/0.011
Φ_0	2.838/0.001	2.767/0.009	2.824/0.034	2.779/0.024
ψ	2.850/0.006	2.786/0.039	2.605/0.140	1.927/0.347
$\cos i$	1.808/0.004	1.807/0.013	1.790/0.010	1.762/0.013
Simulation B2 (90%CI/D _{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.563/0.005
$\cos \theta$	1.797/0.004
$\log d_L$	5.289/0.044
$\log M_{\text{tot}}$	4.033/0.312	0.330/2.911	0.193/1.399	0.264/1.129
$\log f_{\text{gw}}$	1.817/0.023	0.010/3.894	0.010/3.910	0.010/3.915
q	0.906/0.003	0.860/0.086	0.776/0.116	0.826/0.061
Φ_0	2.817/0.001	1.996/0.332	2.142/0.318	1.901/0.368
ψ	2.826/0.003	0.978/1.007	1.115/0.927	0.968/0.992
$\cos i$	1.797/0.005	1.175/0.368	1.189/0.354	1.173/0.378
Simulation C2 (90%CI/D _{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.591/0.012
$\cos \theta$	1.788/0.005
$\log d_L$	5.173/0.061
$\log M_{\text{tot}}$	3.955/0.295	3.686/0.436	0.661/0.288	0.619/0.376
$\log f_{\text{gw}}$	1.754/0.061	1.737/0.189	1.275/1.029	0.406/0.533
q	0.905/0.002	0.903/0.004	0.897/0.009	0.892/0.015
Φ_0	2.835/0.003	2.794/0.010	2.859/0.014	2.862/0.012
ψ	2.854/0.004	2.860/0.034	2.929/0.110	3.001/0.200
$\cos i$	1.789/0.008	1.744/0.039	1.570/0.184	1.478/0.209
Simulation D2 (90%CI/D _{KL})				
	<i>uninformed</i>	<i>ra_dec_z</i>	<i>ra_dec_z_dmtot</i>	<i>ra_dec_z_dmtot_dfgw</i>
ϕ	5.797/0.038
$\cos \theta$	1.822/0.033
$\log d_L$	5.057/0.110
$\log M_{\text{tot}}$	4.103/0.282	4.406/0.204	0.927/0.085	0.567/0.379
$\log f_{\text{gw}}$	1.796/0.147	1.853/0.069	1.822/0.377	0.494/0.674
q	0.902/0.003	0.901/0.002	0.888/0.008	0.893/0.011
Φ_0	2.849/0.008	2.812/0.002	2.818/0.005	2.853/0.006
ψ	2.856/0.018	2.837/0.022	2.780/0.020	2.256/0.142
$\cos i$	1.781/0.021	1.782/0.007	1.781/0.029	1.754/0.017

and *ra_dec_z_dmtot_dfgw* (green and purple histograms), even though we did not apply any priors on the parameter. It is therefore noteworthy that EM-informed priors on a subset of parameters have improved the estimation of the other parameters for which we have no EM information. However, Φ_0 and $\cos i$ remain essentially unconstrained.

Simulation B (Figures 5 and 6): S-B is located at the same sensitive sky location as S-A, but emitting at a higher GW frequency. All parameters are again unconstrained in

uninformed (blue histograms). In the targeted searches, we observe a similar trend of enhanced parameter estimation for all parameters. Furthermore, for each parameter, the posterior distributions are in good agreement with each other and peak near the injected value (orange, green, and purple histograms), except for q , on which there are only lower limits. We note that these improvements are achieved despite the fact that the source is quieter than S-A (in terms of the \mathcal{F}_p statistic or SNR). A comparison of their KL divergence values shows that

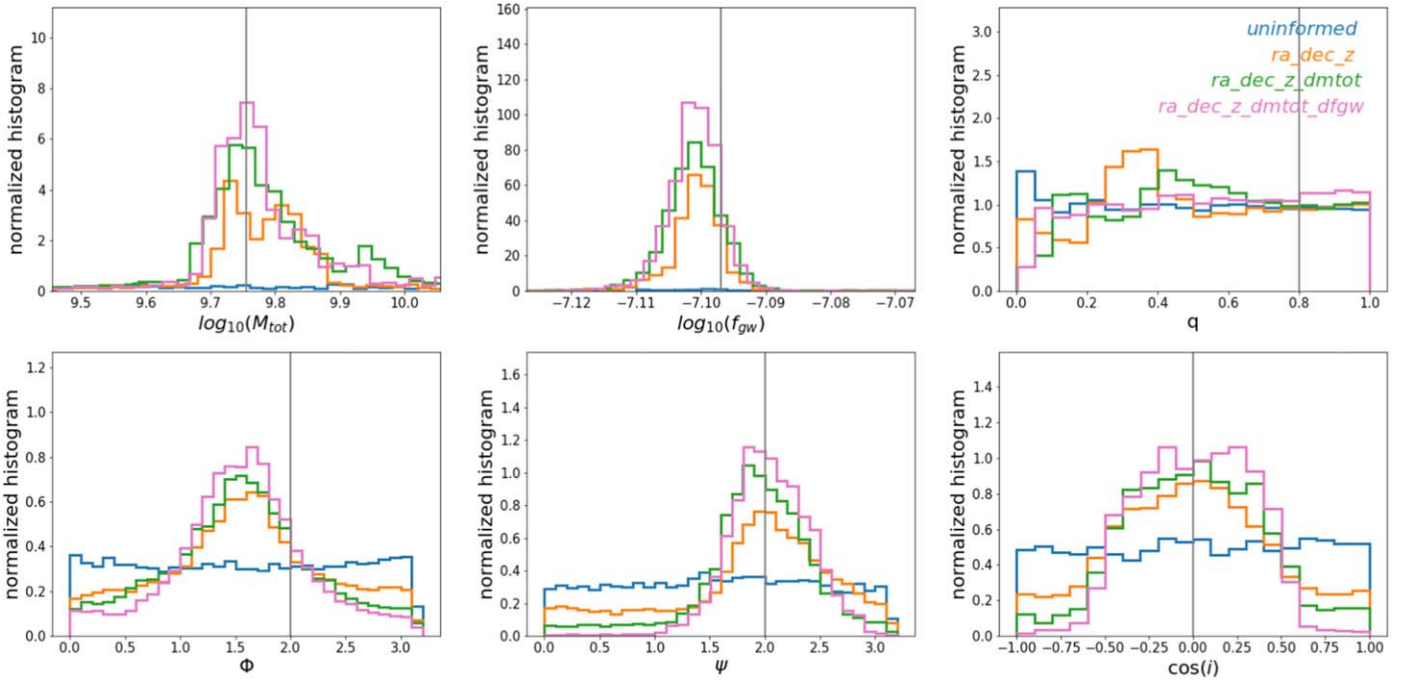


Figure 5. Same as Figure 3, but for Simulation B1. The injected CW signal in S-B1 is located at a favorable sky location and has a f_{gw} of 80 nHz.

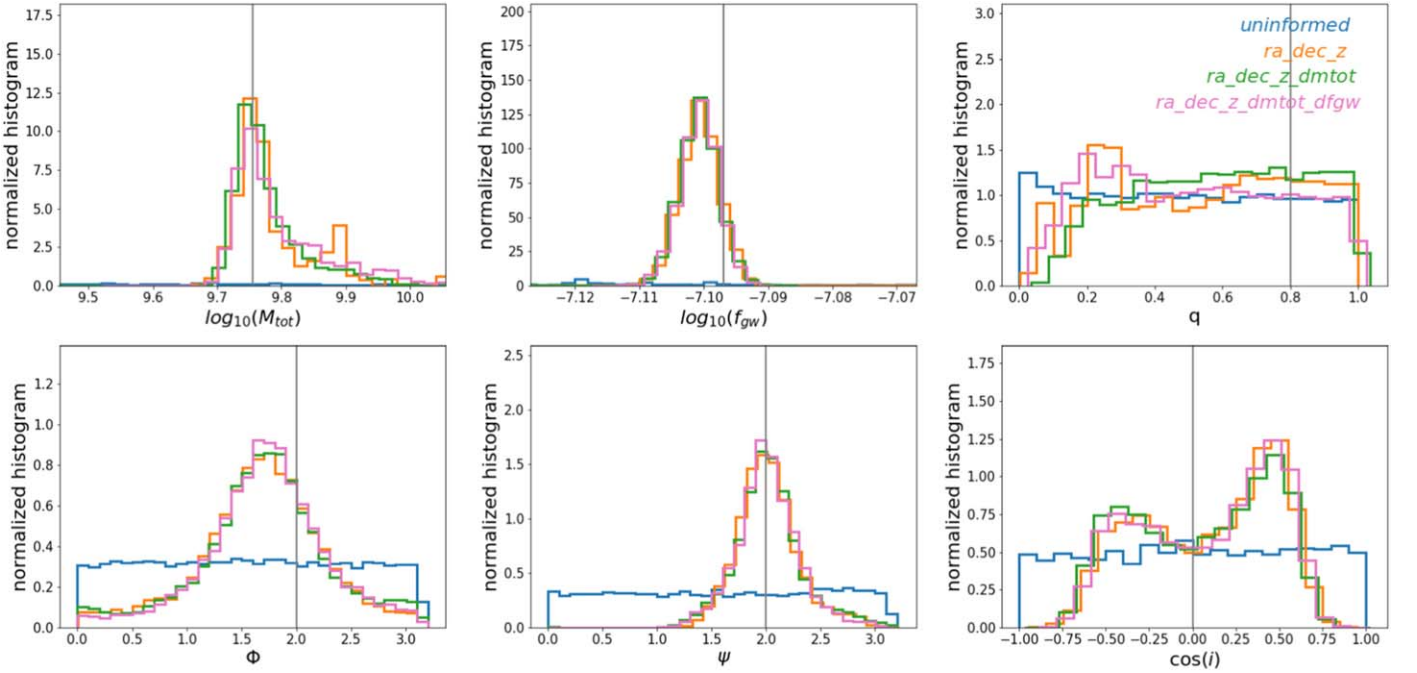


Figure 6. Same as Figure 5, but for S-B2.

parameters in S-B are overall better constrained than S-A, especially the f_{gw} parameter, which has the highest KL divergence. This echoes the better detectability of S-B compared to S-A (Section 3.1). Again this is likely due to the higher frequency of the S-B signal, which appears to be beneficial not only for detecting the signal, but the parameter estimation as well.

Simulation C (Figures 7–8) and Simulation D (Figures 9–10): S-C and S-D are the counterparts of S-A and S-B, respectively, at a less sensitive sky location. While the parameter estimation shares similar overall trends as S-A and S-B, we also observe some differences, which are mainly (1) the improvements are less pronounced overall and (2) S-C1 is the only simulation where the

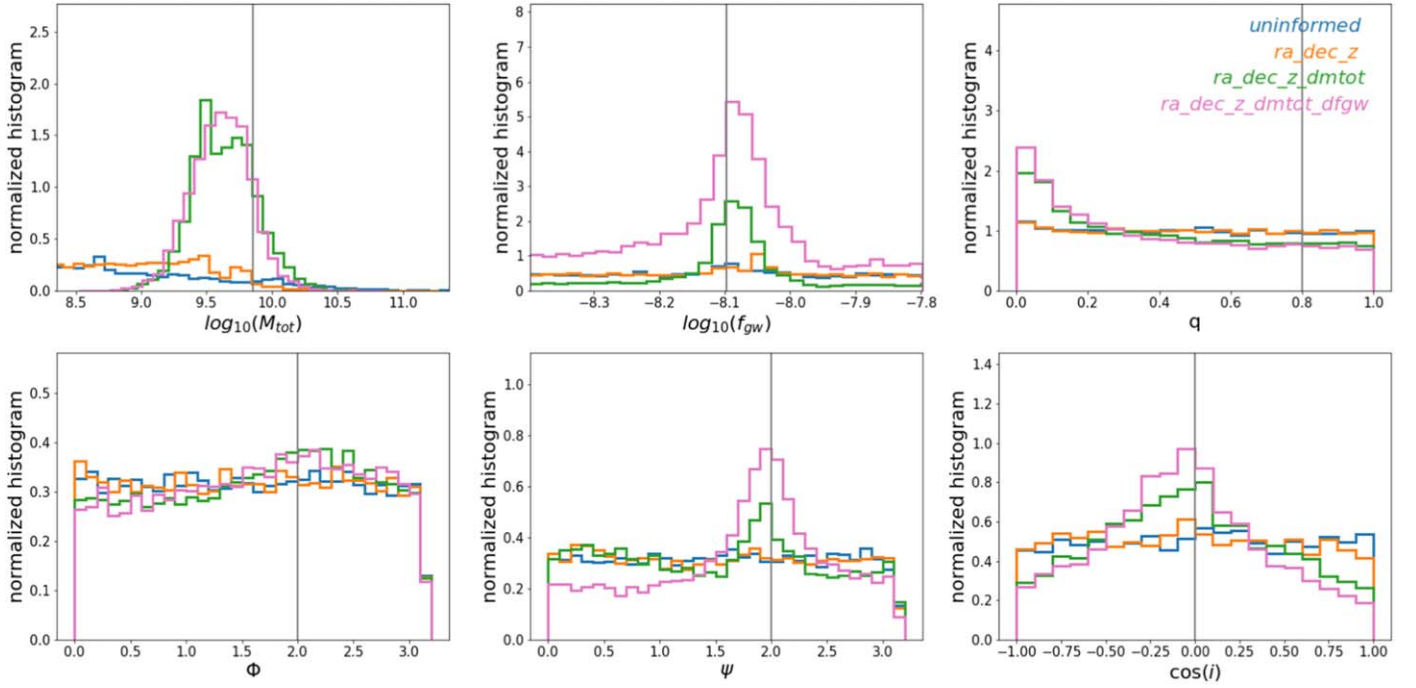


Figure 7. Same as Figure 3, but for Simulation C1. The injected CW signal in S-C1 is located at a less favorable sky location and has a f_{gw} of 8 nHz.

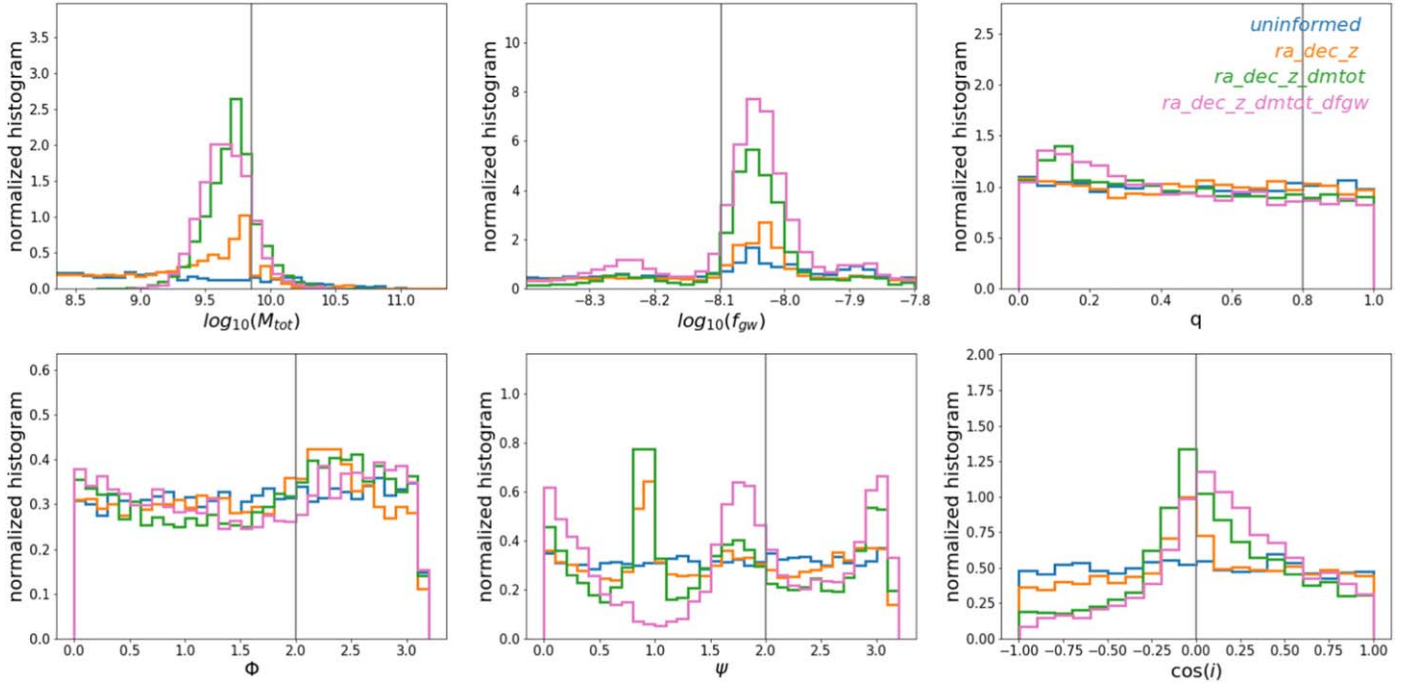


Figure 8. Same as Figure 7, but for S-C2.

EM prior on M_{tot} allows the algorithm to place a lower limit on q (green and purple histograms in the top-right panel of Figure 7).

4. Implications for the Multi-messenger Studies of SMBHBs

PTA searches for GWs from SMBHBs benefit considerably from EM priors, and the resulting posteriors are obtained in a fundamentally different way. This therefore has several interesting applications in the multi-messenger searches for SMBHBs (e.g., through confirming EM-selected SMBHB candidates) and the

studies of these physical systems (e.g., through first estimating the source parameters).

For a weak CW signal, which is likely the case for the first detection, knowing the source's location, redshift, and black hole mass through the observations of its EM counterpart significantly enhances the algorithm's ability to detect the source and estimate its parameters, which are otherwise unobtainable in an uninformed search. Therefore, using EM priors in GW searches could expedite the time-to-detection of individual sources. It remains to be seen, however, whether the

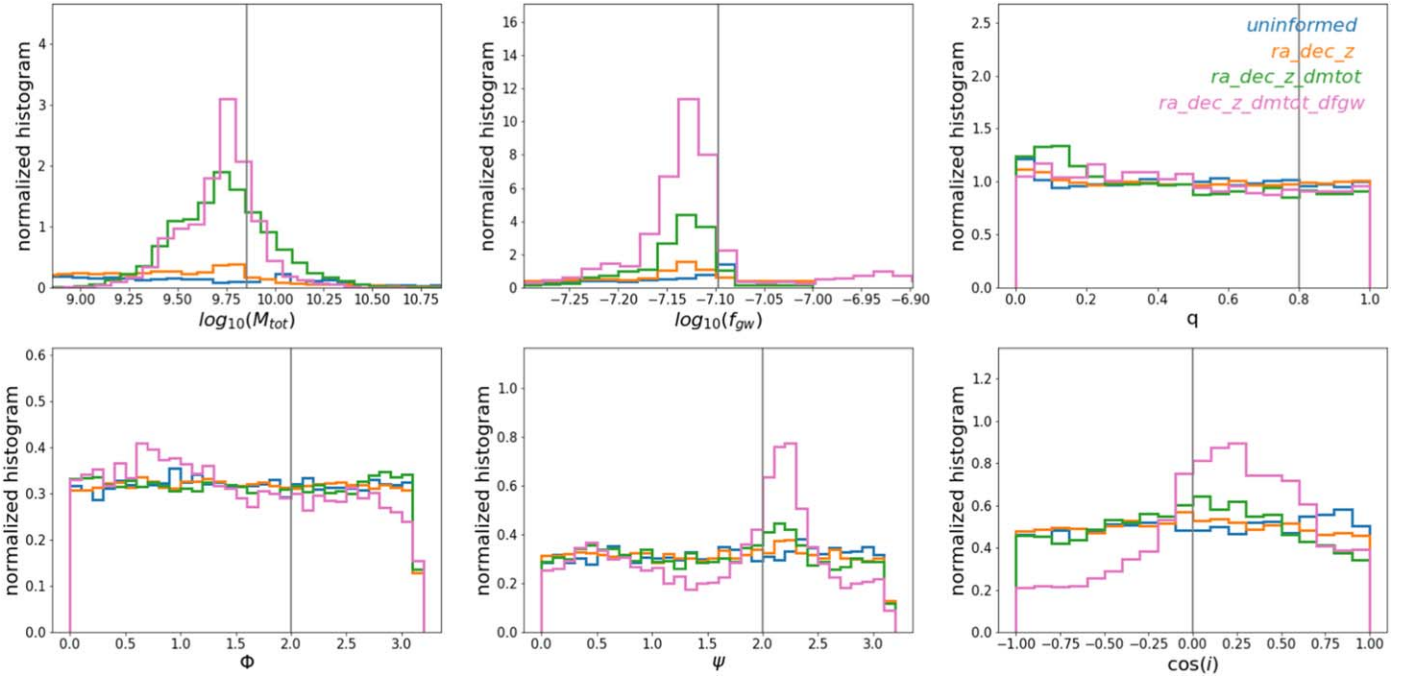


Figure 9. Same as Figure 3, but for Simulation D1. The injected CW signal in S-D1 is located at a less favorable sky location and has a f_{gw} of 80 nHz.

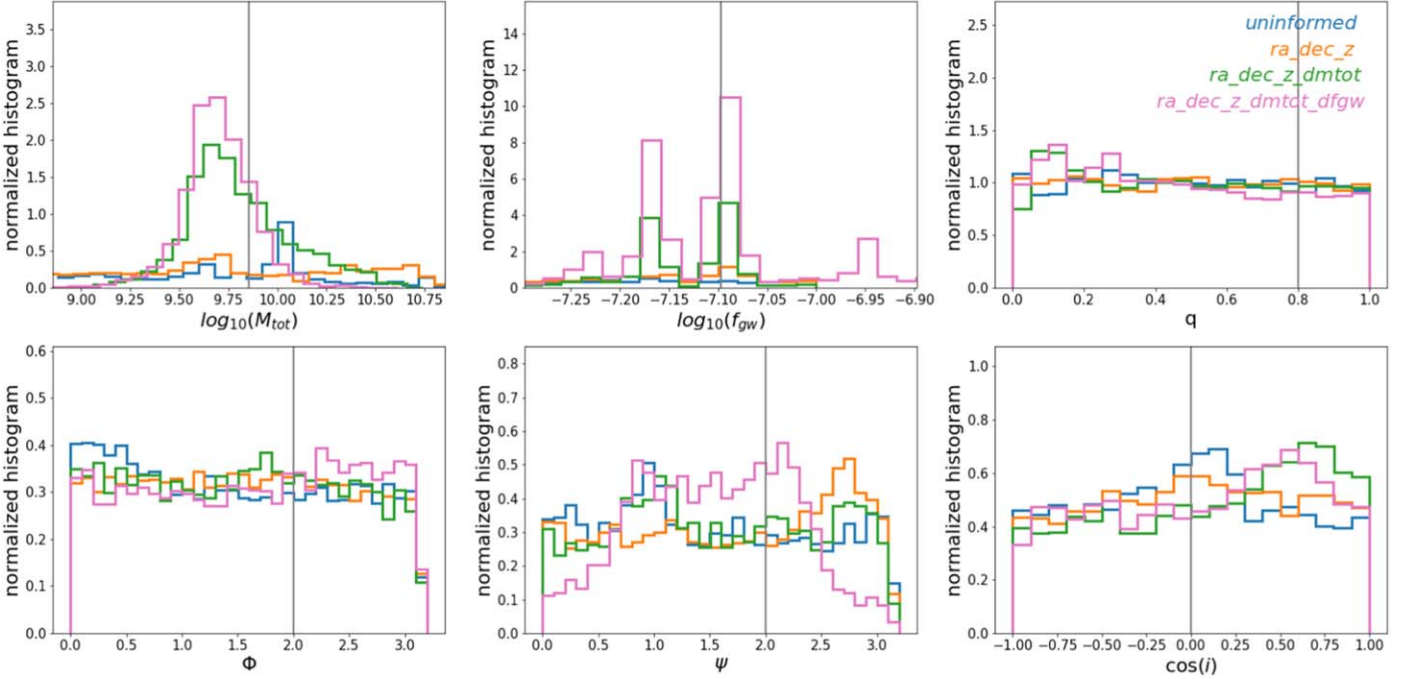


Figure 10. Same as Figure 9, but for S-D2.

same level of improvement can manifest for stronger GW signals.

Our analysis also suggests that the search for SMBHBs will benefit from this improvement regardless of their GW frequencies; there is also an indication that sources at higher frequencies may be more detectable and better characterized than those at lower frequencies. This will be a boon to the synergy and coordination between EM observers and the PTA community in several ways: (1) the time-domain search for GW-emitting SMBHBs typically requires at least a few cycles of quasi-periodicity in order to claim

a possible binary origin and hence is more sensitive to SMBHBs with short periods (i.e., months to years timescales), or those that are emitting GWs at high frequencies (i.e., tens of nanohertz); (2) binary evolution and population estimates predict that long-period SMBHBs are more common (e.g., Kelley et al. 2019; Krolik et al. 2019), thus in general, the observational search for long-period SMBHBs is expected to be more fruitful. Since the improvement in parameter estimation is not sensitive to the GW frequency of the source, as suggested by our results, the multi-messenger search for SMBHBs is promising across the entire PTA band.

By contrast, the improvement level does have a sky location dependence and is less pronounced if the source is located at a less favorable sky location, even though its nominal GW amplitude h_0 is larger. Therefore, the current NANOGrav PTA will have difficulty detecting and characterizing SMBHBs if they happen to be located where few pulsars are being monitored. Coincidentally, many of the current SMBHB candidates, including OJ 287, are located at such sky locations. Fortunately, as more pulsars across the sky are being added to the PTA, either as a regional PTA or under the umbrella of the International Pulsar Timing Array (IPTA; Hobbs et al. 2010), this weakness will be overcome, PTA's synergies with EM facilities will strengthen, and there is great potential that it can operate as a truly all-sky nanohertz GW observatory.

The EM localization of an SMBHB (or candidate) usually implies that its black hole mass can also be measured. This additional EM prior, even with a considerable uncertainty, can enhance the CW search with the PTA even further. The most astrophysically interesting result is the ability to measure the GW frequency f_{gw} with PTA data at this stage, when it is otherwise completely unknown from EM observations alone. For a circular binary, which is assumed throughout the paper, this directly informs the orbital frequency $f_{\text{gw}} = 2f_{\text{orb}}$. If a quasi-periodic variation of the AGN EM flux is observed in the SMBHB system, where the observed quasi-period (or frequency⁶ f_{var}) is interpreted as the electromagnetic imprint of the orbital period, then a direct comparison between f_{orb} and f_{var} is possible. For example, in the scenario that relativistic Doppler boost can be the cause of periodicity in an SMBHB (D'Orazio et al. 2015), which observationally favors an unequal mass ratio, the model has the clear prediction that $f_{\text{var}} = f_{\text{orb}}$. Through time-series analysis methods, f_{var} (or $f_{\text{orb,EM}}$) can be determined to good accuracy, which can then be directly compared with $f_{\text{orb,GW}}$. Through EM light-curve fitting, one additionally obtains an estimate of the mass ratio (albeit with large uncertainties), which can be compared with the lower limit on mass ratio that the PTA may already place (e.g., Figure 7 top-right panel) as additional evidence for the presence of a binary.

An alternative cause of periodicity in an SMBHB is the formation of a lump in the inner circumbinary disk which orbits at the local Keplerian rate f_{lump} , interacts with passing BHs, and modulates the accretion of material at the beat frequency (e.g., Farris et al. 2014; Noble et al. 2021). Since this scenario can produce large-amplitude variations without strongly favoring low mass ratios, high black hole masses, or high inclinations, it is expected to be observationally more common (Kelley et al. 2019). However, translating from the observed f_{var} to the intrinsic f_{orb} is less straightforward than the previous scenario, as it depends on the binary mass ratio. Recent magneto-hydrodynamic simulations of varying mass ratios by Noble et al. (2021) confirm the previous, hydrodynamic-only results that only for $q \gtrsim 0.2$ can the binary's gravitational torque be strong enough to excite a lump in the circumbinary disk, which is at a radius of a few times the binary separation. For these intermediate mass ratios, the less massive, secondary BH predominantly interacts with the lump, therefore producing periodicity at $f_{\text{var}} \approx f_{\text{orb}}$. For an equal mass binary, the two BHs equally interact with the lump at the beat frequency $f_{\text{orb}} - f_{\text{lump}}$; the overall periodicity is thus at the frequency $f_{\text{var}} = 2(f_{\text{orb}} - f_{\text{lump}}) \approx 1.5f_{\text{orb}}$. The EM-observed f_{var}

and the independent GW measurement of the intrinsic f_{orb} (through f_{gw}) would allow one to connect the two timescales, which informs a crude estimate of the mass ratio (i.e., equal mass ratio q or small q) at no additional cost. Note that, again, this is independent from the mass-ratio parameter estimation or limit obtained from the CW search. Even if it is a lower limit in the weak CW signal regime, it nevertheless has a considerable cross-validation value.

Finally, the overall ability of the PTA to cross-validate binary parameters, which is an integral part of confirming an EM-selected binary candidate, is excellent. Depending on the source's sky location and GW frequency, the PTA can estimate the parameters at a level which at least exceeds EM alone (which may only be a factor of a few), and at the few percent level in some parameters. This suggests that even weak CW signals, which are the expected first signals and the majority of most CWs within the reach of a future PTA, have considerable multi-messenger science values. These are realistic but conservative assessment of the parameter estimation ability of PTAs, since the precision level for signals with larger SNRs is expected to be higher, as the SNR of the signal increases with a longer data span. Future work using projected PTA data sets can verify this predication.

5. Summary and Conclusions

In this work, we investigated the ability of the PTA to perform CW searches in the most realistic set-up yet, focusing on SMBHB parameter estimation in an uninformed search with GW data only and EM-informed searches in a multi-messenger scheme. To this end, we first constructed a simulated PTA data set that resembles the 11 yr NANOGrav data set, which takes into account both white noise and red noise of the pulsars as well as their non-uniform timing cadence and sky locations. We injected into the PTA data set CW signals of modest amplitudes as proxies for the first CW signals in a future PTA data set, which are likely to be weak and only marginally detectable. In contrast to previous work that studied CW detection or parameter estimation (e.g., Sesana & Vecchio 2010; Ellis et al. 2012), the injected CW signal contains the full signal of the pulsar term, which allows us to include the astrophysically interesting chirp mass (or equivalently, total black hole mass) parameter in our analysis. We also vary the GW frequency and sky location of the CW signal, in order to study the effects they have on CW searches.

We then attempted to recover the injected signals in a procedure that mimics the search for a CW source in PTA data, using the Bayesian MCMC data analysis algorithm for CW detection. This approach thus gives a more realistic estimate for the precision and accuracy of binary parameter estimation than the lower limit estimates given by the Fisher information matrix approach (Sesana & Vecchio 2010). Our main results can be summarized as the following:

1. In the weak signal regime, which is likely the case for the first CW signals emerging in PTA data, GW data alone struggle to recover the CW signal and its parameters.
2. Searching for the source using EM priors on its sky location, distance, and black hole mass can increase the Bayes factor by a factor of a few up to an order of magnitude if the source is located at a favorable sky location where more pulsars are being monitored. However, we do not observe the same effect if the

⁶ The following discussion assumes f_{var} is in the rest frame.

source is located at a sky location with few pulsars. There are also hints that sources emitting at higher frequencies may be more detectable than those at lower frequencies.

3. By searching for the source at its sky location and distance through the observations of its EM counterpart, we are able to identify the correct parameter values, especially the astrophysically interesting M_{tot} and f_{gw} , if the source is located at a favorable sky location. However, this improvement in binary parameter estimation does not manifest for a source located at a less favorable sky location. Additionally, similar to source detection, the parameter estimation of sources at higher GW frequencies also appears to be better.
4. Further including an EM prior on M_{tot} in the GW search results in the estimation of the parameter that is better than that from EM alone. Additionally, a prior on M_{tot} improves the constraints on the GW frequency, which is otherwise completely unknown from EM observations. Both effects are seen in Simulations A-D, i.e., for both sensitive and less sensitive sky locations and frequencies, although the effects are stronger at the sensitive sky location.
5. If f_{gw} can be estimated within a factor of a few from EM observations, such as through analyzing the source's EM light curve, it can improve the estimation of all parameters even further. It may also permit the extraction of additional information about the system (such as an estimate of the mass ratio), if EM and GW data are interpreted jointly. This is also observed in Simulations A-D, but is more pronounced at the sensitive sky location.
6. Overall, incorporating *any* EM priors on a subset of binary parameters improves the estimation of all or almost all parameters, regardless of sky location and GW frequency.
7. Overall, combining EM and GW data results in parameter estimation at a precision level which is better than the two separately and at a level that permits the cross validation of binary parameters.

These results highlight the complementarity of EM and GW data in searching for SMBHBs, and the feasibility of confirming an SMBHB candidate (by detecting the source and verifying its source parameters) and understanding the physical system in a multi-messenger approach. But here we have only scratched the surface of the potential of combining EM and GW data to advance our study of SMBHBs. Possible future work includes extending to higher SNR signals and applying the parameters of a more astrophysically motivated sample of SMBHBs. These inquiries will together form a more complete picture of the PTA as a GW detector and observatory of SMBHBs and give realistic estimates about its ability to study SMBHBs in a multi-messenger approach.

T.L. thanks members of the NANOGrav Detection and Astrophysics working groups, especially Bence B csy, Steve Taylor, and Caitlin Witt, for helpful discussions. T.L. thanks Luke Kelley for a thorough reading of the manuscript and useful comments. T.L. acknowledges the use of the Geweke diagnostic which is implemented by Nihan Pol. We thank the referee for helpful comments that improved this manuscript.

This work is supported by the NANOGrav National Science Foundation Physics Frontiers Center award no. 1430284. S.J.V. is supported by NSF award no. 2011772. This work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611.

The participation of T.L. at the Aspen Center for Physics was supported by the Simons Foundation.

Appendix A

Bayes Factors of Uninformed and Targeted Searches (S-A-S-D)

We computed the Bayes factors shown in Figure 2 using the Savage–Dickey (SD) formula (Dickey 1971), which can be used to compute the Bayes factor for two nested models. In this case, the two models are a CW signal-plus-noise model \mathcal{H}_1 and a noise-only model \mathcal{H}_0 , which is equivalent to model \mathcal{H}_1 with $h_0 = 0$. The Bayes factor \mathcal{B}_{10} is given by

$$\mathcal{B}_{10} \equiv \frac{\text{evidence}[\mathcal{H}_1]}{\text{evidence}[\mathcal{H}_0]} = \frac{p(h_0 = 0|\mathcal{H}_1)}{p(h_0 = 0|\mathcal{D}, \mathcal{H}_1)}, \quad (\text{A1})$$

where $p(h_0 = 0|\mathcal{H}_1)$ is the prior probability density of $h_0 = 0$ in the embedding model \mathcal{H}_1 , and $p(h_0 = 0|\mathcal{D}, \mathcal{H}_1)$ is the posterior probability density of $h_0 = 0$ in the embedding model \mathcal{H}_1 .

In order to use the SD formula, we first computed the prior and posterior distributions of h_0 using the distributions of M_{tot} , q , f_{gw} , and d_L for each mock search in each simulated data set, using the relationship

$$\log_{10} \mathcal{M} = \frac{3}{5} \log_{10} \left[\frac{q}{(1+q)^2} \right] + \log_{10} M_{\text{tot}}, \quad (\text{A2})$$

$$h_0 = \frac{2\mathcal{M}^{5/3}(\pi f_{\text{gw}})^{2/3}}{d_L}. \quad (\text{A3})$$

The Bayes factor is then computed by taking the ratio of the numbers of samples below some threshold between the prior and the posterior.

We adopt the threshold in terms of the percentile of samples in the posterior. We then use the h_0 value corresponding to this percentile to determine the number of samples below this threshold in the prior. The reason is two-fold: first, it guarantees a sufficient number of samples below the threshold and is more robust against imperfect sampling, and second, it is calculated in a self-consistent manner for all simulations and realizations, while still preserving the principle behind the SD approximation. We then repeat the procedure by varying the threshold between 1% and 5%, both of which amplitudes are sufficiently

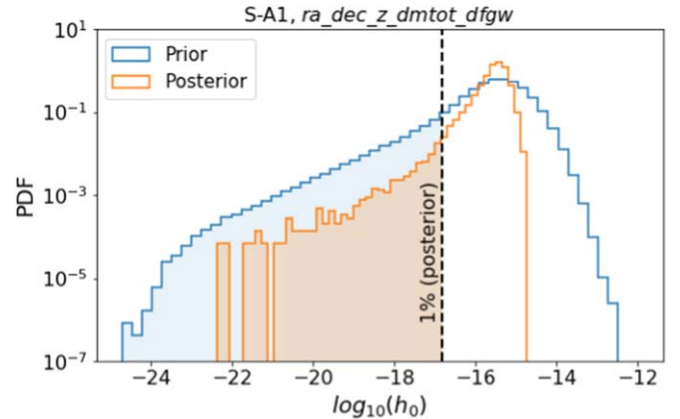


Figure A1. The prior and posterior distributions of h_0 for Simulation A1, $ra_dec_z_dmtot_dfgw$ (blue and orange histograms, respectively). The black dashed line marks the one percentile of samples in the posterior, which is used as the threshold to compute the SD ratio, which is the ratio between the blue- and orange-shaded areas.

low for the SD approximation to be valid. The uncertainty of \mathcal{B}_{10} is computed from the variance in \mathcal{B}_{10} as the threshold is varied. For illustration purposes, Figure A1 shows the prior and posterior distributions for h_0 for S-A1, $ra_dec_z_dmtot_dfgw$. The SD ratio is then the ratio between the blue and orange-shaded areas.

Appendix B KL Divergence

The KL divergence measures the difference between two probability distributions:

$$D_{\text{KL}}(P||Q) = \sum_x P(x) \log \left[\frac{P(x)}{Q(x)} \right], \quad (\text{B1})$$

where $P(x)$ and $Q(x)$ are discrete probability distributions. In Section 3, we take $P(x)$ to be the posterior probability density function, and $Q(x)$ to be the prior probability density function. Then the KL divergence can be interpreted as the information gained from the posterior compared to the prior.

In order to aid in understanding the significance of the values of the KL divergence in Tables 6 and 7, here we work through a toy example. Suppose we are performing parameter estimation for a one-parameter model and obtain a Gaussian posterior

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right], \quad (\text{B2})$$

where μ is the mean and σ is the variance. Assume the prior for this parameter is uniform over the range $[\mu - a, \mu + a]$:

$$q(x) = \begin{cases} \frac{1}{2a} & \mu - a \leq x \leq \mu + a \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B3})$$

For continuous probability distributions, the KL divergence is given by

$$D_{\text{KL}}(P||Q) = \int_{-\infty}^{\infty} dx p(x) \log \left[\frac{p(x)}{q(x)} \right]. \quad (\text{B4})$$

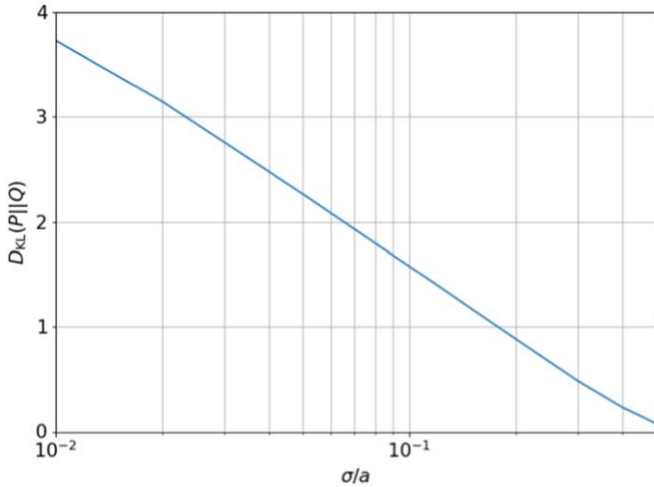


Figure B1. KL divergence $D_{\text{KL}}(P||Q)$ for a posterior p and uniform prior q . The posterior is assumed to be Gaussian with mean μ and variance σ , while the prior is assumed to be uniform centered around μ with width $2a$. The KL divergence measures the additional information in the posterior compared to the prior: larger values correspond to a more constrained posterior.

For the posterior and prior given above, the KL divergence can be computed analytically:

$$D_{\text{KL}}(P||Q) = \frac{a}{\sqrt{2\pi}\sigma^2} \exp \left(-\frac{a^2}{2\sigma^2} \right) + \frac{1}{2} \operatorname{erf} \left(\frac{a}{\sqrt{2}\sigma} \right) \left[\log \left(\frac{2a^2}{\pi\sigma^2} \right) - 1 \right]. \quad (\text{B5})$$

Figure B1 shows the KL divergence as a function of the variance of the Gaussian posterior σ .

We emphasize that this is a toy problem, and the posteriors shown in Section 3 are not simple Gaussian distributions. However, these values may help the reader understand the significance of the values in Tables 6 and 7.

ORCID iDs

Tingting Liu <https://orcid.org/0000-0001-5766-4287>

Sarah J. Vigeland <https://orcid.org/0000-0003-4700-9072>

References

- Aasi, J., Abbott, B. P., Abbott, R., et al. 2015, *ApJ*, **813**, 39
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, *PhRvL*, **119**, 161101
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, *ApJ*, **875**, 122
- Aggarwal, K., Arzoumanian, Z., Baker, P. T., et al. 2019, *ApJ*, **880**, 116
- Aggarwal, K., Arzoumanian, Z., Baker, P. T., et al. 2020, *ApJ*, **889**, 38
- Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2021, *ApJS*, **252**, 4
- Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, arXiv:1702.00786
- Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2020a, *ApJL*, **905**, L34
- Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2018a, *ApJ*, **859**, 47
- Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2020b, *ApJ*, **900**, 102
- Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2021, *ApJ*, **914**, 121
- Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2018b, *ApJS*, **235**, 37
- Babak, S., Gair, J., Sesana, A., et al. 2017, *PhRvD*, **95**, 103012
- Babak, S., & Sesana, A. 2012, *PhRvD*, **85**, 044034
- Barnes, J. E., & Hernquist, L. 1992, *ARA&A*, **30**, 705
- Bécsy, B., & Cornish, N. J. 2020, *CQGrA*, **37**, 135011
- Bowen, D. B., Mewes, V., Campanelli, M., et al. 2018, *ApJL*, **853**, L17
- Charisi, M., Bartos, I., Haiman, Z., et al. 2016, *MNRAS*, **463**, 2145
- Chen, Y.-C., Liu, X., Liao, W.-T., et al. 2020, *MNRAS*, **499**, 2245
- Cornish, N. J., & Porter, E. K. 2007, *CQGrA*, **24**, 5729
- d'Ascoli, S., Noble, S. C., Bowen, D. B., et al. 2018, *ApJ*, **865**, 140
- Dickey, J. M. 1971, *The Annals of Mathematical Statistics*, **42**, 204
- D'Orazio, D. J., & Di Stefano, R. 2018, *MNRAS*, **474**, 2975
- D'Orazio, D. J., Haiman, Z., & MacFadyen, A. 2013, *MNRAS*, **436**, 2997
- D'Orazio, D. J., Haiman, Z., & Schiminovich, D. 2015, *Natur*, **525**, 351
- Earl, D. J., & Deem, M. W. 2005, *PCCP*, **7**, 3910
- Ellis, J., & van Haasteren, R. 2017, jellis18/PTMCMCSampler: Official Release, Zenodo, doi:10.5281/zenodo.1037579
- Ellis, J. A., Siemens, X., & Creighton, J. D. E. 2012, *ApJ*, **756**, 175
- Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2014, *ApJ*, **783**, 134
- Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2015, *MNRAS*, **446**, L36
- Geweke, J. 1992, In *Bayesian Statistics* (Oxford: Clarendon), 169
- Gold, R., Paschalidis, V., Etienne, Z. B., Shapiro, S. L., & Pfeiffer, H. P. 2014, *PhRvD*, **89**, 064060
- Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015a, *MNRAS*, **453**, 1562
- Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015b, *Natur*, **518**, 74
- Green, P. J. 1995, *Biometrika*, **82**, 711
- Gültekin, K., & Miller, J. M. 2012, *ApJ*, **761**, 90
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hobbs, G. 2013, *CQGrA*, **30**, 224007
- Hobbs, G., Archibald, A., Arzoumanian, Z., et al. 2010, *CQGrA*, **27**, 084013
- Hu, B. X., D'Orazio, D. J., Haiman, Z., et al. 2020, *MNRAS*, **495**, 4061
- Iguchi, S., Okuda, T., & Sudou, H. 2010, *ApJL*, **724**, L166
- Ivezic, Z., Tyson, J. A., Abel, B., et al. 2008, arXiv:0805.2366
- Janssen, G., Hobbs, G., McLaughlin, M., et al. 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 37 arXiv:1501.00127
- Jaranowski, P., Królak, A., & Schutz, B. F. 1998, *PhRvD*, **58**, 063001
- Jetten, F. A., Lommen, A., Larson, S. L., & Wen, L. 2004, *ApJ*, **606**, 799

- Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2018, *MNRAS*, **477**, 964
- Kelley, L. Z., Haiman, Z., Sesana, A., & Hernquist, L. 2019, *MNRAS*, **485**, 1579
- Klein, A., Barausse, E., Sesana, A., et al. 2016, *PhRvD*, **93**, 024003
- Kocsis, B., Haiman, Z., & Loeb, A. 2012, *MNRAS*, **427**, 2680
- Kramer, M., & Champion, D. J. 2013, *CQGra*, **30**, 224009
- Krolik, J. H., Volonteri, M., Dubois, Y., & Devriendt, J. 2019, *ApJ*, **879**, 110
- Kullback, S., & Leibler, R. A. 1951, *The Annals of Mathematical Statistics*, **22**, 79
- Lam, M. T., Cordes, J. M., Chatterjee, S., et al. 2017, *ApJ*, **834**, 35
- Lehto, H. J., & Valtonen, M. J. 1996, *ApJ*, **460**, 207
- Liu, F. K., Li, S., & Chen, X. 2009, *ApJL*, **706**, L133
- Liu, T., Gezari, S., Ayers, M., et al. 2019, *ApJ*, **884**, 36
- Liu, T., Gezari, S., Burgett, W., et al. 2016, *ApJ*, **833**, 6
- Liu, T., Gezari, S., Heinis, S., et al. 2015, *ApJL*, **803**, L16
- MacFadyen, A. I., & Milosavljević, M. 2008, *ApJ*, **672**, 83
- McKernan, B., Ford, K. E. S., Kocsis, B., & Haiman, Z. 2013, *MNRAS*, **432**, 1468
- McLaughlin, M. A. 2013, *CQGra*, **30**, 224008
- Mingarelli, C. M. F., Lazio, T. J. W., Sesana, A., et al. 2017, *NatAs*, **1**, 886
- Noble, S. C., Krolik, J. H., Campanelli, M., et al. 2021, arXiv:2103.12100
- Noble, S. C., Mundim, B. C., Nakano, H., et al. 2012, *ApJ*, **755**, 51
- Pol, N. S., Taylor, S. R., Kelley, L. Z., et al. 2021, *ApJ*, **911**, L34
- Ricarte, A., Natarajan, P., Dai, L., & Coppi, P. 2016, *MNRAS*, **458**, 1712
- Roedig, C., Krolik, J. H., & Miller, M. C. 2014, *ApJ*, **785**, 115
- Rosado, P. A., Sesana, A., & Gair, J. 2015, *MNRAS*, **451**, 2417
- Sesana, A., Haiman, Z., Kocsis, B., & Kelley, L. Z. 2018, *ApJ*, **856**, 42
- Sesana, A., & Vecchio, A. 2010, *PhRvD*, **81**, 104008
- Shen, Y. 2013, *BASI*, **41**, 61
- Shi, J.-M., Krolik, J. H., Lubow, S. H., & Hawley, J. F. 2012, *ApJ*, **749**, 118
- Shu, X., Zhang, W., Li, S., et al. 2020, *NatCo*, **11**, 5876
- Sillanpaa, A., Haarala, S., Valtonen, M. J., Sundelius, B., & Byrd, G. G. 1988, *ApJ*, **325**, 628
- Sudou, H., Iguchi, S., Murata, Y., & Taniguchi, Y. 2003, *Sci*, **300**, 1263
- Tanaka, T., Menou, K., & Haiman, Z. 2012, *MNRAS*, **420**, 705
- Tang, Y., Haiman, Z., & MacFadyen, A. 2018, *MNRAS*, **476**, 2249
- Taylor, S. R., Vallisneri, M., Ellis, J. A., et al. 2016, *ApJL*, **819**, L6
- Valtonen, M. J., Lehto, H. J., Nilsson, K., et al. 2008, *Natur*, **452**, 851
- Valtonen, M. J., Zola, S., Ciprini, S., et al. 2016, *ApJL*, **819**, L37
- Vaughan, S., Uttley, P., Markowitz, A. G., et al. 2016, *MNRAS*, **461**, 3145
- Xin, C., Mingarelli, C. M. F., & Hazboun, J. S. 2021, *ApJ*, **915**, 97