Astronomy & Astrophysics

Prospects for future binary black hole gravitational wave studies in light of PTA measurements

John Ellis[1,2,3], Malcolm Fairbairn[2], Gert Hütsi[3], Martti Raidal[3], Juan Urrutia[3,4],
Ville Vaskonen[3,5,6], and Hardi Veermäe[3]

[1] Theoretical Physics Department, CERN, Geneva, Switzerland
[2] King's College London, Strand, London WC2R 2LS, UK
[3] Keemilise ja Bioloogilise Füüsika Instituut, Rävala pst. 10, 10143 Tallinn, Estonia
  e-mail: juan.urrutia@kbfi.ee
[4] Departament of Cybernetics, Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia
[5] Dipartimento di Fisica e Astronomia, Università degli Studi di Padova, Via Marzolo 8, 35131 Padova, Italy
[6] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Via Marzolo 8, 35131 Padova, Italy

ABSTRACT

NANOGrav and other Pulsar Timing Arrays (PTAs) have discovered a common-spectrum process in the nHz range that may be due to gravitational waves (GWs): if so, they are likely to have been generated by black hole (BH) binaries with total masses $>10^9 M_\odot$. Using the Extended Press-Schechter formalism to model the galactic halo mass function and a simple relation between the halo and BH masses suggests that these binaries have redshifts $z = O(1)$ and mass ratios $\gtrsim 10$, and that the GW signal at frequencies above $O(10)$ nHz may be dominated by relatively few binaries that could be distinguished experimentally and would yield observable circular polarization. Extrapolating the model to higher frequencies indicates that future GW detectors such as LISA and AEDGE could extend the PTA observations to lower BH masses $\gtrsim 10^3 M_\odot$.

Key words. black hole physics – gravitational waves – cosmology: theory

1. Introduction

The discovery ([LIGO Scientific & Virgo Collaborations 2016](#)) of gravitational waves (GWs) by the LIGO and Virgo experiments, in the $O(100)$ Hz range of frequencies generated by mergers of black holes (BHs) with masses of $O(10-100) M_\odot$ ([LIGO Scientific, Virgo & KAGRA Collaborations 2021](#)), has opened a new window onto both astrophysics and cosmology. Supermassive black holes (SMBHs) with masses of $O(10^6-10^{10}) M_\odot$ are known to be present in galactic nuclei, and the immediate surroundings of two of them have recently been imaged by the Event Horizon Telescope ([Event Horizon Telescope Collaboration 2019](#), [2022](#)). However, information on intermediate-mass black holes (IMBHs) with masses in the range $O(10^3-10^5) M_\odot$ ([Greene et al. 2020](#)) is less complete. Observations of IMBHs and their mergers would cast light on the uncharted mechanisms that must be presented for SMBHs to form.

A prerequisite for BH mergers is the formation of tightly bound binaries that can radiate GWs efficiently. While there are several well-understood paths for forming compact stellar-mass BHs binaries (e.g., common envelope evolution or dynamical capture in dense stellar environments), the final stage of SMBH binary formation in galaxy mergers is still not fully understood and is commonly known as "the final parsec problem" ([Begelman et al. 1980](#)). However, observations of quasar optical variability provide indirect evidence for the existence of some tight binary SMBHs in the regime where the emission of GWs must already have a noticeable effects on their orbital evolution ([Rieger & Mannheim 2000](#); [De Paolis et al.](#)

[2002](#), [2004](#); [Valtonen 2008](#); [Boroson & Lauer 2009](#); [Iguchi et al. 2010](#); [Graham et al. 2015](#); [O'Neill et al. 2022](#); [Kovačević et al. 2022](#))[1]. The GW data will be essential for a better understanding how SMBHs overcome the "final parsec" obstacle, and how the assembly of the SMBH population proceeds in general.

Pulsar timing arrays (PTAs) are potentially sensitive to GWs in the nHz range, while NANOGrav and other PTAs have recently reported evidence for a common-spectrum stochastic process ([Arzoumanian et al. 2020](#); [Goncharov et al. 2021](#); [Chen et al. 2021](#); [Antoniadis et al. 2022](#)). Their signals have power spectra that are consistent with predictions based on inspiralling binary SMBH models ([Phinney 2001](#))[2], but they have not (yet) detected the Hellings-Downs quadrupolar signature that is characteristic of GWs ([Hellings & Downs 1983](#))[3].

The BH merger interpretation of the PTA measurements does not provide direct information on the masses of the infalling BHs, which requires the modelling and measurements of this stochastic GW background in different frequency

---

[1] However, alternative scenarios for quasar optical variability may also be considered, such as intrinsic variability in the accretion disc ([King et al. 2013](#)).
[2] A similar signal could be generated by primordial BHs ([Vaskonen & Veermäe 2021](#); [De Luca et al. 2021](#); [Kohri & Terada 2021](#); [Ashoorioon et al. 2022](#)), but might require modifications of models based on simple cosmic string networks ([Ellis & Lewicki 2021](#); [Blasi et al. 2021](#); [Buchmuller et al. 2020](#)): see, e.g., [Blanco-Pillado et al. (2021)](#).
[3] We note in passing that binaries capable of generating the PTA nHz background could not explain the year-like quasi-periodicity seen in blazars ([Holgado et al. 2018](#)).

ranges (Sesana et al. 2008). The purpose of this paper is to illustrate how future measurements of the GW spectrum due to unresolved infall sources and individual binary mergers at frequencies between the PTAs and LIGO/Virgo will be able to extend the PTA measurements to lower black-hole mass ranges, probing models for the assembly of SMBHs, and perhaps obtaining direct evidence for the mergers of IMBHs. We take as the amplitude of the common-spectrum process measured by the PTAs $A = 2.8^{+1.2}_{-0.8} \times 10^{-15}$ at a reference frequency of 1/yr, as found by the International Pulsar Timing Array (IPTA; Antoniadis et al. 2022) assuming the spectral index of $\alpha = -2/3$ expected from inspiralling SMBHs, which is consistent with other measurements.

We use examples of next-generation GW detectors the planned LISA laser interferometer (Berti et al. 2006), whose peak sensitivity is at frequencies $O(10^{-4}-10^{-2})$ Hz, the proposed ET laser interferometer (Sathyaprakash et al. 2012), whose peak sensitivity is at frequencies $O(1-10^2)$ Hz and the projected AION-km and AEDGE atom interferometers (Badurina et al. 2020; El-Neaj et al. 2020), whose peak sensitivities would be in the range of $O(10^{-2}-1)$ Hz. We show that, while none of these detectors would observe a signal from the binary black holes weighing $>O(10^9) M_\odot$ that would probably be responsible for most of the PTA signal, LISA would observe a merger signal from binary black holes weighing $O(10^3-10^9) M_\odot$ that might contribute part of the PTA signal, while AEDGE would observe a merger signal if there is a population of binary black holes weighing $O(10^3-10^6) M_\odot$ with a formation history similar to the heavier BHs responsible for the PTA signal.

Throughout this paper, we use natural units with $c = 1$, $G_N = 1$.

2. BH merger rate

As a first step towards estimating the BH merger rate, we use the Extended Press-Schechter (EPS) formalism (Press & Schechter 1974; Bond et al. 1991) to calculate the galactic halo mass function,

$$\frac{\mathrm{d}n(M,t)}{\mathrm{d}\ln M} = \frac{\rho_0}{M} \sqrt{\frac{2}{\pi}} \frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M} \frac{\delta_c(z)}{\sigma(M)} e^{-\frac{\delta_c(z)^2}{2\sigma^2(M)}}, \quad (1)$$

where ρ_0 is the background matter density today, $\sigma^2(M)$ is the variance of the matter fluctuations and $\delta_c(z)$ is the critical overdensity for collapse. The latter is given by $\delta_c(z) \approx 1.686/D(z)$ where $D(z)$ is the linear growth function (Dodelson 2003). We calculate $\sigma^2(M)$ from the cold dark matter power spectrum with the *Planck* 2018 cosmological parameters (Planck Collaboration VI 2020) and the transfer function derived in Eisenstein & Hu (1998). The EPS formalism also gives an estimate for the probability per unit time for a halo of mass M_1 to merge with another one of mass, M_2, at some redshift, z, and become a halo of mass $M_f = M_1 + M_2$ (Lacey & Cole 1993):

$$\frac{\mathrm{d}p(M_1,M_2,t)}{\mathrm{d}t\mathrm{d}M_2} = \frac{1}{M_f} \sqrt{\frac{2}{\pi}} \left| \frac{\dot{\delta}_c}{\delta_c} \right| \frac{\mathrm{d}\ln\sigma}{\mathrm{d}\ln M_f} \left[1 - \frac{\sigma^2(M_f)}{\sigma^2(M_1)} \right]^{-\frac{3}{2}}$$
$$\times \frac{\delta_c}{\sigma(M_f)} \exp\left[-\frac{\delta_c^2}{2}\left(\frac{1}{\sigma^2(M_f)} - \frac{1}{\sigma^2(M_1)} \right) \right]. \quad (2)$$

Defining the halo merger rate kernel:

$$Q(M_1,M_2,t) \equiv \frac{\mathrm{d}p(M_1,M_2,t)}{\mathrm{d}t\mathrm{d}M_2} \left[\frac{\mathrm{d}n(M_2,t)}{\mathrm{d}M_2} \right]^{-1}, \quad (3)$$
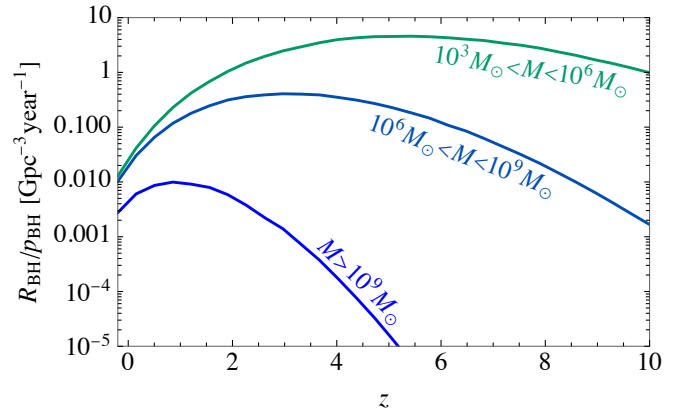
Fig. 1. BH merger rates calculated as functions of the redshift z in three ranges of the total merging BH mass M and normalized relative to the merger probability p_{BH} discussed in the text.

the halo merger rate is

$$\frac{\mathrm{d}R_h}{\mathrm{d}M_1\mathrm{d}M_2} = \frac{\mathrm{d}n(M_1,t)}{\mathrm{d}M_1} \frac{\mathrm{d}n(M_2,t)}{\mathrm{d}M_2} Q(M_1,M_2,t). \quad (4)$$

As discussed by Benson et al. (2005), the merger rate kernel $Q(M_1,M_2,t)$ in the EPS formalism is not symmetric in the exchange of M_1 and M_2. However, it was shown by Erickcek et al. (2006) that taking $M_1 < M_2$ in the merger kernel aptly approximates the result of the Benson-Kamionkowski-Hassani (BKH) merger theory (Benson et al. 2005) where the merger rate kernel is computed from the Smoluchowski coagulation equation and preserves the EPS halo mass function. We will, therefore, require that the first argument of Q is smaller than the second.

In order to estimate the merger rate of central BHs, we must estimate the probability $p_{\mathrm{occ}}(m|M,z)$ that a BH of mass m occupies a halo of mass M, as well as the probability $p_{\mathrm{merg}}(m_1,m_2)$ that the galactic merger leads to a merger of their central BHs. The resulting BH merger rate can be expressed as

$$\frac{\mathrm{d}R_{\mathrm{BH}}}{\mathrm{d}m_1\mathrm{d}m_2} = \int \mathrm{d}M_1\mathrm{d}M_2\, p_{\mathrm{merg}}(m_1,m_2)$$
$$\times p_{\mathrm{occ}}(m_1|M_1,z)p_{\mathrm{occ}}(m_2|M_2,z)\frac{\mathrm{d}R_h}{\mathrm{d}M_1\mathrm{d}M_2}. \quad (5)$$

We assume the following simple redshift-dependent relation between the halo mass and the BH mass (Barkana & Loeb 2001; Wyithe & Loeb 2003):

$$\frac{M_v}{10^{12}M_\odot} = 10.5 \left[\frac{\Omega_M(0)}{\Omega_M(z)} \frac{\Delta_c(z)}{18\pi^2} \right]^{-\frac{1}{2}} (1+z)^{-\frac{3}{2}} \left[\frac{m_{\mathrm{BH}}}{10^8 M_\odot} \right]^{\frac{3}{5}}, \quad (6)$$

where $\Delta_c(z) = 18\pi^2 + 82[\Omega_M(z) - 1] - 39[\Omega_M(z) - 1]^2$ is the critical overdensity at virialization, which corresponds to $p_{\mathrm{occ}}(m|M) = p_{\mathrm{occ}}(m)\delta(m - m_{\mathrm{BH}}(M))$. It is generally expected that SMBHs inhabit most large galaxies, and X-ray observations (Miller et al. 2015) constrain the SMBH occupation fraction $p_{\mathrm{occ}}(m)$ to be $> 20\%$ for early galaxies with lower stellar masses $10^7 < M_*/M_\odot < 10^{10}$. It has been estimated that an initially small occupation fraction may grow at low redshifts (see, e.g., Lippai et al. 2009), which we neglect in this analysis. As we show later, the PTA GW signal is dominated by events in a range $1 \lesssim z \lesssim 3$. The estimated z dependence of p_{occ} would shift this range towards lower redshifts when compared to a constant p_{occ}, increasing the observability of BH mergers.
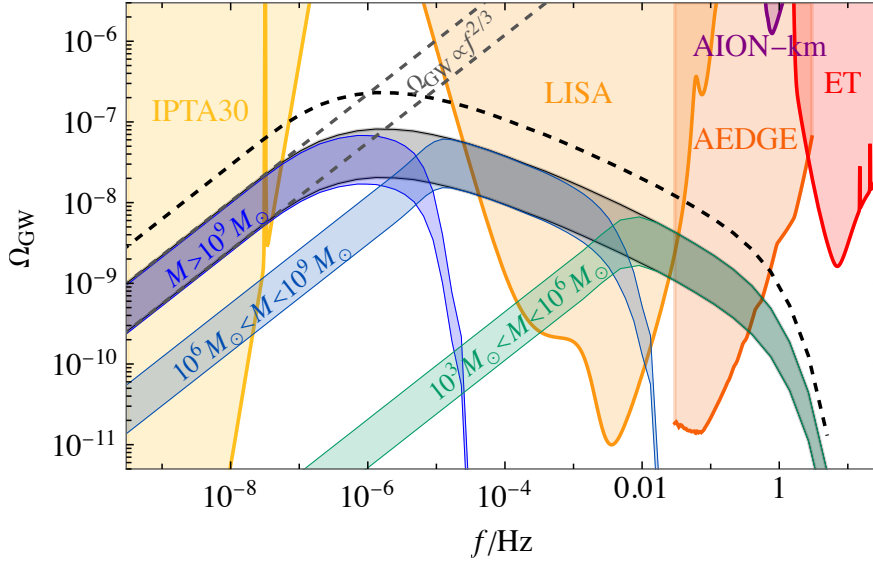
Fig. 2. Mean GW energy density spectrum from massive BH mergers compared with the sensitivities of different experiments. The black dashed curve shows the case where all mergers of galaxies produce a GW signal for BHs heavier than $10^3 \, M_\odot$ ($p_{\mathrm{BH}} = 1$). The colored bands show the spectra from SMBHs heavier than $10^9 \, M_\odot$ (dark blue), from SMBHs in the range $(10^6 \, M_\odot, 10^9 \, M_\odot)$; light blue), and from IMBHs with masses in the range $10^3 - 10^6 \, M_\odot$ (green), assuming a universal efficiency factor $p_{\mathrm{BH}} = 0.17^{+0.18}_{-0.08}$. The shaded regions show the prospective sensitivities of IPTA30 (Chen et al. 2017; Kaiser & McWilliams 2021), LISA (LISA Collaboration 2017), AEDGE (El-Neaj et al. 2020; Badurina et al. 2021), AION-km (Badurina et al. 2020, 2021) and ET (Sathyaprakash et al. 2012).

In view of the paucity of information about the merger probability $p_{\mathrm{merg}}(m_1, m_2)$, we model p_{merg} by a constant for the sake of simplicity. After the galactic merger, various dynamical mechanisms must decrease the size of the SMBH binary below sub-parsec scales in order for a GW emission-driven merger to take place (Begelman et al. 1980). The crossing of the final parsec is determined mostly by binary hardening via stellar loss-cone scattering (Merritt 2013), which can currently be considered the largest source of uncertainty in p_{merg}. The numerical study of Kelley et al. (2017a), based on a population of $10^6 - 10^{10} \, M_\odot$ BH binaries in the Illustris simulation (Vogelsberger et al. 2014), found for a wide range of model parameters that the coalescing fraction is nearly independent of the total mass of the binary but decreases with the mass ratio. Neglecting the latter dependence should not significantly affect our results, since extreme mass ratio inspirals contribute subdominantly to the GW background. With these assumptions, we obtain:

$$\frac{\mathrm{d}R_{\mathrm{BH}}}{\mathrm{d}m_1 \mathrm{d}m_2} \approx p_{\mathrm{BH}} \frac{\mathrm{d}M_1}{\mathrm{d}m_1} \frac{\mathrm{d}M_2}{\mathrm{d}m_2} \frac{\mathrm{d}R_h}{\mathrm{d}M_1 \mathrm{d}M_2}, \qquad (7)$$

where $p_{\mathrm{BH}} \equiv p_{\mathrm{occ}}(m_1) p_{\mathrm{occ}}(m_2) p_{\mathrm{merg}}$ denotes the probability that the halo merger leads to a merger of the central SMBHs.

We show in Fig. 1 the BH merger rates as functions of the redshift z, calculated using this approach in three ranges of the total merging BH mass M and normalized relative to the merger probability p_{BH} discussed above. We see that mergers with total masses $M > 10^9 \, M_\odot$ occur typically at $z = O(1)$, those with total masses $M \in (10^6, 10^9) \, M_\odot$ occur typically at $z = O(3)$, and those with total masses $M \in (10^3, 10^6) \, M_\odot$ occur typically at $z = O(5)$.

In this paper, we use the IPTA measurement (Antoniadis et al. 2022) to normalize p_{BH} for large BH masses, which circumvents the astrophysical uncertainties related to p_{occ} and p_{merg}, and extrapolate it to smaller BH masses. We also comment on the potential mass dependence of p_{BH}, which is relevant for the GW phenomenology when we

extrapolate from the frequency band relevant for PTAs to those to be explored by next-generation GW detectors.

3. Analysis

3.1. The GW energy spectrum

The mean GW energy density spectrum from the SMBH binary population can be estimated as (Phinney 2001):

$$\Omega_{\mathrm{GW}}(f) \equiv \frac{1}{\rho_c} \frac{\mathrm{d}\rho_{\mathrm{GW}}}{\mathrm{d} \ln f} = \int \mathrm{d}\lambda \frac{2\pi}{5} \frac{f^3 |\tilde{h}(f)|^2}{\rho_c}, \qquad (8)$$

where $|\tilde{h}(f)|$ denotes the optimal amplitude of the Fourier transform of the GW strain, $\rho_c = 3H_0^2/8\pi$,

$$\mathrm{d}\lambda = \mathrm{d}m_1 \mathrm{d}m_2 \frac{\mathrm{d}z}{1+z} \frac{\mathrm{d}V_c}{\mathrm{d}z} \frac{\mathrm{d}R_{\mathrm{BH}}(t)}{\mathrm{d}m_1 \mathrm{d}m_2}, \qquad (9)$$

is the differential merger rate and V_c denotes the comoving volume available at redshift z. We compute $|\tilde{h}(f)|$ using the inspiral-merger-ringdown template (Ajith et al. 2008):

$$|\tilde{h}(f)| = \sqrt{\frac{5}{24}} \frac{\mathcal{M}_z^{\frac{5}{6}}}{\pi^{\frac{2}{3}} D_L}$$
$$\times \begin{cases} f^{-\frac{7}{6}} & f < f_{\mathrm{merg}} \\ f_{\mathrm{merg}}^{-\frac{1}{2}} f^{-\frac{2}{3}} & f_{\mathrm{merg}} \le f < f_{\mathrm{ring}} \\ f_{\mathrm{merg}}^{-\frac{1}{2}} f_{\mathrm{ring}}^{-\frac{2}{3}} \frac{\sigma^2}{4(f - f_{\mathrm{ring}})^2 + \sigma^2} & f_{\mathrm{ring}} \le f < f_{\mathrm{cut}}, \end{cases} \qquad (10)$$

where the redshift-dependent chirp mass, $\mathcal{M}_z$, is given in terms of the total mass , M, and the symmetric mass ratio $\eta \equiv m_1 m_2/M^2$ of the binary by $\mathcal{M}_z \equiv (1+z)\mathcal{M}_0 = (1+z)M\eta^{\frac{3}{5}}$, and D_L denotes the luminosity distance of the binary. This implies the canonical $\Omega_{\mathrm{GW}} \propto f^{\frac{2}{3}}$ scaling during the inspiral phase (Phinney 2001). The frequencies $f_{\mathrm{merg}}, f_{\mathrm{ring}}, f_{\mathrm{cut}},$ and σ are parameterised

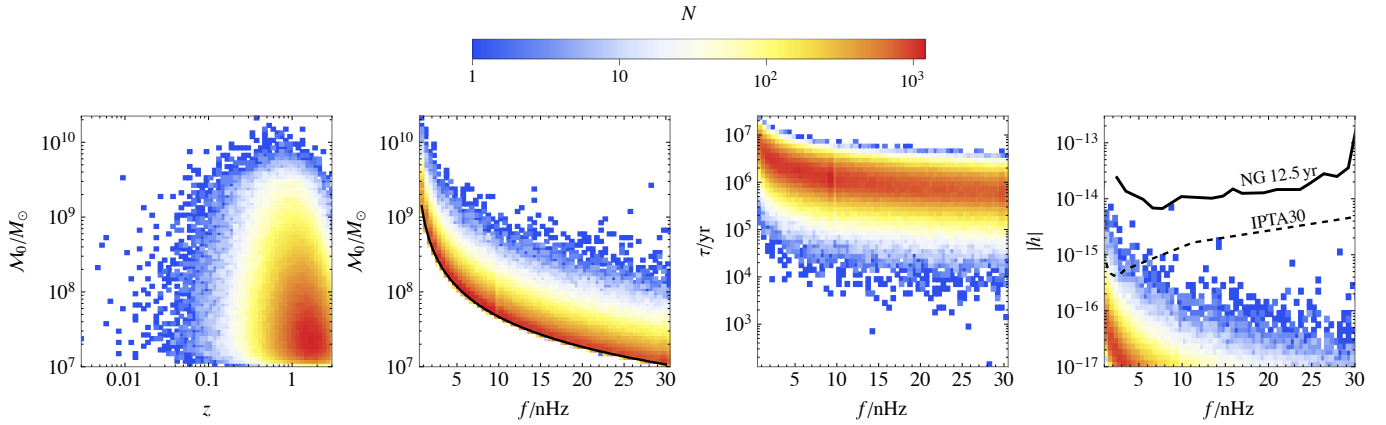**Fig. 3.** Distributions of the binary parameters for a sampling of the BH binary population with $0 < z < 3$, $1\,\text{nHz} < f < 30\,\text{nHz}$ and $\mathcal{M}_0 > 10^9 (f/\text{nHz})^{-\frac{4}{3}}$. The symmetric mass ratio is integrated with the lower bound $\eta > 0.001$. In the right panel, the black curve shows the NANOGrav 95% CL constraint on continuous GW sources (Arzoumanian et al. 2023) and the projected sensitivity of IPTA30 at SNR = 8.

as $f_j = \eta^{\frac{5}{3}}(a_j \eta^2 + b_j \eta + c_j)/(\pi \mathcal{M}_z)$ where a_j, b_j, and c_j are coefficients whose fitted values are given in Table I of Ajith et al. (2008).

We comment briefly on potential uncertainties in the GW energy density spectrum. If orbital decay is driven by other processes in addition to GW emission, such as viscous drag, then the coalescence time is shortened and the total GW spectrum is suppressed by a factor of $T_{\text{tot}}/T_{\text{GW}}$, where T_{tot}, T_{GW} denote the characteristic hardening timescales (Kocsis & Sesana 2011; Kelley et al. 2017a). Here we omit this effect, assuming that the signal is dominated by binaries for which the hardening is driven mainly by GW emission. Additionally, we consider only circular binaries, which, in the inspiral phase, radiate monochromatically at twice the orbital frequency. We note, however, that the GW spectrum of eccentric binaries would contain higher harmonics, with the fundamental harmonic (at the orbital frequency) becoming dominant at eccentricities $e > 0.4$ (Taylor et al. 2016). In addition, the GW luminosity would be enhanced by a factor of $(1 + 73/24e^2 + 37/96e^4)/(1 - e^2)^{\frac{7}{2}}$ (Peters & Mathews 1963; Enoki & Nagashima 2007; Kelley et al. 2017b).

Figure 2 shows the GW energy density spectrum from BH binaries in different mass ranges. The black dashed curve corresponds to $p_{\text{BH}} = 1$, namely, the assumption that each galactic merger would produce a BH merger with $m_2 > m_1 > 10^3 \, M_\odot$. This naive assumption leads to a spectrum exceeding that observed in the PTA band.

We stress that the spectrum in Eq. (8), which is displayed in Fig. 2, does not necessarily correspond to a stochastic GW background as it does not distinguish between individual resolvable events and unresolvable events that would contribute to a GW background. The $\Omega_{\text{GW}} \propto f^{\frac{2}{3}}$ tail at low frequencies arises from nearly monochromatic signals generated by inspiralling BH binaries. Because of this, Ω_{GW} will fluctuate depending on the specific realization of the binary population, so the spectrum in Fig. 2 should be interpreted as the mean Ω_{GW} obtained by averaging over many potential realizations of the binary population. We return to the probability distribution of Ω_{GW} in the next subsection and in Appendix A.

The grey band in Fig. 2 was obtained by choosing a fixed value of p_{BH} so that the amplitude of the total GW spectrum from BH binaries matches the IPTA observations in the nHz range, which requires $p_{\text{BH}} = 0.17^{+0.18}_{-0.08}$. The colored bands show the contributions to the total IPTA spectrum from different ranges of the BH masses. In comparing the dark blue, light blue, and green bands, we can see that the largest contribution to the GW spectrum in the PTA window comes from $M > 10^9 \, M_\odot$ SMBH binaries, consistent with earlier studies (Wyithe & Loeb 2003; Sesana et al. 2004, 2008, 2009; Enoki et al. 2004; Kelley et al. 2017a; Izquierdo-Villalba et al. 2021; Bécsy et al. 2022).

Some characteristics of these binaries are shown in Fig. 3. The left panel shows that the total masses are $\lesssim 10^{10}$, and the redshifts are typically $O(1)$. The middle left panel shows that the typical masses increase with decreasing frequency and the middle right panel shows that the binary lifetimes are only weakly correlated with their emission frequencies, with lifetimes $\gtrsim 10^4$ yr being generally favored. The right panel shows the GW strain from these binaries, together with the upper limit on continuous GW sources from NANOGrav 12.5 years data (Arzoumanian et al. 2023) and the projected sensitivity of IPTA30 (Chen et al. 2017; Kaiser & McWilliams 2021) at a signal-to-noise ratio of SNR = 8, which is similar to that expected for SKA (Janssen et al. 2015).

Returning to Fig. 2, we see that the GW energy spectrum in the LISA frequency range is dominated by lighter BH binaries in the mass range of $M \in (10^6, 10^9) \, M_\odot$, shown in light blue, and the contribution in the AEDGE band is dominated by IMBHs with masses in the range of $M \in (10^3, 10^6) \, M_\odot$, shown in green. The low-frequency GW emissions from the early stages of the inspirals of these binaries also contribute in the lower-frequency bands, but these contributions are subdominant compared to the signals from higher-mass SMBH mergers[4].

We recall that the estimated spectrum in Fig. 2 assumes a constant p_{BH}. Clearly, the relative contribution from binaries with lighter BH masses $< 10^9 \, M_\odot$ could be enhanced (reduced) by increasing (decreasing) p_{BH} for these binaries. In order for mergers of BHs with masses in the range $(10^6, 10^9) \, M_\odot$ to dominate the PTA signal $p_{\text{BH}} \simeq 1$ would be required in this mass

---

[4] In addition to the experiments shown in Fig. 2, we have also considered the prospective sensitivities to GWs using the astrometric data from the *Gaia* and *Nancy Grace Roman* space telescopes (Wang et al. 2021). Their nominal sensitivities lie well above the black dashed line in Fig. 2 corresponding to $p_{\text{BH}} = 1$, but a possible improvement of the Roman sensitivity might enable the detection of SMBH binary signals at frequencies $\in (10^{-7}, 10^{-6})$ Hz. Moreover, while the prospective power-law integrated sensitivity to GWs through binary resonance by laser ranging of the Moon (Blas & Jenkins 2022) reaches the gray band shown in Fig. 2, the most likely signal at $f \sim 10^{-6}$ Hz is well below that (see Sect. 3.2) and consists of a few nearly monochromatic sources.
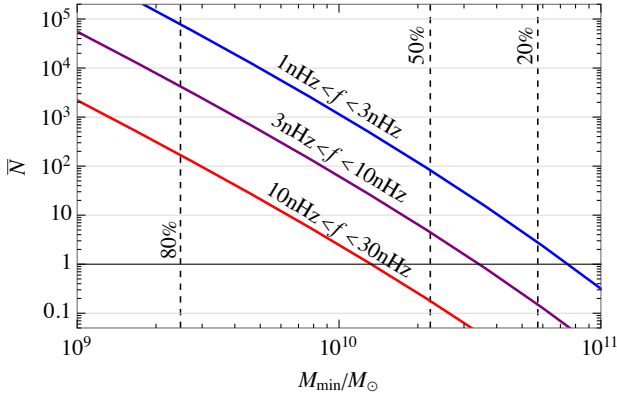
Fig. 4. Expected number of SMBH binaries heavier than M_{min} and with symmetric mass ratio $\eta > 0.01$ emitting GWs in the indicated frequency range for $p_{BH} = 0.17$. The vertical dashed lines indicate the fractions of the total GW signal in each frequency band generated by binaries with larger masses.

range, but even this maximal enhancement would be insufficient for mergers with $M < 10^6 \, M_\odot$ to contribute significantly to the PTA signal. That said, we emphasize that the GW spectrum in either the LISA or AEDGE frequency range could reach the black dashed line if $p_{BH} \simeq 1$ in the relevant mass range, and could also be enhanced if Population III stars make a significant contribution to the BH spectrum. Conversely, the spectrum could be suppressed in these ranges if p_{BH} is smaller than the value that fits the PTA data.

3.2. Distribution of sources

The number of BH binaries at redshift z emitting in a given frequency band can be estimated from the time the binary spends in that frequency band, which gives (Sesana et al. 2008)

$$\frac{\mathrm{d}N}{\mathrm{d}\mathcal{M}_z \mathrm{d}\eta \mathrm{d}z \mathrm{d}\ln f} = \frac{\mathrm{d}\lambda}{\mathrm{d}\mathcal{M}_z \mathrm{d}\eta \mathrm{d}z} \left| \frac{\mathrm{d}\tau}{\mathrm{d}\ln f} \right|, \tag{11}$$

where, assuming circular orbits and orbital decay via GW emission (Peters 1964), the coalescence time of an inspiralling binary emitting GWs with frequency $(1 + z)f$ is[5]

$$\tau(f) = \frac{5}{256 \pi^{\frac{8}{3}} \mathcal{M}_z^{\frac{5}{3}}} f^{-\frac{8}{3}}. \tag{12}$$

We show in Fig. 4 the expected numbers of binaries with total masses of $M \geq M_{min}$ and mass ratios $q \geq 0.01$ that emit GWs in several frequency bands in the range $1 \, \mathrm{nHz} < f < 30 \, \mathrm{nHz}$ (corresponding to the sensitivity ranges of PTAs). As seen in Fig. 2, SMBH binaries generate the dominant contribution to the GW signal in this frequency range. The fractions of the total GW signal in each frequency band that are generated by binaries with larger masses are indicated by the vertical dashed lines. For example, we see that almost 10^5 (10^2) binaries with $M > 3 \times 10^9 \, M_\odot$ generate 80% of the calculated GW signal in the $1 \, \mathrm{nHz} < f < 3 \, \mathrm{nHz}$ ($10 \, \mathrm{nHz} < f < 30 \, \mathrm{nHz}$) frequency band, whereas 50% of the calculated GW signal in the $3 \, \mathrm{nHz} < f < 10 \, \mathrm{nHz}$ band may be generated by just one source.

---

[5] We remark that, in Eq. (11), the observed time spent in a given frequency band is redshifted when compared to the time in the binary rest frame. This redshift has been accounted for in Eq. (12).

These examples indicate that the expected signal, particularly in higher-frequency bins, may be comprised of a limited number of nearly monochromatic signals from heavy SMBH binaries ($>10^9 \, M_\odot$), implying sizeable fluctuations around the smooth $f^{\frac{2}{3}}$ spectrum that would be obtained in the limit of a large population of heavy inspiralling binaries.

In order to study the prospective SMBH binary population in the PTA band via a Monte Carlo approach, we generated realizations of the population from the probability density function of binaries emitting at a given frequency. We divide the frequency range into bins (f_j, f_{j+1}) populated with $N(f_j)$ binaries. A realization of the GW background from each bin can then be obtained as (see Eq. (A.6)):

$$\Omega_{GW}(f_j) = \frac{1}{\ln(f_{j+1}/f_j)} \sum_{k=1}^{N(f_j)} \Omega_{GW}^{(1)}(\theta_k), \tag{13}$$

where the contribution from an individual binary emitting in this frequency band is given by

$$\Omega_{GW}^{(1)}(\theta) = \left| \frac{\mathrm{d}\tau}{\mathrm{d}\ln f} \right|^{-1} \frac{2\pi}{5} \frac{f^3 |\tilde{h}(f)|^2}{\rho_c}, \tag{14}$$

and $\theta \equiv \{\mathcal{M}_z, z, \eta, f\}$ are parameters describing the binary. The number of binaries $N(f_j)$ in this frequency bin is drawn from a Poisson distribution with the expectation value $\bar{N}(f_j) \equiv \int_{f_j}^{f_{j+1}} \mathrm{d}N$ and the binary parameters are generated randomly according to the distribution

$$p(\mathcal{M}_z, \eta, z, f) \propto \frac{\mathrm{d}N}{\mathrm{d}\mathcal{M}_z \mathrm{d}\eta \mathrm{d}z \mathrm{d}\ln f}. \tag{15}$$

This is the expression that was used to calculate the distributions of binary parameters shown in Fig. 3 for a sampling with $p_{BH} = 0.17$ of 4×10^5 SMBH binaries with redshifts $0 < z < 3$, mass ratios $\eta > 0.01$, and GW frequencies $1 \, \mathrm{nHz} < f < 30 \, \mathrm{nHz}$. We display only binaries with chirp mass $\mathcal{M}_0 > 10^9 (f/\mathrm{nHz})^{-\frac{4}{3}}$, so that $\bar{N}(f_j) > O(10^4)$ for each $1 \, \mathrm{nHz}$ frequency bin.

In order to generate the statistical distributions of Ω_{GW}, it is sufficient to consider a smaller set of parameters, and the distribution Eq. (11) for the full population can be reduced to a simpler distribution $P(\Omega^{(1)}, f)$ for individual sources $\Omega_{GW}^{(1)}$ emitting at a frequency f, which can be expressed via two one-parameter functions depending on the merger rate model, as discussed in Appendix A.

The upper left panel of Fig. 5 illustrates the frequency spectra found in 1600 Monte Carlo realizations of the SMBH binary population drawn from a statistical distribution similar to that shown in Fig. 3. The black solid curve shows the mean of the spectra, the grey bands show the 1σ and 2σ CL regions of the GW spectra, and the red line is the spectrum found in one representative realization. We note the importance of fluctuations, particularly at higher frequencies where fewer SMBH binaries contribute, with individual binaries becoming distinguishable at frequencies $f \gtrsim 10 \, \mathrm{nHz}$. We find that the median spectrum, shown by the dashed line, has a somewhat lower slope than the analytic result shown as the green line. Although the numerically obtained mean spectrum (black), lies below the analytic expectation, it will approach it slowly if the number of Monte Carlo realizations is increased. The tendency of the mean spectrum to lie along the upper side of the 1σ CL range is due to the fact that the spectrum has a long tail at high values of Ω_{GW} generated by occasional nearby binaries: the median spectrum always
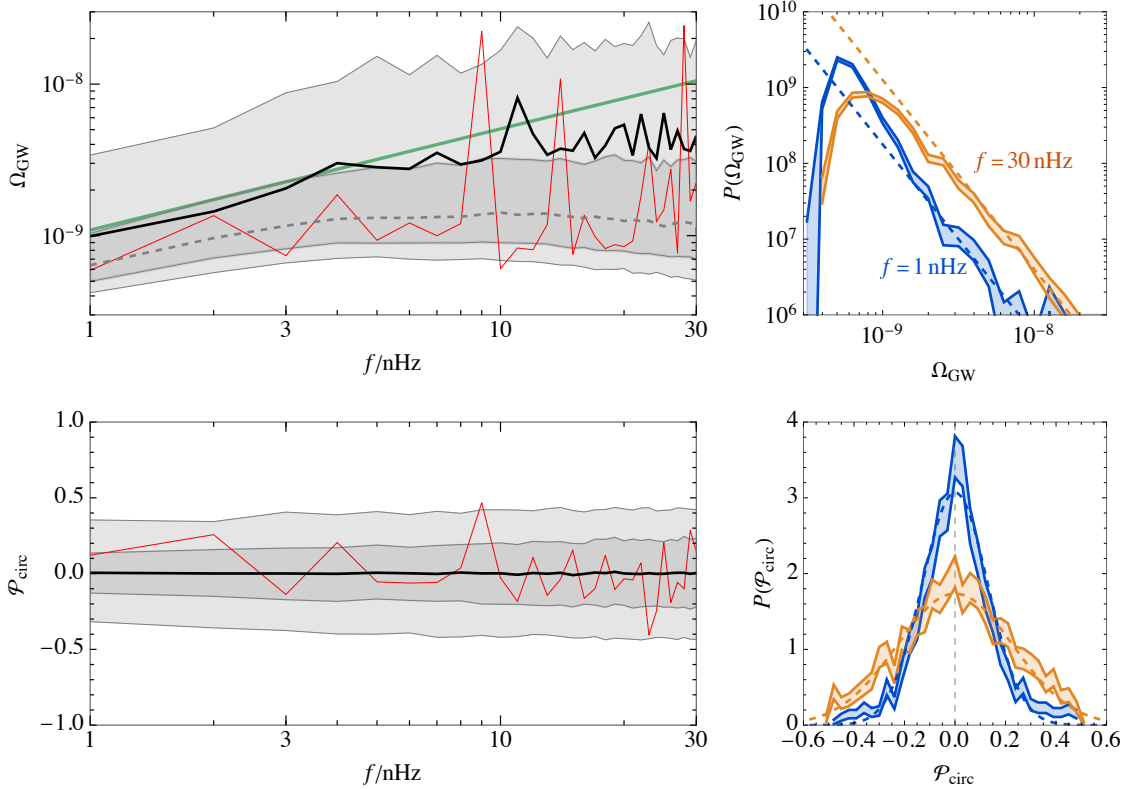
Fig. 5. GW energy spectrum Ω_{GW} and the fractional circular polarization $\mathcal{P}_{circ}$ calculated from 1600 Monte Carlo realizations of the SMBH binary population. In the left panels, the black solid curve shows the mean and the gray bands the 68% and 95% CL regions of the Monte Carlo realizations, and the red curve corresponds to one of the realizations shown in Fig. 3. In the upper left panel the green line shows the mean GW energy density (8). The distributions of Ω_{GW} and $\mathcal{P}_{GW}$ at 1 nHz (blue) and at 30 nHz (orange) calculated from the Monte Carlo realizations are shown in the right panels. In the upper right panel, the dashed lines show $P \propto \Omega^{-\frac{5}{2}}$, and in the lower right panel, the dashed curves show Gaussian fits to these distributions of width $\sigma = 0.13$ (blue) and $\sigma = 0.23$ (orange). In this figure, $p_{BH} = 0.17$ is assumed.

lies within the 1σ CL range. All in all, consistently with earlier Monte Carlo studies of the GW signal (Sesana et al. 2008; Kocsis & Sesana 2011), we find that typical spectra tend to fall below the $\Omega_{GW} \propto f^{\frac{2}{3}}$ expectation in higher frequency bins, while a few bins display sharp peaks.

As there are a finite number of sources, Eq. (8) is subject to statistical fluctuations, which arise mostly from the possibility of having a few strong sources nearby. To obtain an order-of-magnitude estimate, we focus on the closest binaries and ignore the redshift dependence. In this case, $\Omega_{GW}^{(1)} \propto D_L^{-2}$, by Eqs. (8) and (10), and the probability of finding an event at D_L is $P^{(1)}(D_L) \propto D_L^2$. Thus $\langle \Omega_{GW}^{(1)} \rangle \propto \int_{D_{L,min}}^{D_{L,max}} dD_L P(D_L) \Omega_{GW}^{(1)} \propto D_{L,max}$, where, in this simplified approach, $D_{L,max}$ is some large luminosity distance at which $\Omega_{GW}^{(1)}$ gets suppressed, and $D_{L,min}$ is the distance to the nearest possible source. On the other hand, $\langle (\Omega_{GW}^{(1)})^2 \rangle \propto 1/D_{L,min}$. Therefore, we expect the mean of the GW signal to be determined by faraway sources, while the variance is set by a few close-by binaries. Moreover, we can estimate from $P^{(1)}(D_L)$ that $\Omega_{GW}^{(1)}$ has a relatively flat power-law tail at large values,

$$P^{(1)}(\Omega) \propto \Omega^{-\frac{5}{2}}, \quad \text{when} \quad \Omega \to \infty. \tag{16}$$

Since the closest distance to massive BH binaries is constrained, this tail will be cut off at $\Omega_{GW,max}^{(1)} \propto D_{L,min}^{-2}$. Even with this cutoff, the mean and the variance are still not very useful characteristics of the uncertainties, and we find it more illuminating to estimate the confidence intervals around the median value.

The upper right panel of Fig. 5 displays the distributions of Ω_{GW} at 1 nHz (blue) and at 30 nHz (orange) found in 1600 Monte Carlo realizations of the SMBH binary population. We note that the distributions at both frequencies have tails that approach the analytical result $P \propto \Omega^{-\frac{5}{2}}$ (Eq. (16)), as indicated by the dashed lines. These tails exhibit explicitly why the mean (black solid) line in Fig. 5 tends to lie above the 68% CL band, while the median (gray dashed) line lies within it at all frequencies. We note that the overall shapes of the distributions resemble the analytic calculations shown in the left panel of Fig. A.1. In summary, even if a bin has a high number of contributing events (i.e., $\bar{N} \to \infty$), the distribution of Ω_{GW} does not converge to a Gaussian because it retains its $\Omega_{GW}^{-5/2}$ tail. Since the variance is not well behaved, we define the width of the distribution $\Delta_{\Omega_{GW}}$ as the width of the 68% confidence interval, as in Fig. 5. The width-to-mean ratio scales as:

$$\Delta_{\Omega_{GW}} / \langle \Omega_{GW} \rangle \propto \bar{N}^{-\frac{1}{3}}, \tag{17}$$

when $\bar{N} \gg 1$, as is shown in Appendix A. This scaling is slower than the typical $1/\sqrt{\bar{N}}$ scaling predicted by the central limit theorem.

3.3. GW polarization

Recent studies have argued that the circular polarization of the signal can be used to estimate whether the SGWB comes from a handful of sources or a relatively large population of binaries (Kato & Soda 2016; Conneely et al. 2019;

<Hotlinli et al. 2019; Belgacem & Kamionkowski 2020; Sato-Polito & Kamionkowski 2022; Valbusa Dall'Armi et al. 2023>). We recall that the left and right circular GW polarization amplitudes from a binary with inclination angle θ are

$$|h_{L,R}(f)|^2 = \frac{1}{16}(1 \pm \cos \theta)^4 |h(f)|^2, \qquad (18)$$

where

$$|h(f)| = \frac{4 \mathcal{M}_z^{\frac{5}{3}}}{D_L}(\pi f)^{\frac{2}{3}} \qquad (19)$$

is the maximal GW strain from an inspiralling binary, and the gravitational Stokes parameters are defined by

$$\begin{aligned} I(f) &= |h_L(f)|^2 + |h_R(f)|^2, \\ V(f) &= |h_L(f)|^2 - |h_R(f)|^2. \end{aligned} \qquad (20)$$

The fractional amount of circular polarization of the SGWB can be characterized by the quantity

$$\mathcal{P}_{\text{circ}}(f) = \frac{\sum_i V_i(f)}{\sum_i I_i(f)}, \qquad (21)$$

where the sums are over all binaries in a fixed frequency range.

If the GW signal is dominated by a single source, then $\mathcal{P}_{\text{circ}}$ depends only on the inclination angle and large circular polarizations are preferred, with $\mathcal{P}_{\text{circ}} > 0.87(0.2)$ at the 68% CL (95% CL). On the other hand, if the GW signal is dominated by several ($N_{\text{dom}} \gtrsim 10$) sources of comparable strengths, then the $\mathcal{P}_{\text{circ}}(f)$ will be approximately Gaussian with a width determined by fluctuations in σ_V. Since $\sigma_V/\langle I \rangle_\theta = 1.17$ for a single source, we can estimate that (see Appendix A for details)

$$\sigma_{\mathcal{P}_{\text{circ}}} \approx \sigma_V/\langle I \rangle_\theta \approx 1.17/\sqrt{N_{\text{dom}}}. \qquad (22)$$

The lower-left panel of Fig. 5 illustrates the distributions of the circular polarization $\mathcal{P}_{\text{circ}}(f)$ found in the sample of 1600 Monte Carlo realizations of the SMBH binary population with $M > 10^9 M_\odot$ and $\eta > 0.01$ whose frequency spectra were illustrated in Fig. 5. We see that the mean value of $\mathcal{P}_{\text{circ}}(f) \approx 0$, as expected, but large statistical fluctuations are possible even at the 1σ level. This phenomenon is visible in the red curve, which shows results from one of the Monte Carlo realizations. The large fluctuations reflect the fact that the observable GW spectrum could be due to a very limited number of sources, particularly at higher frequencies.

The lower-right panel of Fig. 5 displays the distributions of $\mathcal{P}_{\text{circ}}$ at 1 nHz (blue) and at 30 nHz (orange) found in 1600 Monte Carlo realizations of the SMBH binary population. We note that the statistical distribution is indeed broader at the higher frequency, as expected. The dashed curves are Gaussian fits to polarization distributions with widths $\sigma = 0.13$ (blue) and $\sigma = 0.23$ (orange). These fits are very accurate as also demonstrated in the right panel of Fig. A.1.

As the variance of the circular polarization of the signal depends on the number of binaries that dominate the signal, it would be suppressed if the merger rate of the heaviest binaries were suppressed, which could be the case if p_{BH} is not universal. By suppressing p_{BH}, for instance, at $M > 10^9 M_\odot$ and enhancing p_{BH} for $10^9 M_\odot > M > 10^6 M_\odot$, we can accommodate the IPTA common-spectrum effect. In this case, the GW energy spectrum Ω_{GW} becomes smoother and approaches the naive $\Omega_{\text{GW}} \propto f^{\frac{2}{3}}$ behavior while suppressing the circular polarization $\mathcal{P}_{\text{circ}}$. Moreover, in this case, p_{BH} for the lighter binaries would be enhanced, which would increase the number of signals in the sensitivity range of LISA and possibly AEDGE.

3.4. Prospects for future GW observatories

As discussed earlier, BH binaries can generate two qualitatively distinct classes of signal: long, nearly monochromatic signals from the slowly-evolving inspiralling phase and relatively short signals from the merger and ringdown. The detections of these two types of sources need to be considered separately.

The number of detectable nearly monochromatic GW signals that arise from inspiralling BH binaries is

$$N_{\text{insp}} = \int_{\mathcal{T}+\tau_{\text{min}}}^{\infty} d\tau \int d\lambda \, p_{\text{det}}\left[\frac{\text{SNR}_c}{\text{SNR}(\tau, z)}\right], \qquad (23)$$

where we account only for binaries whose coalescence time τ is longer than $\tau_{\text{min}} = 1$ day and we estimate the SNR for a detector characterized by the noise power spectrum $S_n(f)$ as

$$\text{SNR}(\tau, z) = \sqrt{\int_0^{\mathcal{T}} dt \, \frac{2|h(\tau - t, z)|^2}{S_n(f(\tau - t, z))}}. \qquad (24)$$

The optimal time-dependent inspiral strain $|h(t)| = |h(f(t))|$ is given by Eq. (19) and the GW frequency as a function of time is given by Eq. (12). Analogously, the expected number of BH binary merger events is

$$N_{\text{merg}} = \mathcal{T} \int d\lambda \, p_{\text{det}}\left[\frac{\text{SNR}_c}{\text{SNR}(z)}\right], \qquad (25)$$

where

$$\text{SNR}(z) = \sqrt{\int_{f(\tau_{\text{min}})}^{\infty} df \, \frac{4|\tilde{h}(f)|^2}{S_n(f)}}. \qquad (26)$$

The cut at $f(\tau_{\text{min}})$ here implies that we only consider the SNR only from the last day of the signal. In both cases, the detection probability p_{det} accounts for the detector's antenna patterns and includes the average over the binary inclination, sky location, and polarization (Finn & Chernoff 1993; Gerosa et al. 2019), the noises S_n include the foregrounds from stellar mass BH binaries and white dwarf binaries (see e.g., Lewicki & Vaskonen 2021). Furthermore, we used $\text{SNR}_c = 8$ for the detection threshold and $\mathcal{T} = 1$ yr for the observation time.

In the monochromatic limit, Eq. (24) is simplified to $\text{SNR} = |h(f)|\sqrt{2\mathcal{T}/S_n(f)}$. Using this, we can convert the prospected noise $S_n(f)$ of IPTA30 to a lower bound on the strain $|h(f)|$ for which the SNR exceeds the detection threshold. As indicated by the very long coalescence times in the middle right panel of Fig. 3, the signals in the IPTA30 band can be adequately approximated as monochromatic. In the right panel of Fig. 3 the dashed black curve shows the GW strain $|h(f)|$ that for 20 year observation time with IPTA30 gives $\text{SNR} = 8$. We see that IPTA30 can potentially resolve several, $\mathcal{O}(10)$, monochromatic GW signals from SMBH binaries.

At higher frequencies, $f > 10^{-6}$ Hz, we find that LISA is the only detector capable of measuring the near-monochromatic sources arising in the early stages of the inspiral phase, namely, more than one day from the beginning of the merger phase. It is clear that AEDGE is not able to observe these near-monochromatic sources because the coalescence time of binaries heavier than $10^3 M_\odot$ is less than a day when they enter the AEDGE sensitivity window[6]. Figure 6 illustrates the prospects

---

[6] In this work we have cut the AEDGE sensitivity at 3 mHz. This cut is motivated by the potential Newtonian gravity backgrounds (Hogan et al. 2011; Graham et al. 2017). We note, however, that dedicated studies of these backgrounds for AEDGE have not been performed.
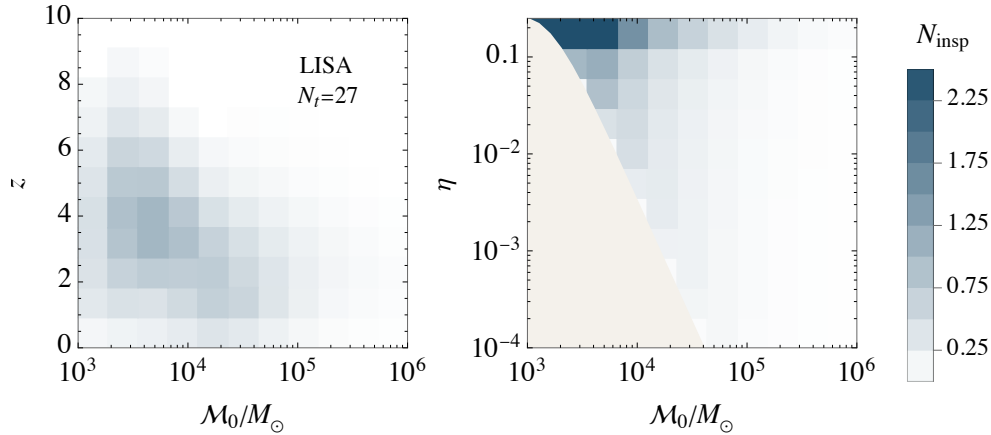
Fig. 6. Expected numbers of near-monochromatic GW signals generated more than a day before the merger that would be detectable in a year of LISA observation, as functions of the chirp mass $\mathcal{M}_0$ and redshift z (left panel) and of the chirp mass and symmetric mass ratio η (right panel). The merger rates assume the abundance of IMBHs described by Eq. (6) for BHs heavier than $10^3 \, M_\odot$ and a universal merger efficiency factor of $p_{BH} = 0.17$ inferred from IPTA data, as discussed in the text.

to detect the inspiral signals with LISA. The shading in the panels correspond to the number of events whose parameters fall within each rectangular bin. The distribution peaks below $\mathcal{M}_0 = 10^4 M_\odot$ reflecting the binary population, that increases towards lower masses, and the cut-off at $10^3 \, M_\odot$. Similarly, because the GW signal (Eq. (10)) from inspirals does not explicitly depend on η, the η distribution in the right panel of Fig. 6 arises solely from the binary population. The binary merger rate peaks at $4 < z < 6$ (see Fig. 1) but LISA cannot spot the majority of the binaries beyond $z \sim 4$.

We find that, at $z < 10$, there are a total of $O(10^7)$ near-monochromatic sources in the frequency range of $10^{-5} \, \text{Hz} < f < 0.1 \, \text{Hz}$[7]. As only a small fraction of these can be resolved by LISA, we expect that the rest constitute a significant stochastic GW background. It should be noted that the potential redshift dependence of p_{BH} introduces additional uncertainties in the merger rate at higher redshifts ($z \gtrsim 3$), which is not well probed by the PTA measurements. We leave a detailed study of this background and its detectability with LISA for future work.

Figure 7 illustrates AEDGE and LISA prospects for observing GW events from BHs less than a day after the merger. The expected total number of detectable events is slightly larger for LISA than for AEDGE. Due to the different frequency ranges, LISA can spot heavier binaries, $\mathcal{M}_0 > 10^4 \, M_\odot$, while AEDGE can see lighter binaries, $\mathcal{M}_0 < 2 \times 10^3 \, M_\odot$. LISA could (in principle) detect even heavier mergers, namely, of $\mathcal{M}_0 \lesssim 10^8 \, M_\odot$, but since such mergers are so rare it is unlikely that LISA will see any mergers above $10^6 \, M_\odot$. The expected number of detectable mergers peaks at $4 < z < 6$ for both of the experiments. For AEDGE, the z distribution of the detectable mergers reflects the binary merger rate (see Fig. 1) while for LISA the z distribution peaks at slightly lower z because the SNR of $\mathcal{M}_0 \sim 10^4$ binaries for LISA is not as high as it is for AEDGE. The lower panels show that, for all detectors, the BH mass ratio peaks at the highest values. For AION-km and ET we find that the total expected number of events per year is less than three.

We note that most of the IMBH events detectable by AEDGE during the last day prior to the merger (see the left panels in Fig. 7) will also have been detected by LISA during the previous infall stage, as seen in Fig. 6. This opens up prospects for

using LISA data to predict when and in what direction AEDGE will observe IMBH mergers, sharpening tests of general relativity and giving advance warnings for searches for possible multi-messenger signals.

We caution that there are observational uncertainties in the low mass cut-off in Eq. (6), due to the difficulty of measuring very faint active galactic nuclei (AGNs) or inactive BHs in dwarf galaxies. As mentioned above, surveys such as eRASS and AMUSE (Miller et al. 2015) are already constraining this low-mass region and are compatible with the assumed cut-off, although a heavier mass cut-off may be favored (Chadayammuri et al. 2023). This could easily be achieved in models with modified initial fluctuation spectra (Hütsi et al. 2023).

The cut-off mass is tightly related to the SMBH formation mechanism since a BH in the halo centre cannot be lighter than the seed that originated the growth. For a full review, we refer to Volonteri et al. (2021). In this sense, the cut-off we have considered assumes the existence of some light seeds, which augment the possible GW signal. Independently of the growth through halo merging, IMBHs can also be born inside dense stellar media like nuclear and globular star clusters. The growth from stellar masses happens because of repeated encounters inside these dense environments. Estimating the GW emission from such encounters is an active field of research (Fragione & Kocsis 2018; Fragione et al. 2019; Fragione & Loeb 2023) and their signals may contribute significantly to the expected number of IMBH mergers.

4. Conclusions

Here, we describe the results from a model for the PTA nHz common-process signal based on a simulation of massive BH mergers. The magnitude of the signal depends on the merger probability, p_{BH}, which is a product of the probabilities that a pair of halos contain massive BHs and the probability that they will merge. We sidestepped the considerable uncertainties in modeling these probabilities by fitting a mass-independent value of p_{BH} to the PTA signal, finding $p_{BH} = 0.17$, with a factor of two uncertainty. With this assumption, the dominant contribution to the PTA signal is made by mergers with total masses $>10^9 \, M_\odot$ and about 10% from masses $<10^9 \, M_\odot$. The PTA mergers would have redshifts $O(1)$ and have mass asymmetries $\gtrsim 10$.

---

[7] We emphasize, however, that there could be lower-mass BHs that are remnants of Population III stars, which are not included in our analysis.
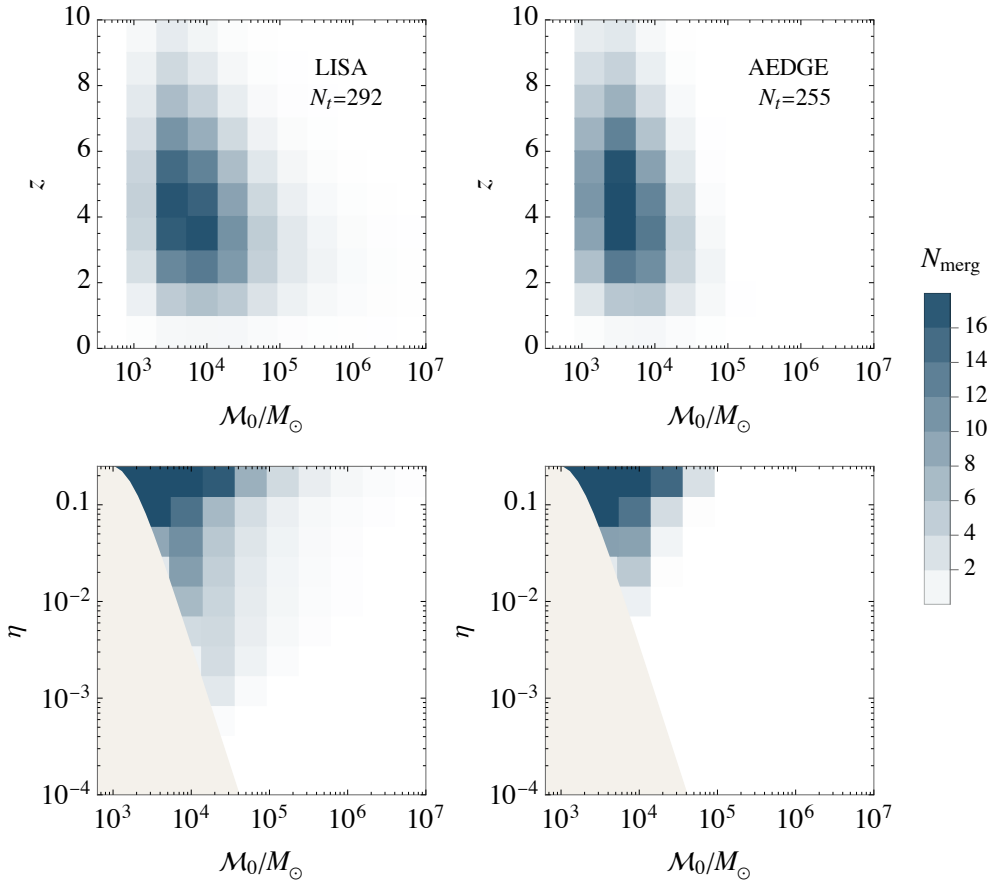
Fig. 7. Expected numbers of detectable GW events generated during the last day before the merger in a year of observation by LISA (left panels) and AEDGE (right panels) as functions of the chirp mass $\mathcal{M}_0$ and redshift z (upper panels) and of chirp mass and symmetric mass ratio η (lower panels). The merger rates assume the abundance of IMBHs described by Eq. (6) for BHs heavier than $10^3 \, M_\odot$ and a universal merger efficiency factor of $p_{BH} = 0.17$ inferred from the IPTA data.

The number of mergers contributing most of the PTA signal at frequencies $O(10)$ nHz is limited. Consequently, the frequency spectrum becomes quite irregular, the spectral index may deviate from the analytic value of 2/3, individual mergers may be distinguished, and there may be detectable circular polarization. These will be interesting targets for future experiments in the nHz range, including PTAs and SKA.

Assuming the same mass-independent value of p_{BH} as for the PTA signal, there would be observable signals from mergers with total masses $\in (10^3, 10^6) \, M_\odot$ in the LISA experiment and from mergers with total masses $\in (10^3, 10^5) \, M_\odot$ in the AEDGE experiment. Data from these experiments will be able to check the accuracy of the constancy of p_{BH}. In principle, their signals could be even larger than our estimates if p_{BH} is closer to unity for masses below the PTA range, or if there is a significant GW contribution seeded by Population III stars. However, a lower value for p_{BH} and, hence, a lower event rate cannot be excluded.

References

Ajith, P., Babak, S., Chen, Y., et al. 2008, Phys. Rev. D, 77, 104017
Antoniadis, J., Arzoumanian, Z., Babak, S., et al. 2022, MNRAS, 510, 4873
Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2020, ApJ, 905, L34
Arzoumanian, Z., Baker, P. T., Blecha, L., et al. 2023, ApJL, submitted [arXiv:2301.03608]
Ashoorioon, A., Rezazadeh, K., & Rostami, A. 2022, Phys. Lett. B, 835, 137542
Badurina, L., Bentine, E., Blas, D., et al. 2020, JCAP, 05, 011
Badurina, L., Buchmueller, O., Ellis, J., et al. 2021, Phil. Trans. A. Math. Phys. Eng. Sci., 380, 20210060
Barkana, R., & Loeb, A. 2001, Phys. Rept., 349, 125
Bécsy, B., Cornish, N. J., & Kelley, L. Z. 2022, ApJ, 941, 119
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307
Belgacem, E., & Kamionkowski, M. 2020, Phys. Rev. D, 102, 023004
Benson, A. J., Kamionkowski, M., & Hassani, S. H. 2005, MNRAS, 357, 847
Berti, E., Cardoso, V., & Will, C. M. 2006, Phys. Rev. D, 73, 064030
Blanco-Pillado, J. J., Olum, K. D., & Wachter, J. M. 2021, Phys. Rev. D, 103, 103512
Blas, D., & Jenkins, A. C. 2022, Phys. Rev. Lett., 128, 101103
Blasi, S., Brdar, V., & Schmitz, K. 2021, Phys. Rev. Lett., 126, 041305
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Boroson, T. A., & Lauer, T. R. 2009, Nature, 458, 53
Buchmuller, W., Domcke, V., & Schmitz, K. 2020, Phys. Lett. B, 811, 135914
Chadayammuri, U., Bogdan, A., Ricarte, A., & Natarajan, P. 2023, ApJ, 946, 51
Chen, S., Middleton, H., Sesana, A., Del Pozzo, W., & Vecchio, A. 2017, MNRAS, 468, 404
Chen, S., Caballero, R. N., Guo, Y. J., et al. 2021, MNRAS, 508, 4970
Conneely, C., Jaffe, A. H., & Mingarelli, C. M. F. 2019, MNRAS, 487, 562
De Luca, V., Franciolini, G., & Riotto, A. 2021, Phys. Rev. Lett., 126, 041303
De Paolis, F., Ingrosso, G., & Nucita, A. A. 2002, A&A, 388, 470
De Paolis, F., Ingrosso, G., & Nucita, A. A. 2004, A&A, 426, 379
Dodelson, S. 2003, Modern Cosmology (Amsterdam: Academic Press)
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
El-Neaj, Y. A., Alpigiani, C., Amairi-Pyka, S., et al. 2020, EPJ Quant. Technol., 7, 6
Ellis, J., & Lewicki, M. 2021, Phys. Rev. Lett., 126, 041304
Enoki, M., & Nagashima, M. 2007, Prog. Theor. Phys., 117, 241

Enoki, M., Inoue, K. T., Nagashima, M., & Sugiyama, N. 2004, ApJ, 615, 19
Erickcek, A. L., Kamionkowski, M., & Benson, A. J. 2006, MNRAS, 371, 1992
Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2019, ApJ, 875, L1
Event Horizon Telescope Collaboration (Akiyama, K., et al.) 2022, ApJ, 930, L12
Finn, L. S., & Chernoff, D. F. 1993, Phys. Rev. D, 47, 2198
Fragione, G., & Kocsis, B. 2018, Phys. Rev. Lett., 121, 161103
Fragione, G., & Loeb, A. 2023, ApJ, 944, 81
Fragione, G., Grishin, E., Leigh, N. W. C., Perets, H. B., & Perna, R. 2019, MNRAS, 488, 47
Gerosa, D., Ma, S., Wong, K. W. K., et al. 2019, Phys. Rev. D, 99, 103004
Goncharov, B., Shannon, R. M., Reardon, D. J., et al. 2021, ApJ, 917, L19
Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, Nature, 518, 74
Graham, P. W., Hogan, J. M., Kasevich, M. A., Rajendran, S., & Romani, R. W. 2017, ArXiv e-prints [arXiv:1711.02225]
Greene, J. E., Strader, J., & Ho, L. C. 2020, ARA&A, 58, 257
Hellings, R. W., & Downs, G. S. 1983, ApJ, 265, L39
Hogan, J. M., Johnson, D. M. S., Dickerson, S., et al. 2011, Gen. Rel. Grav., 43, 1953
Holgado, A. M., Sesana, A., Sandrinelli, A., et al. 2018, MNRAS, 481, L74
Hotinli, S. C., Kamionkowski, M., & Jaffe, A. H. 2019, Open J. Astrophys., 2, 8
Hütsi, G., Raidal, M., Urrutia, J., Vaskonen, V., & Veermäe, H. 2023, Phys. Rev. D, 107, 043502
Iguchi, S., Okuda, T., & Sudou, H. 2010, ApJ, 724, L166
Izquierdo-Villalba, D., Sesana, A., Bonoli, S., & Colpi, M. 2021, MNRAS, 509, 3488
Janssen, G., Hobbs, G., McLaughlin, M., et al. 2015, PoS, AASKA14, 037
Kaiser, A. R., & McWilliams, S. T. 2021, CQG, 38, 055009
Kato, R., & Soda, J. 2016, Phys. Rev. D, 93, 062003
Kelley, L. Z., Blecha, L., & Hernquist, L. 2017a, MNRAS, 464, 3131
Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2017b, MNRAS, 471, 4508
King, O. G., Hovatta, T., Max-Moerbeck, W., et al. 2013, MNRAS, 436, 114
Kocsis, B., & Sesana, A. 2011, MNRAS, 411, 1467
Kohri, K., & Terada, T. 2021, Phys. Lett. B, 813, 136040
Kovačević, A. B., Songsheng, Y.-Y., Wang, J.-M., & Popovic, L. C. 2022, A&A, 663, A99
Lacey, C. G., & Cole, S. 1993, MNRAS, 262, 627
Lewicki, M., & Vaskonen, V. 2021, Eur. Phys. J. C, 83, 168
LIGO Scientific & Virgo Collaborations (Abbott, B. P., et al.) 2016, Phys. Rev. Lett., 116, 061102
LIGO Scientific, Virgo & KAGRA Collaborations (Abbott, R., et al.) 2021, ArXiv e-prints [arXiv:2111.03634]
Lippai, Z., Frei, Z., & Haiman, Z. 2009, ApJ, 701, 360
LISA Collaboration (Amaro-Seoane, P., et al.) 2017, ArXiv e-prints [arXiv:1702.00786]
Merritt, D. 2013, CQG, 30, 244005
Miller, B. P., Gallo, E., Greene, J. E., et al. 2015, ApJ, 799, 98
O'Neill, S., Kiehlmann, S., Readhead, A. C. S., et al. 2022, ApJ, 926, L35
Peters, P. C. 1964, Phys. Rev., 136, B1224
Peters, P. C., & Mathews, J. 1963, Phys. Rev., 131, 435
Phinney, E. S. 2001, ArXiv e-prints [arXiv:astro-ph/0108028]
Planck Collaboration VI. 2020, A&A, 641, A6
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Rieger, F. M., & Mannheim, K. 2000, A&A, 359, 948
Sathyaprakash, B., Abernathy, M., Acernese, F., et al. 2012, CQG, 29, 124013
Sato-Polito, G., & Kamionkowski, M. 2022, Phys. Rev. D, 106, 023004
Sesana, A., Haardt, F., Madau, P., & Volonteri, M. 2004, ApJ, 611, 623
Sesana, A., Vecchio, A., & Colacino, C. N. 2008, MNRAS, 390, 192
Sesana, A., Vecchio, A., & Volonteri, M. 2009, MNRAS, 394, 2255
Taylor, S. R., Huerta, E. A., Gair, J. R., & McWilliams, S. T. 2016, ApJ, 817, 70
Valbusa Dall'Armi, L., Nishizawa, A., Ricciardone, A., & Matarrese, S. 2023, Phys. Rev. Lett., 131, 041401
Valtonen, M. J., et al. 2008, Nature, 452, 851
Vaskonen, V., & Veermäe, H. 2021, Phys. Rev. Lett., 126, 051303
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, Nature, 509, 177
Volonteri, M., Habouzit, M., & Colpi, M. 2021, Nat. Rev. Phys., 3, 732
Wang, Y., Pardo, K., Chang, T.-C., & Doré, O. 2021, Phys. Rev. D, 103, 084007
Wyithe, J. S. B., & Loeb, A. 2003, ApJ, 590, 691

Appendix A: Statistics of the GW energy spectrum and circular polarization

A.1. Distribution of $\Omega_{GW}(f)$

The GW energy spectrum arising from a population of BH binaries can be expressed as the sum of the contributions from individual binaries. We first consider the contribution to Ω_{GW} coming from a single inspiralling binary:

$$\Omega_{GW}^{(1)}(\boldsymbol{\theta}) = \left| \frac{d\tau}{d\ln f} \right|^{-1} \frac{2\pi}{5} \frac{f^3 |\tilde{h}(f)|^2}{\rho_c}, \tag{A.1}$$

where $\boldsymbol{\theta} \equiv \{\mathcal{M}_z, z, f\}$ are parameters describing the binary. The GW energy spectrum arising from the inspiralling binary population (or any of its subpopulations) arises from the sum of nearly monochromatic components

$$\Omega_{GW}(f) \approx \sum_i \Omega_{GW}^{(1)}(\boldsymbol{\theta}_i)\delta(\ln f/f_i). \tag{A.2}$$

This approximation is valid as long as the change in frequency over the observation period is smaller than the spectral resolution. The parameters $\boldsymbol{\theta}_i$ as well as $\Omega_{GW}^{(1)}$ are independent and identically distributed. Thus, the statistical properties of $\Omega_{GW}(f)$ can be inferred from the distribution of $\Omega_{GW}^{(1)}(f)$, which is given by

$$\begin{aligned} P^{(1)}(\Omega|f) &= \frac{1}{n(f)} \int d\lambda \left| \frac{d\tau}{d\ln f} \right| \delta\left(\Omega - \Omega_{GW}^{(1)}\right) \\ &= \frac{\mathcal{F}(\Omega f^{-\frac{10}{3}})}{f\Omega^{\frac{3}{2}} n(f)}, \end{aligned} \tag{A.3}$$

where we define the following function:

$$\begin{aligned} \mathcal{F}(x) \equiv \int dz d\eta &\left[\frac{\partial \lambda(\mathcal{M}_z, z)}{\partial \mathcal{M}_z \partial \eta \partial z} \right. \\ &\left. \times \frac{\mathcal{M}_z}{16\sqrt{10\rho_c}D_L} \right]_{\mathcal{M}_z = \frac{1}{\pi}\left(\frac{5\pi}{8}D_L^3\rho_c x\right)^{\frac{3}{10}}} \end{aligned} \tag{A.4}$$

and the spectral source density:

$$n(f) \equiv \int d\mathcal{M}_z d\eta dz \frac{dN}{d\mathcal{M}_z d\eta dz d\ln f}. \tag{A.5}$$

We see that, for circular inspiralling binaries, the four-parameter distribution (11) can be reduced to two functions of a single parameter. Further reductions are unlikely, as these functions depend on the model of the binary merger rate, which we assume to stem from the merger rate of galaxies.

The average energy spectrum in a frequency bin (f_j, f_{j+1}) can be obtained from

$$\begin{aligned} \Omega_{GW}(f_j) &\equiv \frac{1}{\ln(f_{j+1}/f_j)} \int_{f_j}^{f_{j+1}} \Omega_{GW}(\boldsymbol{\theta}; f) d\ln f \\ &= \frac{1}{\ln(f_{j+1}/f_j)} \sum_{i=1}^{N(f_j)} \Omega_{GW,i}^{(1)}, \end{aligned} \tag{A.6}$$

where the number of binaries $N(f_j)$ is drawn from a Poisson distribution with the expected value,

$$\bar{N}(f_j) = \int_{f_j}^{f_{j+1}} d\ln f\, n(f), \tag{A.7}$$

and the $\Omega_{GW,i}^{(1)}$ are drawn from a random distribution $P^{(1)}(\Omega, f) = P^{(1)}(\Omega|f)n(f)/\bar{N}(f_j)$ with $f \in (f_j, f_{j+1})$. In the following, we suppress f_j.

The moment generating function of Ω_{GW} is given by:

$$\begin{aligned} M_{\Omega_{GW}}(s) &\equiv \langle \exp(s\Omega_{GW}) \rangle \\ &= \sum_{N \geq 0} p_N \prod_{i=1}^N \left\langle \exp\left(s\Omega_{GW}^{(1)}\right) \right\rangle \\ &= \sum_{N \geq 0} \frac{\bar{N}^N}{N!} e^{-\bar{N}} \left(M_{\Omega_{GW}^{(1)}}(s)\right)^{\bar{N}} \\ &= \exp\left[\bar{N}\left(M_{\Omega_{GW}^{(1)}}(k) - 1\right)\right]. \end{aligned} \tag{A.8}$$

Thus, the cumulant generating function of $\Omega_{GW}(f_j)$, that is,

$$\begin{aligned} K_{\Omega_{GW}}(s) &\equiv \ln M_{\Omega_{GW}}(s) \\ &= \int_{f_j}^{f_{j+1}} dN \left[\exp(s\,\Omega_{GW}^{(1)}) - 1\right]. \end{aligned} \tag{A.9}$$

is proportional to the generating function of $M_{\Omega_{GW}^{(1)}}(s)$, implying that the nth cumulant of Ω_{GW} is

$$\kappa_n[\Omega_{GW}] = \bar{N}\langle(\Omega_{GW}^{(1)})^n\rangle = \int_{f_j}^{f_{j+1}} dN(\Omega_{GW}^{(1)})^n. \tag{A.10}$$

Therefore, the mean $\kappa_1(\Omega_{GW}) \equiv \langle\Omega_{GW}\rangle$ matches the expectation value given in Eq. (8), while the variance,

$$\kappa_2(\Omega_{GW}) \equiv \langle(\delta\Omega_{GW})^2\rangle = \bar{N}\langle(\Omega_{GW}^{(1)})^2\rangle \tag{A.11}$$

is divergent due to the long $P^{(1)}(\Omega) \propto \Omega^{-\frac{5}{2}}$ tail (16), unless a minimal distance to the closest BH binary is imposed. In particular, as we will demonstrate briefly, this long tail is preserved in the distribution of Ω_{GW}, i.e., the higher cumulants are not diminished when summing several instances of $P^{(1)}(\Omega) \propto \Omega^{-\frac{5}{2}}$.

The moment-generating function $M_{\Omega_{GW}}(-s)$ is a Laplace transform of the probability distribution. Thus as an alternative to the Monte Carlo approach adopted in the main text, the probability distribution of $\Omega_{GW}(f_j)$ can be obtained by an inverse Laplace transform of the moment generating function given in Eq. (A.8). It can be expressed as

$$\begin{aligned} P(\Omega_{GW}) = &\frac{1}{2\pi\bar{N}} \int_{-\infty}^{\infty} ds\, e^{is\Omega_{GW}/\bar{N}} \times \\ &\times \exp\left(\bar{N}\left\langle \exp\left(-is\Omega_{GW}^{(1)}/\bar{N}\right) - 1\right\rangle\right), \end{aligned} \tag{A.12}$$

where the average is taken over the single event distribution $P^{(1)}$.

The large N limit. — The central limit theorem (and Eq. (A.11)) dictates that when $N \to \infty$, then the distribution (A.6) should approach a Gaussian with its relative width scaling as $\sigma_{\Omega_{GW}}/\langle\Omega_{GW}\rangle \propto 1/\sqrt{N}$. However, this is not the case when the second cumulant diverges. Considering a generic distribution with a tail (droping the "GW" subindex for the sake of brevity):

$$P^{(1)}(\Omega) \overset{\Omega \to \infty}{\sim} C\Omega^{-\frac{5}{2}}, \tag{A.13}$$

where C is a constant. The second moment $\langle\Omega^2\rangle$ diverges. We assume that the small Ω behavior is such that the first moment $\bar{\Omega}^{(1)} \equiv \langle\Omega^{(1)}\rangle$ is finite. The $\bar{N} \to \infty$ asymptotic of the average in Eq. (A.12) can then be computed by separating it into the mean

and a term for which the expectation value can be determined by approximating the distribution by its tail (A.13):

$$
\bar{N}\left\langle \exp\left(-is\frac{\Omega^{(1)}}{\bar{N}}\right) - 1 \right\rangle
$$
$$
= -is\bar{\Omega}^{(1)} + \bar{N}\left\langle \exp\left(-is\frac{\Omega^{(1)}}{\bar{N}}\right) - 1 + is\frac{\Omega^{(1)}}{\bar{N}} \right\rangle
$$
$$
\overset{N\to\infty}{\sim} -is\bar{\Omega}^{(1)} + e^{i\pi\frac{3}{4}}\frac{\sqrt{\pi}4C}{3\sqrt{\bar{N}}}s^{\frac{3}{2}}. \tag{A.14}
$$

In summary, the $N \to \infty$ distribution asymptotes to

$$
P(\Omega) \approx (C\bar{N})^{-\frac{2}{3}}\mathcal{P}\left((C\bar{N})^{-\frac{2}{3}}(\Omega - \bar{\Omega})\right), \tag{A.15}
$$

where $\bar{\Omega} \equiv \bar{N}\bar{\Omega}^{(1)}$ denotes the expectation value of Ω and

$$
\mathcal{P}(x) \equiv \text{Re}\frac{e^{i\pi/6}}{\pi}\int_0^\infty ds\,\exp\left(ie^{i\pi/6}xs - \frac{4\sqrt{\pi}}{3}s^{\frac{3}{2}}\right) \tag{A.16}
$$

is a universal function that is independent of the single event distribution – all information about the latter enters through $\bar{\Omega}^{(1)}$ and C. This limiting case reveals a few crucial features:

1. Since $\mathcal{P}(x) \sim x^{-5/2}$ when $x \to \infty$, the power-law behavior of the tail is preserved in the large $\bar{N}$ limit:

$$
P(\Omega) \overset{\Omega\gg\bar{\Omega}}{\approx} \frac{C\bar{N}}{(\Omega - \bar{\Omega})^{\frac{5}{2}}}. \tag{A.17}
$$

2. When the number of contributing events N is increased by a factor of A to AN, the distribution evolves as

$$
P^{(AN)}(\Omega) = A^{-\frac{2}{3}}P^{(N)}\left(A^{-\frac{2}{3}}(\Omega - \bar{\Omega}^{(AN)}) + \bar{\Omega}^{(N)}\right). \tag{A.18}
$$

3. Although the variance diverges, we can define the width of the distribution $\Delta_\Omega \equiv \Omega_+ - \Omega_-$ from the confidence interval (Ω_-, Ω_+) centred around, e.g., the median. Eq. (A.18) then implies the scaling

$$
\Delta_\Omega \propto \bar{N}^{2/3}. \tag{A.19}
$$

So, the relative width $\Delta_\Omega/\bar{\Omega}$ will approach zero, but at a slower pace than when the central limit theorem applies.

4. Using the method of steepest decent, we find that

$$
\mathcal{P}(x) \overset{\Omega\ll\bar{\Omega}}{\approx} \frac{\sqrt{\bar{\Omega} - \Omega}}{2\pi C\bar{N}}e^{-\frac{(\bar{\Omega}-\Omega)^3}{12\pi(C\bar{N})^2}}, \tag{A.20}
$$

implying values smaller than the mean, that is $\Omega < \bar{\Omega}$, are much less likely than for Gaussian distributions, as they are suppressed by an exponent of a cube.

As an illustration, we consider a population of sources for which individual signals follow the simple long-tailed distribution

$$
P^{(1)}(\Omega) \propto \Omega^{-\frac{5}{2}}\theta(\Omega - \Omega_{\min}), \tag{A.21}
$$

which mimics the statistics of Ω_{GW} from a realistic BH population found in section 3.2. In Fig. A.1, we show the resulting distributions of $\Omega/\bar{\Omega}$ if the signal consists of 10, 100, and 10^5 sources. The mean is given by $\bar{\Omega} \equiv 3N\Omega_{\min}$. The distribution is computed by generating random realizations of the source populations (shown by points) and by using (A.12) (shown by solid lines). These two approaches are in excellent agreement. As expected, one can observe that the total signal will inherit the long $\Omega^{-5/2}$ tail from the distribution of individual sources (shown
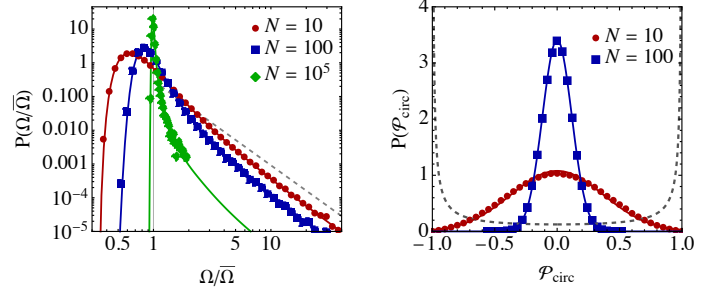
Fig. A.1. Distributions of the total signal strengths from 1, 10, 100, and 10^5 sources whose independent signal strengths have identical $\Omega^{-\frac{5}{2}}$ tails (left). Distribution of $\mathcal{P}_{\text{circ}}$ for 1, 10 and 100 identical sources (right). The dashed line shows the distribution from a single source, the solid lines show the analytic estimates, and the distributions obtained from an explicit Monte Carlo analysis are shown by the points.

by the dashed line). At large Ω, the tail approaches Eq. (A.17) in all cases, while the asymptotic (A.15) works well for $N = 10^5$, but predicts a slightly too wide peak for $N \le 10^5$. As predicted, the peak gets narrower and moves closer to the expectation value, when N is increased. We note, however, that although the shape of the simplified distribution in Fig. A.1 resembles the distribution in Fig. 5 obtained from a model of the heavy BH binary population, the shape around the peak in the latter depends also on the z dependence and the binary mass distribution.

A.2. Distribution of $\mathcal{P}_{\text{circ}}$

Considering now the distribution of the circular polarization

$$
\mathcal{P}_{\text{circ}} \equiv \frac{\sum_{i=1}^N V_i}{\sum_{i=1}^N I_i}, \tag{A.22}
$$

which depends on the distribution and number N of sources. We denote the contributions from individual sources by V_i, I_i given by Eqs. (18-20). As above, we are working in a fixed frequency bin and, thus, suppressing the frequency dependence in the arguments of the functions.

The polarization will vanish on average, that is, at $\langle\mathcal{P}_{\text{circ}}\rangle = 0$, but polarization measurements can provide information from the size of the deviation from that expectation. For a single dominant source, the contribution from the amplitude drops out and we find

$$
\mathcal{P}_{\text{circ}} = \frac{4c_\theta(1 + c_\theta^2)}{1 + 6c_\theta^2 + c_\theta^4}, \tag{A.23}
$$

where $c_\theta \equiv \cos(\theta)$. The corresponding probability distribution is shown as the dashed curve in Fig. A.1 and can be seen to be dominated by the region $|\mathcal{P}_{\text{circ}}| \approx 1$.

If the signal is dominated by several sources, the analytic treatment of $\mathcal{P}_{\text{circ}}$ is complicated due to correlations between V and I. The mean and variance of I and V are given by

$$
\langle I\rangle = \frac{4}{5}\langle|h|^2\rangle, \qquad \langle V\rangle = 0,
$$
$$
\langle\delta I^2\rangle = \frac{284}{315}\langle(\delta|h|^2)^2\rangle, \quad \langle\delta V^2\rangle = \frac{92}{105}\langle(\delta|h|^2)^2\rangle, \tag{A.24}
$$

where we used the fact that the inclination is independent of other parameters of the binary and the fluctuations in $|h|^2$ are computed as in (A.11). The latter is expected to be large. To

estimate the fluctuations in $\mathcal{P}_{\text{circ}}$ in the limit when the number of sources is large, we first assume that I can be replaced by its average over inclinations $\langle I \rangle_\theta \equiv (4/5) \sum_i |h_i|^2$, but not over signal strengths, so that

$$\mathcal{P}_{\text{circ}} \approx V/\langle I \rangle_\theta \equiv \sum_{i=1}^{N} p_i f_V(c_{\theta,i}), \tag{A.25}$$

where $p_i \equiv (4/5)|h_i|^2/\langle I \rangle_\theta \in [0, 1]$ is the fractional contribution to the signal from source i and $f_V(c_\theta) = (5/4)c_\theta(1 + c_\theta^2)$ contains the inclination dependence of V. In this way, it is possible to contain the effect of the large fluctuations in $\langle I \rangle_\theta$ in the random fractions p_i, which vary only in the range $[0, 1]$. We find:

$$\langle \delta \mathcal{P}_{\text{circ}}^2 \rangle = \frac{115}{84} N \langle |p_i|^2 \rangle. \tag{A.26}$$

We note that $N\langle |p_i|^2 \rangle$ takes values in the range $[1/N, 1]$ and could be interpreted as a measure of the effective number of dominant sources $1/N_{\text{dom}}$. For instance, in the limiting case where the signal is sourced by N binaries that contribute equally, that is, $p_i = 1/N$, we would obtain

$$\sigma_{\mathcal{P}_{\text{circ}}} = 1.17/\sqrt{N}, \tag{A.27}$$

as also stated in Eq. (22). In the right panel of Fig. A.1, approximate Gaussian distributions with a width given by (A.27) and shown as solid lines are compared to distributions obtained from explicitly generated random populations of $\mathcal{P}_{\text{circ}}$ arising from exactly 10 and 100 sources of equal strength (shown by points). We see an excellent agreement between the naive Gaussian approximation and the Monte Carlo estimate already for $N = 10$.

Finally, we remark that the approximation (A.25) above can be improved by expanding $\mathcal{P}_{\text{circ}}$ in $\delta I \equiv I - \langle I \rangle_\theta$, that is, $\mathcal{P}_{\text{circ}} = V/\langle I \rangle_\theta(1 - \delta I/\langle I \rangle_\theta + \ldots)$. This allows us to account for correlations between V and δI, for instance, when computing the variance (A.26),

$$\langle \delta \mathcal{P}_{\text{circ}}^2 \rangle = \frac{115}{84} N \langle |p_i|^2 \rangle - \frac{2855}{1386} N \langle |p_i|^3 \rangle + \ldots \tag{A.28}$$

The corrections arising from higher powers of $\delta I/\langle I \rangle_\theta$ are suppressed by increasing powers of $1/N$.