A boosted gravitational wave background for primordial black holes with broad mass distributions and thermal features

Eleni Bagui$^1$ and Sébastien Clesse$^1$

$^1$Service de Physique Théorique, Université Libre de Bruxelles (ULB), Boulevard du Triomphe, CP225, 1050 Brussels, Belgium.

(Dated: January 18, 2023)

Primordial black holes (PBHs) with a wide mass distribution imprinted by the thermal history of the Universe, which naturally produces a high peak at the solar mass scale, could explain the gravitational-wave events seen by LIGO/Virgo and up to the totality of the dark matter. We show that compared to monochromatic or log-normal mass functions, the gravitational wave backgrounds (GWBs) from early PBH binaries and from late binaries in clusters are strongly enhanced at low frequency and could even explain the NANOGrav observations. This enhancement comes from binaries with very low mass ratios, involving solar-mass and intermediate-mass PBHs at low frequency, solar-mass and subsolar-mass at high frequency. LISA could distinguish the various models, while in the frequency band of ground-based detectors, we find that the GWB from early binaries is just below the current LIGO/Virgo limits and above the astrophysical background, if they also explain black hole mergers. The GWB from binaries in clusters is less boosted but has a different spectral index than for neutron stars, astrophysical black holes or early PBH binaries. It is detectable with Einstein Telescope or even with the LIGO/Virgo design sensitivity.

PACS numbers:

I. INTRODUCTION

Since their first detection in September 2015 [1], gravitational waves (GWs) have revealed some intriguing properties of black holes: larger masses than in X-ray binaries, low effective spins, the existence of black holes in the low mass gap or in the pair instability mass gap and binaries with low mass ratios [2, 3]. In particular, GW190521 [4] involved at least one black hole in the pair instability mass gap. It could be a secondary merger from a dense cluster [5] but in such a case, it remains challenging to obtain small enough kick velocities for such a black hole to remain gravitationally bound to the host cluster. Then, GW190814 [4] combines a primary black hole with an almost zero spin ($\chi_1 < 0.07$ at 90\% C.L.) and a probable black hole in the low mass gap. The merging rates for such event are of the same order than for equal-mass binaries, making it hard to explain with the current stellar models. In addition, recent microlensing observations by OGLE and Gaia [6] towards the galactic center have also revealed a possible population of black holes in the low mass gap.

Stellar evolution scenarios may be able to explain all those observations, by invoking different black hole populations and binary formation mechanisms. An alternative attracting a lot of attention is a primordial origin, first suggested in [7–9]. Such primordial black holes (PBHs) may explain a significant fraction or even the totality of the dark matter (DM) in the Universe. PBHs may have formed in the early Universe with almost vanishing spins due to the gravitational collapse of pre-existing large inhomogeneities [10–12], for instance produced during inflation.

Contrary to stellar black holes, PBHs could have any mass above $10^{11}\, \text{kg}$ (otherwise they would have evaported today through Hawking radiation), but the QCD transition provides a natural explanation for a strong peak in their mass function at the solar-mass scale [13–16], as well as a bump around $30\!–\!100\, M_\odot$ [15]. This could explain the rates, masses and spins of the LIGO/Virgo detections [16–19]. These features generically arise from the known thermal history of the Universe and change the equation of state when quarks condensed into protons, neutrons and pions.

If they compose a significant fraction of the DM, the induced Poisson fluctuations in the PBH density must have seeded PBH clusters at high redshifts [20, 21], resulting in strong modifications of the halo mass function. This has four interesting effects. First, PBH clusters induce a suppression of the merging rates of early binaries, below the current limits [22, 23]. Second, such an effect can explain the observed spatial correlations between the infrared and the source subtracted X-ray backgrounds [20, 24]. Third, the merging rate of PBH binaries, formed by tidal capture in those clusters, can be at the same level than the rates inferred from GW observations [17]. Fourth, most of the PBHs in the halos of massive galaxies should be regrouped in $10^9\!–\!10^9\, M_\odot$ clusters, which implies that the point sources monitored in microlensing surveys are gravitationally lensed by these clusters. This reduces the amplitude of microlensing events below the level of detectability, which in turn suppresses the limits on the PBH abundance, allowing $f_{\text{PBH}} = 1$ at the solar mass scale [16]. The clustering of PBHs is therefore unavoidable (Poisson fluctuations are linked to their discrete nature) and essential to evade microlensing and GW limits, if stellar-mass PBHs significantly contribute to the DM.

If some of the detected black holes were primordial, this would suggest a wide mass distribution, including both subsolar and intermediate mass black holes. Be-
sides the dark matter, such wide-mass models imprinted by thermal features could explain a series of other puzzling observations \bib{16}{25}, including microlensing of stars towards the galactic center due to planetary-mass compact objects \bib{26} and of quasars misaligned with the lensing galaxy \bib{27}. This could also explain the critical size of ultra-faint dwarf galaxies, the core profile of dwarf galaxies, the existence of supermassive black holes at high redshifts and even the baryon asymmetry of the Universe \bib{28}{29}. Studying and searching for other signatures of wide-mass PBH models is therefore important to help distinguish the primordial or stellar origin of black holes. The gravitational-wave background (GWB) from PBH binaries is one of them.

In this paper, we compute for the first time the GWB associated to PBH binaries in wide-mass models including thermal features. Our analysis therefore extends a previous work that assumed monochromatic and log-normal mass distributions \bib{30}. We also consider both early binaries and late-time binaries formed by tidal capture in PBH clusters. Compared to previous results, we find important differences that are linked to the existence of a significant fraction of light and heavy PBHs in the mass function. They can combine with PBHs from the QCD peak and form binaries with extreme mass-ratios, contributing to the GWB. This importantly boosts the amplitude of the GWB at frequencies relevant for LISA and Pulsar Timing Arrays (PTAs). We find that it can be compatible with the GWB possibly detected by NANOGrav \bib{31} and more recently by PPTA \bib{32}, for binaries in clusters. A second important finding is the change of the GWB amplitude and spectral index at frequencies relevant for LIGO/Virgo/Kagra (LVK). The total signal, dominated by early binaries, is found to be within the range of the next observing runs by LVK and above expectations for neutron star and astrophysical black hole binaries.

Compared to \bib{22}{30} \bib{33} \bib{36}, our calculations include several improvements: we take into account the rate suppression of early binaries due to Poisson fluctuations and nearby PBHs; we consider merging rates that are consistent with latest GW observations, in particular the ones of GW190425, GW190521 and GW190814. Most importantly, we consider a wide mass function with thermal effects.

The paper is organized as follows: in Section \ref{sec:GWB} we briefly introduce the formalism used to obtain the characteristic strain of the GWB from BH mergers in the Newtonian limit. In Section \ref{sec:PBH} we present the broad PBH mass spectrum impacted by the thermal history of the Universe. Section \ref{sec:merging} focuses on the PBH merging rates of early binaries as well as late-time binaries formed by capture in clusters. In Sections \ref{sec:ground} and \ref{sec:space} we compute the GWB for our two binary formation channels and discuss in Sections \ref{sec:PTA} \ref{sec:LVK} \ref{sec:AAS} its detectability with ground-based interferometers, space-based experiments and PTA’s. The last Section is dedicated to outlining and discussing the main results of our analysis and summarizing our conclusions.

\section{GWB from Black Hole Mergers} \label{sec:GWB}

A GW background (GWB) is produced by a population of BH binaries experiencing merging and is given by the relative GW energy density $\rho_{\text{GW}}$ to the critical density $\rho_c$ today, per unit interval of logarithmic frequency,

$$\Omega_{\text{GW}} = \frac{1}{\rho_c} \frac{d\rho_{\text{GW}}}{d \log f}, \quad \text{(1)}$$

where $f$ is the observed GW frequency. It is linked to the characteristic GW strain amplitude $h_c$ through \bib{37}

$$\frac{d\rho_{\text{GW}}}{d \log f} = \frac{\pi}{4G} f^2 h_c^2(f). \quad \text{(2)}$$

E.S. Phinney \bib{38} established a relationship between the GWB spectrum and the comoving number density of events $N(z)$,

$$\frac{d\rho_{\text{GW}}}{d \log f} = \frac{1}{c^2} \int_0^\infty N(z) \frac{1}{1+z} f \frac{dE_{\text{GW}}}{df} dz, \quad \text{(3)}$$

with $f$ being the frequency of the GWs in a source’s cosmic rest frame, related to $f$ by $f = f(1+z)$, $z$ being the redshift and $dE_{\text{GW}}/df_r$ is the redshifted energy produced by each event, per logarithmic frequency interval. Physically speaking, the GW energy spectrum is a superposition of redshifted GW radiation coming from merging BH binaries over the whole cosmic history. In the Newtonian limit, the energy spectrum associated with two interacting BH binaries is well approximated by \bib{37}

$$\frac{dE_{\text{GW}}}{df} = \frac{\pi^{2/3}}{3G} (GM_c)^{5/3} f_r^{-1/3} F(e), \quad \text{(4)}$$

where $M_c$ denotes the chirp mass. The function of the orbital eccentricity $F(e)$ is given by \bib{30} \bib{37}

$$F(e) = \frac{1}{(1-e^2)^{7/2}} \left( 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right), \quad \text{(5)}$$

and is close to one in the approximation of circular orbits ($e \approx 0$). If one denotes $f_{\text{min}}$ the minimum frequency at which this condition is fulfilled, the approximation is valid in the interval $f_{\text{min}} < f < f_{\text{ISCO}}$, with $f_{\text{min}}$ being derived from Kepler’s law and $f_{\text{ISCO}}$ corresponding to the innermost stable circular orbit (ISCO), derived from the radius $r_{\text{ISCO}} = 6GM_c^2$ below which the Newtonian approximation is no longer valid \bib{37}, that is

$$f_{\text{min}} = \frac{(GM)^{1/2}}{\pi a_0^{3/2}}, \quad f_{\text{max}} \approx 4.4\text{kHz} \left( \frac{M_\odot}{M} \right), \quad \text{(6)}$$

where $M$ is the total mass of the binary and $a_0$ is the semi-major axis of the orbit.
The number density of GW events \( N(z) \) within the redshift interval \([z, z + dz]\) is related to the merging rate \( \tau_{\text{merg}}(z) \) through \(30\)

\[
N(z) = \frac{\tau_{\text{merg}}(z)}{H(z)(1 + z)},
\]

where \( H(z) \) is the Hubble rate, \( H(z) = \frac{\sqrt{\Omega_\Lambda + \Omega_r(1 + z)^4 + (\Omega_{\text{DM}} + \Omega_\nu)(1 + z)^3}}{H_0} \) and \( \Omega_{\text{DM}}, \Omega_b, \Omega_r \) and \( \Omega_\Lambda \) are the usual present-day dark matter (DM), baryons (b), radiation (r) and cosmological constant (\( \Lambda \)) density parameters. Throughout the paper, we will consider broad mass distributions that will be introduced in Section \(11\). In this case, the merging rates depend on the two progenitor masses \( m_1 \) and \( m_2 \), so one must integrate over these masses to get the total GWB. Combining Eqs. \(2\), \(3\), \(4\) and \(7\) (with \( F(e) \approx 1 \)), one gets a characteristic strain today

\[
h_\Delta^2(f) = \frac{4G^{5/3}}{3\pi^{1/3} c^2} f^{-4/3} \int_0^\infty \frac{dz}{(1 + z)^{4/3} H(z)} \times \int \int \, d\ln m_1 \, d\ln m_2 \, \tau_{\text{merg}}(m_1, m_2, z) \, M^5 c^{5/3},
\]

where the rates must be specified per logarithmic mass intervals. The PBH mass function with the thermal features enters in these rates and therefore impacts the overall GW background. Justifications for using \( F(e) \approx 1 \) in the case of PBH binaries can be found in \(30\).

III. BROAD PBH MASS FUNCTIONS

Independently of the DM fraction constituted by PBHs, a primordial origin for some of the observed black hole mergers – in particular GW190814 that has a low mass ratio – would imply that PBHs have a somehow extended mass distribution. Contrary to stellar black holes, their distribution is not restricted by the Chandrasekhar limit on the low-mass side, and by pair-instability on the intermediate-mass side. Indeed, PBHs could have formed at different times in the cosmic history, possibly including the electroweak, QCD and neutrino decoupling epochs. The thermal history of the Universe involves a progressive reduction of the relativistic degrees of freedom at these epochs. This induces slight and transient changes in the equation of state of the Universe and, in turn, in the value of the density threshold for PBH formation. Unavoidably, this effect induces features in the PBH mass function \(15\) \(16\), for the general class of models where PBHs are formed by pre-existing density fluctuations that collapse when re-entering in the horizon during the radiation era (for counter-examples, see e.g. \(30\) \(32\)). These features take the form of a high peak around the solar mass, coming from the formation of protons and neutrons from quarks, a bump between 20 and 100 \( M_\odot \) from the formation of pions, and two other bumps around \( 10^{-5} M_\odot \) and \( 10^6 M_\odot \), respectively coming from the electroweak scale and the neutrino decoupling when electrons and positrons became non-relativistic.

As a result, it becomes possible to obtain a broad and realistic PBH mass function from an almost scale invariant primordial power spectrum, which is a generic outcome of inflation. One still needs a transition in the power spectrum amplitude between the large cosmological scales and the small PBH scales. One should note however that in some models, it is possible to form PBHs from rare fluctuations existing on all scales, in a non-Gaussian tail, without changing much the resulting PBH mass function \(29\), and keeping the primordial power spectrum at the \(10^{-9}\) level. From the point of view of limits on the PBH abundance, using such a mass function is a relatively conservative choice since most models with a peak in the power spectrum could be Taylor expanded on the scales relevant for PBH formation and would then be approximated by an amplitude, a scalar spectral index and a negative running of the scalar spectral index. We point out that a limit obtained for a null running still applies to a more specific model with a negative running. Therefore, assuming a null running is a conservative approach.

For these reasons, we consider in this paper a small-scale power-law power spectrum. We consider that \( \delta_{\text{rms}}(m) \), the root-mean-square amplitude of the Gaussian inhomogeneities from which the PBHs have formed, has the following spectrum \(16\).

\[
\delta_{\text{rms}}(m) = A_s \left( \frac{m}{M_\odot} \right)^{(1-n_s)/4},
\]

where \( n_s \) is a constant small-scale scalar spectral index and \( A_s \) is the spectrum amplitude whose value is fixed in order to give a total PBH abundance \( f_{\text{PBH}} = \rho_{\text{PBH}}/\rho_{\text{DM}} \) today. \( A_s \) and \( n_s \) must not be confused with the usual power spectrum amplitude and scalar spectral index on cosmological scales, inferred from CMB observations.

Introducing the fraction \( \beta \) of the Universe that collapses to form PBHs of mass \( m \) (per unit of logarithmic mass) at the time of formation \(16\)

\[
\beta(m) = \text{erfc} \left( \frac{\delta_c}{\sqrt{2} \delta_{\text{rms}}(m)} \right),
\]

and supposing a radiation-dominated Universe until matter-radiation equality, the normalized \((\int \phi(m) d\ln m = 1)\) broad PBH mass distribution is given by \(15\)

\[
\phi(m) = \frac{1}{\rho_{\text{DM}} f_{\text{PBH}}} \frac{d\rho_{\text{PBH}}}{d\ln m} \approx \frac{2A_s}{f_{\text{PBH}}} \left( \frac{M_{\text{eq}}}{m} \right)^{1/2} \beta(m),
\]

where the numerical factor is \(2(1+\Omega_b/\Omega_{\text{DM}})\) with \( \Omega_{\text{DM}} = 0.245 \) and \( \Omega_b = 0.0456 \), \( M_{\text{eq}} \approx 2.8 \times 10^{17} M_\odot \) is the Hubble mass at matter-radiation equality and \( f_{\text{PBH}} \) is the total dark matter fraction made of PBHs. In Eq. \(10\) \( \delta_c \) is the critical overdensity threshold leading to PBH formation, which varies with the equation of state.
FIG. 1: Top left: Equation-of-state parameter $w$ as a function of the temperature $T$(MeV). The grey vertical lines correspond to the masses of the electron, pion, proton/neutron, W, Z bosons and top quark, respectively. The grey dashed horizontal line indicates the value of $w = 1/3$ (Fig. from [16]). Top right: Critical density contrast $\delta_c$ as a function of $w$ (data from [39]). Bottom left: $\delta_c$ as a function of the PBH mass (assuming $\gamma = 0.8$). Bottom right: Resulting broad PBH mass function $\phi(m)$ for a scalar spectral index $n_s = 0.965$ (dark blue), 0.97 (red), 0.975 (green) and masses between $10^{-9}$ and $10^9 M_\odot$.

Next, we consider the effect of the thermal history and the transient changes in the equation of state parameter $w = p/\rho$ ($p$ being the pressure of the perfect fluid) on $\delta_c$. Typically, it varies everytime a species (electron, pion, proton, bosons, etc.) decouples from the thermal bath (Fig. 1, top left). Through simulations of PBHs in numerical relativity relating $w$ and $\delta_c$ on the one hand [39] (Fig. 1, top right), and lattice QCD relating $w$ and $T$ on the other hand, one can relate $\delta_c$ to the Hubble horizon mass $m_H$ (Fig. 1, bottom left) [15]. The PBH mass is assumed to be $m = \gamma m_H$ with $\gamma = 0.8$, as in [17], such that the QCD peak lies around $2.5 M_\odot$ where the highest merger rate of compact binary coalescences were observed by LIGO/Virgo. The resulting mass functions, for different values of $n_s$ (around 0.97) compatible with LIGO/Virgo observations, are also shown (Fig. 1, bottom right). In the following Sections, we will see that these broad mass functions that take into account the thermal history are included in the expressions of the PBH merging rates, which are used to compute the GWB from PBH mergers.

IV. PBH MERGING RATES

PBH binaries can form through two channels: by capture in dense clusters or in the early Universe, when two PBHs form sufficiently close to each other for their gravitational attraction to counterbalance the expansion of the Universe. This way, they form a binary before the matter-radiation equality. The evolution and mass dependence of the merging rates depend on the formation scenario. These two channels have been invoked to explain the merging rates inferred by LIGO/Virgo (see e.g. [16, 17]). Nevertheless, both are still subject to large uncertainties, essentially linked to PBH clustering. Indeed, clustering either induces a suppression or a boost of the merger rates of early binaries and of binaries in clusters, respectively. Hereafter we summarise some recent prescriptions for these rates, which we assume for the computation of the GWB. We also briefly discuss the possible connection with observations and the limitations of these approaches.
A. Early binaries

An important result of N-body simulations of early PBH binaries is that only a fraction of them are left undisrupted [22] by other PBHs, clusters or matter fluctuations. As a result, the merging rates used in [9] to compute in [51–54]. Beyond these limitations, one should remember that in general, early binaries can be impacted by their environment (e.g. in clusters or not) during the whole cosmic history. In particular, our rates for early binaries implicitly assume that only a fraction of PBHs will be in clusters, as opposed to the rate from capture in clusters, typically formed after matter-radiation equality and not considered in N-body simulations of early binaries. Furthermore, if $f_{\text{PBH}} \approx 1$, evading microlensing limits require that most PBHs are in clusters today. Therefore, in this case the rate model of early binaries

$S_2 \approx \min \left(1, 9.6 \times 10^{-3} f_{\text{PBH}}^{0.65} e^{0.03 \ln^2 f_{\text{PBH}}} \right)$. (16)

That is the approximation we use here. From Eq. [12] one obtains the merging rates of LIGO/Virgo for masses from the QCD peak, with $f_{\text{PBH}}$ close to one. In particular, for equal-mass mergers and $N = 2$, $f_{\text{PBH}} = \phi(m_{\text{PBH}}) = 1$, one gets $S_2 \approx 0.002$ and merging rates $\dot{\tau}_{\text{prim}} \approx 5 \times 10^3 (m_{\text{PBH}}/M_\odot)^{-32/37} \text{Gpc}^{-3} \text{yr}^{-1}$. This is larger than observed around $30 M_\odot$ but compatible with the rates inferred from GW190425 [46] around $2.5 M_\odot$.

Even if the rates given by Eq.[12] are consistent with N-body simulations, there are still a series of uncertainties that may limit this analysis. First, no N-body simulations have been performed in the case of our broad mass function with thermal features, so it is still possible that (tiny or heavy) PBHs far from the QCD peak additionally suppress these merging rates. Second, Eq.[12] does not take into account the merging rates of the perturbed binaries that may become dominant when $f_{\text{PBH}} \gtrsim 0.1$ [47], and so far there are no clear analytical prescriptions for non equal mass binaries. Third, slightly different results have been obtained in [48, 49] using analytical methods. Fourth, it has recently been claimed in [50] that subtle general relativistic effects may highly suppress this PBH binary formation channel, but this has been disputed in [51, 54]. Beyond these limitations, one should remember that in general, early binaries can be impacted by their environment (e.g. in clusters or not) during the whole cosmic history. In particular, our rates for early binaries implicitly assume that only a fraction of PBHs will be in clusters, as opposed to the rate from capture in clusters, typically formed after matter-radiation equality and not considered in N-body simulations of early binaries. Furthermore, if $f_{\text{PBH}} \approx 1$, evading microlensing limits require that most PBHs are in clusters today. Therefore, in this case the rate model of early binaries

Finally, the function $C$ encodes the transition between small and large $N$ limits and can be approximated by

$C \approx f_{\text{PBH}}^2 \left\{ \Gamma(29/37) \left( \frac{21}{64\cdot 2} \frac{5 f_{\text{PBH}}^2}{\sigma_M^2} \right) \right\}^{-\frac{2}{74}} - 1 \right\}^{-1}$ . (15)
FIG. 2: The GW spectrum $\Omega_{GW} h^2$ from late-time PBHs in clusters with a broad mass distribution with scalar spectral index $n_s = 0.965$ (dark blue), 0.970 (red) and 0.975 (dark green). The late binaries spectrum obtained from a log-normal mass distribution with a mean of $\mu_{PBH} = 2.5 M_\odot$ and a standard deviation of $\sigma_{PBH} = 0.1$ is also shown (dotted blue). The numerical spectrum also shows the sensitivities of the ground-based interferometers: LIGO/Virgo O2 (dotted orange) and O3 (solid orange) runs, projections for the HLV (LIGO-Hanford, LIGO-Livingston and Virgo) network at design sensitivity (dashed purple), the mid-scale upgrade of Advanced LIGO (A+ detectors, solid purple) [43] and the Einstein Telescope (light green). The sensitivity of the future space-based interferometer LISA is also shown (best experimental design - solid black, worst experimental design - dashed black). The Pulsar Timing Arrays (PTA’s) considered here are the EPTA (solid pink), NANOGrav (solid light blue), IPTA (solid light pink), PPTA (dotted light pink), and the Square Kilometer Array (SKA, dashed red) [44]. The recent detection of NANOGrav is represented by the blue square, as in [45]. In all Figs., we choose a virial velocity of $v_{vir} = 5$ km/s of the PBHs in the clusters and density contrasts are such that $R_{clust} = 460$. Might be somehow incompatible with the late-time formation of clusters.

Strong claims relying on these merging rates are therefore probably still premature. Nevertheless, Eq. (12) probably gives good estimations for some models or regimes, so it is relevant to use it for estimating the GWB, but keeping in mind the possible limitations of this approach.

B. Capture in clusters

For PBHs in clusters, the merging rate distribution per unit of logarithmic mass can be effectively described by [7, 8, 17]

$$\frac{d\tau_{clust}}{d \ln m_1 d \ln m_2} = R_{clust} \times \phi(m_1)\phi(m_2)f_{PBH}^2$$

$$\times \left(\frac{m_1 + m_2}{m_1 m_2}\right)^{10/7} \text{yr}^{-1}\text{Gpc}^{-3}, \quad (18)$$

where $R_{clust}$ is a scaling factor that depends on the PBH clustering properties and velocity distribution, and $\phi(m)$ is the PBH mass function represented in Fig. 1. Halo mass functions compatible with the standard $\Lambda$CDM cosmological scenario typically lead to $R_{clust} \approx 1 - 10$ [7]. For our mass distribution, this is too low to explain the merging rate inferred from GW190425. This is also too low to explain the rates around $30M_\odot$ [55, 56]. However, if $f_{PBH} \approx 1$ at the solar-mass scale or above, one expects an enhanced clustering due to isocurvature perturbations induced by Poisson fluctuations in the PBH distribution [17, 20, 57], even though PBHs are not clustered at formation time and follow a Poisson distribution [57]. This was invoked to explain some spatial correlations in the X-ray and infrared cosmic backgrounds [20, 21, 24], and implies that most PBHs reside in dense clusters. Several physical processes impact the PBH clustering properties: isocurvature perturbations from Poisson fluctuations, the primordial power spectrum, the dynamical heating and absorption of less massive clusters into larger ones, etc. These effects naturally lead to a clustering scale around $10^6 M_\odot$ for $m_{PBH} \sim O(M_\odot)$, $f_{PBH} \approx 1$ and $R_{clust} \sim O(10^2)$, compatible with GW observations. We assumed that $R_{clust} = 460$ as in [17], but it is also possible that PBHs are directly formed in clusters, which could induce even larger values of $R_{clust}$. Finally, let us mention that if most PBHs are regrouped in clusters, the microlensing limits on their abundance can be evaded, allowing $f_{PBH} = 1$.

V. GWB FROM PBH CLUSTERS

In this Section, we compute the GW spectrum from late-time PBHs clustered in virialized subhalos of galaxies. To do so, one needs to inject Eq. (18) into Eq. (8), and the calculation of the characteristic strain is performed by numerical integration on the masses by using a Monte Carlo method, following [30]. This allows us to only consider GW frequencies between $f_{min}$ and $f_{max}$ for our synthetic population of binaries. Since the integration on the masses is performed numerically, one can
FIG. 3: The GW spectrum $\Omega_{GW} h^2$ for early PBH binaries with a thermal broad mass distribution. Upper panel: scalar spectral index $n_s = 0.97$ and a PBH abundance $f_{PBH} = 1$ (solid dark red), 0.1 (dashed dark red), 0.01 (dotted dark red). Lower panel: scalar spectral index $n_s = 0.965$ (dark blue), 0.970 (red) and 0.975 (dark green) and an abundance of $f_{PBH} = 1$. The early binaries spectrum obtained from a log-normal mass distribution with a mean of $\mu_{PBH} = 2.5 M_\odot$ and a standard deviation of $\sigma_{PBH} = 0.1$ is also shown (dotted blue). For sensitivity curves, refer to Fig. 2.
write the expression for the characteristic strain without explicitly writing the mass integral,

\[ h_c^2(f) = \frac{4\tau_{\text{clust}}}{3\pi^{1/3} c^2} \left( \frac{GM_c}{c^3} \right)^{5/3} f^{-4/3} \]

\[ \times \int_0^\infty \frac{dz}{(1+z)^{4/3}H(z)} \tag{19} \]

The redshift integral is performed separately, since Eq. (19) does not contain any other redshift dependence. Expressing the total merging rate in yr\(^{-1}\)Gpc\(^3\), the chirp mass in solar masses \(M_\odot\) and the frequency in Hz, one gets

\[ h_c(f) \simeq 1.15 \times 10^{-25} \left( \frac{\tau_{\text{clust}}}{\text{yr}^{-1}\text{Gpc}^{-3}} \right)^{1/2} \]

\[ \times \left( \frac{f}{\text{Hz}} \right)^{-2/3} \left( \frac{M_c}{M_\odot} \right)^{5/6} \tag{20} \]

where we used the following result, obtained by numerical integration,

\[ \int_0^\infty \frac{dz}{(1+z)^{4/3}H(z)} = 0.7642 H_0^{-1}. \tag{21} \]

Considering a synthesized population of \(10^6\) binaries with given velocities and impact parameters, the marginalized GW spectrum over PBH masses is then obtained. The spectrum is shown in Fig. 2 where we have chosen the virial velocity in those clusters to be \(v_{\text{vir}} = 5\) km/s (which is compatible with velocity dispersions in faint dwarf galaxies).

VI. GWB FROM EARLY PBH BINARIES

Early PBH binaries have a merging rate distribution that depends not only on the masses, but also on the redshift \(z\) (see Eq. (12)). Consequently, one also has to take into account the redshift term present in the merger rate of early binaries, and perform the redshift integral separately, as in the previous Section. Inserting the primordial rate \(\tau_{\text{prim}}(z)\) in the integral \(\mathbf{[8]}\), one obtains

\[ h_c^2(f) = \frac{4}{3\pi^{1/3} c^2} \left( \frac{GM_c}{c^3} \right)^{5/3} f^{-4/3} \]

\[ \times \int_0^\infty \frac{dz}{(1+z)^{4/3}H(z)} \tau_{\text{prim}}(z), \tag{22} \]

where \(\tau_{\text{prim}}(z) = \tau_{\text{prim}} \times \left( \frac{t(z)}{t_0} \right)^{-34/37-0.29}\). One has to perform the following numerical calculation:

\[ \int_0^\infty \frac{t(z)/t_0}{(1+z)^{4/3}H(z)} \frac{dz}{1+z} = 5.92 H_0^{-1}. \tag{23} \]

Using the same units as before, one gets

\[ h_c(f) \simeq 3.21 \times 10^{-25} \left( \frac{\tau_{\text{prim}}}{\text{yr}^{-1}\text{Gpc}^{-3}} \right)^{1/2} \]

\[ \times \left( \frac{f}{\text{Hz}} \right)^{-2/3} \left( \frac{M_c}{M_\odot} \right)^{5/6}. \tag{24} \]

The GWB produced by these early binaries is shown in Fig. 3.

VII. DETECTABILITY WITH GROUND-BASED INTERFEROMETERS

As shown in Figs. 2 and 3, the \(2\sigma\) power-law integrated sensitivity of LIGO/Virgo O2 and O3 observing runs does not yet constrain the expected GW background for PBH clusters. However, the GW spectrum clearly reaches the sensitivity of Einstein Telescope (ET), as well as the design sensitivity of upgraded LIGO facilities (A+) in the PBH model with the lowest value of \(n_s\) (0.965). In Fig. 4, we also show an updated estimate of the combined GWB coming from the merging of three classes of compact binaries: binary black holes (BBHs), binary neutron stars (BNSs), and neutron star-black hole binaries (NSBHs), from \(\mathbf{[8]}\). The combined background coming from BBHs and BNSs is predicted to be \(\Omega_{\text{BBH}+\text{BNS}} h^2(25\text{ Hz}) = 3.3^{+3.3}_{-2.3} \times 10^{-10}\). Furthermore, combining the upper limit on \(\Omega_{\text{NSBH}}(f)\) with the upper 95% bound on the contributions from BBH and BNS mergers, we set an upper bound on the total expected GWB, \(\Omega_{\text{Total}} h^2(25\text{ Hz}) \leq 0.86 \times 10^{-9}\). The GWB from PBHs in clusters is approximately at the same level as the one coming from sources of astrophysical origin, though the former is significantly lower around 100 Hz due to merging of PBHs with low mass ratios. As a result, at the frequency at which LIGO will have its best sensitivity, the GWB spectral index (defined as \(d \log \Omega_{\text{GW}}/d \log f\)) from PBHs in clusters is slightly negative, whereas the prediction for astrophysical sources is around 2.3. Probing this GWB spectral index with ground-based detectors can therefore be used to distinguish an astrophysical from a primordial origin if they have a broad mass function, when one includes thermal features.

For early PBH binaries, the GWB is represented in Figs. 3 and 4 for different values of \(n_s\) and \(f_{\text{PBH}}\). We observe that if PBHs contribute to the dark matter with more than \(f_{\text{PBH}} \approx 0.01\), the GW spectrum is above the projected limits for the HLV (LIGO-Hanford, LIGO-Livingston and Virgo) and A+ detector designs. For \(f_{\text{PBH}} = 1\), it is just below the current limits imposed by the third observing run of LIGO/Virgo, in particular in the model with the lowest value of \(n_s\) (0.965). Therefore, the next observing runs (O4 and O5) of LIGO/Virgo may be crucial for its detection and for the determination of its properties. The spectrum is also well above expectations for astrophysical sources if the PBH constitute almost all of the dark matter, whereas for lower values of \(f_{\text{PBH}}\), the GW spectrum crosses the upper limit on the GWB from BBHs, BNSs and NSBH mergers, and has the same spectral index. In this case, differentiating the two backgrounds in future detections could be challenging.

Our analysis also reveals the important contribution to the GWB of PBH binaries with very low mass ratios, combining a solar-mass PBH from the QCD peak with a
FIG. 4: Model predictions for the GWB from BBHs, BNSs and NSBHs, with present and projected sensitivity curves in the frequency range $5 - 500$ Hz. We have included the combined BNS and BBH energy density spectra, $2\sigma$ power-law integrated curves for O2 and O3 observing runs (orange lines), projections for the HLV network at design sensitivity and the A+ detectors (purple). The solid gray line shows the median estimate of $\Omega_{BBH+BNS}h^2(f)$, whereas the shaded gray band represents $90\%$ credible Poisson uncertainties. The black dashed line indicates the projected upper limit on the total GWB, which includes the upper limit on the contribution from NSBH mergers [43]. The GW spectrum $\Omega_{GW}h^2$ from PBH clusters is shown in dark blue ($n_s = 0.965$), red ($n_s = 0.970$) and dark green ($n_s = 0.975$). We choose an abundance of $f_{PBH} = 1$.

FIG. 5: Model predictions for the GWB from BBHs, BNSs and NSBHs, with current and projected sensitivity curves in the frequency range $5 - 500$ Hz (see caption of Fig. 4). The GW spectrum $\Omega_{GW}h^2$ from early PBH binaries is shown: (left panel) $f_{PBH} = 1$ (solid dark red), 0.1 (dashed dark red), 0.01 (dotted dark red); (right panel) $n_s = 0.965$ (dark blue), 0.970 (red) and 0.975 (dark green).

subsolarmass PBH. In order to quantify the contribution from the different masses, we have computed the GWB amplitude $\Omega_{GW}h^2$ as a function of $m_1$ and $m_2$, both for clusters and early binaries, shown in Fig. 6. This way, we explain why the dark green curve corresponding to $n_s = 0.975$ (Figs. 2 and 3) is lower than for the other values of $n_s$, given that massive black holes are less numerous (see Fig. 1) and so contribute less to the GWB. Moreover, as shown in Fig. 3, increasing the PBH abundance $f_{PBH}$ increases the amplitude of the GW spectrum. This can be seen directly from Eq. (12), and comes from the fact that a higher PBH abundance makes a greater contribution to the spectrum. At frequencies above 100 Hz, binaries with subsolar black holes dominate the signal, leading to an opposite conclusion. This particular behaviour could be probed by ET. Finally, we point out that the relative contribution of binaries with low-mass ratios and a subsolar black hole is more important for early binaries than for PBHs in clusters.

VIII. DETECTABILITY WITH LISA

Contrary to the case of a log-normal, close to monochromatic mass distribution ($\sigma_{PBH} = 0.1$), the case of a broad PBH mass distribution with thermal features exhibits a GWB with a negative spectral index ($d \log \Omega_{GW}/d \log f < 0$) down to a frequency around
$5 \times 10^{-4}$ Hz, leading to a strong enhancement (by up to 6 orders of magnitude compared to the log-normal mass function) in the frequency range of LISA, for late-time PBHs in clusters (Fig. 2). It would be detectable in a broad range of frequencies ($10^{-6}$ to 0.1 Hz) even for the worst experimental design, allowing for a precise reconstruction. This strong increase in amplitude comes from the merging of binaries with large mass ratios involving an intermediate-mass and a solar-mass (QCD) PBH, as shown in Fig. 6. This will help differentiate this background from the one of astrophysical origin. A similar feature is observed in the GWB for early binaries, shown in the lower panel of Fig. 3 except that the amplitude of the background is up to two orders of magnitude lower than for PBHs in clusters, and that the GWB spectral index remains overall positive. Hence, probing the GWB spectral index at these frequencies with LISA could in principle help us differentiate late-time PBH binaries in clusters from early PBH binaries. Finally, one can notice a bump around $10^{-3}$ Hz, coming from the bump in the PBH mass function around $10^6 M_\odot$. The shape of the GWB could therefore be used to reconstruct the exact PBH mass function at large mass.

IX. DETECTABILITY WITH PTA’S

Another difference coming from the above-mentioned effect is that the GWB induced by PBHs in clusters with a broad mass distribution and consistent with LIGO/Virgo merging rates is now detectable by the PTA’s (Fig. 2), in particular by future SKA and IPTA observations, as opposed to the log-normal case where the amplitude is very strongly suppressed at low frequency. We also find that the change of the GWB spectrum with $n_s$ is more important than at higher frequencies. This is due to the fact that the QCD peak remains unchanged as we vary $n_s$ (see Fig. 1) while the abundance of intermediate-mass PBHs strongly changes. The latter dominate the GWB signal at nanoHertz and therefore the three curves are shifted away from each other.

X. A HINT OR A LIMIT FROM NANOGRAV?

The NANOGrav collaboration recently published an analysis of 12.5 years of pulsar timing (PT) data which provides evidence of a stochastic process that can be interpreted as a background of GWs [31]. A similar observation has been recently reported using data from Parkes PTA’s (PPTA) [32] and a possible non-GW origin was also explored. The detected signal is within the interval $10^{-9} - 10^{-8}$ Hz with $\Omega_{GW} h^2 \sim 10^{-9}$ (blue window). The density fluctuations at the origin of solar-mass or planetary-mass PBHs have already been considered as a source of such a GWB [33, 59, 61].

Our analysis reveals that the signal, when interpreted as a GWB, could also be explained by binaries of PBHs with an extended mass function. While the NANOGrav signal does not exactly coincide with the curves shown in Fig. 2 for PBH binaries in clusters, this would have been the case for a scalar spectral index smaller than 0.965. As for the early PBH binaries, the amplitude of the GWB is always much lower than the NANOGrav signal. Consequently, on the basis of our results, the observed signal could be explained by PBH binaries in clusters with very small mass ratios (with a primary black hole of intermediate mass) and a small enough spectral index ($n_s < 0.965$), if PBHs constitute the totality or a large fraction of the DM.

If the GW nature of the NANOGrav signal were to be confirmed, the measurement of its spectral index $^1$ could distinguish between a background from PBH binaries or from second order perturbations (almost flat in our model) [62].

XI. CONCLUSION

We have studied the GWB induced by the formation and merging of binaries from a PBH population. In this work, we only considered PBH-PBH mergers, but one could also explore the GW background from PBH-neutron star mergers. Nevertheless, their rates are expected to be subdominant for the models we considered here [63]. For the first time, we took into consideration wide PBH mass distributions with features from the thermal history. We calculated and characterized the GWB produced by a superposition of GW sources coming from the merging of PBHs binaries in clusters and early PBH binaries, and compared it to the astrophysical GWB produced by astrophysical black holes, binary neutron stars, and neutron star-black hole binaries. We then confronted it to the latest limits of LIGO/Virgo and PTA’s, and to the forecasted limits of the Einstein Telescope and LISA.

For binaries in clusters, we found that the GWB amplitude is greatly enhanced than in previous estimations for monochromatic or log-normal mass functions at frequencies below 10 Hz. We identified the origin of this boosted GWB, coming from the numerous binaries with low mass ratios. This GWB will be probed by future ground and space-based facilities such as A+, ET or LISA. We expect that this feature will allow us to distinguish this background from the astrophysical GWB because of the specific frequency dependence. Interestingly, the GWB at nanoHertz frequency could coincide with a signal from NANOGrav for a small enough spectral index.

For early binaries, we used the most recent merger rate prescriptions, and found that it is also possible to have a wide mass function that can explain GW events and, at

$^1$ Encoded in the so-called $\gamma = 5 - n = 3 - 2\alpha$ parameter in [31] where $n$ is the GWB spectral index and $\alpha$ is the strain spectral index.
FIG. 6: The mass contribution to the GWB from late-time and early PBHs binaries. The color bar indicates the values of the quantity $\log_{10}(\Omega_{GW} h^2)$, which is represented as a function of the logarithm of the masses $m_1$ and $m_2$ of the PBHs following a mass distribution from the thermal history of the Universe (with $n_s = 0.97$ and $v_{\text{vir}} = 5 \text{ km/s}$) for the frequencies $f = 100 \text{ Hz}$ (top), $10^{-3} \text{ Hz}$ (middle) and $10^{-8} \text{ Hz}$ (bottom). Left column: late-time PBHs binaries in clusters. Right column: early PBHs binaries (with $f_{\text{PBH}} = 1$).
the same time, pass the most recent LIGO/Virgo limits on the GWB. This background typically dominates by one order of magnitude the ones from PBH clusters and astrophysical binaries. Nevertheless, the complexity of the phenomena still poses important challenges and one probably still needs to better characterize the merging phenomena still poses important challenges and one probably still needs to better characterize the merging processes in binary systems, gas accretion, relativistic effects, etc., and verify their consistency with N-body simulations.

To summarize, our work has revealed strong differences between PBH models and the possibility to discriminate the different mass functions and binary origins with the observations of the GWB. In this context, multi-wavelength observations are essential. Future instruments like Einstein Telescope, LISA and SKA will play a major role and will be able to confirm or exclude a primordial origin of black holes.

Acknowledgments

The authors warmly thank Juan García-Bellido, Matteo Braglia, Sachiko Kuroyanagi, Mairi Sakellariadou, Alex Jenkins and Federico De Lillo for useful comments and discussion. E.B. is supported by the FNRS-IISN (under grant number 4.4501.19). S.C. acknowledges support from the Francqui Foundation.
[32] B. Goncharov et al., (2021), 10.3847/2041-8213/ac17f1, arXiv:2107.12112 [astro-ph.HE].
[33] V. Mandic, S. Bird, and I. Cholis, Phys. Rev. Lett. 117, 201102 (2016), arXiv:1606.06699 [astro-ph.CO].
[34] M. Raidal, V. Vaskonen, and H. Veermäe, JCAP 09, 037 (2017), arXiv:1707.01480 [astro-ph.CO].
[35] S. Mukherjee, M. S. P. Meinema, and J. Silk, (2021), arXiv:2107.02181 [astro-ph.CO].
[36] S. Mukherjee and J. Silk, Mon. Not. Roy. Astron. Soc. 506, 3977 (2021), arXiv:2105.11139 [gr-qc].
[37] M. Maggiore, Gravitational Waves. Vol. 1: Theory and Experiments, Oxford Master Series in Physics (Oxford University Press, 2007).
[38] E. S. Phinney, (2001), arXiv:astro-ph/0108028.
[39] I. Musco and J. C. Miller, Class. Quant. Grav. 30, 145009 (2013), arXiv:1201.2379 [gr-qc].
[40] E. Cotner, A. Kusenko, and V. Takhistov, Phys. Rev. D 98, 083513 (2018), arXiv:1801.03321 [astro-ph.CO].
[41] E. Cotner, A. Kusenko, M. Sasaki, and V. Takhistov, JCAP 10, 077 (2019) arXiv:1907.10613 [astro-ph.CO].
[42] M. M. Flores and A. Kusenko, Phys. Rev. Lett. 126, 041101 (2021) arXiv:2008.12456 [astro-ph.CO].
[43] R. Abbott, T. D. Abbott, et al. (LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration), Phys. Rev. D 104, 022004 (2021)
[44] K. Schmitz, J. High Energ. Phys. 97 (2021), 10.1007/JHEP01(2021)097, arXiv:2002.04615 [hep-ph].
[45] V. De Luca, G. Franciolini, and A. Riotto, Phys. Rev. Lett. 126, 041303 (2021) arXiv:2009.08268 [astro-ph.CO].
[46] B. P. Abbott et al. (LIGO Scientific, Virgo), Astrophys. J. Lett. 892, L3 (2020) arXiv:2001.01761 [astro-ph.HE].
[47] V. Vaskonen and H. Veermäe, Phys. Rev. D 101, 043015 (2020) arXiv:1908.09752 [astro-ph.CO].
[48] B. Koceis, T. Suyama, T. Tanaka, and S. Yokoyama, Astrophys. J. 854, 41 (2018) arXiv:1709.00007 [astro-ph.CO].
[49] L. Liu, Z.-K. Guo, and R.-G. Cai, Phys. Rev. D 99, 063523 (2019) arXiv:1812.05376 [astro-ph.CO].
[50] C. Boehm, A. Kobakhidze, C. A. O’Hare, Z. S. Picker, and M. Sakellariadou, (2020), arXiv:2008.10743 [astro-ph.CO].
[51] V. De Luca, V. Desjacques, G. Franciolini, and A. Riotto, JCAP 11, 028 (2020), arXiv:2009.04731 [astro-ph.CO].
[52] G. Hütsi, T. Koivisto, M. Raidal, V. Vaskonen, and H. Veermäe, (2021), arXiv:2105.09328 [astro-ph.CO].
[53] G. Hütsi, T. Koivisto, M. Raidal, V. Vaskonen, and H. Veermäe, (2021), arXiv:2106.02007 [astro-ph.CO].
[54] T. Harada, H. Maeda, and T. Sato, (2021), arXiv:2106.06651 [gr-qc].
[55] B. Abbott et al. (LIGO Scientific, Virgo), Astrophys. J. Lett. 833, L1 (2016) arXiv:1602.03842 [astro-ph.HE].
[56] B. Abbott et al. (LIGO Scientific, Virgo), Astrophys. J. Lett. 882, L24 (2019) arXiv:1811.12940 [astro-ph.HE].
[57] A. Moradinezhad Dizgah, G. Franciolini, and A. Riotto, JCAP 11, 001 (2019) arXiv:1906.08978 [astro-ph.CO].
[58] R. Abbott et al. (KAGRA, Virgo, LIGO Scientific), Phys. Rev. D 104, 022004 (2021) arXiv:2101.12130 [gr-qc].
[59] Z.-C. Chen, C. Yuan, and Q.-G. Huang, Phys. Rev. Lett. 124, 251101 (2020) arXiv:2010.12239 [astro-ph.CO].
[60] K. Kohri and T. Terada, Phys. Lett. B 813, 136040 (2021) arXiv:2009.11853 [astro-ph.CO].
[61] G. Domènech and S. Pi, (2020) arXiv:2010.03976 [astro-ph.CO].
[62] S. Clesse, J. García-Bellido, and S. Orani, (2018), arXiv:1812.11011 [astro-ph.CO].
[63] M. Sasaki, V. Takhistov, V. Vardanyan, and Y.-I. Zhang, (2021), arXiv:2110.09509 [astro-ph.CO].