Tracking the origin of black holes with the stochastic gravitational wave background popcorn signal

Matteo Braglia,1,2* Juan García-Bellido,1† Sachiko Kuroyanagi1,3‡
1Instituto de Fisica Teorica, Universidad Autonoma de Madrid, Madrid, 28049, Spain
2INAF/OAS Bologna, via Gobetti 101, I-40129 Bologna, Italy
3Department of Physics and Astrophysics, Nagoya University, Nagoya, 464-8602, Japan

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Unresolved sources of gravitational waves (GWs) produced by the merger of a binary of black holes at cosmological distances combine into a stochastic background. Such a background is in the continuous or popcorn regime, depending on whether the GW rate is high enough so that two or more events overlap in the same frequency band. These two regimes respectively correspond to large and small values of the so-called duty cycle. We study the detection regime of the background in models of Primordial Black Holes (PBHs) and compare it to the one produced by black holes of stellar origin. Focusing on ground-based detectors, we show that the duty cycle of the PBH-origin background is larger than that of astrophysical black holes because of differences in their mass function and the merger rate. Our study opens up the possibility to learn about the primordial or astrophysical nature of black hole populations by examining the statistical properties of the stochastic background.

Key words: black hole physics – gravitational waves

1 INTRODUCTION
The Stochastic Gravitational Wave Background (SGWB) is one of the most interesting targets of current and future gravitational wave (GW) observatories. A detection of a SGWB from the Early Universe would provide a breakthrough in our understanding of the origin of the Universe. In addition to those of primordial origin, we expect SGWBs composed of many astrophysical sources at cosmological distances that cannot be individually resolved (Regimbau 2011; Rosado 2011). Example of sources contributing to such an astrophysical background are core-collapse supernovae (Ferrari et al. 1999a), neutron stars (both during their formation (Coward et al. 2001) and rotation (Ferrari et al. 1999b; Regimbau & de Freitas Pacheco 2001)), magnetars (Regimbau & de Freitas Pacheco 2006a; Chowdhury & Kholopov 2021), the early inspiral phase of compact binaries (Farmer & Phinney 2003; Regimbau & de Freitas Pacheco 2006b) or the coalescence of very massive BHs (Sesana et al. 2005, 2008). Such astrophysical backgrounds can be used as a tool to constrain the properties of such sources. Moreover, each of these signals is associated with a characteristic frequency and spectral shape, which can help infer which astrophysical sources are contributing to the background.

Besides the frequency dependence, there are other properties that distinguish the nature of different backgrounds, such as their anisotropies (Cusin et al. 2018; Jenkins et al. 2018, 2019; Cusin et al. 2019b; Bertacca et al. 2020; Pitrou et al. 2020), polarizations (Cusin et al. 2019a), and popcorn (sometimes referred as non-Gaussian or time-dependent) signal (Coward & Regimbau 2006; Regimbau & Mandic 2008; Wu et al. 2012; Mukherjee & Silk 2020). Those characteristics would be essential for disentangling different possible sources and identifying the origin of the SGWB. The popcorn signal, which is the focus of this paper, could be seen in a SGWB originating from overlapped GW sources (Coward & Regimbau 2006; Regimbau et al. 2012). It depends on the relative duration of the transient signal and the time interval between successive events. If such interval is small compared to the duration of a single event and/or the number of sources is very large, the GW events overlap, and the background is in the so-called continuous regime, characterized by Gaussian statistics, as a consequence of the central limit theorem. On the other hand, if the interval between events is comparable or larger than the typical duration of the signal, the waveforms may or may not overlap, and the statistical properties are strongly non-Gaussian. We denote this regime as popcorn background. A useful quantity to distinguish between the two regimes is the so-called astrophysical duty cycle, which represents the average number of events present in a given frequency band. Continuous and popcorn regimes correspond to large and small duty cycles, respectively (i.e. values of the duty cycle larger or smaller than one).

Our goal in this paper is to show that the popcorn signature of the SGWB can be used to learn about the nature of BHs. The cumulative detection of GWs from binary black

© 2021 The Authors
hole (BBH) inspirals has attracted wide attention to Primordial Black Hole (PBH) as a possible source (Bird et al. 2016; Sasaki et al. 2016; Clesse & García-Bellido 2017a). Contrary to standard astrophysical black holes (ABHs), which form by the collapse of massive stars at the end of their life cycle, PBHs could have been formed by the collapse of very large density perturbations during the radiation era (Zeldovich & Novikov 1967; Hawking 1971). The amplification of curvature perturbations produces the seeds of such perturbations during inflation at scales much smaller than those tested by Cosmic Microwave Background (CMB) observations (Carr & Lidsey 1993; Carr et al. 1994; García-Bellido et al. 1996). Being almost collisionless and dark by definition, PBHs cosmologically behave as Cold Dark Matter (CDM) and make up a fraction of it. They can take masses well below Chandrasekhar mass, with the only constraint being that PBHs lighter than $10^{15}$ g would have already evaporated by today.

Besides their mass function, PBHs also differ in their merger rate, which is expected to grow with redshift, unlike that of ABHs, which follow the star formation rate (Cholis 2017; Raidal et al. 2017; Vaskonen & Veermäe 2020; Atal et al. 2020; De Luca et al. 2020a; Mukherjee & Silk 2021).

The observation of BBHs indicates that the SGWB from their superposition could be detected in the near future by an upgraded ground-based detector network (Abbott et al. 2016b, 2021b; Mandic et al. 2016; Clesse & García-Bellido 2017b; Wang et al. 2018). Once detected, one of the primary challenges is to identify whether its origin is astrophysical or primordial. In this paper, we explore the possibility of using the duty cycle for distinguishing BBH formation mechanisms. One approach to address this issue is to use the spectral shape (Mukherjee & Silk 2021; Bavera et al. 2021), while this is the first study where the SGWB from PBHs is studied using the duty cycle and going beyond its spectral shape. Focusing on ground-based interferometers, we find that the duty cycle for PBH is generically higher than the one of astrophysical populations. We develop a simple procedure to compute the duty cycle, taking into account the sensitivity of the detectors. This allows selecting, from all the events contributing to the duty cycle, only those seen by a given detector with a Signal-to-Noise-Ratio (SNR) exceeding a certain threshold. Our results are suitable to be used in population searches with unresolved events (Smith et al. 2020; Biscoveanu et al. 2020).

Our paper is organized as follows. In Sec. 2, we describe the PBH models that we use in our work. We consider not only the widely used log-normal mass function, but also the mass function motivated by the thermal history of the universe. The theoretical framework is reviewed in Secs. 3, 4 and 5 which are dedicated to the merger rate, the SGWB and the duty cycle respectively. We finally present our results in Sec. 6 and conclude in Sec. 7.

2 PRIMORDIAL BLACK HOLE MODELS

PBHs form when the overdensities produced during inflation are larger than a critical threshold $\delta_c = \delta_\rho/\rho$ at horizon reentry. The latter is sensitive to the energy content of the cosmic fluid and its Equation of State (EOS) (Musco & Miller 2013), which is usually taken to be \( w = 1/3 \) during radiation dominated era. The fraction of horizon patches that collapse into PBHs is given by (Carr 1975)

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

where ‘erfc’ is the complementary error function. Note that we have explicitly expressed the dependence of the critical density in terms of \( m \), the mass of the formed PBHs.

In Eq. (1), the root-mean-squared amplitude of density perturbations $\delta_{\text{rms}}$ is given in terms of the primordial power spectrum of the curvature perturbations $P_\zeta(k)$ produced during inflation as:

$$\delta_{\text{rms}}^2[m(k)] = \int d \ln k W^2(kR) \left( k R^2 \right)^4 P_\zeta(k),$$
where, for simplicity, we have assumed that curvature perturbations from inflation are Gaussian distributed, and \( W \) is a window function smoothing over a comoving scale \( R \approx 2Gm/\Omega_{\text{CDM}}\gamma^{-1} \). Here, \( \gamma \) parametrizes the ratio between the PBH mass and the mass of the collapsing horizon-sized region at PBH formation, which typically takes values \( \gamma \in [0.1, 1] \). Note that the scale of perturbations \( k \) correspond to the mass of PBHs as

\[
m(k) \sim 30M_\odot \left( \frac{\gamma}{0.2} \right) \left( \frac{g_*}{10.75} \right)^{-1/6} \left( \frac{k}{2.9 \times 10^5 \text{ Mpc}^{-1}} \right)^{-2}
\]

where \( g_* \) is the number of relativistic degrees of freedom at formation.

From Eq. (1), we can compute the mass fraction of PBHs per logarithmic interval of masses as

\[
f_{\text{PBH}}(m) \equiv \left( \frac{\Omega_{\text{PBH}}}{\Omega_{\text{CDM}}} \right) \approx 2 \left( 1 + \frac{\Omega_b}{\Omega_{\text{CDM}}} \right) \left( \frac{M_{\text{cd}}}{m} \right) \beta(m)\sqrt{\frac{M_{\text{cd}}}{m}},
\]

where \( \Omega_{\text{CDM}} = 0.245 \) and \( \Omega_b = 0.0456 \) are CDM and baryon density parameters, and \( M_{\text{cd}} = 2.8 \times 10^{17} M_\odot \) is the horizon mass at matter-radiation equality. The total fraction of PBHs is obtained by integrating \( f_{\text{PBH}}(m) \) over its full support

\[
f_{\text{PBH}}^{\text{tot}} \equiv \int d\ln m f_{\text{PBH}}(m),
\]

which should satisfy \( f_{\text{PBH}}^{\text{tot}} \leq 1 \).

As clear from Eq. (1), the PBH mass function is affected by the shape of the primordial power spectrum and the EOS parameter\(^2\). In this paper, we consider two different PBH models. One is the widely used mass function motivated by inflationary models producing a peaky primordial power spectrum. The other is the mass function that carries the effect of the changes of \( w \) due to the thermal history evolution.

As a first example, we consider the Lognormal (LN) mass function (Dolgov & Silk 1993)

\[
f_{\text{PBH}}(m) = \frac{f_{\text{PBH}}^{\text{tot}}}{\sqrt{2\pi} \sigma} \exp \left[ -\frac{\ln^2 m/\mu}{2\sigma^2} \right],
\]

which is the most widely considered mass function and arises in many inflationary models featuring a peak in the power spectrum (Clesse & García-Bellido 2015; Braglia et al. 2020).

As a reference, we adopt the values \( \mu = 30M_\odot \) and \( \sigma = 1.5 \). In order to avoid astrophysical constraints, we assume \( f_{\text{PBH}}^{\text{tot}} = 1 \).

We refer to the second model as Thermal History (TH) model. This mass function has a rich structure induced by the thermal history evolution. As the temperature of the Universe decreases, Standard Model particles become non-relativistic when the temperature becomes comparable to the mass scale of the particles. This and the QCD phase transition cause small drops in \( w \) and induce pronounced features in the PBH mass function as a result of the exponential dependence of the PBH formation probability on the critical threshold \( \delta_c(w) \) (Jedamzik 1997; Byrnes et al. 2018; Carr et al. 2021).

For the primordial power spectrum, we assume a nearly scale-invariant spectrum of the form \( P_c(k) = A_s(k/k_{\text{QCD}})^{n_s - 1} \) at small scales. Here \( k_{\text{QCD}} \) is the pivot scale and is the mode that re-enters the horizon at the QCD transition. Note that \( A_s \) and \( n_s \) take values that are independent of the ones measured at CMB scales. This spectrum is a good approximation to the one produced in inflationary models featuring the second stage of slow-roll with first slow-roll parameter \( \epsilon_{\text{QCD}} \ll \epsilon_{\text{CMB}} \) (García-Bellido & Ruiz Morales 2017; Ezquiaga et al. 2018).

As a reference, we take \( n_s = 0.97 \), and \( A_s \) is fixed to satisfy the condition of \( f_{\text{PBH}}^{\text{tot}} = 1 \).

The mass functions of the two models are shown in Fig. 1. As can be seen, the TH mass function shows two pronounced peaks at \( m \sim 2M_\odot \) and \( \sim 70M_\odot \). This is the result of taking into account the variations of the equation of state, which deviates from \( w = 1/3 \) at the QCD transition in the Thermal Model of particle physics\(^3\) and when pions become non-relativistic. Although the plot in Fig. 1 is restricted to PBH masses of interest for ground-based detectors, we note that the mass function shows secondary peaks also at \( m \sim 10^{-3}M_\odot \) and \( \sim 10^9M_\odot \).

Note that functional form of the LN mass function in Eq. 6 would be also modulated by the variations of \( w \) at the QCD transition and the pion mass scale (Carr & Kuhnel 2019). However, being the mass function very peaked, our results are almost unaffected by this.

As a result, the CMB constraints (dashed line in Fig. 1) at face value, the presence of the secondary peak around \( \sim 10^6 M_\odot \) would rule out the TH mass function. We stress, however, that the results presented in this paper are mainly affected by the masses in the range \( [\mathcal{O}(0.1)M_\odot, \mathcal{O}(100)M_\odot] \) because of the limited sensitivity and frequency range of the ground-based experiments, and larger masses are almost irrelevant for our purposes. We also note that a possible way to avoid such constraints is to add a small amount of running to the spectral index in the simple parameterization for \( P_c(k) \)\(^4\), which is well motivated since it mimics the rise of the power spectrum from the large scales probed by the CMB (Ezquiaga et al. 2018). Furthermore, the constraints shown in Fig. 1 have multiple caveats, some due to the assumption of single-mass (so-called monochromatic) mass functions or that of uniformly distributed PBH. Relaxing those assumptions, many of the constraints either go away or shift around. A reanalysis of some of the bounds can be found in (Carr et al. 2021).

\(\text{3 MERGER RATES}\)

The next step toward a computation of the SGWB is to build a model for the merger rate of the black hole binaries.

**Primordial Black Holes.** For PBHs, two formation channels exist, both in principle contributing to the total merger rate. In the first channel, so-called early binary, binaries are

---

\(^2\) Non-Gaussianities of primordial perturbations also affect the PBH mass function. In this paper, for simplicity, we restrict to Gaussian perturbations. For the effects of non-Gaussianities, see e.g. (Young & Byrnes 2013; García-Bellido et al. 2017; Franciolini et al. 2018; Atal & Germani 2019; De Luca et al. 2019; Yoo et al. 2019; Ezquiaga et al. 2020; Kitajima et al. 2021).

\(^3\) Phase transitions around the QCD epoch and/or Lepton Flavour Asymmetries also modify the PBH mass function (Bödeker et al. 2021; García-Bellido et al. 2021). We have tested that the results in this paper are qualitatively not very different from the TH case.

\(^4\) Private communication from one of the authors (JGB) to the author of Hasinger (2020).
formed during the radiation era by the tidal torques from other PBHs (Nakamura et al. 1997; Sasaki et al. 2016). The other channel, called late binary, PBH binaries are created by tidal capture of PBHs in dense halos (G.D. Quinn et al. 1989; Mouri & Taniguchi 2002; Clesse & García-Bellido 2017a). It is still unclear which gives the dominant contribution; especially the PBH clustering is important as it suppresses the former and enhances the latter (Raidal et al. 2017; Bagui & Clesse 2021). In this paper, we consider the late binary formation channel because the merger rate of early binaries is estimated under the assumption that the PBH mass function does not extend over many orders of magnitude (Kocsis et al. 2018; Gow et al. 2020), and the application to a broad mass function, such as the one of the TH model, is not clearly understood. We quickly comment on that in the Conclusions 7.

We assume that the differential merger rate per unit time, comoving volume, and mass interval takes the following form:

\[ \frac{d^2 n_{\text{merg}}(z, m_1, m_2)}{d \log_{10} m_1 d \log_{10} m_2} = R(0, m_1, m_2) (1 + z)^{\alpha_z} \text{yr}^{-1} \text{Gpc}^{-3}, \]

where, in the case of late binaries, \( R \) is given explicitly by the following expression (Clesse & García-Bellido 2022)

\[ R(0, m_1, m_2) = R_{\text{clust}} f_{\text{PBH}}(m_1) f_{\text{PBH}}(m_2) \left( \frac{m_1 + m_2}{m_1 m_2} \right)^{10/7} \left( \frac{m_1 m_2}{m_1 + m_2} \right)^{5/7}, \]

and the total merger rate is the integral of Eq. (7) over \( dz d \log_{10} m_1 d \log_{10} m_2 \). The clustering dynamics of PBHs and the time dependence of the merger rate are not clearly understood, and we parameterize them with the two constants \( R_{\text{clust}} \) and \( \alpha_z \) (Clesse & García-Bellido 2022; Mukherjee & Silk 2021; Mukherjee et al. 2021) (see also Atal et al. (2022) for a different parameterization of the merger rate redshift dependence).

In practice, we use \( R_{\text{clust}} \) to normalize the total merger rate to the value of \( 45 \text{yr}^{-1} \text{Gpc}^{-3} \), consistently with the upper bound of the 90% credible interval on the local merger rate inferred from GWTC-3 (Abbott et al. 2021a). This amounts to the assumption that all the observed binary BH events during the second observation run of LIGO-Virgo consist of PBH binaries. It requires the value of \( R_{\text{clust}} \) to be of the order of \( 10^{2} - 10^{3} \), depending on the specific value of \( n_s \) (Braglia et al. 2021). Such large values of \( R_{\text{clust}} \) are expected if PBHs are strongly clustered (Clesse & García-Bellido 2022).

The mass distribution of the merger rate for our PBH models is shown in the left panel of Fig. 2. We see that the two models show quite different features. For the LN case, it peaks at the mass \( m_1 \), and quickly decays moving away from it. For the TH case, the merger rate is relatively large even away from the peaks induced by the thermal history of the Universe. Indeed, since Eq. (8) increases as the mass ratio \( q = m_2 / m_1 \) gets small, we expect many events involving secondary masses \( m_2 < 3 \text{M}_{\odot} \). The fact that many of such events are predicted in the TH model has important implications for the computation of the duty cycle in the next Sections.

Since the merger rate of PBHs is expected to grow with redshift, we take \( \alpha_z \) to be a positive index. As a reference value, we use \( \alpha_z = 0 \), corresponding to a constant merger rate as assumed in (Clesse & García-Bellido 2017a,b; Clesse & García-Bellido 2022). We note that, besides \( \alpha_z = 0 \), another typical value often assumed in the literature is \( \alpha_z = 1.3 \) (Raidal et al. 2017, 2019).

**Astrophysical Black Holes.** The phenomenological expression for the merger rate is (Safarzadeh et al. 2020;
Mukherjee & Silk 2021):

\[
\frac{d^2 \tau_{\text{merg}}(z_m, m_1, m_2)}{dm_1 dm_2} = N P(m_1, m_2) \\
\times \int_{z_m}^{\infty} \frac{dz_f}{dz_m} \frac{1}{(1+z_f)^{2.7}} \left(1 + \frac{1+z_f}{2.9}\right)^{5.6}.
\tag{9}
\]

In the equation above, \( N \) is a normalization constant and we fix it in the same way as PBHs by requiring that \( d^2 \tau_{\text{merg}}(z_m, m_1, m_2) \) integrated over the component masses and evaluated at \( z = 0 \) gives a total rate of 45 yr\(^{-1}\) Gpc\(^{-3}\). \( z_f \) and \( z_m \) denote the redshift at which the binaries form and merge respectively and we use the Madau-Dickinson relation (Madau & Dickinson 2014) for the star-formation rate (\( \text{SFR}(z) = (1+z)^2/\left[1 + \left(\frac{1+z}{2.9}\right)^5\right] \)). Finally, \( t_d(z_m, z_f) \equiv t(z_m) - t(z_f) \) is the time delay between formation and merger. As can be seen from the right panel of Fig. 2, the time-delay governs the redshift evolution of the merger rate. Since models of stellar population synthesis predict different values of \( t_d \) that can range from hundreds of Myr up to the age of the Universe, the time-delay is the main source of uncertainty in the computation of the merger rate (Safarzadeh et al. 2020). Following (Mukherjee & Silk 2021), we choose a probability distribution of the time delay that scales as \( 1/t_d \). The joint population distribution on \( (m_1, m_2) \) follows \( P(m_1, m_2) \propto \Theta(m_1 - m_2) m_1^{-2.3}/(m_1 - M_{\text{min}}) \), where the Hagnostides theta ensures that the primary mass is always larger than the secondary one (Abbott et al. 2016a, 2019).

The cutoff in the GW energy spectrum is roughly determined by the total mass of the merger. Since the mass function of the TH model contains PBHs with very small masses, this cutoff is pushed to very high frequencies outside the plot. The single source energy spectrum is given by the following expression in the non-spinning limit (Ajith et al. 2011)

\[
\frac{dE_{\text{GW}}(f)}{df} = \frac{\pi^{2/3}}{3} (GM_{\odot})^{5/3} f^{-1/3} G(f)
\tag{11}
\]

where the function \( G(f) \) describes the frequency dependence during the inspiral, merger, and ringdown phases (see e.g. Eqs. (3.2)-(3.4) of (Braglia et al. 2021)). Since the function \( G(f) \) has a cutoff at the maximal emission frequency of the ringdown phase (\( f_{\text{in}} \) in the notation of (Braglia et al. 2021)), the upper limit of the redshift integration in Eq. (10) is given by \( z_{\text{max}} = \frac{d\tau}{dz} - 1 \).

In Fig. 3, we plot some examples of the SGWB spectra from PBH binaries, together with the noise curves of LIGO O3, O4, O5, and Einstein Telescope (ET). The sensitivity curves for the cross-correlation analysis have been computed using the threshold of SNR = 1 and a frequency resolution of \( \Delta f/f = 0.1 \). We have taken the observation time to be 1 year, except for O3 for which we have chosen the actual observation time described in (Abbott et al. 2021b). For comparison, we also plot the SGWB produced by ABHs, calculated by Eq. (10) using the merger rate in Eq. (9). We see that different population models show different amplitude and spectral shapes of the SGWB.

The amplitude of the SGWB in PBH models is generally larger than that of ABHs if they are to explain all of the LVK observed events\(^5\). Both the LN and TH models have similar amplitudes at the peak sensitivity of LIGO, but their spectral shapes are significantly different. This will be crucial for ET, which has a broader frequency sensitivity, to discriminate between these models (and ABHs). The reason for the different spectral shapes is that the TH population contains binaries with a broad range of total masses and mass ratios. The cutoff in the GW energy spectrum is roughly determined by the total mass of the merger. Since the mass function of the TH model contains PBHs with very small masses, this cutoff is pushed to very high frequencies outside the plot. Similarly, for the LN case, we observe a broadening of the peak for larger values of \( \sigma \).

The redshift dependence of the merger rate affects mainly the amplitude of \( \Omega_{\text{GW}} \) and does not significantly alter its spectral shape. Such redshift-dependent effects are seen in the plots for all the population models. For PBHs, an increasing \( \alpha_s \) increases the amplitude of the SGWB. On the other hand, a larger time delay between the formation and merger of astrophysical binaries implies fewer events that contribute to the background, which is therefore characterized by a smaller amplitude.

4 STOCHASTIC GRAVITATIONAL WAVE BACKGROUND

We now discuss the SGWB generated by BH mergers. The energy density of GWs is expressed as an integral over redshift and masses as follows:

\[
\Omega_{\text{GW}}(f, z_{\text{min}}, z_{\text{max}}) = \frac{\rho_c}{\rho_c} \int_{z_{\text{min}}}^{z_{\text{max}}} dz' \int dm_1 dm_2 \frac{1}{(1+z')^2} H(z')
\times \frac{d^2 \tau_{\text{merg}}(z', m_1, m_2)}{dm_1 dm_2} \frac{dE_{\text{GW}}(f_s)}{df_s},
\tag{10}
\]

where \( f_s \) is the redshifted source frame frequency \( f_s = f(1+z) \). \( \rho_c = 3H_0^2/8\pi G \) is the critical energy density of the Universe. We note that the mass integral is performed over the mass function support, while the redshift integral is between \( z_{\text{min}} = 0 \) and \( z_{\text{max}} \) defined below. Note that, by setting \( z = 0 \), we keep all the binaries in the computation of \( \Omega_{\text{GW}} \), including also those detectable as single events.

5 DUTY CYCLE

5.1 Standard definition

Another property of the SGWB, besides its amplitude and spectral dependence, is its statistical behavior. The duty cycle

\(^5\) The possibility of having mixed populations of PBHs and ABHs, as considered in e.g. (Mukherjee & Silk 2021; De Luca et al. 2021a; Franciolini et al. 2021), would also be very interesting, but for simplicity, we do not consider it here. In that case, also the contribution from PBH-ABH mergers should be taken into account (Kritos et al. 2021; Cui et al. 2021)
is one possible observable to characterize the non-Gaussianity of the SGWB. It is defined by the ratio of the duration of the signal emitted between frequency \( f \) and \( f + df \) and the time interval between two events (Coward & Regimbau 2006):

\[
\frac{dD}{df} = \int dz \frac{dR}{dz} \frac{d\tau}{df}.
\]  

(12)

In the equation above, the event rate per redshift slice \( dR/dz \) is given in terms of the comoving distance \( r = \int_0^z d' \frac{dz'}{H(z')} \) and the merger rate as

\[
\frac{dR}{dz} = \frac{1}{(1 + z)} \frac{4\pi r^2}{H} \int dm_1 dm_2 \frac{d^2 \tau_{\text{merge}}(z, m_1, m_2)}{dm_1 dm_2},
\]  

(13)

and the duration of the signal at frequency \( f \) can be written in terms of the chirp mass as

\[
\frac{d\tau}{df} = \frac{5}{96\pi^{4/3}} (G M_\odot^3)^{-5/3} f^{-11/3},
\]  

(14)

and is also integrated over \( dm_1 dm_2 \) together with the integrand in Eq. (13). The total duty cycle in the frequency range \([f_{\text{min}}, f_{\text{max}}]\) is simply obtained by integrating over frequencies as

\[
\xi = \int_{f_{\text{min}}}^{f_{\text{max}}} df \frac{dD}{df}.
\]  

(15)

With the help of the duty-cycle, we are now in the position to define the popcorn and continuous contributions to the SGWB as follows (Regimbau et al. 2012):

\[
\Omega_{\text{GW}}(f) = \Omega_{\text{GW}}^{\text{pop}}(f) + \Omega_{\text{GW}}^{\text{cont}}(f),
\]  

(16)

where

\[
\Omega_{\text{GW}}^{\text{pop}}(f) \equiv \Omega_{\text{GW}}(f, 0, z^*(f)),
\]  

(17)

\[
\Omega_{\text{GW}}^{\text{cont}}(f) \equiv \Omega_{\text{GW}}(f, z^*(f), z_{\text{max}}),
\]  

(18)

and, at given frequency \( f \), the function \( z^*(f) \) is defined as the redshift that solves \( dD/df(f) = N \). The authors of (Coward & Regimbau 2006; Regimbau et al. 2012) use \( N = 10 \) to define the boundary between continuous and popcorn background whereas a less stringent \( N = 1 \) is used in (Damour & Vilenkin 2000, 2005). We adopt the latter in the following, which physically means that the time interval between two consecutive GW events is the same as the duration of the signal emitted in the frequency bin \([f, f + df]\). In other words, \( N = 1 \) corresponds to the situation where we always find one GW event on average in the frequency bin.

The continuous background consists of overlapping signals at a given frequency, either because the number of sources is very large or because the signal duration is long compared to the time interval between consecutive events. Because of the central limit theorem, the background is well described by Gaussian statistics, implying that the optimal analysis method is the cross-correlation statistic (Allen & Romano 1999). On the other hand, the popcorn background is strongly non-Gaussian and arises from unresolvable signals whose duration is comparable to the distance between consecutive events, in a way that the signal at the given frequency is not continuously present. Because of its non-Gaussian properties, data analysis techniques beyond cross-correlation need to be adopted in searches of popcorn backgrounds (Drasco & Flanagan 2003; Seto 2008, 2009; Thrane 2013; Martellini & Regimbau 2014, 2015; Cornish & Romano 2015; Smith & Thrane 2018; Smith et al. 2020; Yamamoto et al. 2022).

---

**Figure 3.** The spectral energy density of the SGWB \( \Omega_{\text{GW}} \) from PBH binaries for different BH populations. [Top] LN model. [Center] TH model. In both panels, solid, dashed, and dotted lines correspond to \( \alpha_z = 0, 0.5, 1.28 \), respectively. In the LN case, we also vary the width of the mass function \( \sigma \). [Bottom] SGWB of astrophysical origin, for which we vary the time-delay parameter \( t_d \).
5.2 Definition with the horizon distance

The standard expression of the duty cycle, Eq. (12), only depends on the properties of the source population, and it is obtained by integrating the source redshift from zero to infinity. This quantity is useful to describe whether the SGWB is in the continuous or popcorn regime in general discussion. However, when we consider real GW data of a popcorn background, very low SNR (≪ 1) signals would be lost among the detector noise and would not contribute as popcorn. Such events should be removed from the computation when we want to make a theoretical prediction for the purpose of comparing it with an observed value of the duty cycle, possibly achievable by future experiments. Note that this argument applies only when events are sparse and the low SNR signal is isolated from the other events. When low SNR events are numerous, their accumulation could form a continuous SGWB with a detectable amplitude. In that case, we should not remove low SNR events as they are making an essential contribution, although in that case, we have dD/df ≫ 1 and the value of the duty cycle no longer provides useful information on the statistical property as it just becomes Gaussian.

We will make a special note for such a case.

We now derive a detector-dependent duty cycle by replacing the upper limit of the redshift integration with the horizon distance of the detector. This quantity is clearly not an intrinsic property of the background itself, but it rather depends on the detector sensitivity and can be used to compare with future data analysis. As noted above, this detector-dependent duty cycle would provide useful information only for dD/df ≪ 1.

The horizon distance of the detector at a given frequency f depends on the chirp mass and is given by (Abadie et al. 2012; Chen et al. 2021; Carr et al. 2021)

\[ R_{\text{det}}(f, m_1, m_2, z) = \frac{c\pi^{2/3}}{\text{SNR}_{\text{rh}}} \times 2.26 \sqrt{\frac{5}{6}} \left( \frac{GM^2}{c^3} \right)^{5/6} \int f_{\text{max}, \text{det}} \frac{d\tau}{d\ln f} \left( \frac{f}{f_{\text{ISCO}}} \right)^{1/2} , \]

where the detector sensitivity is in the range \([f_{\text{min}, \text{det}}, f_{\text{max}, \text{det}}]\) and

\[ f_{\text{max}}(f) = \max (f_{\text{min}, \text{det}, \text{SNR}}, \min (f, \min (f_{\text{max}, \text{det}}, f_*))) \],

\[ f_* = 2f_{\text{ISCO}} \frac{4m_1 m_2}{(m_1 + m_2)^2} , \]

\[ f_{\text{ISCO}} = 4.4\text{kHz} \frac{M_\odot}{m_1 + m_2} . \]

The horizon distance represents the distance of the furthest detectable source with a signal-to-noise ratio exceeding a given value of SNR.

The generalization of Eq. (12) that includes the sensitivity of the horizon is straightforward. For each frequency f and masses \((m_1, m_2)\), we solve the equation \(R_{\text{det}}(f, m_1, m_2, z_{\text{det}}) = d_0(z_{\text{det}})\) and obtain the redshift corresponding to the horizon distance of the detector \(z_{\text{det}} = z_{\text{det}}(m_1, m_2, f)\). The new expression for the detector-dependent duty cycle is:

\[ \xi_{\text{det}} = \int_{f_{\text{min}}}^{f_{\text{max}}} df \frac{dD_{\text{det}}}{df} = \int_{f_{\text{min}}}^{f_{\text{max}}} df \int_0^{z_{\text{up}}} \frac{dR}{dz} \frac{df}{d\tau} , \]

Table 1. Total duty cycle \(\xi_{\text{det}}\) for different models and detectors.

| O3   | O4   | AdvLIGO | ET   |
|------|------|---------|------|
| LN   | 1.6 \times 10^{-5} | 10^{-4} | 6 \times 10^{-4} | 0.33 |
| TH   | 1.7 \times 10^{-3} | 0.02    | 0.08  | 532.6 |
| ABHs | 2.0 \times 10^{-7} | 1.5 \times 10^{-6} | 7.0 \times 10^{-6} | 2.2 \times 10^{-3} |

where now

\[ z_{\text{up}} = \min \left[ z_{\text{max}}(m_1, m_2, f) ; z_{\text{det}}(m_1, m_2) \right] . \]

Note that, by definition, we always have \(\xi_{\text{det}} \leq \xi\).

6 RESULTS

6.1 Benchmark models

Having introduced the necessary theoretical formalism, we now investigate the detection regimes of the BH populations introduced above. In Fig. 4, we show their duty cycle for the baseline models described in Section 2.

Let us start by commenting on the top-left panel, which compares the detector-independent duty cycle of the three different models (PBHs with LN and TH mass functions, and ABHs). Interestingly, both PBH models predict the duty cycle larger than that of ABHs. This remains the same even for the detector-dependent duty cycle, as can be observed in the other panels, where we fix SNR_{rh} = 1 in Eq. (19) and compute \(dD_{\text{det}}/df\) for the O3, O4, Advanced-LIGO design sensitivity, and ET.

As anticipated in the previous sections, there are two reasons why this happens. The first is the different redshift dependence of the merger rate between PBHs and ABHs (see right panel of Fig. 2). Since Eqs. (12) and (21) are essentially integrals of the merger rate over redshift, many distant PBH events with a very small SNR contribute to the integral, yielding a large \(\xi\) and \(\xi_{\text{det}}\). As explicitly reported in Table 1, the duty cycle of PBHs is generically two orders of magnitude larger than that of ABHs. The second reason is that the duty cycle crucially depends on the mass function of the BH population. This is particularly relevant for the TH PBH model, which is described by a very broad mass function. The model thus predicts many events with several combinations of the total mass and mass ratio, but the largest contribution comes from \(2M_\odot\) BHs. Compared to the astrophysical case or the LN case with a peak at \(30M_\odot\), such small mass BHs typically create events with a longer duration, thus resulting to larger \(dD/df\).

Note that, in Table 1, we show the detector-dependent duty cycle for all models. For the TH model, we see from Fig. 4 that \(\xi \gg 1\) and the SGWB is continuous, and we should look at the detector-independent one \(\xi\). However, we still show the values of \(\xi_{\text{det}}\) since it can be useful information for analyses like (Smith et al. 2020) to know the number of sub-threshold events in each experiment.

In Fig. 5, we show the results by changing the different
redshift dependence of the merger rate $\alpha_z$. As expected, the duty cycle increases for larger values of $\alpha_z$, which corresponds to the growth of the PBH merger rate with redshift. Although it is not easy to appreciate it from the figures plotted in log-scale, the difference in the duty cycle is of $\mathcal{O}(1)$. For the LN model, we also show the effect of increasing or decreasing the width of the LN mass function. Higher values of $\sigma$ increase the duty cycle since a broader tail of the mass function leads to more events with small mass, with large event duration. The variation is more significant than the one for $\alpha_z$.

As the TH model shows quite a different behavior compared to the other models, let us look into some details with the help of Fig. 6. The first two figures show how different mass bins contribute to the SGWB amplitude and the duty cycle. The top-left panel is obtained by integrating Eq. (10) over smaller mass bins. Note that the sum of $\Omega_{GW}$ for each mass bin is not equal to $\Omega_{GW}$ integrated over the full range because mergers with large mass ratio (i.e. BHs paired beyond the mass bins) are not included, but the figure is useful for illustrative purposes. As seen in the figure, the mass bin of [0.1, 1] $M_\odot$ is giving the dominant contribution on the SGWB amplitude, and they are essentially in their inspiral phase as we see the $f^{2/3}$ dependence of the spectrum. The top-right panel of Fig. 6 shows the duty cycle calculated for different mass bins. As can be seen, the mass bin of [0.1, 1] $M_\odot$ is giving the dominant contribution also to the duty cycle. Many of them have a very small SNR, but they sum up as incoherent noise in the detector and produce continuous SGWB, dominating the popcorn component. This can be seen in the bottom-left panel of Fig. 6, which is produced by using Eqs. (16) – (18). As mentioned in Sec. 5.2, in such a case, we should look at the detector-independent duty cycle, but in the figure, we also plot the detector-dependent duty cycle, which helps to see that most of the events indeed have SNR $\ll 1$ and each event cannot be resolved as a single event.

Let us now comment on the bottom-right panel of Fig. 6. We show the redshift corresponding to the horizon distance of the detector $z_{\text{det}}(f, m_1 = 2M_\odot, m_2 = q_m)$ for both ET (solid lines) and AdvLIGO (dashed lines). Together, we also show the popcorn and continuous regimes using a threshold of $dD/df = 1$. As can be seen, only very nearby binaries contribute to the popcorn regime and a very large number of
5

![Graph](image)

Figure 5. Duty cycle for different PBH model parameters. We vary the redshift dependence of the merger rate, parameterized by $\alpha_z$, in the LN and TH models in the top and bottom panels respectively. In the center panel, we vary the width of the LN mass function.

Finally, let us note that in the last part of this Section, we focused on the TH model because the differences with respect to ABHs are maximized in this model. However, the statistical properties of LN PBHs are also different as their duty cycle differs by two orders of magnitude. The next generation of ground-based detectors should be sensitive to such a large difference (Smith et al. 2020), suggesting that the duty cycle could be used to distinguish the primordial and astrophysical nature of BHs, as we elaborate in the next Subsection.

6.2 Distinguishing between PBHs and ABHs using the duty cycle

In the previous subsection, we normalized the amplitude of the SGWB for each model by fixing the merger rate to explain the rate of the observed BBHs in the GWTC-3 catalog. Here, we consider a different setup. Suppose that ET telescope measures a SGWB, but different models can explain the observed amplitude of SGWB as shown in the left panel of Fig. 7. The question we would like to address is: can we use its statistical properties, i.e., the duty cycle, to infer which sources produced it? This information would be highly complementary to the spectral shape of the SGWB.

For this purpose, in this subsection, we normalize the amplitude of the SGWB for all the models to be the same at $f = 10$ Hz, roughly the frequency of the best sensitivity of ET. In this case, the merger rate of PBHs is lower than the one used in the previous subsection and cannot explain the merger rate of observed BBHs. Thus, we consider the situation where the LVK rate is essentially explained by ABHs and PBHs only give a subdominant contribution. We note that the normalization of PBH merger rate has nothing to do with the fraction of CDM as our merger rate model described in Eq. (7) depends on the combination $\rho_{\text{cDM}}/\rho_{\text{PBH}}$, which is the known degeneracy between clustering and abundance of PBHs (Raidal et al. 2017; Clesse & Garcia-Bellido 2022; Vaskonen & Veermäe 2020; Young & Byrnes 2020; Trashorras et al. 2021; Atal et al. 2020; De Luca et al. 2020).

The plot on the right panel of Fig. 7 shows that, indeed, we could pin down the source of the background by looking at its duty cycle. We see that, even if the backgrounds produced from PBHs have amplitudes smaller than those in Fig. 3, as we lowered $\Omega_{\text{GW}}$ to match the one from ABHs at $f = 10$ Hz, their duty cycle is still larger than that of ABHs, and we find an order of magnitude difference between the total duty cycle of ABHs ($\xi_{\text{ET}} = 2.2 \times 10^{-5}$) and of LN PBHs ($\xi_{\text{ET}} = 0.10$). Previous studies (Smith & Thrane 2018; Yamamoto et al. 2022) have simulated the parameter estimation of the non-Gaussian background and showed that the error on $\xi$ could be $10\%$ if the background is detected with a certain SNR. Thanks to the large difference between the scenarios considered, even if we take into account uncertainties that could affect the estimation of the duty cycle, such as glitch noise, we could still expect the next generation of ground-based detectors to tell apart the astrophysical and primordial channels for the formation of BHs.

We end by noting that in the case of the TH model, we have $dD/df \gg 1$, which indicates that the background is always in the continuous regimes and, restricting to this model, standard cross-correlation searches for the SGWB are already ad-
Figure 6. Figures to show some details about the TH model. [Top-Left] Contribution of different mass bins to the energy density of the SGWB. [Top-Right] Contribution of different mass bins to the duty cycle. Solid and dashed lines represent the detector-independent and detector-dependent duty cycle (ET is assumed), respectively. The mass bins are divided as follows: Bin 1 $= [10^{-5}, 10^{-3}] M\odot$, Bin 2 $= [10^{-3}, 10^{-1}] M\odot$, Bin 3 $= [10^{-1}, 10] M\odot$ and Bin 4 $= [10, 10^3] M\odot$. [Bottom-Left] Comparison of the continuous (solid) and popcorn (dashed) contributions, separated by the threshold $dD/df = 1$. [Bottom-Right] We plot $z_{det}(f)$ for different values of the mass ratio $q = m_2/2M\odot$ (solid curves for ET and dashed curves for Advanced-LIGO) on top of the redshift bands separated by the continuous and popcorn regimes.

7 CONCLUSIONS

In this paper, we have studied the popcorn signature as a tool to understand the origin of the SGWB produced by unresolved BBHs at the frequency of ground-based interferometers. A popcorn background is characterized by GW signals that may or may not overlap in the frequency band of the detector, unlike a continuous background where the sensitivity band is constantly occupied by the superposition of GW signals which effectively act as Gaussian noise. In order to characterize the two regimes of the background, we have made use of the duty cycle, which represents the average number of events present in a given frequency band, and is larger (smaller) than one for a continuous (popcorn) background.

Our main finding is that populations of binaries formed from PBHs predict a duty cycle that is generically orders of magnitude larger than the one for astrophysical binaries. This opens up the possibility to discriminate between astrophysical and primordial BHs using the statistical properties of the SGWB. Our results are especially relevant to population analyses of sub-threshold events.

To this purpose, we have also proposed a new way to compute the duty cycle, which takes into account the sensitivity of the detector. Our procedure selects only the events contributing to the duty cycle with SNR larger than a certain threshold, which we take as $\text{SNR} = 1$. This quantity would be useful when comparing the theoretical prediction with an observed duty cycle. Note that the value of the SNR threshold should be determined by how much the data analysis can be sensitive to sub-threshold events and how it affects...
the measurement of the duty cycle. It should be investigated in more detail.

We have shown that each model predicts different values of the duty cycle. Our results indicate that not only are ABHs distinguishable from primordial ones based on their duty cycle, but we can also tell apart different PBH models. In fact, while the LN model (or any other PBH model with a mass function peaked around $\mathcal{O}(10) M_\odot$) is characterized by popcorn signatures that differ from ABHs, the TH model, whose mass function is very broad, does not show popcorn signatures at all. Rather, the background they produce is continuous. The peculiar spectral shape of the SGWB, almost flat in the [1, 1000] Hz range, is yet another observable to distinguish PBHs with very broad mass functions to LN PBHs and astrophysical ones.

We also would like to point out an interesting application of the duty cycle. It can also be used to estimate the number of unresolved multiple simultaneous events in the LVK band. Especially, for the TH model, 2 $M_\odot$ PBHs merge at high frequencies where LVK has a poor sensitivity due to quantum shot noise, and they leave only a signal of the inspiral phase lasting a few minutes inside the detector sensitivity, making it difficult to distinguish them from each other. If their merger rate is high enough, as in the TH model, such events may overlap with other individual events and cause wrong parameter estimation. Eventually, with the much better sensitivity of the Einstein Telescope, we may be able to separate the different components into individual events, but otherwise, sub-threshold events could act as correlated non-Gaussian noise. So, even before the detection of the SGWB, the first indication of the large peak at $\sim 2 M_\odot$ in the TH mass function could be the presence of multiple simultaneous sub-threshold events. The values of the detector-dependent duty cycle calculated in this paper can be used to estimate the probability of such events.

Although our results are very optimistic and show the promising prospects of testing PBHs with the popcorn signature, there are some caveats in our analysis related to the modeling of the merger rate in the clustered PBHs scenario. Indeed, the clustering of PBHs enhances both the probability of forming binaries, but also the one of disrupting them before their merger (Raidal et al. 2019; Trashorras et al. 2021), which is not taken into account in our analysis. Another issue would be to include the early binary formation channel with a better understanding of the effect of a broad mass function. In that case, supposing that we normalize the local merger rate to be $45 \text{yr}^{-1} \text{Gpc}^{-3}$, as done in Sec. 6.1, the difference appears only in the mass dependence of the merger rate. As we have seen in the comparison between the LN and TH models, the major factor affecting the value of the duty cycle is the difference in the mass function. Also, we have seen that the effect of the difference in the merger rate evolution is relatively small. Therefore, we do not expect the duty cycle to change significantly, even if we change the merger rate model to the early binary formation scenario. This should be true in the case of the LN mass function, which is very peaked, and the mass dependence of the merger rate becomes less important. On the other hand, the TH mass function has a broad mass spectrum, and it may cause a non-trivial difference in the value of the duty cycle. However, the duty cycle in this model is much larger than unity, so we do not expect this to change the conclusion that the stochastic background is continuous for the TH model.

For both PBHs and ABHs, the details for the merger rate estimation still need to be thoroughly investigated and should be prepared before the detection of the SGWB. Nevertheless, this work has provided an important step toward using the duty cycle as a new observable to discriminate the primordial and astrophysical origin of the SGWB.

**ACKNOWLEDGEMENTS**

This work is supported by the Spanish Research Projects PGC2018-094773-B-C32 (MINECO) and PID2021-123012NB-C43 (MICINN-FEDER) and the Centro de Excelencia Severo Ochoa Program CEX2020-001007-S. MB and SK are supported by the Spanish Atracción de Talento contract no. 2019-T1/TIC-137177 granted by Comunidad de Madrid, the I+D grant PID2020-118159GA-C42 of the Spanish Ministry of Science and Innovation and the i-LINK 2021 grant LINKA20416 of CSIC. SK is partially supported...
This paper has been typeset from a TeX/LaTeX file prepared by the author.