



Publication Year	2022
Acceptance in OA	2025-01-30T17:39:17Z
Title	Dancing in the dark: detecting a population of distant primordial black holes
Authors	MARTINELLI, Matteo, Scarcella, Francesca, Hogg, Natalie B., Kavanagh, Bradley J., Gaggero, Daniele, Fleury, Pierre
Publisher's version (DOI)	10.1088/1475-7516/2022/08/006
Handle	http://hdl.handle.net/20.500.12386/35743
Journal	JOURNAL OF COSMOLOGY AND ASTROPARTICLE PHYSICS
Volume	08

Dancing in the dark: detecting a population of distant primordial black holes

Matteo Martinelli,^{a,b} Francesca Scarcella,^b Natalie B. Hogg,^{c,b}
Bradley J. Kavanagh,^d Daniele Gaggero,^{e,f,b} and Pierre Fleury^{c,b}

^aINAF - Osservatorio Astronomico di Roma,
via Frascati 33, 00040 Monteporzio Catone (Roma), Italy

^bInstituto de Física Teórica UAM-CSIC, Universidad Autónoma de Madrid,
Cantoblanco, 28049 Madrid, Spain


^cUniversité Paris-Saclay, CNRS, CEA,
Institut de physique théorique, 91191, Gif-sur-Yvette, France

^dInstituto de Física de Cantabria (IFCA), UC-CSIC,
Avenida de Los Castros s/n, 39005 Santander, Spain

^eInstituto de Física Corpuscular, Universidad de Valencia and CSIC,
Edificio Institutos de Investigación, Calle Catedrático José Beltrán 2, 46980 Paterna, Spain

^fDipartimento di Fisica, Università di Torino, via P. Giuria 1, I-10125 Torino, Italy

E-mail: matteo.martinelli@inaf.it, francesca.scarcella@uam.es, natalie.hogg@ipht.fr,
kavanagh@ifca.unican.es, daniele.gaggero@ifc.uv.es, pierre.fleury@ipht.fr

Abstract. Primordial black holes (PBHs) are compact objects proposed to have formed in the early Universe from the collapse of small-scale over-densities. Their existence may be detected from the observation of gravitational waves (GWs) emitted by PBH mergers, if the signals can be distinguished from those produced by the merging of astrophysical black holes. In this work, we forecast the capability of the Einstein Telescope, a proposed third-generation GW observatory, to identify and measure the abundance of a subdominant population of distant PBHs, using the difference in the redshift evolution of the merger rate of the two populations as our discriminant. We carefully model the merger rates and generate realistic mock catalogues of the luminosity distances and errors that would be obtained from GW signals observed by the Einstein Telescope. We use two independent statistical methods to analyse the mock data, finding that, with our more powerful, likelihood-based method, PBH abundances as small as $f_{\text{PBH}} \approx 7 \times 10^{-6}$ ($f_{\text{PBH}} \approx 2 \times 10^{-6}$) would be distinguishable from $f_{\text{PBH}} = 0$ at the level of 3σ with a one year (ten year) observing run of the Einstein Telescope. Our mock data generation code, `darksirens`, is fast, easily extendable and publicly available on GitLab .

Preprint: IFT-UAM/CSIC-22-50

Contents

1	Introduction	2
2	Modelling the black hole merger rate	5
2.1	Preliminary definitions	5
2.2	Primordial black holes	6
2.3	Astrophysical black holes	9
2.4	Normalisation of the redshift distributions	10
3	Generating mock GW catalogues	11
3.1	Parameters	12
3.2	Sketch of the generation algorithm	13
3.3	Distance uncertainty	14
4	Detecting PBHs: the cut-and-count method	15
4.1	Description of the method	16
4.2	Optimal distance threshold and smallest detectable PBH fraction	17
5	Quantifying the PBH fraction: a likelihood-based method	20
5.1	Description of the method	20
5.2	Future constraints on f_{PBH}	21
6	Discussion	23
7	Conclusions	25
A	Merger rate suppression due to PBH clustering	27
A.1	PBH cluster formation	27
A.2	Relaxation time	28
A.3	Cluster collapse	29
A.4	Merger rate suppression factor	29
B	Obtaining the signal-to-noise ratio of an event	31
B.1	Gravitational waves	31
B.2	Luminosity distance	32
B.3	Uncertainty on the lensed luminosity distance	32
C	Lensing	34
C.1	Magnification of gravitational waves	34
C.2	Variance of the magnification in weak lensing	36
C.3	Lognormal probability distribution of the magnification	37

1 Introduction

The idea that a population of primordial black holes (PBHs) may exist in the Universe and constitute a significant portion of the dark matter (DM) has been debated in the literature since the pioneering works of the late 1960s and early 1970s [1–3] (see e.g. [4] for a recent review). The discovery of such a population would have profound implications for fundamental physics, even if it were to represent only a subdominant portion of the DM. The formation of PBHs could reveal precious hints about inflationary and early Universe physics [5–11], and their subsequent evolution could impact structure formation [12] and solve long-standing puzzles related to the early formation of supermassive black holes [13]. Moreover, their detection could help to exclude the presence of thermal weakly interacting massive particles and could strongly constrain models of particle physics which invoke new physics at the Weak scale [14–17].

PBHs whose masses lie between $1 M_{\odot}$ and $100 M_{\odot}$ are particularly promising candidates for study, given the now routine detection of gravitational wave (GW) signals from the mergers of compact objects in this mass range. Gravitational waves from the inspiral, merger and ringdown of a pair of BHs were first detected by the LIGO Scientific Collaboration (henceforth LIGO) and the Virgo Collaboration (henceforth Virgo) in 2015 [18] and the detection of a binary neutron star merger with an electromagnetic counterpart signal was made in 2017 by LIGO and Virgo [19]. In 2020, the KAGRA observatory in Japan [20] joined the global network of GW detectors, reporting its first observations in conjunction with the GEO600 instrument in March 2022 [21].

The analysis of the combined information about the merger rate, mass and spin distribution of low redshift merger events collected by LIGO, Virgo and KAGRA (LVK) since 2015 have allowed upper limits on the PBH abundance $f_{\text{PBH}} = \Omega_{\text{PBH}}/\Omega_{\text{DM}}$ in the aforementioned mass window to be set [22, 23]. This has also triggered a debate about the possibility of identifying a subdominant population of PBHs on top of the (likely dominant) contribution of astrophysical black holes (ABHs) [24–26], motivated in particular by the detected substructure in the mass distribution of these objects [27, 28]. Definitively detecting such a population of PBHs will become a concrete possibility with the advent of the third generation of GW observatories, such as the Einstein Telescope (ET) [29] and Cosmic Explorer (CE) [30].¹ Assessing the potential constraining power of the ET on PBHs is the focus of this work.

The most up-to-date configuration of the planned ET facility is known as ET-D [34], which proposes a three-armed observatory consisting of three interferometers arranged in an equilateral triangle. In the ET-D configuration, each detector is in fact made up of a *pair* of detectors – one sensitive to a lower frequency range and the other to a higher frequency range – thus greatly increasing the overall sensitivity of the instrument with respect to the current generation of GW observatories. Furthermore, the triangular shape of the observatory will enable improved sky localisation of GW events [35]. Lastly, the interferometers and detectors will all be constructed underground, in an effort to reduce seismic noise [36].

The less noise in the detector, the smaller the amplitude of GWs – or *strain* – that can be detected. This noise is typically quantified by the strain amplitude spectral density (ASD),

¹Other proposed third-generation observatories include LISA [31], DECIGO [32], TianQin and Taiji [33], which will all be space-based rather than terrestrial like the ET and CE, allowing for far longer interferometer arms than is possible on Earth. These instruments will hence be sensitive to a very different GW frequency range to that of ET and CE, which means that the latter will have far better prospects for detecting PBHs in the mass range $1\text{--}100 M_{\odot}$.

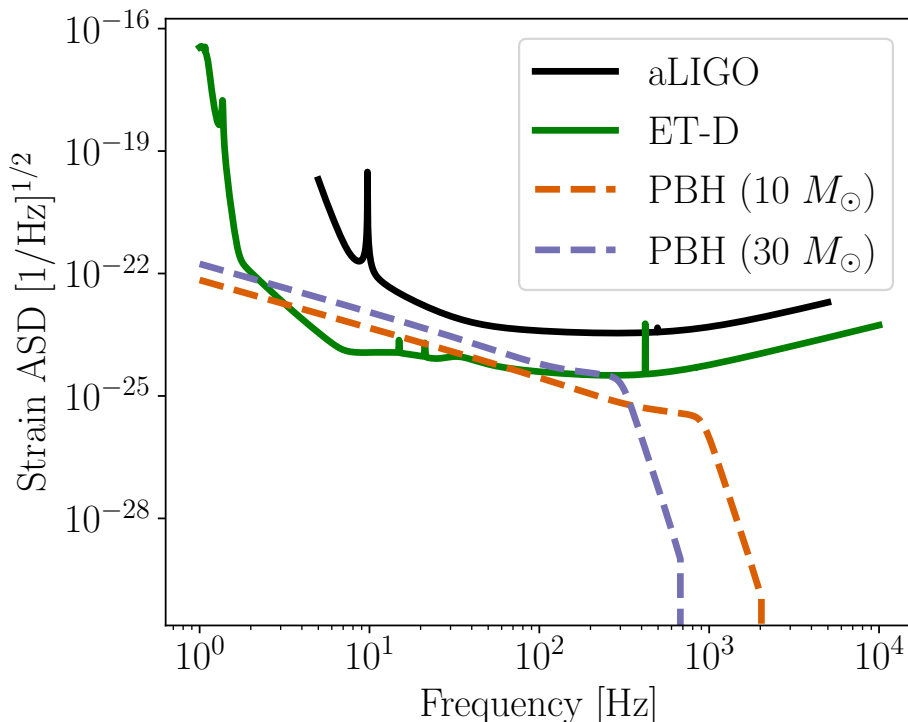


Figure 1. The strain amplitude spectral densities (ASD) of advanced LIGO (black) and the ET-D configuration of the Einstein Telescope (green), along with the characteristic strain of two PBH merger events, with component masses of $10 M_{\odot}$ (orange) and $30 M_{\odot}$ (purple) taking place at $z = 1$, that is, with a luminosity distance $D \approx 6.8$ Gpc.

which we show for advanced LIGO² (aLIGO) and ET-D³ in Fig. 1. These ASD curves effectively show the lowest GW strain that can be detected and we show for comparison the characteristic strains of two PBH merger events with the example masses we use in the rest of this work: $10 M_{\odot}$ and $30 M_{\odot}$. From this plot, we can see that ET-D will be sensitive to strains around two orders of magnitude smaller than what aLIGO can currently detect, and will also probe a much broader range of frequencies.

With all these factors taken together, the ET-D design is expected to yield many more observations of GWs at ever-greater cosmic distances (or redshifts) than the current generation of terrestrial detectors.

A crucial difference between ABH mergers and PBH mergers is the *redshift evolution* of their merger rates. While there is broad consensus on a steeply decreasing rate for ABHs beyond $z \simeq 2$ (see e.g. [37] and references therein), the PBH merger rate is expected instead to be a *monotonically increasing* function of redshift. In fact, a significant number of PBH binaries are expected to form by gravitational decoupling from the Hubble flow before matter–radiation equality. As shown in [22, 38, 39], the distribution of the orbital parameters for these primordial binaries peaks at low values of both semi-major axis and angular momentum, and therefore low values of the merger time (compared to the Hubble time scale). Hence, even a small population of PBHs is expected to dominate the merger rate at early times.

²<https://dcc.ligo.org/LIGO-T1800042/public>.

³<http://www.et-gw.eu/index.php/etsensitivities>.

This idea naturally suggests a focus on high redshift GWs for the purposes of detecting PBHs [40]. They may manifest as resolved individual events [41–44], but also as a measurable contribution to the stochastic GW background, as recently pointed out in [45, 46]. In this article we focus on the former case: the prospects for detecting an anomalously large number of *individual resolved events* at large cosmic distances. Specifically, we propose (i) a careful theoretical modelling of both the PBH and ABH merger rates; and (ii) a realistic assessment of the ability of future observatories – specifically, the ET – to disentangle the two populations.

As far as modelling is concerned, a useful starting point is the formalism for the PBH merger rate developed in [22, 38, 47], based on the assumption of an initially uniform and isotropic distribution of PBHs all having the same mass (in other words, a “monochromatic” mass function). In this scenario, neighbouring PBHs may decouple from the Hubble flow before matter–radiation equality, forming binaries which merge with an ever decreasing rate throughout the age of the Universe. We extend this formalism to include the effects of the early-time formation of PBH clusters [48–50] on the evolution of the binaries. The “background” to these PBH merger events is provided by the mergers of ABHs. The ABH merger rate traces the binary ABH birth rate, with some delay, which in turn is expected to trace the star formation rate, again with some delay [37]. A careful description of PBH and ABH merger rates as a function of redshift is therefore crucial in understanding how well future GW observatories can distinguish these two populations.


In this work, we make the conservative imposition of monochromatic mass functions for both the ABH and PBH populations – in other words, all the BHs in a given population have the same mass – and focus on the redshift dependence of the merger rates as the only discriminant between the two. We also focus on PBHs in the $1 M_{\odot} - 100 M_{\odot}$ mass range, but we note that several other investigations of the capability of future observatories to detect PBHs selected the sub-solar mass range as an interesting avenue for study [51–53].

The uncertainty in the measured luminosity distances of events at high redshift is also expected to play an important role in separating the two populations of BHs. The authors of [44] simulate the response of different detector networks located in Europe and the USA, and assess whether a single-event-based PBH identification can be unambiguous, if the redshift is large enough. The authors conclude that the typical redshift measurement is not precise enough to conclude with certainty that a single source is of primordial origin.

Motivated by these results, we present a framework that implements a statistically sound assessment of the capability of ET to:

- *Detect* an excess of merger events at high redshift with respect to the astrophysical expectation;
- *Measure* the f_{PBH} associated with the detection, if present, or constrain f_{PBH} to lie within some range.

To this aim, we generate a mock data set associated to the null hypothesis of $f_{\text{PBH}} = 0$, and several data sets associated to different PBH fractions. We analyse these data using two different methods. First, we use an intuitive two-bin approach to determine the minimum abundance of PBHs that can be detected as a significant “excess” with respect to the astrophysical background, identifying the optimal binning in redshift. Next, we develop a parameter estimation pipeline, in which we estimate the posterior distribution of f_{PBH} , by comparing the mock data generated for different fiducial values of f_{PBH} with the theoretical distribution of event distances.

The article is structured as follows: in Sec. 2 we discuss our theoretical and phenomenological models for the redshift evolution of the merger rate of primordial and astrophysical BHs; in Sec. 3 we describe how we generate our mock GW catalogues; in Sec. 4 we present the results from the “cut-and-count” method, which assesses ET’s ability to detect PBHs; in Sec. 5 we present the results from the likelihood-based method, which further analyses ET’s ability to measure the PBH fraction; in Sec. 6 we discuss the implications and possible shortcomings of these results and finally in Sec. 7 we conclude. Appendices A to C provide more information on the merger rate suppression due to PBH clustering, the method for computing the signal-to-noise ratio of an event in our mock catalogues, and the effect of lensing on GW measurements. The code associated with the article, `darksirens`,⁴ is publicly available on GitLab .

Conventions, notation, cosmology. We assume a spatially flat Λ CDM cosmology throughout, using $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_m = 0.315$ as reported by the *Planck* collaboration [54]. A three-bar equality sign (\equiv) indicates a definition; an upper-case P indicates a probability, while a lower-case p indicates a probability density function (PDF); bold symbols ($\mathbf{x}, \boldsymbol{\theta}, \mathcal{A}, \dots$) stand for vectors or matrices. The luminosity distance is simply denoted with a D throughout, because we use no other notion of distance. We endeavour with all heroism to keep explicit factors of G and c where relevant.

2 Modelling the black hole merger rate

This section summarises our assumptions regarding the modelling of the merger rate of both ABHs and PBHs, which will then be used to produce mock catalogues of GWs.

2.1 Preliminary definitions

Correctly interpreting a population of observed merger events relies on our expectations for the distribution of mergers. We therefore need to accurately model this quantity. Physical models for the formation and mergers of BHs typically yield the merger rate density $\mathcal{R}(z)$; that is, the number of mergers per unit comoving volume and per unit time in the rest frame of the source, as a function of redshift z . However, from an observational point of view, the relevant quantity is rather the number of mergers that would be observed by an ideal detector per unit redshift and per unit time in the rest frame of the detector. This is what we shall refer to as the *merger rate* [55, 56],

$$R(z) = \frac{\mathcal{R}(z)}{1+z} \frac{dV_c}{dz}. \quad (2.1)$$

The factor of $(1+z)^{-1}$ serves to convert a source-frame rate to a detector-frame rate, and V_c denotes the comoving volume; dV_c/dz then represents the comoving volume of a spherical shell between z and $z+dz$ around the detector,

$$\frac{dV_c}{dz} = 4\pi r^2(z) \frac{c}{H(z)}, \quad \text{with} \quad r(z) \equiv \int_0^z d\zeta \frac{c}{H(\zeta)}, \quad (2.2)$$

the comoving distance at redshift z . Both $r(z)$ and $H(z)$ are computed using the public cosmology solver CAMB [57, 58].⁵ Given a merger rate $R(z)$, the total theoretical number of

⁴<https://darksirens.readthedocs.io>

⁵<https://camb.info>.

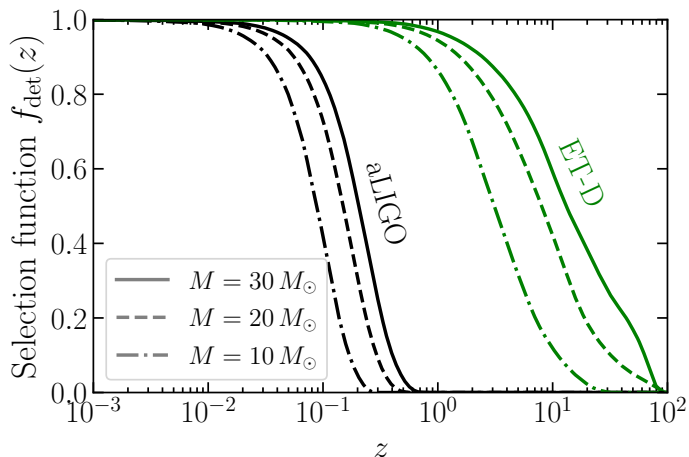


Figure 2. The selection functions $f_{\text{det}}(z)$ for aLIGO (black) and ET-D (green) for three different total masses M : $10 M_{\odot}$ (dot-dashed lines), $20 M_{\odot}$ (dashed lines) and $30 M_{\odot}$ (solid lines). The selection function is defined as the fraction of mergers which are detectable at a given redshift.

events that would be detected by an infinitely sensitive instrument during an observation time T_{obs} reads

$$\bar{N}_{\text{tot}} = T_{\text{obs}} \int_{z_{\text{min}}}^{z_{\text{max}}} dz R(z), \quad (2.3)$$

where we limit ourselves to mergers in a redshift range $z \in [z_{\text{min}}, z_{\text{max}}]$. Their redshift distribution, still in the case of an ideal detector, is given by

$$p(z) = \frac{T_{\text{obs}} R(z)}{\bar{N}_{\text{tot}}} = \frac{R(z)}{\int_{z_{\text{min}}}^{z_{\text{max}}} d\zeta R(\zeta)}. \quad (2.4)$$

Importantly, not all of these mergers will have a signal-to-noise ratio (SNR) large enough to be detectable. We therefore define the selection function $f_{\text{det}}(z)$, which encodes the fraction of mergers that are detectable at a given redshift. We compute this explicitly for aLIGO and ET-D by sampling over binary orientations and sky positions, following the approach of Ref. [59], and requiring a threshold of $\text{SNR}_{\text{min}} = 8$ for an event to qualify as detected. We illustrate these selection functions in Fig. 2 for a range of BH masses. With this, the rate of *detectable* events is given by

$$R_{\text{det}}(z) = f_{\text{det}}(z) R(z). \quad (2.5)$$

The expected number of detected events, \bar{N}_{det} , and their redshift distribution, $p_{\text{det}}(z)$, are defined analogously to Eqs. (2.3) and (2.4), respectively, by substituting $R(z)$ with $R_{\text{det}}(z)$.

In the following subsections, we outline our computation of the merger rate density $\mathcal{R}(z)$ for ABH and PBH mergers. In Fig. 3, we show examples of these merger rate densities (left panel), along with the corresponding redshift distribution $p(z)$ (right panel). These illustrate why high redshift mergers are so crucial to discovering and constraining a PBH population; above $z \sim 30 - 40$ the ABH rate quickly becomes negligible and mergers from a potential PBH population come to rapidly dominate.

2.2 Primordial black holes

PBH binaries are efficiently formed deep in the radiation era, in what is known as the early-time formation channel. Binaries can also be formed in dense environments through

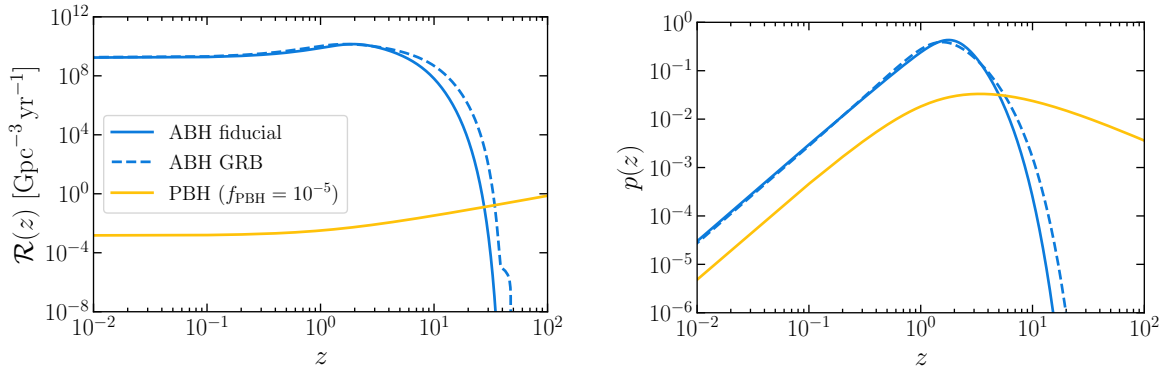


Figure 3. Distribution of merger events as a function of redshift for ABHs and PBHs. We show results for our fiducial ABH model (solid blue), the more optimistic ABH model (dashed blue) based on GRB data (Sec. 2.3), and our PBH model (Sec. 2.2) with $f_{\text{PBH}} = 10^{-5}$ (yellow). **Left:** Comoving merger rate density per unit volume per unit time $\mathcal{R}(z)$. **Right:** Redshift probability distribution $p(z)$, defined in Eq. (2.4). This corresponds to the distribution of observed event redshifts for an ideal detector, ignoring the effects of a selection function $f_{\text{det}}(z)$.

gravitational capture or three-body interactions [60, 61]. These late-time formation channels are expected to become increasingly important for large PBH abundances and low redshifts. In this work, we focus on high redshift observations and our analysis is restricted to $f_{\text{PBH}} \lesssim 10^{-3}$ (in agreement with current constraints, see Sec. 2.4). In this case, we can consider the merger rate to be dominated by the contribution of binaries formed in the early Universe. We compute it following the derivation of [22], which builds on several earlier works [38, 39, 47].

The PBHs are assumed to be initially Poisson distributed (we neglect any possible initial clustering, see e.g. [62–64]) with negligible initial velocity dispersion. Under mutual gravitational attraction, PBH pairs decouple from the Hubble flow and form binary systems. Head-on collisions are avoided thanks to the small angular momentum provided by the presence of surrounding PBHs and smooth density perturbations. After formation, the binaries slowly lose energy through GW emission and eventually merge. The coalescence time (for small values of j) is given by [65]

$$t_{\text{merger}} = \frac{3}{170} \frac{c^5}{G^3 M_{\text{PBH}}^3} a^4 j^7, \quad (2.6)$$

where M_{PBH} is the PBH mass, a is the binary’s initial semi-major axis and j is the initial dimensionless angular momentum, related to the eccentricity e through $j \equiv \sqrt{1 - e^2}$. As the binaries which merge within a Hubble time typically form early in the radiation era [22], we identify the coalescence time with the cosmic time at which the merger occurs.

Given Eq. (2.6), we can estimate the merger rate density from the initial distribution of the orbital parameters (a, j) as

$$\mathcal{R}_{\text{PBH}}[z(t)] = \frac{1}{2} n_{\text{PBH}} \int da dj p(j|a) p(a) \delta[t - t_{\text{merger}}(a, j)], \quad (2.7)$$

where $n_{\text{PBH}} = f_{\text{PBH}} \rho_{\text{DM},0} / M_{\text{PBH}}$ is the comoving number density of PBHs, and $\rho_{\text{DM},0}$ is the energy density of cold dark matter at redshift zero. We obtain distributions for the semi-major axis $p(a)$ and for the angular momentum $p(j|a)$ following [22, 23]. The latter distribution

is computed taking into account the torques generated by all the surrounding PBHs and from the density perturbations in the matter fluid. Most binaries are born with very high eccentricities and small semi-major axes, which correspond to short times to merger, see Eq. (2.6). It follows that the PBH merger rate increases with redshift, with a larger number of binaries merging at early times, as shown in the left panel of Fig. 3.

In estimating the initial distribution of orbital parameters, we do not take into account the early-time disruption of binaries by neighbouring PBHs, which ultimately results in a suppression of the merger rate, as shown in [50]. However, in this work we restrict ourselves to densities $f_{\text{PBH}} \lesssim 10^{-3}$ (see Sec. 2.4 for details) for which the effect is negligible [26].

The expression for the merger rate given in Eq. (2.7) assumes that the orbital parameters evolve exclusively through GW emission. In fact, a and j can be altered via a number of mechanisms. Apart from the early-time disruption by neighbours discussed above, one of these is accretion, which can modify the PBH mass and shrink the binary [66–70]. The effect of accretion is particularly relevant for high PBH masses and results in an enhancement of the high mass tail of the mass function: we leave the study of this effect to a future work where extended mass functions will be considered. Furthermore, if PBHs exist side-by-side with particle DM, the latter forms mini-haloes around the PBHs. In this case, another perturbation to the binary can come from the dynamical friction that the mini-haloes induce on the PBHs when they approach each other [23]. While the binaries are dramatically perturbed by this interaction, the shrinking of the binary compensates the increase of angular momentum in such way that the coalescence time is almost unaffected.

The simplified picture of isolated binaries which we have presented so far fails to capture an important aspect: under the action of gravity, PBHs form bound clusters, where complex N -body interactions take place. Various works have been dedicated to studying the effect of clustering on the evolution of binaries [50, 71–74], showing that it can have a significant impact on the merger rate. In the most extreme scenarios, binaries can be easily disrupted in these dense environments. But even if the binary is not completely disrupted, it must be extremely eccentric in order to merge within a Hubble time; given the strong dependence of the time of merger on the angular momentum, $t_{\text{merger}} \propto j^7$, even a small increase in j due to interactions within the cluster is sufficient to delay the merger beyond our time.

In this work, we quantify the suppression of the merger rate due to clustering following the semi-analytical modelling presented in [72]. This calculation is based on estimating the fraction of clusters and sub-clusters that undergo core collapse following gravo-thermal instability. It is assumed that all binary systems within these structures do not contribute to the merger rate: they end up being perturbed in the high density cores in such a way that their coalescence time exceeds the age of the Universe. In this sense, the calculation can be considered an over-estimation of the effect (while it is very likely that a binary is perturbed in a dense environment, it is not necessarily so; furthermore, new binaries can be created in the cores). However, this may be partially compensated by the fact that the formalism does not consider perturbations to binaries in stable clusters (i.e. those not affected by core collapse). The calculation of the suppression factor is detailed in Appendix A.

In Fig. 4, we plot the PBH merger rate for different values of M_{PBH} and f_{PBH} , with and without the effects of clustering. The suppression factor increases going towards low redshifts, as clusters of increasingly larger size have sufficient time to undergo core collapse. As expected, it also increases with f_{PBH} . As we explain in Appendix A, for small PBH fractions $f_{\text{PBH}} \lesssim 10^{-2}$, the effect becomes negligible, as most clusters in this case form late in the Universe and do not have sufficient time to undergo gravo-thermal collapse before today.

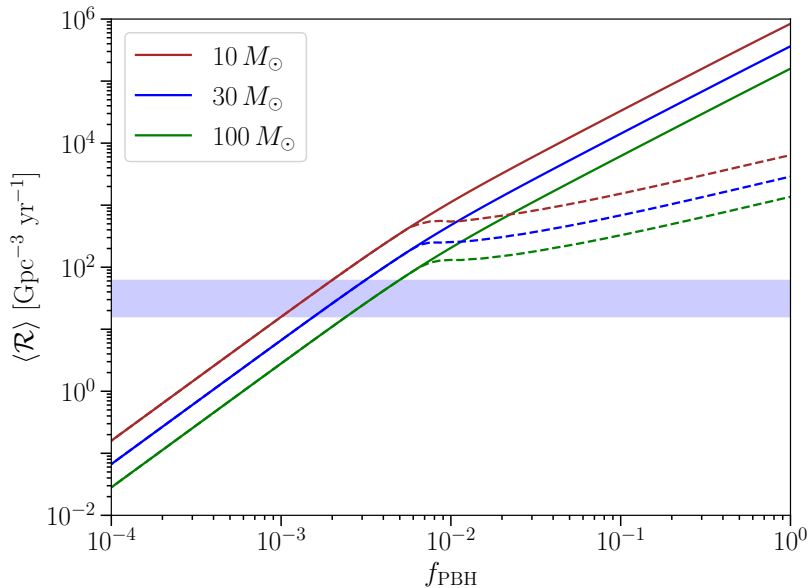


Figure 4. Average merger rate density of PBH binaries formed in the early Universe, see Eq. (2.10), as a function of f_{PBH} , for $M_{\text{PBH}} = 10 M_{\odot}$, $M_{\text{PBH}} = 30 M_{\odot}$ and $M_{\text{PBH}} = 100 M_{\odot}$. The dashed lines are obtained considering the disruption in clusters as described in Sec. 2.2. The coloured band indicates for comparison the latest measure of the total binary BH merger rate as reported in [28].

As the disruption in clusters becomes greater with time, it causes an overall enhancement of the slope of the function $\mathcal{R}(z)$, making small values at $z = 0$ compatible with larger overall PBH abundances. Given that current bounds from GWs lie in the range $f_{\text{PBH}} \sim 10^{-3} - 10^{-2}$, where this effect may still be relevant, we update these to incorporate clustering in Sec. 2.4.

2.3 Astrophysical black holes

The ABH merger rate depends on the rate of formation of ABHs from stars via supernovae. We note that there are two channels through which ABH binaries can form: directly from binary stellar systems or from pairs of originally isolated black holes. In both scenarios, the ABH merger rate density \mathcal{R}_{ABH} can be expressed as [37, 75]

$$\mathcal{R}_{\text{ABH}}(t, M_{\text{ABH}}) = \mathcal{N}_{\text{ABH}} \int_{\Delta t_{\text{min}}}^{\Delta t_{\text{max}}} d\Delta t p(\Delta t) \mathcal{R}_{\text{birth}}(t - \Delta t, M_{\text{ABH}}), \quad (2.8)$$

where $\mathcal{R}_{\text{birth}}(t, M_{\text{ABH}})$ is the birth rate density of the stellar remnants as a function of redshift and mass, and $p(\Delta t)$ is the distribution function of the time delay between ABH formation and merger. In general, this time delay depends on the detailed rate of binary formation and on the orbital parameters of the resulting binaries (semi-major axis and eccentricity). In our model we follow the simplified approach of [37] and express this distribution as $p(\Delta t) \propto 1/\Delta t$ for $\Delta t_{\text{min}} \leq \Delta t \leq \Delta t_{\text{max}}$ with $\Delta t_{\text{min}} = 50 \text{ Myr}$ and $\Delta t_{\text{max}} = H_0^{-1}$. This expression is motivated by the high-resolution numerical simulations of binary BH formation via the evolution of isolated binary stars presented in [76]. However, there are a number of uncertainties in the rate of ABH mergers [77–83]. We therefore include an overall normalising factor \mathcal{N}_{ABH} , which we fix based on the observed low redshift merger rate, as we will describe in Sec. 2.4.

We assume that the ABH birth rate density $\mathcal{R}_{\text{birth}}(t, M_{\text{ABH}})$ is proportional to the star formation rate (SFR) $\psi_{\text{SFR}}(z)$. Galaxy surveys performed in the ultraviolet band and in

the far-infrared band can directly measure the instantaneous star formation rate density at different redshifts, with the former range of wavelengths being a direct tracer of short-lived massive stars, and the latter being a signature of UV light emitted by the same population of stars and subsequently absorbed and re-emitted by dust. In-depth analyses of these data allow for modelling of the cosmic history of the SFR, unambiguously highlighting a rising trend at low redshift up to a peak of very intense star formation in the range $1 \leq z \leq 2$, followed by a fall-off at large redshift [84, 85].

However, the high redshift behaviour of the SFR may be affected by significant uncertainties and biases, mainly due to dust obscuration and to the fact that early star formation took place in very faint galaxies which are typically missed in existing surveys [86]. Therefore we also consider an alternative tracer of the high redshift SFR, via the rate of gamma ray bursts (GRBs). Several studies (see e.g. [87] and references therein) based on this observable have indicated a higher SFR at high redshift compared to studies based on galaxy counts. With this method, the main uncertainty is the (model-dependent) relation between the rate of GRBs and the star formation rate itself.

We follow [88] and adopt the following parametrisation for the SFR,

$$\psi_{\text{SFR}}(z) = k \frac{a e^{b(z-z_m)}}{a - b + b e^{a(z-z_m)}}. \quad (2.9)$$

The free parameters k, a, b, z_m in this expression are usually fitted by comparison to existing catalogues of different tracers of star formation. Aiming to bracket the uncertainties associated with the astrophysical rate, we consider two different choices for these parameters. The reference set of parameters corresponds to the *fiducial model* quoted in [41] based on a fit to observations of bright galaxies. The values are $k = 0.178 M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$, $z_m = 2$, $a = 2.37$, $b = 1.8$. We also consider a GRB-based fit as a *maximal model*, with the values taken from [87]: $k = 0.146 M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}$, $z_m = 1.72$, $a = 2.8$, $b = 2.46$.

We stress that, unlike in the PBH case, the overall normalisation factor \mathcal{N}_{ABH} of the astrophysical rate is not determined *a priori*. Instead, we set this normalisation using the low redshift binary BH merger rate data provided by GW observatories currently in operation. Since in this work we are considering a hybrid scenario where a portion of the events is ascribed to a PBH population, each value of f_{PBH} actually corresponds to a different normalisation factor for the rate associated to the ABH population, so that the integrated rate in the redshift range $z \leq 1$ matches the observed one. The procedure to match the data naturally provides an upper limit on f_{PBH} , set by the requirement not to overshoot the rate measurement. The details of the procedure, and the corresponding bound on f_{PBH} , are described next.

2.4 Normalisation of the redshift distributions

Given the models for the redshift evolution of the ABH and PBH rate density described above, we now aim to: *(i)* obtain an expression for the actual merger rate by convolving the theoretical rate estimates with the detector space-time sensitivity; *(ii)* compare the predicted merger rate for each value of f_{PBH} with the recent estimates provided by LVK. In this way, we will simultaneously set the normalisation of the astrophysical component for each value of f_{PBH} , and obtain an upper limit on this quantity.

We compare the predicted PBH merger rate to the observational data from the second and third Gravitational-Wave Transient Catalogs (GWTC-2 and GWTC-3, produced by LIGO–Virgo and LVK respectively) for different PBH mass intervals and values of f_{PBH} . In particular, we adopt the merger rate density estimates reported in Table 4 of [27] for GWTC-2

and Table 4 of [28] for GWTC-3.⁶ These are reported assuming a non-evolving rate, while our models for the ABH and PBH merger rates evolve with redshift. We therefore compute the average merger rate density $\langle \mathcal{R} \rangle$, defined as

$$\langle \mathcal{R} \rangle = \frac{\int_0^{z_{\max}} R_{\text{det}}(z) dz}{\int_0^{z_{\max}} \frac{f_{\text{det}}(z) dV_c}{1+z} dz}. \quad (2.10)$$


This corresponds to the non-evolving merger rate density which would give rise to the same number of expected detectable events as our redshift-dependent model, over the redshift range $z \in [0, z_{\max}]$. For $f_{\text{det}}(z)$ we use the selection function for aLIGO shown in Fig. 2 and we fix $z_{\max} = 2$.⁷ For each mass bin considered in these catalogues, we make the conservative choice to exclude the values of f_{PBH} that would correspond to an average merger rate density $\langle \mathcal{R} \rangle$ that exceeds the measured one by 2σ . For values of f_{PBH} smaller than this bound, we adjust the ABH merger rate normalisation \mathcal{N}_{ABH} so that the sum of the averaged rates corresponds to the reported one, that is

$$\langle \mathcal{R}_{\text{PBH}}(f_{\text{PBH}}) \rangle + \langle \mathcal{R}_{\text{ABH}}(\mathcal{N}_{\text{ABH}}) \rangle = \mathcal{R}_{\text{obs}}^{\text{LVK}}. \quad (2.11)$$

We have hence defined an allowed region for f_{PBH} and properly set the normalisation of the astrophysical rate for each value of f_{PBH} . We show the resulting new upper bounds we obtain on f_{PBH} in Fig. 5. These bounds are compatible with previously reported bounds from GWTC-2 including the contribution of ABH mergers [26].

We find that the suppression due to late-time clustering, described in Sec. 2.2 and Appendix A, does not impact this upper limit. Its effect is relevant only for values of f_{PBH} larger than $\mathcal{O}(10^{-2})$, as shown in Fig. 4. Though the rate is significantly suppressed by clustering for large values of f_{PBH} , the predicted PBH merger rate would still exceed the merger rate observed by LVK, and large values of f_{PBH} remain excluded. Around our reported bound of $f_{\text{PBH}} \lesssim 10^{-3}$, the suppression becomes negligible. Even so, we emphasise that the upper limit has to be taken *cum grano salis*, given the strong assumption of a monochromatic mass function. We cannot exclude scenarios featuring non-trivial mass functions and non-negligible initial clustering, which may potentially evade the bound. The study of such scenarios are beyond the scope of the current work.

3 Generating mock GW catalogues

Having presented our merger rate calculation and the subsequent updated upper bounds on f_{PBH} from current data, we are ready to turn to the core of this work: assessing the ability of the ET to disentangle ABH mergers from PBH mergers using their redshift distributions. This section presents our method to produce mock data for the ET, accounting for the instrumental response and other observational effects such as lensing of the signal. Since we focus on the information contained in the redshift distribution of the GW events – or, more accurately, their luminosity distance distribution which is the actual observable – in the remainder of this article, a *data set* \mathcal{D} will refer to N_{det} luminosity distance measurements D_i with their uncertainty σ_i , $\mathcal{D} = \{(D_i, \sigma_i)\}_{i=1, \dots, N_{\text{det}}}$. Our mock generation code, **darksirrens** , is publicly available and could be easily modified to include other observables.

⁶For the results presented in Sec. 4 and Sec. 5, we fix the normalisation specifically using GWTC-3, from which the best estimate of merger rate density of BH binaries is $\mathcal{R} = 22 \text{ Gpc}^{-3} \text{ yr}^{-1}$.

⁷In practice, this integral is largely insensitive to the precise value of z_{\max} , as the selection function drops rapidly above $z \sim 0.1$.

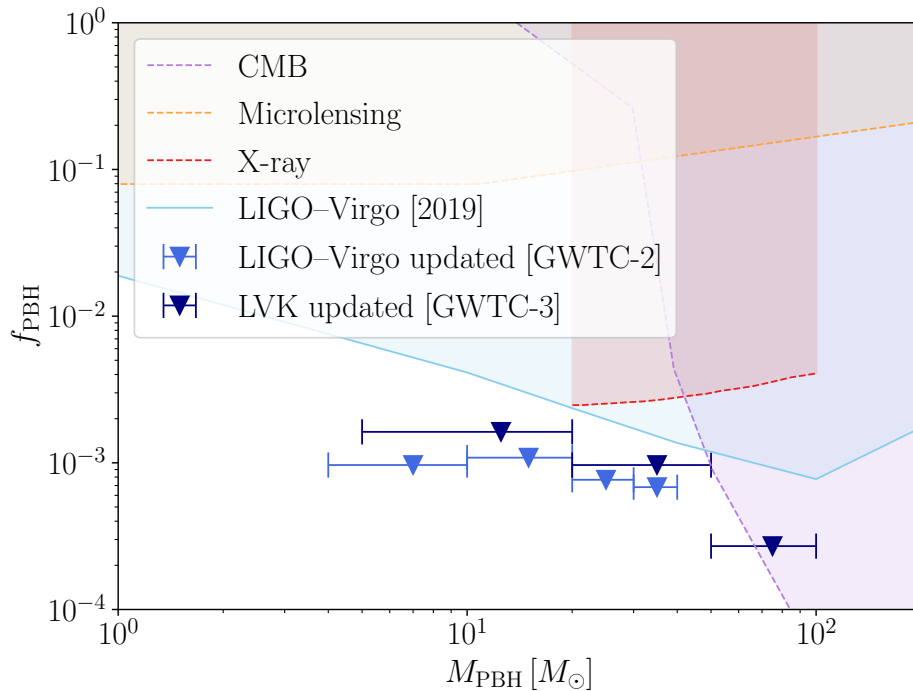


Figure 5. Updated LIGO–Virgo and LVK bounds on f_{PBH} , based on the observed low redshift merger rate reported in recent GW transient catalogues. The various bins in M_{PBH} correspond to the reported upper limits on the merger rate across different mass bins. For comparison, we also show bounds from the CMB due to PBH accretion [89], microlensing constraints from the high redshift star Icarus [90], and X-ray observations of the Milky Way [91]. We obtained these bounds from the [PBHbounds repository](#) [92].

3.1 Parameters


The parameters that must be set to produce a mock catalogue of GW distance measurements can be divided into four categories:

- **Cosmology.** We assume a spatially flat homogeneous and isotropic Λ CDM cosmological background. The parameters to be set are the Hubble–Lemaître constant H_0 and the total matter density parameter Ω_m .
- **Primordial black holes.** Following Sec. 2.2, the PBH population is characterised by the masses of the objects, M_{PBH} , the fraction of DM made up of PBHs, f_{PBH} , and whether or not clustering of PBHs is considered (we remark once again that the effect of clustering is negligible within the current implementation).
- **Astrophysical black holes.** Following the model outlined in Sec. 2.3, the ABH population is characterised by the masses of the objects M_{ABH} and the SFR parameters z_m , a and b which enter into Eq. (2.9).
- **Specifications.** This class of parameters allows us to set the observational time T_{obs} of the survey and the SNR threshold, SNR_{min} , which determines if a candidate event is detected or not. It also allows us to customise details of the mock generation, i.e. to specify the redshift range over which calculations are done ($z \in [z_{\text{min}}, z_{\text{max}}]$) and whether or not to include the effect of lensing.


We remind the reader that we assume for simplicity a single mass for PBHs, M_{PBH} , and a single mass for ABHs, M_{ABH} . Recall also that the normalisation of the ABH merger rate \mathcal{N}_{ABH} is not set as an external parameter but is fixed by comparison with low redshift GW observations, as detailed in Sec. 2.4.

For the results presented in Secs. 4 and 5, we fix the parameters used to generate the mock data to the fiducial values listed in Table 1, unless otherwise specified in the text. We report the survey specifications assumed for the ET observations in Table 2.

Cosmology		ABH parameters				PBH parameters	
Ω_m	H_0 [km s ⁻¹ Mpc ⁻¹]	$M_{\text{ABH}} [M_\odot]$	z_m	a	b	$M_{\text{PBH}} [M_\odot]$	clustering
0.315	67.4	7	2	2.37	1.8	10	yes

Table 1. Cosmological and BH related parameters accessible in `dark sirens` . The values shown here represent our baseline settings used throughout the article. We do not report a value of f_{PBH} as this will be changed depending on the analysis done with the data.

Specifications				
T_{obs} [yrs]	SNR_{min}	lensing	z_{min}	z_{max}
1	8	yes	0.001	100

Table 2. Parameters available in `dark sirens`  to specify the characteristics of the survey and the redshift range over which the mock is constructed. The values shown here represent our baseline settings used throughout the article.

3.2 Sketch of the generation algorithm

Once the parameters are chosen, `dark sirens` produces a mock catalogue \mathcal{D} of GW distances with their uncertainties as they would be measured by ET. For $f_{\text{PBH}} \neq 0$, the data is a mix of ABH and PBH mergers, which takes into account the potential clustering of the latter, the effect of lensing on distance measurements, and ET’s instrumental uncertainties. We now explain exactly how the mock data are generated.

The very first step consists in computing the total number N_{tot} of mergers occurring in the redshift range $[z_{\text{min}}, z_{\text{max}}]$. We draw a random number from a Poisson distribution whose mean, \bar{N}_{tot} , is given by Eq. (2.3), in which we consider the total merger rate $R(z) = R_{\text{ABH}}(z) + R_{\text{PBH}}(z)$ as described in Sec. 2. Each data point $i = 1, \dots, N_{\text{tot}}$ is then produced as follows:

1. Randomly draw the “true” redshift z_i of the merger from the redshift distribution $p(z)$ given in Eq. (2.4), once again considering the total merger rate.
2. Convert z_i into the “true” unlensed luminosity distance $\bar{D}_i \equiv \bar{D}(z_i)$ of the merger, using the background cosmological model and CAMB.
3. Compute the unlensed signal-to-noise ratio (SNR) $\bar{\rho}_i$ of the event. The position, polarisation and inclination of the event are drawn randomly and the SNR computed based on these quantities along with the specifications of the ET. The full details of this calculation are given in Appendix B.

4. Determine the weak gravitational lensing magnification μ_i of the event by randomly drawing it from the theoretical PDF $p(\mu)$ – see Appendix C for details on lensing and its statistics. The lensing magnification enhances (or reduces) the SNR as $\rho_i = \sqrt{\mu_i}\bar{\rho}_i$, and reduces (or increases) the luminosity distance as $\tilde{D}_i = \bar{D}_i/\sqrt{\mu_i}$.
5. Determine whether the SNR is large enough for the event to be properly detected: if $\rho_i < 8$, then the event is considered to be too faint to be a true GW candidate and is removed from the catalogue. The choice of discarding events with an SNR smaller than eight follows the approach taken by the LIGO collaboration as an estimate for the detection threshold [93].
6. Compute the measured luminosity distance, $D_i = \tilde{D}_i + \Delta D_i$, where ΔD_i represents the instrumental error on the measurement. The latter is drawn from a Gaussian distribution $\mathcal{N}(0, \sigma_i^{\text{inst}})$, with error inversely proportional to the SNR, $\sigma_i^{\text{inst}} = 2\tilde{D}_i/\rho_i$ [94].
7. Compute the total uncertainty σ_i on this data point as the quadratic sum of the instrumental and lensing uncertainties,

$$\sigma_i^2 = (\sigma_i^{\text{inst}})^2 + (\sigma_i^{\text{lens}})^2 = \left(\frac{2\tilde{D}_i}{\rho_i}\right)^2 + \sigma_\kappa^2(z_i)\bar{D}_i^2, \quad (3.1)$$

where $\sigma_\kappa^2(z)$ is the variance of the weak-lensing convergence – see Appendix C for details.

The end product is a catalogue of distance measurements with their uncertainties, $\mathcal{D} = \{(D_i, \sigma_i)\}_{i=1, \dots, N_{\text{det}}}$, where $N_{\text{det}} \leq N_{\text{tot}}$ is the number of events that survive the SNR cut of step 5. We have checked that N_{det} agrees on average with the theoretical expectation \bar{N}_{det} . Figure 6 shows an example of a catalogue that can be obtained using this approach. The figure shows the distance D_i of the events that survive the SNR cut,⁸ together with their relative error, for both ABH (blue) and PBH (yellow). The data shown here are obtained setting $f_{\text{PBH}} = 10^{-5}$.

3.3 Distance uncertainty

Correctly modelling uncertainties in the luminosity distance of the GW events is crucial to disentangling high redshift and low redshift populations of mergers. A full analysis pipeline based on real GW data would provide as an output a probability distribution for the true luminosity distance $P(\bar{D})$ of the merger. Instead, in our catalogues, the distance uncertainty is described by only a single number σ_i , given in Eq. (3.1). We therefore model the probability distribution for the luminosity distance as a Gaussian, given explicitly as:

$$p(\bar{D}|D_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(\bar{D} - D_i)^2}{2\sigma_i^2}\right]. \quad (3.2)$$

In the statistical modelling which we will describe in Sec. 4 and Sec. 5, it is important to distinguish between $P(D_i|\bar{D})$ and $p(\bar{D}|D_i)$. In particular, some of the calculations presented in those sections are simplified if we identify the right hand side of Eq. (3.2) as $p(D_i|\bar{D})$, rather than $p(\bar{D}|D_i)$. However, this does not follow an intuitive definition of what is typically meant by the measurement error. We would typically consider that D_i is our best estimate of the

⁸A plot including the events which would not survive the SNR cut is shown in Fig. 13.

luminosity distance \bar{D} , with σ_i parametrising our uncertainty on \bar{D} , matching the definition given in Eq. (3.2). We also find that interpreting the right hand side of Eq. (3.2) as $p(D_i|\bar{D})$ leads to inconsistencies. Consider the case where both D_i and σ_i are large, in which case the right hand side of Eq. (3.2) is a relatively flat function of \bar{D} . If we interpret this as $p(D_i|\bar{D})$, it would imply that a wide range of \bar{D} values would all give rise to the same observed D_i with similar probabilities. However, we know that the measurement error becomes smaller as we reduce \bar{D} , in which case large D_i values should be very unlikely. The only consistent definition for the error is therefore given by Eq. (3.2).

We will now present two different analysis methods of the mock data we have generated: an intuitive binned approach, and a more complete Bayesian analysis of the data, in Sec. 4 and Sec. 5 respectively. With the first method we aim to estimate the detectability threshold of a PBH population, while with the second method we also assess the capability of the ET to measure the PBH fraction. We refer to these methods as “cut-and-count” and “likelihood-based” respectively.

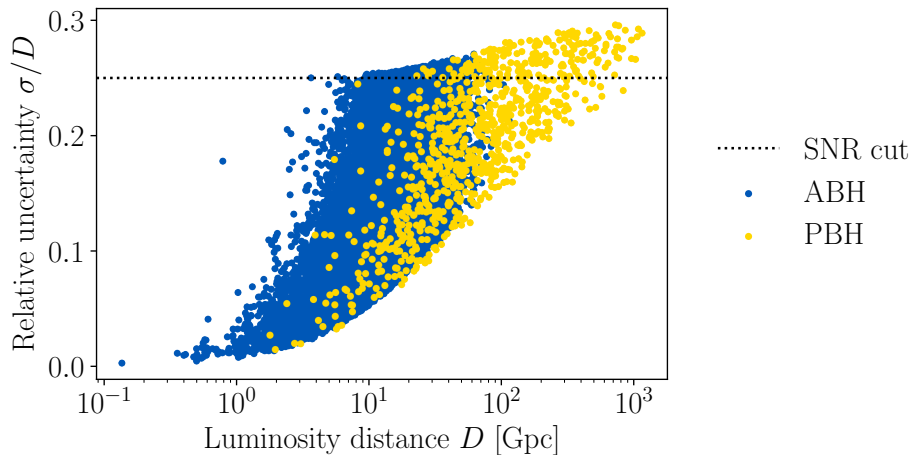


Figure 6. Data set generated using `dark sirens` with the baseline parameters of Table 1 and Table 2, with $f_{\text{PBH}} = 10^{-5}$. The blue dots show the observed ABH events, while the yellow dots represent the observed PBH events. At a fixed distance D , the typical uncertainty is smaller for PBHs than for ABHs due to our choice of a larger PBH mass ($M_{\text{PBH}} = 10 M_{\odot}$, $M_{\text{ABH}} = 7 M_{\odot}$). The dotted horizontal line shows the level of instrumental uncertainty corresponding to our SNR cut of $\rho_i = 8$. The events that lie above this line still satisfy the SNR condition we impose, but receive an extra contribution to the error from lensing.

4 Detecting PBHs: the cut-and-count method

In this section, we present the “cut-and-count” method which we use to determine the smallest PBH fraction that could be detected by the ET. The idea is simple and intuitive: since we expect no ABHs to be formed beyond some high redshift (as there is a significant delay between the beginning of the Universe, the birth and death of the first stars and hence the formation and merging of ABHs), any sufficiently high redshift GW event produced by a BH merger should be the result of merging *primordial* BHs.

In practice, however, large distances are also the most uncertain ones, since the GW signal is typically much fainter than closer events. Hence the question we aim to answer is

better phrased as: what is the lowest value of f_{PBH} that would still produce a sufficient number of large-distance events so that they would be statistically distinguishable from $f_{\text{PBH}} = 0$ when observing with the ET? We hence *cut* the data set into two subsets containing the small- and large-distance events, *count* the number of events in the large-distance subset, and compare to the expected number for $f_{\text{PBH}} = 0$. This method only exploits a fraction of the available information in the data set, but it is nevertheless a natural starting point.

4.1 Description of the method

Let $D_* \equiv \bar{D}(z_*)$ be an arbitrary distance threshold (z_* is the corresponding background redshift threshold). Given a data set \mathcal{D} , we divide it into two subsets: a small-distance subset \mathcal{D}_{\leq} on the one hand, made of the events with distances beneath the threshold ($D \leq D_*$); and a large-distance subset $\mathcal{D}_{>}$ on the other hand, with events beyond the threshold ($D > D_*$). We call $N_{>}$ the cardinal of $\mathcal{D}_{>}$, i.e., the number of events above the threshold D_* ; the larger f_{PBH} , the larger the expected $N_{>}$.

To be more specific, the computation of $N_{>}$ does not simply consist in counting the number of events whose best-fit distance is above the threshold. Due the measurement errors, it may happen that an event truly lies beneath the distance threshold ($\bar{D}_i < D_*$) but is actually measured beyond it ($D_i > D_*$), or vice versa. In order to account for this, we calculate the probability that the true distance lies beyond the arbitrary distance threshold as

$$P_i = \frac{1}{\sqrt{2\pi}\sigma_i} \int_{D_*}^{\infty} d\bar{D}_i \exp\left[-\frac{(\bar{D}_i - D_i)^2}{2\sigma_i^2}\right], \quad (4.1)$$

using the definition of the uncertainty from Eq. (3.2). We can therefore obtain the value of $N_{>}$ by summing over all the events the probability of falling within the bin being considered,

$$N_{>} \equiv \sum_{i=1}^{N_{\text{det}}} P_i. \quad (4.2)$$

The uncertainty on $N_{>}$ is twofold. First, since the data set is discrete, it inevitably comes with a Poisson uncertainty with variance

$$\sigma_{\text{P}}^2 = \bar{N}_{>}, \quad (4.3)$$

where $\bar{N}_{>}$ is the theoretical expectation value of $N_{>}$.⁹ On the other hand, the number $N_{>}$ of large-distance events may be seen as the sum of N_{det} independent Bernoulli variables, the i^{th} one having a probability P_i of being equal to 1, and a probability $1 - P_i$ of being equal to zero. This observational contribution to the variance of $N_{>}$ is the sum of the individual Bernoulli variances,

$$\sigma_{\text{B}}^2 = \sum_{i=1}^{N_{\text{det}}} P_i(1 - P_i). \quad (4.4)$$

Summing the two sources of uncertainty in quadrature yields the total error on $N_{>}$,

$$\sigma_{>} = \sqrt{\sigma_{\text{P}}^2 + \sigma_{\text{B}}^2}. \quad (4.5)$$

⁹In practice, we estimate $\bar{N}_{>}$ by generating a number of mock data sets and taking the mean of the $N_{>}$ values obtained for each.

As an example, we consider the null case of a data set \mathcal{D}_0 containing no PBH ($f_{\text{PBH}} \rightarrow 0$). Figure 7 shows the expected value of $N_{>}$ as a function of the redshift threshold z_* , to which the $N_{>}$ of an observed data set should be compared. The error bars of that figure are computed following the approach described above, and they represent the uncertainty $\sigma_{>}$ on $N_{>}$.

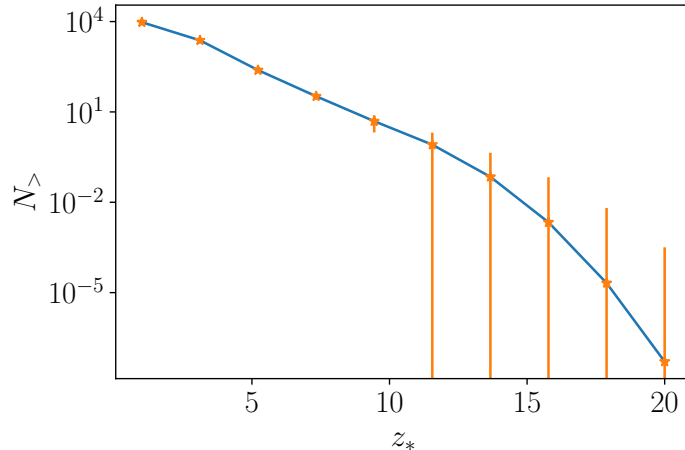


Figure 7. Value of $N_{>}$ as a function of the redshift threshold z_* , for a data set with a negligible number of PBH. The error bars represent the uncertainty $\sigma_{>}$ on $N_{>}$.

4.2 Optimal distance threshold and smallest detectable PBH fraction

Let us now determine what is the smallest detectable PBH fraction f_{PBH} that could be detected using the cut-and-count method. This will also be the occasion to determine the optimal choice for the arbitrary distance threshold, that is, the value of D_* leading to the maximal sensitivity in f_{PBH} .

For that purpose, we generate mock data sets $\mathcal{D}_{f_{\text{PBH}}}$ for 100 values of f_{PBH} ranging from 10^{-6} to 10^{-2} . For each of these data sets, we apply the cut-and-count method for 10 different values of the redshift threshold z_* , taken as equispaced in the range $[5, 50]$. This yields a number $N_{>}(\mathcal{D}_{f_{\text{PBH}}}, z_*)$ for each mock data set and each value of z_* . Those numbers come with their own uncertainty $\sigma_{>}(\mathcal{D}_{f_{\text{PBH}}}, z_*)$, computed as described in Sec. 4.1 for the null case \mathcal{D}_0 . We estimate the statistical significance of the detection of a non-zero f_{PBH} as

$$\mathcal{S}(\mathcal{D}_{f_{\text{PBH}}}, z_*) \equiv \frac{|N_{>}(\mathcal{D}_{f_{\text{PBH}}}, z_*) - N_{>}(\mathcal{D}_0, z_*)|}{\sqrt{\sigma_{>}^2(\mathcal{D}_{f_{\text{PBH}}}, z_*) + \sigma_{>}^2(\mathcal{D}_0, z_*)}}, \quad (4.6)$$

which represents the “number of sigmas” with which we could claim a detection.

In the following, we consider a “significant” detection to be one made at 3σ ; in other words, for each value of z_* , the smallest detectable f_{PBH} is determined by finding the data set $\mathcal{D}_{f_{\text{PBH}}}$ such that $\mathcal{S}(\mathcal{D}_{f_{\text{PBH}}}, z_*) > 3$. In the top left panel of Fig. 8, both curves (for $M_{\text{PBH}} = 10 M_{\odot}$ in red and $M_{\text{PBH}} = 30 M_{\odot}$ in yellow) show the evolution of the smallest detectable f_{PBH} with z_* . The optimal value for the redshift threshold, leading to the best sensitivity in f_{PBH} , is found to be $z_* \approx 10$, corresponding to a distance of $D_* \approx 106$ Gpc.

The existence of an optimal value for z_* arises from a trade-off between two effects: on the one hand, increasing z_* reduces the contamination due to ABH events in the high redshift subset; on the other hand, reducing z_* increases the number of events in that high redshift

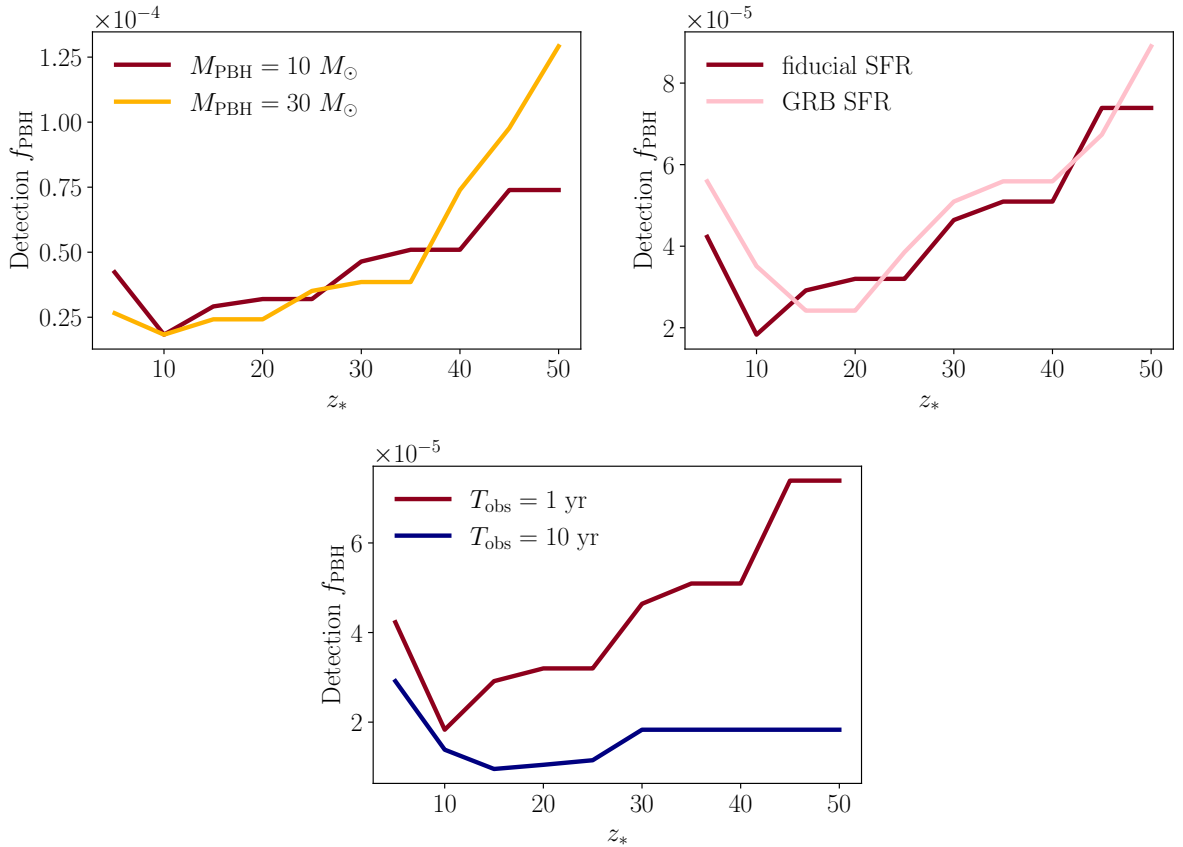


Figure 8. **Top left:** Lowest 3σ -detectable f_{PBH} as a function of the redshift z_* discriminating between the low redshift and high redshift subsets, for $M_{\text{PBH}} = 10 M_{\odot}$ (red lines) and $M_{\text{PBH}} = 30 M_{\odot}$ (yellow lines). **Top right:** Lowest 3σ -detectable f_{PBH} as a function of z_* , for $M_{\text{PBH}} = 10 M_{\odot}$ for the fiducial ABH merger rate (red line) and for the one obtained using the GRB SFR (pink line). **Bottom:** Lowest 3σ -detectable f_{PBH} as a function of the chosen z_* , for $M_{\text{PBH}} = 10 M_{\odot}$ for an observation time $T_{\text{obs}} = 1 \text{ yr}$ (red line) and for $T_{\text{obs}} = 10 \text{ yr}$ (blue line).

subset and hence the significance of the detection. Our results suggests that choosing $z_* < 10$ implies too much ABH contamination to easily assess whether or not PBHs are present in the data set, while $z_* > 10$ worsens the statistics and only provides a detection for large values of f_{PBH} . These conclusions are valid for both $M_{\text{PBH}} = 10 M_{\odot}$ and $M_{\text{PBH}} = 30 M_{\odot}$, although the exact value of the optimal z_* does change when the mass of the PBHs is changed.

We expect the SFR model to affect the results, as it determines how fast the probability of ABH mergers vanishes with redshift, see Fig. 3. We verify the impact of this uncertainty by repeating our analysis using the merger rate obtained assuming the GRB SFR model. In the top right panel of Fig. 8, we compare the GRB SFR results with the baseline case ($M_{\text{PBH}} = 10 M_{\odot}$, fiducial SFR and $T_{\text{obs}} = 1 \text{ yr}$). We can see how in the GRB case, the optimal threshold redshift z_* increases due to the ABH merger probability being non-vanishing up to higher redshifts than in the fiducial case. We also explore the impact of the observation time of the survey on these results; we compare our baseline results with an extended survey time for ET, setting $T_{\text{obs}} = 10 \text{ yr}$. As can be seen in the bottom panel of Fig. 8, increasing the time of the survey generally leads to a lower detection threshold for f_{PBH} , while also

decreasing the dependency of such a threshold on the choice of the binning strategy. With a greater survey time, more high redshift events are observed, meaning that a large detection significance is still achievable even as the high redshift bin is moved to larger values of z_* .

Overall, we find that a fraction $f_{\text{PBH}} \approx 2 \times 10^{-5}$ is the lowest value detectable by the ET with this method for $T_{\text{obs}} = 1$ year, with a weak dependency on the mass of the progenitor systems and on the SFR model chosen. In our baseline case, for the optimal $z_* \approx 10$, such a fraction of PBH corresponds to $N_{>} = 16 \pm 4.6$, while the no PBH case yields $N_{>} = 1 \pm 1.7$, a result that puts this f_{PBH} over the 3σ threshold we consider for detection.

Figure 9 shows more details for the results obtained in our baseline case, when taking $z_* = 10$. The left panel shows in red the results for $N_{>}$ for the sampled values of f_{PBH} , together with their error bars and the uncertainty on $N_{>}$ in the no PBH case (grey band). In this plot, we also show the trend of the detection significance \mathcal{S} with f_{PBH} (blue line). The right panel shows the counts in the two redshift bins for $f_{\text{PBH}} = 2 \times 10^{-5}$ in red, i.e. the first value for which the counts in the large-distance bin result in $\mathcal{S} \geq 3$. Together with the counts we also show the distance and its error for the events contained in this mock data set, as a function of the event redshifts (which would not be observed in reality).

The conclusions we have found for our baseline case using this cut-and-count approach are compatible with those that can be obtained using the approach of [44], where the discriminant for the presence of PBH in the observed data set is the observation of events at $z > 30$, where the contribution of ABHs is negligible. With our simulated data we indeed find that $f_{\text{PBH}} \approx 2 \times 10^{-5}$ is the lowest value in the baseline case for which we have at least one event above $z = 30$ with 99.7% confidence level.

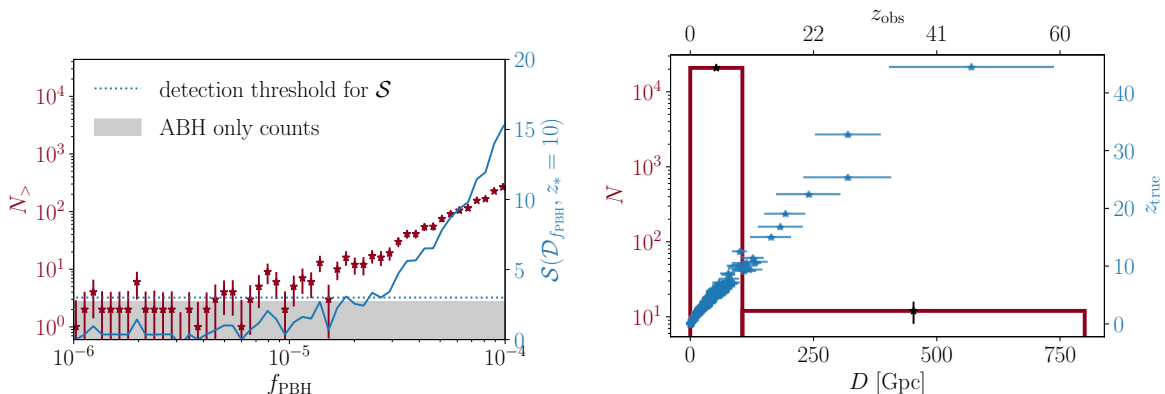


Figure 9. Left: Number $N_{>}$ of events in the high redshift subset, $z > z_* = 10$, for each value of f_{PBH} in the baseline case (see Table 1 and Table 2) (red points). The grey band shows the uncertainty for $N_{>}$ in the no-PBH case. The blue line shows instead the trend of the detection significance for a departure from the no-PBH case as a function of f_{PBH} . **Right:** Counts in the small- and large-distance bins are shown by the red bar plot for the baseline case, with $z_* = 10$ and $f_{\text{PBH}} = 2 \times 10^{-5}$, together with the errors we determine on the value of the counts (left y -axis). In blue, the distances and their errors for all the events contained in the data set are shown as a function of the true redshift z_{true} (right y -axis), while the top x -axis shows the redshift z_{obs} that could potentially be inferred from the distance measurements.

5 Quantifying the PBH fraction: a likelihood-based method

Having followed our simple cut-and-count method to quantify the lowest value of f_{PBH} that would be statistically distinguishable from $f_{\text{PBH}} = 0$ (i.e. the ability of the ET to *detect* PBHs), we now utilise a more powerful, likelihood-based method to assess the ET’s potential capacity to both detect PBHs *and* measure the f_{PBH} associated with a detection. The analysis that follows assumes fixed values of the cosmological parameters Ω_{m} and H_0 , considering them as priors set from independent experiments. We thus neglect the potential degeneracies that could exist between f_{PBH} and the cosmological model. This approach is justified by the fact that even future GW experiments are expected to yield much looser cosmological constraints than established probes such as *Planck*.

5.1 Description of the method

We describe an unbinned likelihood-based approach to comparing the data to a given model for the merger rate, which may include a contribution from both ABH and PBH mergers. To do this, we must compute the likelihood $\mathcal{L}(f_{\text{PBH}})$, which is the probability to observe the data set \mathcal{D} given the PBH fraction f_{PBH} .

We write the probability distribution for the *true* luminosity distances of ABH and PBH mergers as $p_{\text{ABH}}(\bar{D})$ and $p_{\text{PBH}}(\bar{D}|f_{\text{PBH}})$ respectively, where we have indicated explicitly that the distribution for PBH mergers depends on f_{PBH} . The probability that a BH merger has a true luminosity distance in the range $[\bar{D}, \bar{D} + d\bar{D}]$ is given by

$$p(\bar{D}|f_{\text{PBH}}) d\bar{D} = \frac{N_{\text{ABH}}}{\bar{N}_{\text{obs}}} p_{\text{ABH}}(\bar{D}) d\bar{D} + \frac{N_{\text{PBH}}}{\bar{N}_{\text{obs}}} p_{\text{PBH}}(\bar{D}|f_{\text{PBH}}) d\bar{D}, \quad (5.1)$$

where $\bar{N}_{\text{obs}}(f_{\text{PBH}}) = N_{\text{ABH}} + N_{\text{PBH}}(f_{\text{PBH}})$ is the total *expected* number GW events and N_{ABH} and N_{PBH} are the *expected* numbers of ABH and PBH mergers respectively. These are the numbers of events expected to be observed above the SNR threshold, taking into account the detection efficiency of the observatory, as detailed in Sec. 2.1.

We do not observe the true distance \bar{D} but instead an estimate of the luminosity distance D . The probability that we observe an event with estimated luminosity distance D can be written as

$$p(D|f_{\text{PBH}}) = \int p(D|\bar{D}) p(\bar{D}|f_{\text{PBH}}) d\bar{D}. \quad (5.2)$$

The distribution of “true” luminosity distance $p(\bar{D}|f_{\text{PBH}})$ is given by the theoretical expectation in Eq. (5.1). The term $p(D|\bar{D})$ can be obtained from the measurement uncertainty in Eq. (3.2) via Bayes’ theorem [95],

$$p(D|\bar{D}) = \frac{p(\bar{D}|D)}{\tilde{p}(\bar{D})} \tilde{p}(D). \quad (5.3)$$

Here, $\tilde{p}(D)$ and $\tilde{p}(\bar{D})$ are the *overall* probability distributions of \bar{D} and D , where by *overall* we mean that they are marginalised over all theory parameters (which in our case consists of only f_{PBH}). The term $\tilde{p}(D)$ enters as an overall normalisation which can be pulled out of the integral in Eq. (5.2) and does not depend on the theory parameters. This can therefore be safely neglected in the likelihood. The final task is then to compute $\tilde{p}(\bar{D})$, for which we need a prior on the PBH fraction $\text{Pr}(f_{\text{PBH}})$:

$$\tilde{p}(\bar{D}) = \int_0^1 p(\bar{D}|f_{\text{PBH}}) \text{Pr}(f_{\text{PBH}}) df_{\text{PBH}}. \quad (5.4)$$

We assume an uninformative log-flat prior, $\Pr(f_{\text{PBH}}) \propto 1/f_{\text{PBH}}$, with $f_{\text{PBH}} \in [10^{-9}, 10^{-3}]$.

For a sample of N_{obs} observed merger events, the likelihood can then be written as

$$\mathcal{L}(\mathcal{D}|f_{\text{PBH}}) = \frac{\bar{N}_{\text{obs}}(f_{\text{PBH}})^{N_{\text{obs}}} e^{-\bar{N}_{\text{obs}}(f_{\text{PBH}})}}{N_{\text{obs}}!} \times \prod_{i=1, N_{\text{obs}}} p(D_i|f_{\text{PBH}}). \quad (5.5)$$

The first term is the Poisson probability to observe N_{obs} merger events, given that we expect to observe $\bar{N}_{\text{obs}}(f_{\text{PBH}})$. The second term accounts for the contribution of each observed event to the likelihood, where D_i are the estimated luminosity distances. Combining the results above, the probability of observing a merger at an estimated distance D_i is given by

$$p(D_i|f_{\text{PBH}}) \propto \int \frac{p(\bar{D}_i|D_i)}{\tilde{p}(\bar{D}_i)} p(\bar{D}_i|f_{\text{PBH}}) d\bar{D}_i, \quad (5.6)$$

which depends on the uncertainty σ_i through $p(\bar{D}_i|D_i)$ in Eq. (3.2).

We then adopt a Bayesian approach in order to construct the posterior distribution function $p(f_{\text{PBH}}|\mathcal{D})$ and thus determine projected constraints on the PBH fraction,

$$p(f_{\text{PBH}}|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|f_{\text{PBH}})\Pr(f_{\text{PBH}}). \quad (5.7)$$

We sample $\log_{10} f_{\text{PBH}}$ from a uniform prior $\log_{10} f_{\text{PBH}} \in [-9, -3]$ using the public cosmological sampling code `Cobaya` [96], while fixing all other parameters to the values used to generate mock data. We analyse the results using `GetDist` [97] to obtain the bounds achievable on f_{PBH} using this approach.

5.2 Future constraints on f_{PBH}

In this section we discuss the bounds on f_{PBH} that ET can potentially obtain, as a function of the “true” value of this parameter, using the likelihood-based method. In order to estimate the constraints, we generate 10 mock data sets with fiducial values $f_{\text{PBH}}^{\text{fid}}$ logarithmically distributed in the range $[10^{-6}, 10^{-4}]$. For each fiducial value of $f_{\text{PBH}}^{\text{fid}}$, we compute the likelihood associated to each data set, and obtain the posterior distribution function applying Eq. (5.7). We define the mean of the distribution as the “measured value” $f_{\text{PBH}}^{\text{meas}}$. We show the results of this analysis in Fig. 10. In the top left panel, we visualise the 68% and 99.7% confidence level bounds (i.e. 1σ and 3σ) on the measured PBH fraction for each of the chosen $f_{\text{PBH}}^{\text{fid}}$ in our baseline settings, together with the mean value obtained for $f_{\text{PBH}}^{\text{meas}}$. We interpolate between the results obtained for our ten values of $f_{\text{PBH}}^{\text{fid}}$, in order to visualise the trend of these bounds.

We notice how the qualitative behaviour of these results is similar to that of the previous method; the analysis highlights how for low fiducial values ($f_{\text{PBH}} \lesssim 10^{-5}$) the ET is not able to detect the presence of PBHs, and only an upper bound can be placed on f_{PBH} . Choosing a threshold of 3σ for a non-vanishing value of $f_{\text{PBH}}^{\text{meas}}$ to be considered a detection, as we did in Sec. 4, we find that ET will be able to detect the presence of PBHs for $f_{\text{PBH}}^{\text{fid}} \approx 7 \times 10^{-6}$. Such a value is the result of interpolating between the 99.7% confidence level lower bounds obtained for the different data sets, and then finding for which value of $f_{\text{PBH}}^{\text{fid}}$ this function would exclude the no PBHs case (set to be at $f_{\text{PBH}} = 10^{-7}$). Such a value is roughly three times lower than what we found with the cut-and-count method, highlighting how using the full amount of information present in the data set helps to boost the survey sensitivity.

For higher values of $f_{\text{PBH}}^{\text{fid}}$, a detection is possible and the constraining power of ET increases with higher fiducial values. In the right panel of Fig. 10, we show the posterior

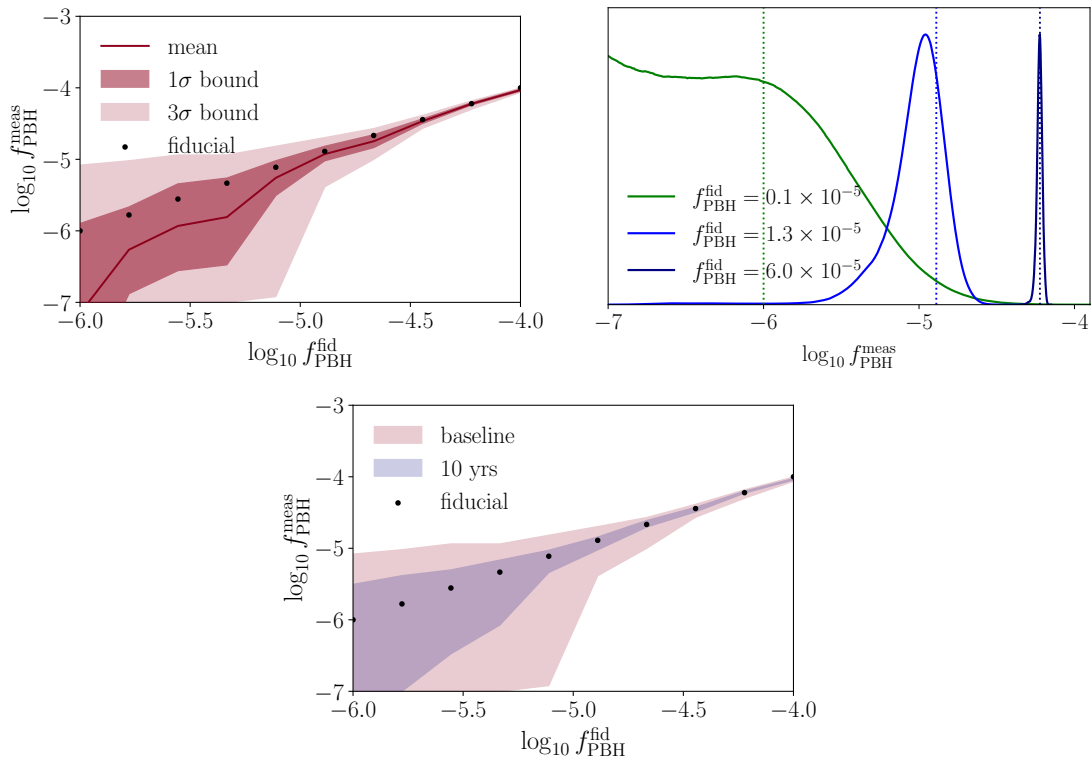


Figure 10. **Left:** recovered mean value $f_{\text{PBH}}^{\text{meas}}$ (solid line) and 68% and 99.7% confidence level limits (red bands) as a function of the fiducial $f_{\text{PBH}}^{\text{fid}}$ (black dots). **Right:** posterior distributions obtained on $f_{\text{PBH}}^{\text{meas}}$ for low (green), intermediate (blue) and high (navy) values of $f_{\text{PBH}}^{\text{fid}}$, with the vertical dotted line showing the value of $f_{\text{PBH}}^{\text{fid}}$ corresponding to each color. **Bottom:** Comparison between the 99.7% (3 σ) confidence regions for the baseline case (outer band) and for an observation time of $T_{\text{obs}} = 10$ yrs (inner band).

distribution of $f_{\text{PBH}}^{\text{meas}}$ for three different $f_{\text{PBH}}^{\text{fid}}$; here we highlight how for low $f_{\text{PBH}}^{\text{fid}}$, the posterior distribution is extremely flat at low f_{PBH} , as the ET cannot distinguish between fractions of PBHs that produce a very low number of events. Moving towards higher $f_{\text{PBH}}^{\text{fid}}$, the posterior becomes increasingly peaked, showing how these PBH abundances could be measured with very high precision.

These results highlight how, should the Universe contain a high enough number of PBHs, the method proposed in this work will be able to constrain the value of f_{PBH} . For example, considering $f_{\text{PBH}}^{\text{fid}} = 1.3 \times 10^{-5}$, i.e. the first generated data set above the detection threshold of $f_{\text{PBH}} \approx 7 \times 10^{-6}$ found above, we find that one can obtain a measurement of $f_{\text{PBH}}^{\text{meas}}$ with a precision of 35% at 68% confidence level.

Finally, we compare the baseline results with what can be achieved increasing the observation time to $T_{\text{obs}} = 10$ yrs. This comparison is depicted in the bottom panel of Fig. 10, where the baseline bound is shown in red, while the result for the extended observation time is shown in navy. Our results confirm that extending the survey time allows the detection of lower fractions of PBH; specifically, the smallest detectable $f_{\text{PBH}}^{\text{fid}}$ in the ten year case is about five times smaller than in the one year case. For a given detectable value of f_{PBH} , the uncertainty on its measurement is also reduced; for instance, with $f_{\text{PBH}}^{\text{fid}} = 1.3 \times 10^{-5}$ we find that $f_{\text{PBH}}^{\text{meas}}$ can be obtained with a precision of 11% with 68% confidence in the ten year case.

6 Discussion

The rationale of the two statistical frameworks presented here is to present complementary approaches, and emphasise different aspects of the analysis of GWs from BH merger events as a tool to detect and measure a PBH population. In particular, the cut-and-count method is especially designed to intuitively capture and visualise the essential aspect of the problem, namely the potential discovery of an anomalous excess of high redshift events beyond some reference distance. On the other hand, the likelihood-based approach exploits the capability of a Bayesian framework while using all the available information encoded in the redshift distribution of the events. In both cases, the characterisation of the instrument response plays a crucial role, and the main results we obtained revealed the capacity of the ET to disentangle astrophysical and primordial BH merger events on solid statistical grounds.

However, when searching for signatures of new physics against a highly uncertain astrophysical background, a key question is whether such signatures remain undetected simply due to the limitations of instrumental sensitivity or instead due to poor understanding of either the background or the signal itself, and the *systematic* uncertainties associated with each. In the search for PBHs we are considering here, the modelling of both the background and the signal poses a clear challenge.

Beginning with the signal, our model relies on a set of simplifying assumptions which includes: *(i)* a statistically *uniform* initial spatial distribution of PBHs (i.e. no initial clustering); and *(ii)* a *monochromatic* mass function (i.e. all the PBHs have the same mass). Regarding *(i)*, as pointed out in [43], initial clustering may significantly change the picture: the authors of that study aimed to find the minimum PBH abundance testable by future detectors and concluded that abundances as small as $f_{\text{PBH}} \sim 10^{-10}$ can be probed, if PBHs are highly clustered at formation; in the case of no initial clustering, they estimated that values of $f_{\text{PBH}} \sim 10^{-5}$ can provide at least one event per year at the ET. Our results based on the mock data generation and careful statistical analysis are compatible with the latter estimate. We leave the assessment of the role of the initial clustering, as well as a more refined analysis of the late-time clustering, to future work.

Regarding *(ii)*, realistic PBH production scenarios generally predict an extended mass function rather than a monochromatic one as assumed here. Assessing the role of the mass function requires us to understand the role of the PBH mass itself in the merger rate. On the one hand, the larger the chirp mass of a PBH binary, the higher the amplitude of the corresponding GW signal, which may result in a higher SNR, and hence an increased detectability of the PBH population. However, changing the chirp mass also affects the frequency range of the GW emitted by the merger. Our choice for the PBH mass, $M_{\text{PBH}} = \mathcal{O}(10 M_{\odot})$, roughly coincides with an optimal detectability of high-redshift events, given the ET noise curve.¹⁰ On the other hand, for a fixed f_{PBH} , the larger the PBH mass, the fewer PBHs, and hence the rarer the merger events. Therefore, if we consider a broad mass function, we expect that a high-mass tail would be associated with rare but (potentially) high-SNR events, while a low-mass tail would correspond to abundant but low-SNR events. Overall, for a relatively narrow mass function we do not expect our main conclusions to dramatically change, but we leave the study of more complex mass functions for future work.

Related to the above point, we chose to keep the masses of the ABH and PBH populations fixed in the statistical analysis. We have not used any information about the inferred masses of the mergers in this work, instead focusing on how the distance information can disentangle

¹⁰Note however that the optimal PBH mass depends on the precise value of the event's redshift.

the two populations. Of course, including mass information is likely to enhance our ability to detect a PBH population, since detecting a deviation from the conventional ABH mass distribution model would be a hint at a possible population of PBHs, but that would require further assumptions about the ABH mass distribution.

The possibility to detect a subdominant population of PBHs by analysing the mass function of binary BH mergers has progressively gained momentum with the detection of many BH merger events with masses around $30 M_{\odot}$ [27, 28]. For example, the authors of [42] claim that there is decisive evidence in favour of a two-population model from these data, with a second, subdominant, population peaked around $30 M_{\odot}$. This population displays a significant high-mass tail, hence its presence is further supported by the detection of a BH with mass in the *pair-instability gap* region [98] (several astrophysical models of BH formation in the pair-instability mass gap also exist, see e.g. [99]).

Motivated by these considerations, we emphasise that an exploration of the impact of broad mass distributions for the PBH population (possibly with multiple peaks on different mass scales [100]) on the PBH detection prospects with the ET is an important future goal in this research line. This would require a more careful treatment of late-time clustering than that presented in this work, given the complex phenomenology that may arise (for instance, mass segregation in clusters [101]).

We have assumed throughout this work that PBH binaries formed in the early Universe dominate the merger rate. Dense environments, such as PBH clusters, can potentially lead to the disruption of early Universe binaries. We have considered the effect of late-time clustering (see Appendix A), which seems to be negligible for the small PBH abundances considered in this work; the impact of initial PBH clustering is left for future work. PBH binaries could also be disrupted by interactions with other objects; while it is unlikely that they will be disrupted by stellar encounters in the disks of typical galaxies [39], it seems feasible that they may be disrupted in sufficiently dense environments, such as close to the centre of the host galaxy. This is unlikely to affect the high redshift signals we study here, with the first haloes having little effect on early-formed binaries [22]. On the other hand, dense environments can lead to the formation of a population of late-time PBH binaries. This is typically expected to be subdominant [22, 60], in particular for the small values of f_{PBH} considered in this work. These late-time binaries may be more difficult to distinguish from ABH ones, as they are expected to form only at low redshift, in which case alternative approaches are necessary, such as a measurements of the clustering bias [102]. However, as long as the population of early-time binaries remains relatively unperturbed, high redshift observations would be able to probe it, and hence detect the presence of PBHs, even if this population were to be subdominant with respect to the late-time one. Regarding the measurability discussion, a sizable population of late-time binaries merging at relatively high redshifts could lead us to overestimate the abundance of PBHs. In neglecting this, we rely on the expectation that the merger rate due to such binaries is only relevant at low redshifts and for large PBH abundances.

There are also a number of ways in which the modelling of the astrophysical background could alter the results presented here. In particular, we emphasise that our model for the ABH evolution, characterised by a peak at $z \sim 2$ and a monotonic decrease at higher redshifts, is designed to capture the rate associated to binary systems made of second- and third-generation stars (usually called Population I and II). These systems may form via different channels, such as binary stellar evolution in galactic fields, or dynamical formation through multi-body interactions in star clusters.

However, a significant contribution to the binary BH merger rate, including a peak at high

redshift, $z \sim \mathcal{O}(10)$, can be expected from the first-generation (Population III) stars, which formed out of the pristine gas left over after cosmological nucleosynthesis, and generated the first heavy elements in the Universe [87, 103]. Taking into account this additional component would certainly require, again, to address the mass dependence in more detail. In fact, this population would be modelled as a high (or intermediate) mass model of star formation at high redshift, to be added to the background model [104]. An even more detailed treatment would imply a marginalisation over the parameters that describe the metallicity dependence of the SFR, treated as nuisance parameters. We leave this more detailed treatment to future studies, and emphasise that the numerical and statistical tools presented here are the ideal framework to consistently address uncertainties in both the signal and background models.

As we were finalising this work, Ng et al. [105] proposed an independent analysis on the same topic. Reference [105] explores how a network consisting of two third generation detectors (ET and CE) can be used to distinguish between ABH and PBH populations using high redshift mergers. Their analysis differs from ours in that they consider exclusively population III ABHs and restrict their analysis to events above $z > 8$, as well as working directly with measurement errors on the event redshift, rather than on the luminosity distance, as we do here. Their results are broadly consistent with ours, highlighting that a PBH population with $f_{\text{PBH}} \sim \mathcal{O}(10^{-5})$ should be within reach of third generation detectors. We find a one year sensitivity estimate for f_{PBH} roughly a factor of three stronger, which may be partly due to differences in the distance uncertainties and population models assumed. This emphasises the importance of careful modelling of these details in quantifying the PBH detection and measurement potential of future GW detectors.

7 Conclusions

In this article we have assessed the capability of the forthcoming Einstein Telescope (ET) gravitational wave (GW) observatory to detect and measure a subdominant population of primordial black hole (PBH) mergers using a novel statistical framework.


We have described a procedure that computes the redshift evolution of both the expected background associated with the astrophysical black hole (ABH) merger events, and the signal under consideration associated with the PBH population (assumed to be characterised by masses of the same order of magnitude). In our modelling, we paid particular attention to the impact of late-time clustering of PBHs. We set the relative normalisation of the two contributions by comparison with the latest (low redshift) data released by the LIGO, Virgo and KAGRA collaborations. This procedure naturally provided an updated upper limit on the fraction of dark matter in the form of PBHs (quantified using the parameter f_{PBH}), based on the third Gravitational Wave Transient Catalog (GWTC-3) – see Fig. 5.

The key feature of our merger rate models is the different behaviour at high redshift of signal (i.e. PBH mergers) and background (i.e. ABH mergers), with the former monotonically increasing with increasing distance, and the latter steadily decreasing. Motivated by this qualitative aspect, and taking into account the expected high sensitivity of ET, we presented two methods to assess the potential of discovering a distant PBH population, which both accurately take into account the experimental sensitivity. Both methods are based on the generation and analysis of a set of mock data catalogues, associated to the null hypothesis (no PBHs) and to different values of the fraction of DM in the form of PBHs.

The first method (dubbed “cut-and-count”) relies on a two-redshift-bin data analysis strategy that highlights the PBH contribution as a significant excess in the high redshift bin.

We demonstrated that a PBH fraction as low as $f_{\text{PBH}} \approx 2 \times 10^{-5}$ can be detected with a 3σ significance, given a proper choice of the cut position – see Fig. 8.

The second method takes a Bayesian approach to the problem, and fully exploits the information on the redshift dependence of both the signal and background. Within this framework, we provided an assessment of the posterior probability density function for different fiducial values of f_{PBH} . We found that fractions of DM in the form of PBHs as low as $\mathcal{O}(10^{-5})$ can be measured by ET with one year of data taking, with a precision of $\sim 35\%$ at 68% confidence level – see Fig. 10. These results demonstrate a well-defined avenue towards a discovery or, in the absence of a discovery, the prospect of setting a significantly improved upper limit on the existence of PBHs in this mass window.

Furthermore, in the process of this work, we have developed the `darksirens`  code, which allows for the fast generation of realistic mock GW catalogues sourced by ABH and PBH merger events. We have made the code publicly available for community use, with extensive scope for development and application to other investigations.

The final message of this work is clear: the possibility of detecting and measuring a distant population of PBHs is well within the capabilities of future GW observatories such as the ET. The exciting prospect that these exotic objects could constitute a fraction of the dark matter, a substance that despite its fundamental importance in the Universe we still know so little about, makes the pursuit of this goal nothing short of vital. In the rigorous analysis we have here presented, based on the twin pillars of robust statistics and a thorough treatment of the modelling of both signal and noise, we have shown that the ET will be able to measure a fraction of PBHs as low as $\mathcal{O}(10^{-5})$, through observations of their luminosity distances alone. If a high redshift population of PBHs exists in our Universe, the detection of these distant dark sirens is therefore something we no longer need to hope for, but can begin to expect.

CRedIT statement

Matteo Martinelli: Conceptualization; Methodology; Software; Writing – original draft; Writing – review & editing; Supervision; Funding acquisition. **Francesca Scarcella:** Software; Writing – original draft; Writing – review & editing. **Natalie B. Hogg:** Software; Data Curation (code documentation); Writing – original draft; Writing – review & editing. **Bradley J. Kavanagh:** Methodology; Software (likelihoods); Writing – original draft; Writing – review & editing; Project administration. **Daniele Gaggero:** Conceptualization; Formal analysis (astrophysical background, PBH upper limit); Software (astrophysical background, PBH upper limit); Writing – original draft, review & editing; Supervision; Funding acquisition. **Pierre Fleury:** Conceptualization; Formal analysis (lensing); Software (lensing); Writing - review & editing; Supervision; Funding acquisition.

Acknowledgments

We are very grateful to Evan Hall for making the `gw-horizon` code public and open. We thank Tjonnie Li, Michele Maggiore and Bangalore Sathyaprakash for useful discussions, and Cristina Fernández Suárez for contributions in the early stages of this work.

MM acknowledges funding by the Agenzia Spaziale Italiana (ASI) under agreement n. 2018-23-HH.0. FS was supported by the Spanish Agencia Estatal de Investigación through the grants IFT Centro de Excelencia Severo Ochoa CEX2020-001007-S and PGC2018-095161-B-I00, and, during the early stages of this work, through the grants Severo Ochoa SEV-2016-0597

and Red Consolider MultiDark FPA2017-90566-REDC. FS has received financial support through la Caixa Banking Foundation (grant n. LCF/BQ/LI18/11630014) during the early stages of the project. NBH is supported by a postdoctoral position previously funded through two “la Caixa” Foundation fellowships (ID00010434), with codes LCF/BQ/PI19/11690015 and LCF/BQ/PI19/11690018 respectively, and currently funded by the French Commissariat à l’énergie atomique et aux énergies alternatives (CEA). BJK thanks the Spanish Agencia Estatal de Investigación (AEI, MICIU) for the support to the Unidad de Excelencia María de Maeztu Instituto de Física de Cantabria, ref. MDM-2017-0765. DG has received financial support through the Postdoctoral Junior Leader Fellowship Programme from la Caixa Banking Foundation (grant n. LCF/BQ/LI18/11630014) during the early stage of the project. DG was also supported by the Spanish Agencia Estatal de Investigación through the grants PGC2018-095161-B-I00, IFT Centro de Excelencia Severo Ochoa SEV-2016-0597, and Red Consolider MultiDark FPA2017-90566-REDC during the early stages of the project. DG acknowledges funding from the “Department of Excellence” grant awarded by the Italian Ministry of Education, University and Research (MIUR) in October-December 2021. DG also acknowledges support from the INFN grant “LINDARK,” and the project “Theoretical Astroparticle Physics (TAsP)” funded by the INFN in October-December 2021. DG acknowledges support from Generalitat Valenciana through the plan GenT program (CIDE-GENT/2021/017) starting from 01/01/2022. In the early stages of this work, MM and PF received the support of a fellowship from “la Caixa” Foundation (ID 100010434). The fellowship codes are LCF/BQ/PI19/11690015 and LCF/BQ/PI19/11690018 for MM and PF respectively.

A Merger rate suppression due to PBH clustering

In this appendix, we discuss the suppression of the PBH merger rate due to PBH clustering. We follow the numerical implementation given in Ref. [72] but here we provide additional details which are useful for understanding why the suppression of the merger rate is negligible for the ranges of f_{PBH} allowed by present bounds.

A.1 PBH cluster formation

Even if PBHs are uniformly distributed in the early Universe (i.e. according to a Poisson distribution), their discrete nature means that the mean PBH density in a given region can be subject to large fluctuations. Regions with overdensities of PBHs begin to collapse early, forming PBH *clusters* [12, 48, 49]. Here, we briefly review this mechanism, before commenting on the relevance for the survival of PBH binaries.

At matter–radiation equality, $t_i = t_{\text{eq}}$, we consider a spherical region of radius r_i , containing an overdensity $\delta_i = (\rho_i - \rho_{\text{eq}})/\rho_{\text{eq}}$. This overdense region will evolve as an isolated matter-dominated Universe with effective curvature [106–108],

$$\mathcal{K} = \frac{8\pi G\rho_{\text{eq}}\delta_i r_i^2}{3} = \frac{2G\bar{M}_i\delta_i}{r_i}, \quad (\text{A.1})$$

where $\bar{M}_i = (4\pi/3)G\rho_{\text{eq}}r_i^3$ would be the initial mass of the region in the unperturbed

background. The evolution of the radius of this region $r(t)$ can be solved parametrically as

$$r(\theta) = \frac{GM_i}{\mathcal{K}}(1 - \cos \theta), \quad (\text{A.2})$$

$$t(\theta) = \frac{GM_i}{\mathcal{K}^{3/2}}(\theta - \sin \theta). \quad (\text{A.3})$$

The region collapses then at

$$t_{\text{coll}} = t(\theta = 2\pi) = \frac{2\pi GM_i}{\mathcal{K}^{3/2}}, \quad (\text{A.4})$$

or, using the fact that $z \propto t^{-2/3}$ in a matter-dominated Universe:

$$z_{\text{coll}} = z_{\text{eq}} \left(\frac{t_{\text{eq}}}{t_{\text{coll}}} \right)^{2/3} = z_{\text{eq}} \mathcal{K} \left(\frac{t_{\text{eq}}}{2\pi GM_i} \right)^{2/3}. \quad (\text{A.5})$$

We can relate t_{eq} and r_i by assuming that the size of the region corresponds to its comoving size at matter–radiation equality. With this, we find the redshift of collapse as

$$z_{\text{coll}} = \frac{2}{(18\pi^2)^2} z_{\text{eq}} \frac{\delta_i}{(1 + \delta_i)^{2/3}} \approx 0.36 z_{\text{eq}} \delta_i, \quad (\text{A.6})$$

for $\delta_i \ll 1$. The redshift at which the expansion of the region turns around is a little larger, $z_{\text{ta}} \approx 0.56 z_{\text{eq}} \delta_i$.

Consider now a region of the Universe which we expect to contain, on average, N PBHs with total PBH mass $M = NM_{\text{PBH}}$. If a cluster forms from this collection of PBHs, it is likely to form from a typical overdensity in PBHs of $\delta_{\text{PBH}} = \sqrt{N}/N$, due to Poisson fluctuations. If PBHs make up only a fraction f_{PBH} of the total DM density, then the total overdensity will be $\delta_i = f_{\text{PBH}}/\sqrt{N}$ (assuming that the DM fluctuations are subdominant). Dropping some order one factors, we therefore expect that the PBH cluster to form at a redshift

$$z_{\text{coll}} \approx z_{\text{eq}} f_{\text{PBH}} / \sqrt{N}. \quad (\text{A.7})$$

While this expression is a good estimate of collapse time for large values of N , it fails for smaller values. In our numerical analysis, we rely on the more complete approach discussed in [12]. However, in the following sections we will make use of Eq. (A.7) to estimate the typical size of clusters which are disrupted at late times.

A.2 Relaxation time

We will assume that a PBH binary is completely disrupted if the PBH cluster it resides in undergoes core collapse due to gravo-thermal instability and that all other PBH binaries are unperturbed. We take the characteristic core-collapse timescale to be $t_{\text{cc}} \sim 18 t_{\text{r}}$ [109], where t_{r} is the relaxation time.

The relaxation time t_{r} for a PBH cluster containing N PBHs. For a system with density ρ and component masses m , this relaxation time can be estimated as [109]

$$t_{\text{r}} = 0.065 \frac{v^3}{G^2 m \rho \ln \Lambda}, \quad (\text{A.8})$$

where $\ln \Lambda \approx \ln(N/f_{\text{PBH}})$ is the Coulomb logarithm associated with interactions between component masses. Assuming that the fraction of PBHs in clusters matches the mass fraction in the Universe, then the total mass of each cluster is $M = mN/f_{\text{PBH}}$.

The mean density of a cluster will be

$$\bar{\rho} \equiv \frac{3M}{4\pi R^3} \approx 18\pi^2 \rho_c a^{-3} = 18\pi^2 \rho_c a_{\text{eq}}^{-3} \frac{f_{\text{PBH}}^3}{N^{3/2}}, \quad (\text{A.9})$$

where the second equality follows from the theory of spherical collapse [110] and from setting $a = 1/z_{\text{coll}}$ (the scale factor at which the cluster forms). Here, ρ_c is the critical density.

Next, we set the PBH velocity v in the cluster equal to some typical velocity dispersion $v \approx \sigma_v \approx \sqrt{GM/R}$. Using the definition of $\bar{\rho}$ in Eq. (A.9), we then have

$$\sigma_v^3 = \left(\frac{GM}{R}\right)^{3/2} = \sqrt{\frac{4\pi}{3}} G^{3/2} M \sqrt{\bar{\rho}}. \quad (\text{A.10})$$

Assuming that there is no substantial growth or evaporation of the cluster after formation, the relaxation timescale can then be written as

$$t_r = 0.065 \frac{\sigma_v^3}{G^2 m \bar{\rho} \ln \Lambda} = 0.065 \sqrt{\frac{4\pi}{3}} \frac{M}{\sqrt{G \bar{\rho}} m \ln \Lambda} \quad (\text{A.11})$$

$$= 0.065 \sqrt{\frac{2}{27\pi}} \frac{1}{\sqrt{G \rho_{\text{m,eq}}}} \frac{N^{7/4}}{f_{\text{PBH}}^{5/2} \ln \Lambda} = 2.1 \text{ kyr} \frac{N^{7/4}}{f_{\text{PBH}}^{5/2} \ln \Lambda}. \quad (\text{A.12})$$

Here, we have defined $\rho_{\text{m,eq}} = \rho_c a_{\text{eq}}^{-3}$ (the matter density at matter–radiation equality). We have also taken the numerical values $\rho_c = 2.78 \times 10^{11} h^2 M_{\odot} \text{ Mpc}^{-3}$ [111] and $H_0 = h \times 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, with $h = 0.673$ [54]. This value for the relaxation time t_r matches the result given in Eq. (7) of [72].

A.3 Cluster collapse

In order to undergo core collapse before redshift z , a PBH cluster must be formed with fewer than a critical number of members $N_c(z)$. This critical number $N_c(z)$ is obtained by equating the core-collapse timescale to the time between formation and collapse,

$$18 t_r(N_c) = t(z) - t(z_c), \quad (\text{A.13})$$

where z_c is the redshift at which clusters of size N_c are formed. Here, $t_r(N_c)$ is the relaxation time for clusters of size N_c , defined in Eq. (A.12). Recall also from the previous sections that

$$z_c \approx z_{\text{eq}} \frac{f_{\text{PBH}}}{\sqrt{N_c}}. \quad (\text{A.14})$$

With these definitions, we can calculate $N_c(z)$ from Eq. (A.13). This is illustrated in Fig. 11. There, we see that for $f_{\text{PBH}} = 1$, the critical size of clusters is $N_c \sim 5000$ at $z = 0$. This decreases with decreasing f_{PBH} . Below $f_{\text{PBH}} \sim 0.01$, core-collapse becomes more or less irrelevant, as the critical number of PBHs tends to one.

A.4 Merger rate suppression factor

Following Ref. [72], the suppression factor is expressed as the probability of a binary not belonging to a cluster that undergoes gravo-thermal collapse. Estimating this factor requires knowledge of the PBH halo mass function, which we describe using the analytical model discussed in [12]. We indicate the probability of a PBH belonging to a cluster of N elements at

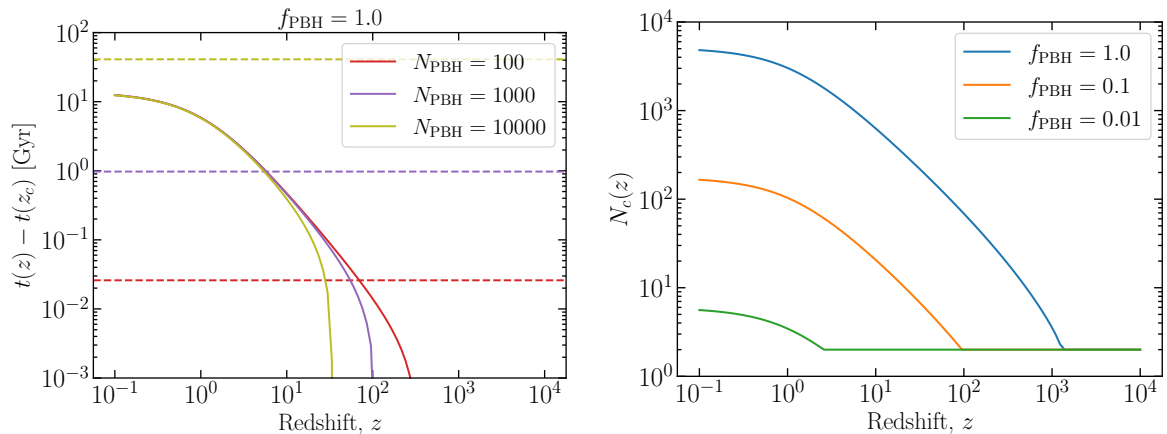


Figure 11. Left: Time-scales associated with PBH clusters. Solid lines show the time between the redshift of cluster formation z_c and a given redshift z , for clusters containing different numbers of PBHs N_{PBH} . Horizontal dashed lines show the core-collapse timescale for clusters with N_{PBH} members. Where the solid and dashed lines (of a given colour) cross then $N_{\text{PBH}} = N_c(z)$. **Right:** Critical number of PBHs per cluster N_c as a function of redshift z . Clusters with $N < N_c(z)$ will be disrupted due to core-collapse before redshift z .

given z and f_{PBH} with $P(N|z, f_{\text{PBH}})$. We thus have $P(N|z, f_{\text{PBH}}) \propto \exp[-N/N^*(z, f_{\text{PBH}})]$, where $N^*(z, f_{\text{PBH}})$ is the characteristic number of PBHs in clusters forming at redshift z .

The probability of a binary being disrupted by redshift z is given, according to [72], by the sum of two terms:

1. The probability of the binary belonging to a cluster which has reached instability before redshift z , i.e. a cluster with $N \leq N_c(z)$,

$$P_p^{(1)} = \sum_{N=3}^{N_c} P_{\text{bin}}(N|z_c), \quad (\text{A.15})$$

where the probabilities of binaries belonging to a cluster $P_{\text{bin}}(N)$ are taken to be $\propto P(N)$ but normalised to exclude $P(N=1)$:

$$\sum_{N=2}^{\infty} P_{\text{bin}}(N) = 1. \quad (\text{A.16})$$

2. The probability of a binary belonging to a sub-cluster of $N \leq N_c$ within a larger cluster ($N > N_c$),

$$P_p^{(2)} = \sum_{N > N_c} \left(\sum_{N'=3}^{N_c} P_{\text{sub}}(N'|z_c) \right) P_{\text{bin}}(N|z_c), \quad (\text{A.17})$$

where the probabilities of belonging to a sub-cluster $P_{\text{sub}}(N)$ are taken to be proportional to $P(N)$, but normalised summing up to the size of the cluster that contains them

$$\sum_{N'=2}^N P_{\text{sub}}(N') = 1. \quad (\text{A.18})$$

All the probabilities in the expressions above are evaluated at $z = z_c$, which is the typical redshift of formation of the clusters with N_c elements. Then z_c is obtained by requiring that $N^*(z_c) = N_c$. Finally, the probability of a binary not being perturbed is given by $P_{\text{np}} = 1 - P_{\text{p}}^{(1)} - P_{\text{p}}^{(2)}$. This suppression factor is shown in Fig. 12 for different values of z and f_{PBH} .

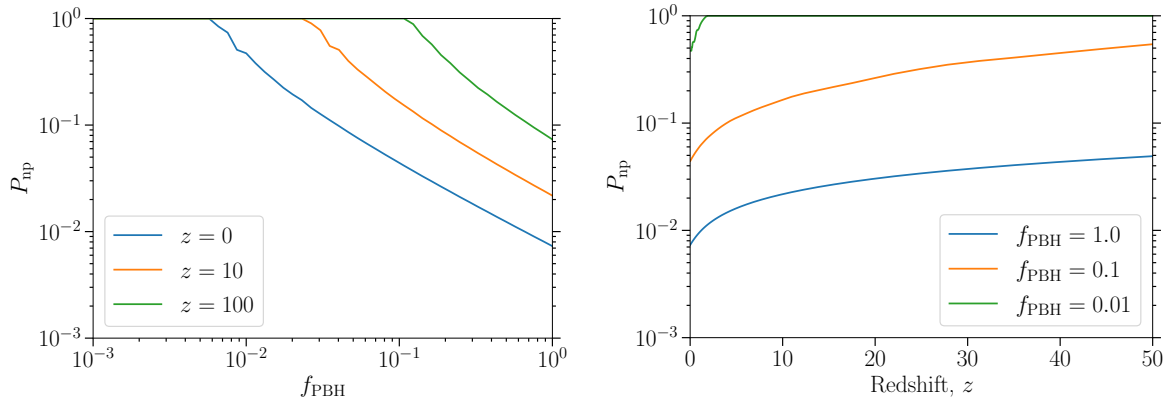


Figure 12. Merger rate suppression factor due to interactions in gravo-thermally unstable clusters, computed following [72]. **Left:** Probability of a binary *not* being disrupted as a function of f_{PBH} for different redshifts. Notice that, depending on the value of z , there is a threshold value of f_{PBH} below which no binaries are perturbed. **Right:** Same probability as a function of redshift, for different values of f_{PBH} .

B Obtaining the signal-to-noise ratio of an event

In this appendix, we provide a detailed description of how our mock catalogues are constructed. We start by introducing GW signals in Appendix B.1, move to a discussion of the luminosity distance from GWs in Appendix B.2 and finally the computation of the uncertainty on the luminosity distance in Appendix B.3.

B.1 Gravitational waves

The merging of massive compact objects results in the emission of GWs. These can be detected by the strain they produce in GW interferometers, housed in observatories like LIGO and Virgo. Let us note the different terms we use in the following discussion: the term GW *detector* refers to the equipment used to measure the strain itself; the term GW *interferometer* refers to the combination of two separate vacuum tubes or “arms” housing the laser and mirrors with a detector at the junction of the arms; and the term GW *observatory* refers to an entire facility, which may contain multiple detectors, or as in the case of LIGO, multiple interferometers at different physical locations (LIGO Livingston and LIGO Hanford).

The GW strain is given in the transverse traceless gauge as

$$h(t) = F_+(\theta, \phi, \psi) h_+(t) + F_\times(\theta, \phi, \psi) h_\times(t), \quad (\text{B.1})$$

where $h_{+,\times}$ are the plus and cross polarisations of the metric perturbation $h_{\alpha\beta}$, $F_{+,\times}$ are the corresponding antenna pattern functions which describe the angular dependence of the sensitivity of the interferometers. The angles (θ, ϕ) are the spherical polar coordinates of the

wave source in the celestial sphere, $\theta = \pi/2$ being the detector plane; ψ denotes the angle between the two-dimensional basis of the plane orthogonal to the line of sight with respect to which the plus and cross polarisations are defined, and the natural basis $(\mathbf{e}_\theta, \mathbf{e}_\phi)$ associated with the celestial coordinates (θ, ϕ) [112]. More simply put, ψ is the polarisation angle.

For the results presented in this work, we generate mock catalogues of GW events based on the configuration and sensitivity of the ET, a proposed third-generation GW observatory. Specifically, we consider the ET-D configuration, which involves three interferometers arranged in an equilateral triangle shape and two detectors at each vertex of the triangle, one sensitive to higher frequencies and one to lower, i.e. six detectors in total. The antenna patterns for a single interferometer with a 60 degree opening angle are given by

$$F_+ = \frac{\sqrt{3}}{2} \left[\frac{1}{2} (1 + \cos^2 \theta) \cos(2\phi) \cos(2\psi) - \cos \theta \sin(2\phi) \sin(2\psi) \right], \quad (\text{B.2})$$

$$F_\times = \frac{\sqrt{3}}{2} \left[\frac{1}{2} (1 + \cos^2 \theta) \cos(2\phi) \sin(2\psi) + \cos \theta \sin(2\phi) \cos(2\psi) \right]. \quad (\text{B.3})$$

The antenna patterns for the two other interferometers in the ET-D configuration are then given by $F_{+, \times}(\theta, \phi + 2\pi/3, \psi)$ and $F_{+, \times}(\theta, \phi + 4\pi/3, \psi)$.

We assume that the basic information available from the detection of GWs from merging compact objects by the ET will be the luminosity distance to the merger event, D , and the uncertainty on that luminosity distance, σ , which has contributions from the instrumental noise and from the uncertainty due to weak lensing of the GW.

B.2 Luminosity distance

To obtain the luminosity distance for a given simulated event, we firstly draw a redshift from a probability distribution defined based on the merger rate of the progenitor system being considered (binary PBHs or binary PBHs), which we previously described in Section 2.

However, unlike binary neutron star mergers which produce an electromagnetic counterpart, the redshift of an individual¹¹ binary BH merger is always unknown – hence the name *dark siren*. We therefore convert the simulated redshift to a luminosity distance in the standard Λ CDM cosmology using CAMB. All of the analysis presented in the main text is based solely on the luminosity distance information, rather than the redshift information, in order to accurately simulate the actual analysis that could be done with a catalogue which contains only binary BHs.

Finally, we rescale the luminosity distance to account for weak lensing of the signal. The full computation is shown in Appendix C.

B.3 Uncertainty on the lensed luminosity distance

The second step in the generation of the mock catalogue is then to compute the uncertainty on the lensed luminosity distance for a given event. As we mentioned, this has two components: the instrumental noise and the noise due to weak lensing of the GW. The instrumental noise depends on the signal-to-noise ratio for a given event.

¹¹There have been proposals to cross-correlate GW events with galaxy catalogues (see e.g. [113, 114]), allowing the redshift of the event to be estimated, but this introduces a good deal of uncertainty, especially with the current small number of events and relatively poor sky localisation.

The unlensed signal-to-noise ratio for a single detector is given by

$$\rho_{\text{det}} = \frac{\int_{-\infty}^{\infty} df h(f) K^*(f)}{\sqrt{\int_{-\infty}^{\infty} df \frac{1}{2} S_n(f) |K(f)|^2}}, \quad (\text{B.4})$$

where $h(f)$ is the waveform of the GW (the strain as a function of the frequency f), $K(f)$ is the filter used in the *matched filtering* process which maximises the signal-to-noise ratio and S_n is the noise power spectral density [94]. Asterisks denote the complex conjugate. The optimal unlensed signal-to-noise ratio is obtained using a Wiener filter, and is given by

$$\rho_{\text{opt}} = \left[4 \int_{f_{\text{lower}}}^{f_{\text{upper}}} df \frac{h(f)h^*(f)}{S_n(f)} \right]^{\frac{1}{2}}, \quad (\text{B.5})$$

where f_{upper} and f_{lower} are the cutoff frequencies for the strain, beyond which it is assumed to be zero. The waveform $h(f)$ is the Fourier transform of the strain $h(t)$ given by Eq. (B.1). We randomly draw the angles ϕ and ψ from a uniform distribution between 0 and 2π . However, the angles θ and ι (the inclination of the event i.e. the angle between the source plane and the detector plane) should have their cosine uniformly distributed, meaning that for these quantities we draw their values randomly from a uniform distribution between -1 and 1 and then take the arccosine of the result. The inclination of the event enters into the function used to weight the final signal-to-noise ratio sum, as we will see in a moment.

We use the publicly available package PyCBC¹² to generate a mock waveform $h(f)$ for a given event, using the IMRPhenomD waveform model, and by inputting the masses and simulated luminosity distance of the event. We keep the spins fixed to zero. We also generate the frequency range, and hence f_{upper} and f_{lower} , using PyCBC.

Assuming that the noise can be described by a Gaussian stochastic process, the power spectral density is given by

$$S_n(f) = 2 \int_{-\infty}^{\infty} d\tau R(\tau) e^{i2\pi f\tau}, \quad (\text{B.6})$$

where

$$R(\tau) = \langle n(t + \tau) n(t) \rangle \quad (\text{B.7})$$

is the autocorrelation of the noise $n(t)$ [94]. In practice, we use the publicly available power spectral density data for ET-D¹³.

With the waveform and power spectral density in hand, the signal-to-noise ratio for an event seen in a single interferometer can be computed. The total signal-to-noise ratio for an event i seen by the three-armed ET-D observatory is obtained by summing the squares of the individual interferometer signal-to-noise ratios,

$$\bar{\rho}_i = \sqrt{w \left(\rho_{i, \text{opt}, 1^{\text{st}} \text{ arm}}^2 + \rho_{i, \text{opt}, 2^{\text{nd}} \text{ arm}}^2 + \rho_{i, \text{opt}, 3^{\text{rd}} \text{ arm}}^2 \right)}, \quad (\text{B.8})$$

where the weighting factor w is given by

$$w = F_+^2(\theta, \phi, \psi)(1 + \cos^2 \iota)^2 + 4F_\times^2(\theta, \phi, \psi) \cos^2 \iota, \quad (\text{B.9})$$

¹²<https://pycbc.org/>.

¹³<http://www.et-gw.eu/index.php/etsensitivities>.

where F_+, F_\times are the antenna patterns and ι is the inclination. Lastly, following [94], we approximate the instrumental uncertainty on the lensed luminosity distance \tilde{D}_i for a given event i as

$$\sigma_i^{\text{inst}} \approx \frac{2\tilde{D}_i}{\rho_i}, \quad \rho_i = \sqrt{\mu_i} \bar{\rho}_i. \quad (\text{B.10})$$

Here, $\bar{\rho}_i$ is the unlensed signal-to-noise ratio for the event determined by the observatory in question and μ_i is the magnification of the signal-to-noise ratio due to weak lensing. The factor 2 in Eq. (B.10) accounts for the contribution of the uncertainty on the inclination ι to the instrumental noise [94]. Once the rescaling of the unlensed signal-to-noise ratio by the lensing magnification has been carried out, it is then trivial to compute the instrumental uncertainty for the event from Eq. (B.10).

However, we note that a precise modelling of the uncertainty on the distance measurement would require access to the full parameter estimation pipeline of the experiment. Furthermore, the uncertainty actually depends on the full network of observatories that may be operating in collaboration with the ET, see [35]. Extending the uncertainty computation in this way is beyond the scope of this work.

Furthermore, our analysis also rests on the assumption that the expression for the instrumental uncertainty given in Eq. (B.10) is correct. The signal-to-noise ratio in this expression depends on the noise being stationary and Gaussian. In a real detector this is not exactly the case. Depending on the specific template used in the matched filtering process, so-called glitches – artifacts with very high signal-to-noise ratios – can be seen in the detector [115]. The identification and removal of glitches and other noise artifacts through proper characterisation of the detectors is an important part of current gravitational wave analysis pipelines [116]. However, since we are working solely with mock data, we generate our catalogues using the assumption of stationary, Gaussian noise. This is effectively equivalent to using a real catalogue of events with glitches removed.

Finally, as mentioned in Sec. 3, after computing the *lensed* signal-to-noise ratio for all the events, we remove all the events with a lensed signal-to-noise ratio of less than eight from the catalogue. This serves to exclude events which may or may not be true detections of GWs. The signal-to-noise ratio threshold that we use follows that of the LIGO collaboration [93]. In Fig. 13 we show the computed signal-to-noise ratio as a function of redshift for ABH and PBH events, highlighting the significant amount of events which are discarded due to the cut in the signal-to-noise ratio.

C Lensing

This appendix provides details about our modelling of the weak lensing of GW signals. Appendix C.1 is a short theoretical reminder where we state our assumptions and define the relevant quantities to be used. Appendix C.2 indicates how the variance of the magnification is estimated. Appendix C.3 deals with our model for the magnification PDF in the mocks.

C.1 Magnification of gravitational waves

Consider a GW modelled as a small perturbation $h_{\mu\nu}$ of an arbitrary background geometry $g_{\mu\nu}$. If the wave is propagating in vacuum, in the sense that we can neglect its direct interaction with matter, then its equation of motion reads [117]

$$\nabla_\rho \nabla^\rho h_{\mu\nu} = 0, \quad (\text{C.1})$$

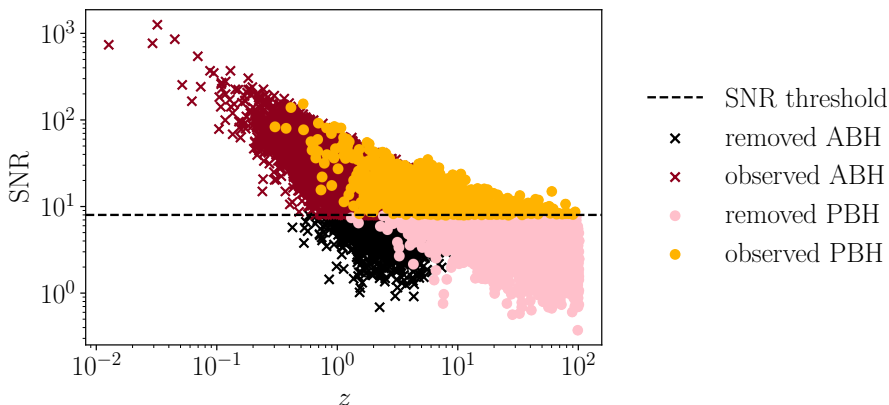


Figure 13. Trend of the signal-to-noise ratio with event redshift for ABHs (crosses) and PBHs (dots). Events that are above the signal-to-noise ratio threshold (black dashed line) are considered to be properly observed GW events (red for ABHs and yellow for PBHs), while the black and pink points are respectively the ABHs and PBHs with signal-to-noise ratios that make them too faint to be observed.

in the linear regime, harmonic transverse-traceless gauge, and assuming that the GW’s wavelength is much shorter than the typical curvature radii of the background spacetime geometry, whose covariant derivative is denoted with ∇_ρ in Eq. (C.1).

Introducing the wave ansatz $h_{\mu\nu} = H_{\mu\nu} \exp(iw)$, where $H_{\mu\nu}$ and w respectively denote the amplitude and the phase of the GW, Eq. (C.1) implies (i) that the wave follows null geodesics of the background spacetime; and (ii) that its amplitude satisfies

$$(1+z)^{-1}DH_{\mu\nu} = \text{const.}, \quad (\text{C.2})$$

where z is the source’s redshift and D is the electromagnetic luminosity distance – see, e.g., [118] for further details. It follows that $H_{\mu\nu} \propto D^{-1}$, that is, the energy of a GW “dilutes” just like the energy of an electromagnetic wave as it propagates. Since this is valid in any background spacetime, we conclude that GWs experience the same gravitational lensing effects as light.

In the following, we only consider *weak-lensing* effects, in the sense that we neglect the possibility that a GW source might be multiply imaged. We also assume that the luminosity distance can be computed from the propagation of an infinitesimal beam of null geodesics. In that framework, the distortions of the beam with respect to the homogeneous-isotropic FLRW case are customarily encoded in the so-called distortion matrix. If $\boldsymbol{\theta}$ is the observed incoming direction of GWs and $\boldsymbol{\beta}$ the direction in which they would be observed in FLRW, then the distortion matrix is defined as the Jacobian matrix of $\boldsymbol{\theta} \mapsto \boldsymbol{\beta}(\boldsymbol{\theta})$,

$$\mathcal{A} \equiv \frac{d\boldsymbol{\beta}}{d\boldsymbol{\theta}} = \begin{bmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 + \omega \\ -\gamma_2 - \omega & 1 - \kappa + \gamma_1 \end{bmatrix}, \quad (\text{C.3})$$

where κ is called the convergence of the beam, $\gamma = \gamma_1 + i\gamma_2$ its shear distortion, and $\omega = \mathcal{O}(\gamma^2)$ its rotation.

The correction to the observed luminosity distance, D , relative to the FLRW case, \bar{D} , is quantified by the *magnification* μ , which is related to the distortion matrix as follows,

$$\mu^{-1} \equiv \left(\frac{D}{\bar{D}}\right)^2 = \frac{d^2\boldsymbol{\beta}}{d^2\boldsymbol{\theta}} = \det \mathcal{A} = (1 - \kappa)^2 - |\gamma|^2 + \omega^2. \quad (\text{C.4})$$

In particular, at lowest order in κ, γ, ω , the magnification only depends on the convergence, $\mu \approx 1 + 2\kappa$. We shall use this approximation for the estimate of the variance of the magnification – see next subsection. It follows from Eq. (C.4) that the absolute uncertainty on the luminosity distance due to lensing may be estimated as

$$\sigma_{\text{lens}} = \frac{1}{2} \bar{D} \sigma_{\mu}, \quad (\text{C.5})$$

where σ_{μ} denotes the standard deviation of the magnification.

C.2 Variance of the magnification in weak lensing

We aim to model the PDF of the magnification of the observed GW signals in a realistic inhomogeneous universe. As a first step, we shall estimate the variance of that distribution. Previous analyses, both in the context of supernova cosmology [119] or GW cosmology [120–125], considered a dispersion of magnitude due to that grows linearly with redshift [126], $\Delta m = 0.055 z$, which corresponds to a magnification dispersion of $\sigma_{\mu} = 4 \ln(10) \Delta m / 5 = 0.10 z$. While this linear approximation may be valid at low redshift, we expect it to fail at the high redshifts ($z \sim 10$ to 100) considered here. The intuition is that at high redshift the Universe is increasingly homogeneous, thereby reducing the growth of the lensing dispersion.

In order to get a more accurate estimate of σ_{μ} across a wide range of redshifts, we shall use perturbation theory at second order. In that framework, we first notice that

$$\sigma_{\mu}^2 \equiv \langle \mu^2 \rangle - \langle \mu \rangle^2 = 4\sigma_{\kappa}^2 + \mathcal{O}(\kappa^4), \quad (\text{C.6})$$

where σ_{κ}^2 is the variance of the convergence, because $\langle \kappa \rangle = 0$ at linear order and $\omega \sim \kappa^2 \sim \gamma^2$. Hence, we may estimate σ_{μ} at second order from the linear-order results on σ_{κ} .

In the flat-sky and Limber approximations, the variance of the convergence for a source at redshift z is related to the convergence power spectrum $P_{\kappa}(\ell, z)$ as

$$\sigma_{\kappa}^2(z) = \int_0^{\infty} \frac{\ell d\ell}{2\pi} P_{\kappa}(\ell, z), \quad (\text{C.7})$$

which is itself related to the power spectrum P_W of the Weyl potential $W = (\Phi + \Psi)/2$, Φ and Ψ being the Bardeen potentials, via


$$P_{\kappa}(\ell, z) = \ell^2 (\ell + 1)^2 \int_0^r dr' \left(\frac{r - r'}{r} \right)^2 \left(\frac{1}{\ell + 1/2} \right)^4 P_W \left(\eta_0 - r', \frac{\ell + 1/2}{r'} \right), \quad (\text{C.8})$$

where η_0 denotes today's conformal time, and $r = r(z)$ is the comoving distance to the source.

We compute P_W and the associated quantities using **CAMB**. For a Λ CDM cosmology with $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\text{m}} = 0.315$, $\Omega_{\text{b}} = 0.05$, $A_{\text{s}} = 2 \times 10^{-9}$, $n_{\text{s}} = 0.965$, we find that the standard deviation of the magnification as a function of redshift is very well fit (with sub-percent accuracy) by

$$\sigma_{\kappa}(z) = a \arctan \left[(1 + b z^c)^d - 1 \right] \quad (\text{C.9})$$

with $a = 0.116$, $b = 1.26$, $c = 1.46$, $d = 0.268$.¹⁴ The result is depicted in Fig. 14, where we also indicate the linear ansatz of [126] for comparison. The latter is a reasonably good approximation of $\sigma_{\kappa}(z)$ up to $z \approx 3$, but it highly overestimates it at high z . At $z = 100$, we find $\sigma_{\kappa} \approx 16\%$, which is more than an order of magnitude below the prediction of the linear ansatz. This emphasises the importance of a careful modelling of lensing at high redshift.

¹⁴These numerical values depend on the cosmology; our code includes a notebook  capable of generating such a fitting function for other values of the cosmological parameters.

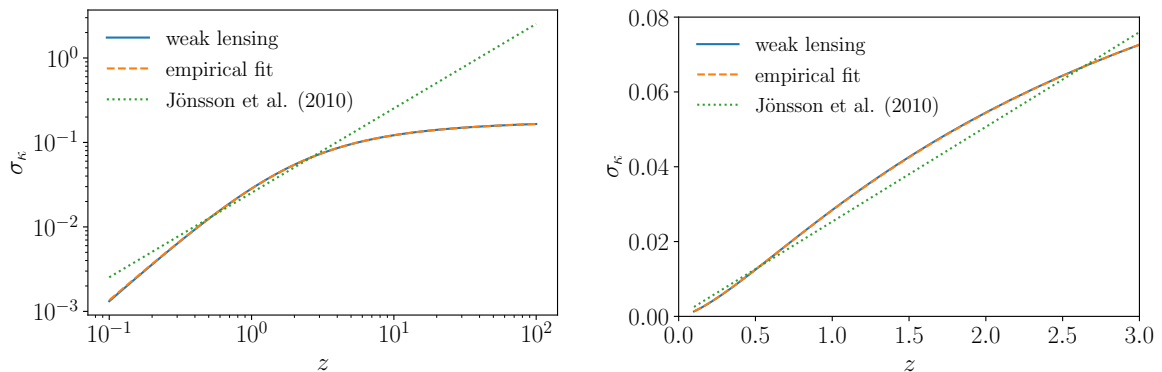


Figure 14. Standard deviation of the convergence σ_κ as a function of the redshift z of the source. Blue solid lines indicate the numerical results from CAMB; dashed orange lines indicate the empirical fitting function (C.9); green dotted lines show the linear prescription of [126]. The left panel shows the entire redshift range $z \in [0, 100]$, while the right panel focuses on $z < 3$.

C.3 Lognormal probability distribution of the magnification

We now turn to the full PDF $p(\mu)$ of the magnification. This distribution must satisfy three physical requirements:

1. It must vanish for $\mu < \mu_{\min}$, where μ_{\min} corresponds to the (de)magnification of Zel'dovich's empty beam [127]. This is due to the fact that a bundle of null geodesics cannot be less focused than a beam propagating through pure vacuum. In that case, the angular diameter distance coincides with the affine parameter λ along the bundle; the minimum magnification thus reads

$$\mu_{\min}(z) = \left[\frac{\bar{D}(z)}{(1+z)^2 \lambda(z)} \right]^2 < 1, \quad \lambda(z) = \int_0^z \frac{d\zeta}{(1+\zeta)^2 H(\zeta)}. \quad (\text{C.10})$$

The evolution of μ_{\min} with redshift z is depicted in the left panel of Fig. 15.

2. The magnification averaged over sources must be unity [128]

$$\langle \mu \rangle = \int_{\mu_{\min}}^{\infty} d\mu \mu p(\mu) = 1. \quad (\text{C.11})$$

This magnification theorem assumes that sources are homogeneously distributed in space. Note that the underlying averaging procedure – over sources rather than over random directions in the sky – is essential to this result. See [129] and references therein for detailed discussions.

3. The variance of the magnification should coincide with the one evaluated in Appendix C.2,

$$\langle \mu^2 \rangle - 1 = \int_{\mu_{\min}}^{\infty} d\mu \mu^2 p(\mu) - 1 = 4\sigma_\kappa^2. \quad (\text{C.12})$$

There are, of course, many models that would satisfy the above three requirements. We adopt, for simplicity, the following shifted lognormal model,

$$p(\mu) = \frac{1}{\sqrt{2\pi}\sigma(\mu - \mu_{\min})} \exp \left\{ -\frac{[\ln(\mu - \mu_{\min}) - m]^2}{2\sigma^2} \right\}, \quad (\text{C.13})$$

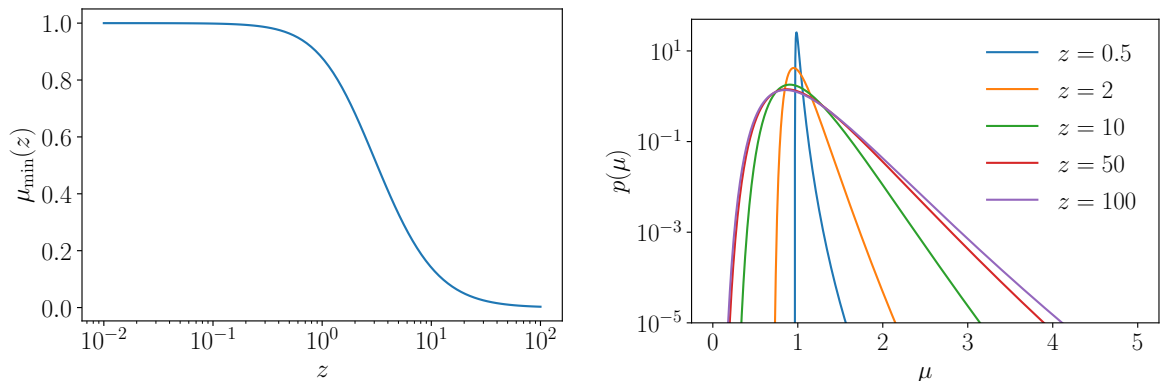


Figure 15. *Left:* minimum magnification corresponding to Zel’dovich’s empty-beam case, as defined in Eq. (C.10). *Right:* PDF of the magnification determined by the shifted lognormal model (C.13) for different values of the source redshift z .

where μ_{\min} is given by Eq. (C.10), and the two free parameters m, σ are fixed by the two conditions (C.11) and (C.12),

$$\sigma = \sqrt{\ln \left[1 + \frac{4\sigma_{\kappa}^2}{(1 - \mu_{\min})^2} \right]}, \quad (\text{C.14})$$

$$m = \ln(1 - \mu_{\min}) - \frac{\sigma^2}{2}. \quad (\text{C.15})$$

The resulting magnification PDF is depicted in the right panel of Fig. 15 for various values of the source redshift. Despite its simplicity, the shifted lognormal model mimics important properties of the expected magnification distribution in the inhomogeneous Universe. On the one hand, $p(\mu)$ peaks at $\mu < 1$, thereby encoding the fact that most lines of sight have a lower column density than average, because voids occupy more volume in the Universe. On the other hand, $p(\mu)$ exhibits a longer tail towards high magnifications, which encodes that, albeit rare, overdensities can produce large magnifications. Note however that our model tends to underestimate the probability of those high magnifications compared to what is obtained with ray tracing in N -body simulations.

References

- [1] Y. B. Zel'dovich, I. D., *The Hypothesis of Cores Retarded during Expansion and the Hot Cosmological Model*, *Soviet Astron. AJ (Engl. Transl.)*, **10** (1967) 602.
- [2] S. Hawking, *Gravitationally collapsed objects of very low mass*, *Mon. Not. Roy. Astron. Soc.* **152** (1971) 75.
- [3] G. F. Chapline, *Cosmological effects of primordial black holes*, *Nature* **253** (1975), no. 5489 251–252.
- [4] A. M. Green and B. J. Kavanagh, *Primordial Black Holes as a dark matter candidate*, *J. Phys. G* **48** (2021), no. 4 043001, [[arXiv:2007.10722](#)].
- [5] A. Dolgov and J. Silk, *Baryon isocurvature fluctuations at small scales and baryonic dark matter*, *Phys. Rev. D* **47** (May, 1993) 4244–4255.
- [6] K. Jedamzik, *Primordial black hole formation during the QCD epoch*, *Phys. Rev. D* **55** (1997) 5871–5875, [[astro-ph/9605152](#)].
- [7] J. Garcia-Bellido, A. D. Linde, and D. Wands, *Density perturbations and black hole formation in hybrid inflation*, *Phys. Rev. D* **54** (1996) 6040–6058, [[astro-ph/9605094](#)].
- [8] M. Y. Khlopov, *Primordial Black Holes*, *Res. Astron. Astrophys.* **10** (2010) 495–528, [[arXiv:0801.0116](#)].
- [9] J. Garcia-Bellido and E. Ruiz Morales, *Primordial black holes from single field models of inflation*, *Phys. Dark Univ.* **18** (2017) 47–54, [[arXiv:1702.03901](#)].
- [10] G. Ballesteros and M. Taoso, *Primordial black hole dark matter from single field inflation*, *Phys. Rev. D* **97** (2018), no. 2 023501, [[arXiv:1709.05565](#)].
- [11] C. Pattison, V. Vennin, H. Assadullahi, and D. Wands, *Quantum diffusion during inflation and primordial black holes*, *JCAP* **10** (2017) 046, [[arXiv:1707.00537](#)].
- [12] D. Inman and Y. Ali-Haïmoud, *Early structure formation in primordial black hole cosmologies*, *Phys. Rev. D* **100** (2019), no. 8 083528, [[arXiv:1907.08129](#)].
- [13] M. Volonteri, M. Habouzit, and M. Colpi, *The origins of massive black holes*, *Nature Rev. Phys.* **3** (2021), no. 11 732–743, [[arXiv:2110.10175](#)].
- [14] B. C. Lacki and J. F. Beacom, *Primordial Black Holes as Dark Matter: Almost All or Almost Nothing*, *Astrophys. J. Lett.* **720** (2010) L67–L71, [[arXiv:1003.3466](#)].
- [15] J. Adamek, C. T. Byrnes, M. Gosenca, and S. Hotchkiss, *WIMPs and stellar-mass primordial black holes are incompatible*, *Phys. Rev. D* **100** (2019), no. 2 023506, [[arXiv:1901.08528](#)].
- [16] G. Bertone, A. M. Coogan, D. Gaggero, B. J. Kavanagh, and C. Weniger, *Primordial Black Holes as Silver Bullets for New Physics at the Weak Scale*, *Phys. Rev. D* **100** (2019), no. 12 123013, [[arXiv:1905.01238](#)].
- [17] B. Carr, F. Kuhnel, and L. Visinelli, *Black holes and WIMPs: all or nothing or something else*, *Mon. Not. Roy. Astron. Soc.* **506** (2021), no. 3 3648–3661, [[arXiv:2011.01930](#)].
- [18] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016), no. 6 061102, [[arXiv:1602.03837](#)].
- [19] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, *Phys. Rev. Lett.* **119** (2017), no. 16 161101, [[arXiv:1710.05832](#)].
- [20] **KAGRA** Collaboration, T. Akutsu et al., *KAGRA: 2.5 Generation Interferometric Gravitational Wave Detector*, *Nature Astron.* **3** (2019), no. 1 35–40, [[arXiv:1811.08079](#)].

- [21] **LIGO Scientific, VIRGO, KAGRA** Collaboration, R. Abbott et al., *First joint observation by the underground gravitational-wave detector, KAGRA, with GEO600*, [arXiv:2203.01270](#).
- [22] Y. Ali-Haïmoud, E. D. Kovetz, and M. Kamionkowski, *Merger rate of primordial black-hole binaries*, *Phys. Rev. D* **96** (2017), no. 12 123523, [[arXiv:1709.06576](#)].
- [23] B. J. Kavanagh, D. Gaggero, and G. Bertone, *Merger rate of a subdominant population of primordial black holes*, *Phys. Rev. D* **98** (2018), no. 2 023536, [[arXiv:1805.09034](#)].
- [24] G. Franciolini, V. Baibhav, V. De Luca, K. K. Y. Ng, K. W. K. Wong, E. Berti, P. Pani, A. Riotto, and S. Vitale, *Quantifying the evidence for primordial black holes in LIGO/Virgo gravitational-wave data*, [arXiv:2105.03349](#).
- [25] G. Franciolini, R. Cotesta, N. Loutrel, E. Berti, P. Pani, and A. Riotto, *How to assess the primordial origin of single gravitational-wave events with mass, spin, eccentricity, and deformability measurements*, [arXiv:2112.10660](#).
- [26] G. Hütsi, M. Raidal, V. Vaskonen, and H. Veermäe, *Two populations of LIGO-Virgo black holes*, *JCAP* **03** (2021) 068, [[arXiv:2012.02786](#)].
- [27] **LIGO Scientific, Virgo** Collaboration, R. Abbott et al., *Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog*, *Astrophys. J. Lett.* **913** (2021), no. 1 L7, [[arXiv:2010.14533](#)].
- [28] **LIGO Scientific, VIRGO, KAGRA** Collaboration, R. Abbott et al., *The population of merging compact binaries inferred using gravitational waves through GWTC-3*, [arXiv:2111.03634](#).
- [29] M. Maggiore et al., *Science Case for the Einstein Telescope*, *JCAP* **03** (2020) 050, [[arXiv:1912.02622](#)].
- [30] D. Reitze et al., *Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO*, *Bull. Am. Astron. Soc.* **51** (2019), no. 7 035, [[arXiv:1907.04833](#)].
- [31] P. Amaro-Seoane et al., *Laser Interferometer Space Antenna*, *arXiv e-prints* (Feb., 2017) [arXiv:1702.00786](#), [[arXiv:1702.00786](#)].
- [32] S. Kawamura et al., *Current status of space gravitational wave antenna DECIGO and B-DECIGO*, *PTEP* **2021** (2021), no. 5 05A105, [[arXiv:2006.13545](#)].
- [33] Y. Gong, J. Luo, and B. Wang, *Concepts and status of Chinese space gravitational wave detection projects*, *Nature Astron.* **5** (2021), no. 9 881–889, [[arXiv:2109.07442](#)].
- [34] S. Hild, *Beyond the Second Generation of Laser-Interferometric Gravitational Wave Observatories*, *Class. Quant. Grav.* **29** (2012) 124006, [[arXiv:1111.6277](#)].
- [35] S. Vitale and M. Evans, *Parameter estimation for binary black holes with networks of third generation gravitational-wave detectors*, *Phys. Rev. D* **95** (2017), no. 6 064052, [[arXiv:1610.06917](#)].
- [36] M. G. Beker, G. Cella, R. Desalvo, M. Doets, H. Grote, J. Harms, E. Hennes, V. Mandic, D. S. Rabeling, J. F. J. van den Brand, and C. M. van Leeuwen, *Improving the sensitivity of future GW observatories in the 1-10 Hz band: Newtonian and seismic noise*, *General Relativity and Gravitation* **43** (Feb., 2011) 623–656.
- [37] I. Dvorkin, E. Vangioni, J. Silk, J.-P. Uzan, and K. A. Olive, *Metallicity-constrained merger rates of binary black holes and the stochastic gravitational wave background*, *Mon. Not. Roy. Astron. Soc.* **461** (2016), no. 4 3877–3885, [[arXiv:1604.04288](#)].
- [38] T. Nakamura, M. Sasaki, T. Tanaka, and K. S. Thorne, *Gravitational waves from coalescing black hole MACHO binaries*, *Astrophys. J. Lett.* **487** (1997) L139–L142, [[astro-ph/9708060](#)].

- [39] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, *Primordial Black Hole Scenario for the Gravitational-Wave Event GW150914*, *Phys. Rev. Lett.* **117** (2016), no. 6 061101, [[arXiv:1603.08338](#)]. [Erratum: *Phys.Rev.Lett.* 121, 059901 (2018)].
- [40] S. M. Koushiappas and A. Loeb, *Maximum redshift of gravitational wave merger events*, *Phys. Rev. Lett.* **119** (2017), no. 22 221104, [[arXiv:1708.07380](#)].
- [41] Z.-C. Chen and Q.-G. Huang, *Distinguishing Primordial Black Holes from Astrophysical Black Holes by Einstein Telescope and Cosmic Explorer*, *JCAP* **08** (2020) 039, [[arXiv:1904.02396](#)].
- [42] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *Bayesian Evidence for Both Astrophysical and Primordial Black Holes: Mapping the GWTC-2 Catalog to Third-Generation Detectors*, *JCAP* **05** (2021) 003, [[arXiv:2102.03809](#)].
- [43] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *The Minimum Testable Abundance of Primordial Black Holes at Future Gravitational-Wave Detectors*, [arXiv:2106.13769](#).
- [44] K. K. Y. Ng, S. Chen, B. Goncharov, U. Dupletsa, S. Borhanian, M. Branchesi, J. Harms, M. Maggiore, B. S. Sathyaprakash, and S. Vitale, *On the single-event-based identification of primordial black hole mergers at cosmological distances*, [arXiv:2108.07276](#).
- [45] V. Atal, J. J. Blanco-Pillado, A. Sanglas, and N. Triantafyllou, *Constraining changes in the merger history of (P)BH binaries with the stochastic gravitational wave background*, [arXiv:2201.12218](#).
- [46] M. Braglia, J. Garcia-Bellido, and S. Kuroyanagi, *Tracking the origin of black holes with the stochastic gravitational wave background popcorn signal*, [arXiv:2201.13414](#).
- [47] K. Ioka, T. Chiba, T. Tanaka, and T. Nakamura, *Black hole binary formation in the expanding universe: Three body problem approximation*, *Phys. Rev. D* **58** (1998) 063003, [[astro-ph/9807018](#)].
- [48] J. R. Chisholm, *Clustering of primordial black holes: basic results*, *Phys. Rev. D* **73** (2006) 083504, [[astro-ph/0509141](#)].
- [49] J. R. Chisholm, *Clustering of Primordial Black Holes. II. Evolution of Bound Systems*, *Phys. Rev. D* **84** (2011) 124031, [[arXiv:1110.4402](#)].
- [50] M. Raidal, C. Spethmann, V. Vaskonen, and H. Veermäe, *Formation and Evolution of Primordial Black Hole Binaries in the Early Universe*, *JCAP* **02** (2019) 018, [[arXiv:1812.01930](#)].
- [51] S. Bhatiani, X. Dai, and E. Guerras, *Confirmation of Planet-mass Objects in Extragalactic Systems*, *ApJ* **885** (Nov., 2019) 77, [[arXiv:1909.11610](#)].
- [52] S. Mukherjee, M. S. P. Meinema, and J. Silk, *Prospects of discovering subsolar primordial black holes using the stochastic gravitational wave background from third-generation detectors*, *Mon. Not. Roy. Astron. Soc.* **510** (2022), no. 4 6218–6224, [[arXiv:2107.02181](#)].
- [53] O. Pujolas, V. Vaskonen, and H. Veermäe, *Prospects for probing gravitational waves from primordial black hole binaries*, *Phys. Rev. D* **104** (2021), no. 8 083521, [[arXiv:2107.03379](#)].
- [54] **Planck** Collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, [arXiv:1807.06209](#).
- [55] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Supplement: The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914*, *Astrophys. J. Suppl.* **227** (2016), no. 2 14, [[arXiv:1606.03939](#)].
- [56] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott et al., *Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO*, *Phys. Rev. D* **96** (2017), no. 2 022001, [[arXiv:1704.04628](#)].

- [57] A. Lewis, A. Challinor, and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473–476, [[astro-ph/9911177](#)].
- [58] C. Howlett, A. Lewis, A. Hall, and A. Challinor, *CMB power spectrum parameter degeneracies in the era of precision cosmology*, *JCAP* **2012** (Apr., 2012) 027, [[arXiv:1201.3654](#)].
- [59] E. Hall, “gw-horizon-plot [Git Repository].”
<https://git.ligo.org/evan.hall/gw-horizon-plot>, 2019.
- [60] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli, and A. G. Riess, *Did LIGO detect dark matter?*, *Phys. Rev. Lett.* **116** (2016), no. 20 201301, [[arXiv:1603.00464](#)].
- [61] G. Franciolini, K. Kritos, E. Berti, and J. Silk, *Primordial black hole mergers from three-body interactions*, [arXiv:2205.15340](#).
- [62] T. Bringmann, P. F. Depta, V. Domcke, and K. Schmidt-Hoberg, *Towards closing the window of primordial black holes as dark matter: The case of large clustering*, *Phys. Rev. D* **99** (2019), no. 6 063532, [[arXiv:1808.05910](#)].
- [63] G. Ballesteros, P. D. Serpico, and M. Taoso, *On the merger rate of primordial black holes: effects of nearest neighbours distribution and clustering*, *JCAP* **10** (2018) 043, [[arXiv:1807.02084](#)].
- [64] S. Young and C. T. Byrnes, *Initial clustering and the primordial black hole merger rate*, *JCAP* **03** (2020) 004, [[arXiv:1910.06077](#)].
- [65] P. C. Peters, *Gravitational radiation and the motion of two point masses*, *Phys. Rev.* **136** (Nov, 1964) B1224–B1232.
- [66] A. Caputo, L. Sberna, A. Toubiana, S. Babak, E. Barausse, S. Marsat, and P. Pani, *Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electromagnetic counterpart*, *Astrophys. J.* **892** (2020), no. 2 90, [[arXiv:2001.03620](#)].
- [67] M. Ricotti, J. P. Ostriker, and K. J. Mack, *Effect of Primordial Black Holes on the Cosmic Microwave Background and Cosmological Parameter Estimates*, *Astrophys. J.* **680** (2008) 829, [[arXiv:0709.0524](#)].
- [68] G. Hütsi, M. Raidal, and H. Veermäe, *Small-scale structure of primordial black hole dark matter and its implications for accretion*, *Phys. Rev. D* **100** (2019), no. 8 083016, [[arXiv:1907.06533](#)].
- [69] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *The evolution of primordial black holes and their final observable spins*, *JCAP* **04** (2020) 052, [[arXiv:2003.02778](#)].
- [70] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *Primordial Black Holes Confront LIGO/Virgo data: Current situation*, *JCAP* **06** (2020) 044, [[arXiv:2005.05641](#)].
- [71] K. M. Belotsky, V. I. Dokuchaev, Y. N. Eroshenko, E. A. Esipova, M. Y. Khlopov, L. A. Khromykh, A. A. Kirillov, V. V. Nikulin, S. G. Rubin, and I. V. Svadkovsky, *Clusters of primordial black holes*, *Eur. Phys. J. C* **79** (2019), no. 3 246, [[arXiv:1807.06590](#)].
- [72] V. Vaskonen and H. Veermäe, *Lower bound on the primordial black hole merger rate*, *Phys. Rev. D* **101** (2020), no. 4 043015, [[arXiv:1908.09752](#)].
- [73] K. Jedamzik, *Primordial Black Hole Dark Matter and the LIGO/Virgo observations*, *JCAP* **09** (2020) 022, [[arXiv:2006.11172](#)].
- [74] V. De Luca, V. Desjacques, G. Franciolini, and A. Riotto, *The clustering evolution of primordial black holes*, *JCAP* **11** (2020) 028, [[arXiv:2009.04731](#)].
- [75] S. Mukherjee and J. Silk, *Can we distinguish astrophysical from primordial black holes via the stochastic gravitational wave background?*, *Mon. Not. Roy. Astron. Soc.* **506** (2021), no. 3 3977–3985, [[arXiv:2105.11139](#)].

- [76] K. Belczynski, D. E. Holz, T. Bulik, and R. O’Shaughnessy, *The first gravitational-wave source from the isolated evolution of two 40-100 Msun stars*, *Nature* **534** (2016) 512, [[arXiv:1602.04531](#)].
- [77] M. Mapelli and N. Giacobbo, *The cosmic merger rate of neutron stars and black holes*, *Mon. Not. Roy. Astron. Soc.* **479** (2018), no. 4 4391–4398, [[arXiv:1806.04866](#)].
- [78] M. Mapelli, N. Giacobbo, F. Santoliquido, and M. C. Artale, *The properties of merging black holes and neutron stars across cosmic time*, *Mon. Not. Roy. Astron. Soc.* **487** (2019), no. 1 2–13, [[arXiv:1902.01419](#)].
- [79] L. Graziani, R. Schneider, S. Marassi, W. Del Pozzo, M. Mapelli, and N. Giacobbo, *Cosmic archaeology with massive stellar black hole binaries*, *Mon. Not. Roy. Astron. Soc.* **495** (2020), no. 1 L81–L85, [[arXiv:2004.03603](#)].
- [80] F. Santoliquido, M. Mapelli, Y. Bouffanais, N. Giacobbo, U. N. Di Carlo, S. Rastello, M. C. Artale, and A. Ballone, *The cosmic merger rate density evolution of compact binaries formed in young star clusters and in isolated binaries*, *Astrophys. J.* **898** (2020), no. 2 152, [[arXiv:2004.09533](#)].
- [81] M. Mapelli, F. Santoliquido, Y. Bouffanais, M. A. Sedda, M. C. Artale, and A. Ballone, *Mass and Rate of Hierarchical Black Hole Mergers in Young, Globular and Nuclear Star Clusters*, *Symmetry* **13** (2021), no. 9 1678, [[arXiv:2007.15022](#)].
- [82] M. Mapelli, *Formation Channels of Single and Binary Stellar-Mass Black Holes*, [arXiv:2106.00699](#).
- [83] M. A. Sedda, M. Mapelli, M. Benacquista, and M. Spera, *Population synthesis of black hole mergers with B-POP: the impact of dynamics, natal spins, and intermediate-mass black holes on the population of gravitational wave sources*, [arXiv:2109.12119](#).
- [84] P. Madau, L. Pozzetti, and M. Dickinson, *The Star Formation History of Field Galaxies*, *ApJ* **498** (May, 1998) 106–116, [[astro-ph/9708220](#)].
- [85] P. Madau and M. Dickinson, *Cosmic Star Formation History*, *Ann. Rev. Astron. Astrophys.* **52** (2014) 415–486, [[arXiv:1403.0007](#)].
- [86] M. D. Kistler, H. Yuksel, and A. M. Hopkins, *The Cosmic Star Formation Rate from the Faintest Galaxies in the Unobservable Universe*, [arXiv:1305.1630](#).
- [87] E. Vangioni, K. A. Olive, T. Prestegard, J. Silk, P. Petitjean, and V. Mandic, *The Impact of Star Formation and Gamma-Ray Burst Rates at High Redshift on Cosmic Chemical Evolution and Reionization*, *Mon. Not. Roy. Astron. Soc.* **447** (2015) 2575, [[arXiv:1409.2462](#)].
- [88] K. Nagamine, V. Springel, and L. Hernquist, *Abundance of damped Lyman-alpha absorbers in cosmological SPH simulations*, *Mon. Not. Roy. Astron. Soc.* **348** (2004) 421, [[astro-ph/0302187](#)].
- [89] P. D. Serpico, V. Poulin, D. Inman, and K. Kohri, *Cosmic microwave background bounds on primordial black holes including dark matter halo accretion*, *Phys. Rev. Res.* **2** (2020), no. 2 023204, [[arXiv:2002.10771](#)].
- [90] M. Oguri, J. M. Diego, N. Kaiser, P. L. Kelly, and T. Broadhurst, *Understanding caustic crossings in giant arcs: characteristic scales, event rates, and constraints on compact dark matter*, *Phys. Rev. D* **97** (2018), no. 2 023518, [[arXiv:1710.00148](#)].
- [91] J. Manshanden, D. Gaggero, G. Bertone, R. M. T. Connors, and M. Ricotti, *Multi-wavelength astronomical searches for primordial black holes*, *JCAP* **06** (2019) 026, [[arXiv:1812.07967](#)].
- [92] B. J. Kavanagh, “PBHbounds [Code v1.0, accessed 03/03/2022].” <https://github.com/bradkav/PBHbounds>, DOI:10.5281/zenodo.3538998, 2022.

- [93] B. P. Abbott et al., *Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy*, *Phys. Rev. D* **93** (2016), no. 11 112004, [[arXiv:1604.00439](#)]. [Addendum: *Phys.Rev.D* 97, 059901 (2018)].
- [94] T. G. F. Li, *Extracting Physics from Gravitational Waves: Testing the Strong-field Dynamics of General Relativity and Inferring the Large-scale Structure of the Universe*. PhD thesis, Vrije U., Amsterdam, 2013.
- [95] T. Bayes, *An essay toward solving a problem in the doctrine of chances*, *Phil. Trans. Roy. Soc. Lond.* **53** (1764) 370–418.
- [96] J. Torrado and A. Lewis, *Cobaya: Code for Bayesian Analysis of hierarchical physical models*, *JCAP* **05** (2021) 057, [[arXiv:2005.05290](#)].
- [97] A. Lewis, *GetDist: a Python package for analysing Monte Carlo samples*, [arXiv:1910.13970](#).
- [98] **LIGO Scientific, Virgo** Collaboration, R. Abbott et al., *GW190521: A Binary Black Hole Merger with a Total Mass of $150M_{\odot}$* , *Phys. Rev. Lett.* **125** (2020), no. 10 101102, [[arXiv:2009.01075](#)].
- [99] **LIGO Scientific, Virgo** Collaboration, R. Abbott et al., *Properties and Astrophysical Implications of the $150 M_{\odot}$ Binary Black Hole Merger GW190521*, *Astrophys. J. Lett.* **900** (2020), no. 1 L13, [[arXiv:2009.01190](#)].
- [100] B. Carr, S. Clesse, J. García-Bellido, and F. Kühnel, *Cosmic conundra explained by thermal history and primordial black holes*, *Phys. Dark Univ.* **31** (2021) 100755, [[arXiv:1906.08217](#)].
- [101] M. Trashorras, J. García-Bellido, and S. Nesseris, *The clustering dynamics of primordial black holes in N -body simulations*, *Universe* **7** (2021), no. 1 18, [[arXiv:2006.15018](#)].
- [102] G. Cañas Herrera, O. Contigiani, and V. Vardanyan, *Learning How to Surf: Reconstructing the Propagation and Origin of Gravitational Waves with Gaussian Processes*, *Astrophys. J.* **918** (2021), no. 1 20, [[arXiv:2105.04262](#)].
- [103] R. S. de Souza, N. Yoshida, and K. Ioka, *Populations III.1 and III.2 gamma-ray bursts: constraints on the event rate for future radio and X-ray surveys*, *A&A* **533** (Sept., 2011) A32, [[arXiv:1105.2395](#)].
- [104] K. K. Y. Ng, S. Vitale, W. M. Farr, and C. L. Rodriguez, *Probing multiple populations of compact binaries with third-generation gravitational-wave detectors*, *Astrophys. J. Lett.* **913** (2021), no. 1 L5, [[arXiv:2012.09876](#)].
- [105] K. K. Y. Ng, G. Franciolini, E. Berti, P. Pani, A. Riotto, and S. Vitale, *Constraining high-redshift stellar-mass primordial black holes with next-generation ground-based gravitational-wave detectors*, [arXiv:2204.11864](#).
- [106] J. E. Gunn and J. R. Gott, III, *On the Infall of Matter into Clusters of Galaxies and Some Effects on Their Evolution*, *Astrophys. J.* **176** (1972) 1–19.
- [107] J. A. Fillmore and P. Goldreich, *Self-similar gravitational collapse in an expanding universe*, *ApJ* **281** (June, 1984) 1–8.
- [108] E. Bertschinger, *Self-similar secondary infall and accretion in an Einstein-de Sitter universe*, *ApJS* **58** (May, 1985) 39–65.
- [109] G. D. Quinlan, *The Time scale for core collapse in spherical star clusters*, *New Astron.* **1** (1996) 255, [[astro-ph/9606182](#)].
- [110] P. J. E. Peebles, *The large-scale structure of the universe*. Princeton University Press, 1980.
- [111] **Particle Data Group** Collaboration, P. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020), no. 8 083C01.
- [112] M. Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments*. Oxford University Press, 2007.

- [113] G. Cañas Herrera, O. Contigiani, and V. Vardanyan, *Cross-correlation of the astrophysical gravitational-wave background with galaxy clustering*, *Phys. Rev. D* **102** (2020), no. 4 043513, [[arXiv:1910.08353](#)].
- [114] S. Mukherjee, B. D. Wandelt, S. M. Nissanke, and A. Silvestri, *Accurate precision cosmology with redshift unknown gravitational wave sources*, *Phys. Rev. D* **103** (Feb, 2021) 043520, [[arXiv:2007.02943](#)].
- [115] L. K. Nuttall, *Characterizing transient noise in the LIGO detectors*, *Philosophical Transactions of the Royal Society of London Series A* **376** (May, 2018) 20170286, [[arXiv:1804.07592](#)].
- [116] **LIGO** Collaboration, D. Davis et al., *LIGO detector characterization in the second and third observing runs*, *Class. Quant. Grav.* **38** (2021), no. 13 135014, [[arXiv:2101.11673](#)].
- [117] N. Straumann, *General Relativity*. Graduate Texts in Physics. Springer, Dordrecht, 2013.
- [118] C. Dalang, P. Fleury, and L. Lombriser, *Horndeski gravity and standard sirens*, *Phys. Rev. D* **102** (2020), no. 4 044036, [[arXiv:1912.06117](#)].
- [119] M. Betoule et al., *Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples*, *Astronomy and Astrophysics* **568** (Aug., 2014) A22, [[arXiv:1401.4064](#)].
- [120] B. S. Sathyaprakash, B. F. Schutz, and C. Van Den Broeck, *Cosmography with the Einstein Telescope*, *Class. Quant. Grav.* **27** (2010) 215006, [[arXiv:0906.4151](#)].
- [121] W. Zhao, C. Van Den Broeck, D. Baskaran, and T. G. F. Li, *Determination of Dark Energy by the Einstein Telescope: Comparing with CMB, BAO and SNIa Observations*, *Phys. Rev.* **D83** (2011) 023005, [[arXiv:1009.0206](#)].
- [122] R.-G. Cai and T. Yang, *Estimating cosmological parameters by the simulated data of gravitational waves from the einstein telescope*, *Physical Review D* **95** (2016) [[arxiv:1608.08008v2](#)].
- [123] M. Du, W. Yang, L. Xu, S. Pan, and D. F. Mota, *Future constraints on dynamical dark-energy using gravitational-wave standard sirens*, *Phys. Rev.* **D100** (2019), no. 4 043535, [[arXiv:1812.01440](#)].
- [124] S.-J. Jin, D.-Z. He, Y. Xu, J.-F. Zhang, and X. Zhang, *Forecast for cosmological parameter estimation with gravitational-wave standard siren observation from the Cosmic Explorer*, [[arXiv:2001.05393](#)].
- [125] N. B. Hogg, M. Martinelli, and S. Nesseris, *Constraints on the distance duality relation with standard sirens*, *JCAP* **12** (2020) 019, [[arXiv:2007.14335](#)].
- [126] J. Jönsson, M. Sullivan, I. Hook, S. Basa, R. Carlberg, A. Conley, D. Fouchez, D. A. Howell, K. Perrett, and C. Pritchett, *Constraining dark matter halo properties using lensed Supernova Legacy Survey supernovae*, *MNRAS* **405** (June, 2010) 535–544, [[arXiv:1002.1374](#)].
- [127] Y. B. Zel'dovich, *Observations in a Universe Homogeneous in the Mean*, *Sov. Astron. Lett.* **8** (Aug., 1964) 13.
- [128] O. Wucknitz, *From planes to spheres: about gravitational lens magnifications*, *MNRAS* **386** (May, 2008) 230–244, [[arXiv:0801.3758](#)].
- [129] M.-A. Breton and P. Fleury, *Theoretical and numerical perspectives on cosmic distance averages*, *A&A* **655** (Nov., 2021) A54, [[arXiv:2012.07802](#)].