From inflation to black hole mergers and back again: Gravitational-wave data-driven constraints on inflationary scenarios with a first-principle model of primordial black holes across the QCD epoch

Gabriele Franciolini[1,2] Ilia Musco,[2] Paolo Pani,[1,2] and Alfredo Urbano[1,2]

[1]*Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Roma, Italy*
[2]*INFN, Sezione di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy*

Recent population studies have searched for a subpopulation of primordial black holes (PBHs) in the gravitational-wave (GW) events so far detected by LIGO/Virgo/KAGRA (LVK), in most cases adopting a phenomenological PBH mass distribution. When deriving such a population from first principles in the standard scenario, however, the equation of state of the Universe at the time of PBH formation may strongly affect the PBH abundance and mass distribution, which ultimately depend on the power spectrum of cosmological perturbations. Here we improve on previous population studies on several aspects: (i) we adopt state-of-the-art PBH formation models describing the collapse of cosmological perturbations across the QCD epoch; (ii) we perform the first Bayesian multipopulation inference on GW data including PBHs and directly using power spectrum parameters instead of phenomenological distributions; (iii) we critically confront the PBH scenario with LVK phenomenological models describing the GWTC-3 catalog both in the neutron-star and in the BH mass ranges, also considering PBHs as a subpopulation of the total events. Our results confirm that LVK observations prevent the majority of the dark matter to be in the form of stellar mass PBHs. We find that the best-fit PBH model can comprise a small fraction of the total events, in particular it can naturally explain events in the mass gaps. If the lower mass-gap event GW190814 is interpreted as a PBH binary, we predict that LVK should detect up to a few subsolar mergers and one to ≈ 30 lower mass-gap events during the upcoming O4 and O5 runs. Finally, mapping back the best-fit power spectrum into an ultra-slow-roll inflationary scenario, we show that the latter predicts detectable PBH mergers in the LVK band, a stochastic GW background detectable by current and future instruments, and may include the entirety of dark matter in asteroid-mass PBHs.

I. INTRODUCTION

Primordial black holes (PBHs) [1–4] might have formed in the early Universe after inflation from the collapse of large amplitude cosmological perturbations [5–8] or by other mechanisms. In the standard formation scenario, their characteristic mass depends mostly on the time these inhomogeneities reenter the cosmological horizon, whereas their abundance and mass distribution depend strongly on the equation of state (EOS) of the Universe at that epoch [4,9–16], and it is ultimately controlled by the power spectrum of cosmological curvature perturbations. In particular, the mass of PBHs can span several orders of magnitude and is not bounded from below ($M \gtrsim 2M_\odot$) as in the case of stellar-origin BHs, providing one of the key distinctive features [17] of this scenario.

Besides being unique messengers of the early-time cosmology and inflationary models, in certain mass ranges PBHs could comprise the entirety of the dark matter, and could seed supermassive BHs at high redshift [18–20].

These tantalizing possibilities have motivated the recent growing interest in searching for PBHs (see [21] for a recent review), especially using gravitational-wave (GW) data.

PBHs could contribute to at least a fraction of the BH merger events detected by the LIGO-Virgo-KAGRA (LVK) collaboration [22–24] so far [25–47], and to those that will be detected by future GW instruments [17,42,48–53] (see Refs. [54–56] for reviews on PBHs as GW sources). In addition to outstanding events such as GW190425 [with a total mass that exceeds that one of known galactic neutron star (NS) binaries] and the mass-gap events (such as GW190814 [57], GW190521 [58], and GW190426_190642) which do not fit naturally in the standard astrophysical scenarios and might have a different origin, a subpopulation of PBHs may be competitive with certain astrophysical population models for explaining a fraction of events [45]. Population studies [59] will inevitably become very relevant as the number of detections increases, both during future LVK runs and especially in the era of next-generation detectors [60,61].

So far population studies aimed at identifying a (sub)population of PBHs in LVK data (e.g., [38–40,42,45]) have adopted phenomenological PBH mass distributions, such as a log-normal or a power-law function, which should approximately capture different underlying formation mechanisms. However, in a realistic setting starting from first principles, the computation of the mass distribution should take into account several aspects: a given underlying early Universe model directly determines the power spectrum of primordial curvature perturbations, which in turn affects the collapse and eventually the PBH mass function. The latter might show several features which are not necessarily captured by simple parametrizations.

For example, the QCD phase transition of the early Universe, when free quarks are confined within hadrons, strongly affects the EOS of the cosmological fluid at energy scales corresponding to the formation of solar-mass PBHs [15,62–65]. As a rule of thumb, any drop of the EOS parameter $w = p/\rho$ (being p and ρ the pressure and energy density of the cosmological fluid, respectively) relative to the radiation-dominated case ($w = 1/3$) is associated with an enhancement of PBH production, since the pressure contribution to balance gravity is weaker. Thus, as a consequence of the QCD phase transition at few hundred MeV, one would generically expect a peak of the PBH mass function in the solar mass range, provided the power spectrum is sufficiently large at those specific scales. However, since the gravitational collapse is a nonlinear process, several details of the initial power spectrum might affect the final PBH mass function, also providing characteristic tails and subtle correlations between different mass scales that, as we shall discuss, should be taken into account.

In this paper we go beyond phenomenological models and build a framework to link the formation of PBH binaries and their GW signatures from first principles. Our final goal is to use GW data to constrain *ab initio* models and inform inflationary dynamics. This allows us to build a self-consistent scenario which, on the one hand, is compatible with current constraints and, on the other hand, makes concrete predictions across a wide range of PBH masses. Indeed, owning to the specific shape of the mass distribution arising from an *ab initio* model, constraints on a given mass range can percolate on different mass scales, making *ab initio* models much more predictive (and hence falsifiable) than generic parametrizations.

One of the key novel ingredients of our framework is the inclusion of state-of-the-art PBH formation models describing the collapse of radiation across the QCD epoch, incorporating the effect of critical collapse in shaping the QCD enhancement [16]. A scenario in which the QCD era was deemed responsible for shaping the mass distribution of PBHs in the solar mass range was devised in Refs. [64,65] (see also [57,66–68]), where the power spectrum of curvature perturbations was specifically tuned to be nearly, but not exactly, scale invariant, which

enhances the relevance of the QCD peak around the solar mass scale. However, the physics of the collapse across the QCD epoch alone does not determine the entire PBH mass function, which chiefly depends also on the shape of the curvature perturbation spectrum. It follows that the ratio between the abundance of PBHs at $\mathcal{O}(M_\odot)$ and $\mathcal{O}(30M_\odot)$ (relevant for LVK detections) cannot be predicted by the QCD effect alone, unless strong assumptions on the spectral amplitudes at those two scales are made. Reference [69] specifically analyzed such a scenario, and concluded that the GW bound in the subsolar mass range (from the absence of subsolar events during O1/O2/O3 LVK runs) sets the most important constraint. However, bounds on subsolar PBHs rely on assuming a specific PBH mass distribution [70], which is not necessarily the one assumed to come from the QCD phase transition in previous works and by the *ab initio* model considered here. Furthermore, Ref. [69] concluded that PBH mergers shaped by the QCD EOS may not contribute to current LVK observations, unless an *ad hoc* mass evolution for the PBH mass function and a cutoff in the power spectrum very close to the QCD scale are artificially introduced by hand.

We will extend the scope of these analyses, by exploring the role of the spectral tilt, which was previously fixed to a specific value (and is not *a priori* related to the one constrained by cosmic microwave background (CMB) observations at much larger scales). As we shall later discuss in detail, we will leave the tilt as a free parameter of the model, which is eventually inferred from the data. Due to the exponential dependence of the PBH abundance on the density variance, small modifications to the tilt (around 10%) greatly reduce the QCD solar mass peak and render the scenario insensitive to the high-scale (i.e., low-mass) spectrum cutoff.

We revisit previous constraints by performing the first Bayesian population inference on GW data including a subpopulation of PBHs and directly using *ab initio* power spectrum parameters (including the tilt and the effect of the QCD phase) instead of phenomenological distributions, and confronting the PBH scenario with the most recent GWTC-3 dataset [24,59]. We allow the PBH model to produce subsolar merger events, and the constraint deriving from the absence of such binaries in LVK data is consistently included in our analysis by construction. This constraint was not included in Ref. [71], where the fit was arbitrarily cut at $\approx M_\odot$ and no constraining power from the absence of subsolar mergers is included in the inference. Another important addition of our analysis relative to [71] is the inclusion of a phenomenological fit describing the population [59], which is crucial to assess the nature of events in the solar-mass range.

Finally, employing the reverse engineering approach devised in Ref. [72], we show how the GW data-driven power spectrum can be naturally accommodated into an ultra-slow-roll (USR) inflationary scenario [73–83].

Remarkably, a single USR model informed by current observational constraints may explain the entirety of the dark matter in asteroid-mass PBHs while also allowing [72] for detectable PBH mergers in the LVK band due to the enhancement of the PBH distribution around the solar-mass range induced by both the QCD phase transition and spectral features, and for detectable stochastic GW background (SGWB) signals from the nano-Hertz to the kilo-Hertz band.

Throughout all this paper we assume geometrical units with $c = G = 1$.

II. PBH FORMATION ACROSS THE QCD EPOCH

Within the standard PBH formation scenario, which assumes PBHs form out of the collapse of large amplitude cosmological perturbations in the radiation dominated early Universe, a crucial role is played by the power spectrum of primordial curvature perturbations, $P_\zeta(k)$, and the corresponding value of the threshold δ_c for PBH formation. In this section, we summarize the results of [16], where a state-of-the-art derivation of the threshold for the formation of PBHs during the QCD epoch is given. This is obtained using detailed general relativistic numerical simulations, assuming spherical symmetry.

We start with a brief introduction of the QCD phase transition, followed by a quick review of the mathematical formalism one needs to describe consistently the initial condition for PBH formation, clarifying the key ingredients used in the computation of the threshold. This allows us to discuss the impact on the formation of PBHs characterized by a solar-mass range of scale, and how we can include the effects of this phase within the computation of the PBH mass distribution.

A. The QCD phase transition

During the confinement of quarks into hadrons the particle degrees of freedom are varying with the temperature T. This results in a ratio between the pressure p and the total energy density ρ of the medium being not constant—as in the case of a gas of ultrarelativistic particles—but varying with time according to

$$w(T) \equiv \frac{p}{\rho} = \frac{4g_{*,s}(T)}{3g_*(T)} - 1. \tag{1}$$

The functions $g_*(T)$ and $g_{*,s}(T)$ denote the two relevant measures of the effective number of relativistic degrees of freedom, defined as

$$g_*(T) = \frac{30\rho}{\pi^2 T^4} \quad \text{and} \quad g_{*,s}(T) = \frac{45s}{2\pi^2 T^3}, \tag{2}$$

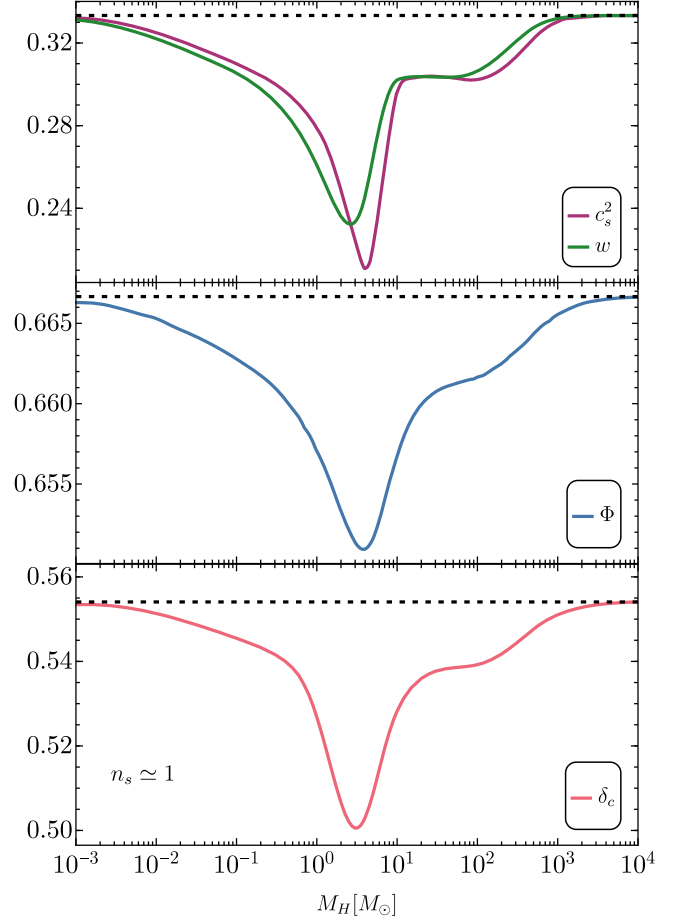where s is the entropy density of the medium and the pressure p is given by

FIG. 1. Top panel: the EOS parameter $w = p/\rho$ (red) and squared speed of sound (blue) as functions of the cosmological horizon mass M_H. Central panel: evolution of the EOS dependent parameter Φ, relating the density contrast to the curvature perturbation as functions of the cosmological horizon mass M_H. Bottom panel: same as above but showing the threshold for PBH formation. The dashed horizontal lines refer to the values obtained in the perfect radiation-fluid case.

$$p = sT - \rho = w(T)\rho. \tag{3}$$

In the top panel of Fig. 1 we show the behavior of w and the sound speed squared $c_s^2 \equiv \partial p/\partial \rho$ during the QCD phase transition, obtained from lattice QCD simulations [84,85], using the cosmological horizon mass M_H as a measure of the fluid temperature.[1] As we will discuss later, the non-negligible change of these two quantities during the QCD epoch, with respect the constant value ($w = c_s^2 = 1/3$) they have when the Universe is radiation dominated, plays a crucial role during the collapse of cosmological perturbations, and gives rise to a reduction of the threshold for the formation of PBHs (see the bottom panel of Fig. 1).

---

[1]The cosmological horizon R_H is a marginally trapped surface [86], as the apparent horizon of a black hole, with $R_H = 2M_H$.

B. Gradient expansion

The threshold δ_c for PBH formation is defined as the critical value of the cosmological perturbation amplitude δ, such that for $\delta > \delta_c$ an apparent horizon appears during the collapse and a PBH is formed, while for $\delta < \delta_c$ the collapse bounces and the cosmological perturbation is dispersed into the surrounding medium.

To compute the value of the threshold one needs to specify initial conditions of the numerical simulations on the superhorizon scale, when the asymptotic form of the space-time metric is given by

$$ds^2 = -dt^2 + a^2(t)e^{2\zeta(r)}[dr^2 + r^2 d\Omega^2], \quad (4)$$

where $a(t)$ is the scale factor, while $\zeta(r)$ is the conserved comoving curvature perturbations defined on a super-Hubble scale, converging to zero at infinity where the Universe is taken to be unperturbed and spatially flat.

In this regime, using the so-called gradient expansion or long wavelength approximation [87–89], the energy density contrast $\delta\rho/\rho_b$ for adiabatic perturbations [the ones generated by a curvature profile $\zeta(r)$] can be written as [90]

$$\frac{\delta\rho}{\rho_b}(r,t) = -\frac{4}{3}\Phi\left(\frac{1}{aH}\right)^2 e^{-5\zeta(r)/2}\nabla^2 e^{\zeta(r)/2}, \quad (5)$$

where $H \equiv \dot{a}/a$ is the Hubble parameter, while the function $\Phi(t)$ depends on the equation of state of the Universe and is obtained by solving the following equation [88]:

$$\frac{1}{H}\frac{d\Phi(t)}{dt} + \frac{5+3w(t)}{2}\Phi(t) - \frac{3}{2}(1+w(t)) = 0 \quad (6)$$

integrated from past infinity to the time when the amplitude of the perturbation is computed. In standard models of the very early Universe (i.e., just after inflation) this is assumed to be dominated by a radiation dominated medium, with EOS $p = w\rho$ and $w = 1/3$.

When a constant $w(t) = \bar{w}$ characterizes the fluid dominating the energy budget of the Universe, we have $d\Phi(t)/dt = 0$ and one obtains

$$\bar{\Phi} = \frac{3(1+\bar{w})}{(5+3\bar{w})}, \quad (7)$$

yielding $\bar{\Phi} = 2/3$ for a radiation fluid with $\bar{w} = 1/3$. Equation (7) is an attractor solution of Eq. (6), i.e., if $w(t)$ slowly varies in time, $d\Phi(t)/dt \simeq 0$ and the evolution of Φ approaches the value given by Eq. (7). The behavior of Φ across the QCD phase transition, obtained by solving Eq. (6), differs from the average $\bar{\Phi}$, particularly in the region where w and c_s^2 are quickly varying with respect to M_H. This is shown in the middle panel of Fig. 1.

It was shown that a consistent way to define the threshold for PBH formation is in terms of the smoothed density contrast δ_m computed at horizon crossing time, i.e., $aH = 1/r_m$. Using a top-hat window function with areal radius $R = a(t)\exp[\zeta(r_m)]r_m$, where r_m indicates the location of the maximum of the mass excess, also called compaction function, the amplitude of spherically symmetric peaks in the smoothed density field is related to the curvature perturbation as [91]

$$\delta_m = -\Phi r_m\zeta'(r_m)[2 + r_m\zeta'(r_m)]. \quad (8)$$

Although strictly speaking the gradient expansion approach is valid only on superhorizon scales, to compute the perturbation amplitude δ_m it is useful to extend this approach up to the cosmological horizon crossing time. Since then the region involved in the formation of a PBH becomes causally connected, and the collapse starts shortly afterwards. This gives a well defined criterion to quantify the amplitude of cosmological perturbations, comparing different initial configuration collapsing at different epochs.

C. The threshold for PBHs

The spherically symmetric numerical simulations used to compute the threshold δ_c, and the mass distribution discussed in the next section, have been performed with a numerical code developed in [12] that has been widely used and tested [13,14], including also an adaptive mesh refinement scheme, which makes the code very flexible. This allows to compute the threshold with very high accuracy, a crucial point for calculating the mass spectrum discussed in Sec. II D.

The behavior of the threshold δ_c during the QCD phase transition, computed when a nearly scale invariant power spectrum (see Sec. II E) is assumed, is shown in the bottom panel of Fig. 1. Looking at the top panel we can appreciate the varying EOS during this epoch: the value of threshold δ_c is affected by the change of both w and c_s^2, with a minimum value $\delta_c \simeq 0.5$ reached at $M_H = 3M_\odot$ (between the minimum of these two quantities), about 10% less than the value $\delta_c \simeq 0.55$ one has during the standard radiation scenario (dashed line). The effect of Φ shown in the central panel is to give an additional lowering, accounting up to 25% of the total decrement of the threshold, smoothing the whole behavior of δ_c, monotonically decreasing for $M_H \lesssim 3M_\odot$, and monotonically increasing afterwards, for $M_H \gtrsim 3M_\odot$.

This is quite different from the behavior of δ_c obtained in [15] where the variation of the threshold during the QCD epoch was obtained simply from a fit of the numerical results given in [14] where only w is varying.[2] This neglects

---

[2]In [14] the threshold was not defined at the maximum of the compaction function but at the edge of the overdensity, as it was used to be done in past works [11–13]. This does not allow to make a direct comparison with the new numerical results without a proper rescaling (see [16] for more details).
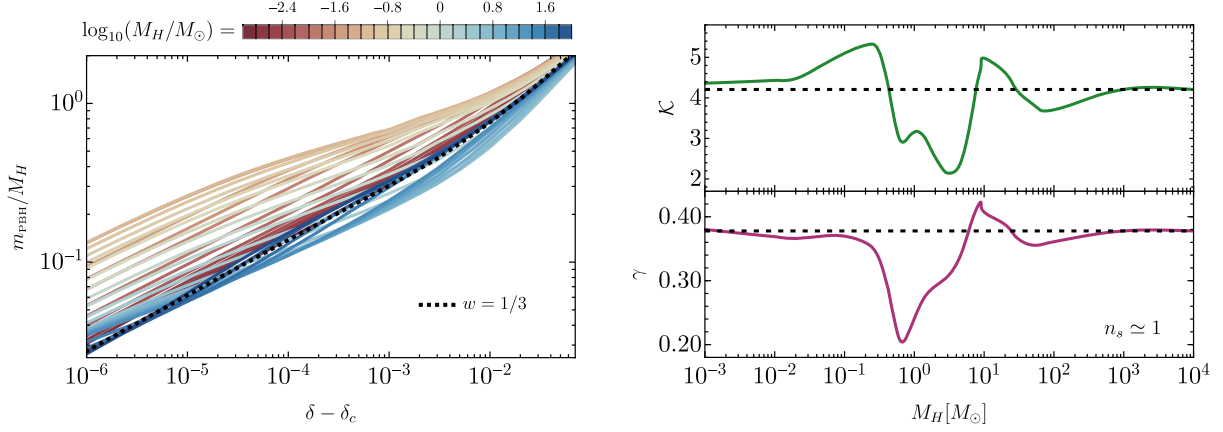
FIG. 2.   Left panel: PBH mass m_{PBH} plotted as a function of $\delta - \delta_c$ computed at the cosmological horizon crossing (see Ref. [16] for more details). The behavior for a radiation dominated medium is plotted with a black dashed line. Right panel: the values of the power-law coefficients in Eq. (9) found by fitting the results of numerical simulations shown in the left panel.

completely the effects of c_s^2 during the dynamics of the collapse, and a correct computation of Φ entering in the definition of δ_m given in (8).

More recently an attempt to improve the calculation, including also the effects of the sound speed, has been investigated in [92], computing an analytic estimation of the threshold based on the three zone model used in [93]. This however has the well known drawback of not being able to include the effects of the pressure gradients during the collapse, corresponding to an underestimation of the threshold which is strongly shape dependent [91], varying with the initial curvature power spectrum of cosmological perturbation [94].

Even looking at the qualitative behavior of [92], one can see a nonmonotonic behavior in δ_c in the two key regions (M_H smaller/larger than $3M_H$) which do not appear in the full numerical results shown in the bottom panel of Fig. 1. This is a clear evidence of the intrinsic limit of making an analytic approximation of a nonlinear collapse process, as in [92] it was also pointed out, which is not able to include properly all the combined nonlinear effects related to the behavior of w and c_s^2 when a cosmological perturbation is collapsing during the QCD phase. In general a proper computation of the threshold δ_c, to be used in precise estimation of the abundance of PBHs, requires necessarily fully relativistic numerical simulations as the ones used here (see [16] for more details).

D. The mass spectrum for PBHs

In Fig. 2 we show the resulting mass spectrum of PBHs obtained from the numerical simulations of [16] obtained after the computation of the threshold, plotting m_{PBH}/M_H against $(\delta - \delta_c)$ during the QCD phase transition. As it is well known, in the standard scenario of a radiation dominated medium a critical collapse arises [95,96] and the mass spectrum of PBHs is characterized by a scaling law [10–14,97] given by

$$m_{\mathrm{PBH}}(\delta) = \mathcal{K} M_H (\delta - \delta_c)^\gamma, \qquad (9)$$

where for $\delta - \delta_c \lesssim 10^{-2}$ the critical exponent γ depends only on the parameter of the equation of state, i.e the value of w, completely independent on the initial configuration of the initial conditions, given by the initial profile of $\zeta(r)$, which affect instead the value of $\mathcal{K}$. This is shown on both plots of Fig. 2 with a dashed line when $w = 1/3$, which gives $\gamma \simeq 0.36$ and $\mathcal{K} \simeq 4$ for a nearly scale invariant curvature power spectrum, as the one considered here.

The QCD phase transition introduces an additional degree of freedom into the problem, which is the characteristic scale of the horizon crossing of the cosmological perturbation. This makes δ_c, γ and $\mathcal{K}$ to depend also on M_H, i.e., when the perturbation is crossing the cosmological horizon. The different lines shown in Fig. 2 with a color varying between red, for smaller values of M_H, and blue for larger values, shows how the scaling law is modified by the characteristic scale of the problem.

For the calculation of the mass distribution (see Sec. II F) it is important to include these effects due to the variation of γ and $\mathcal{K}$ in terms of M_H: an exact power-law critical behavior is only obtained close enough to the density threshold $(\delta - \delta_c \lesssim 10^{-5})$, where the PBH masses are significantly smaller than the cosmological horizon mass, not able to affect significantly the collapse, while for larger values the EOS during the QCD epoch induces further modifications. We fit the relation between the PBH and horizon mass using the power-law template (9) in the range of δ which most contributes to the abundance, i.e., $(\delta - \delta_c) \in [10^{-5}, 2 \times 10^{-2}]$, and find that deviations from the functional form used in Eq. (9) would only induce a small correction which we can neglect.

The resulting values of $\mathcal{K}(M_H)$ and $\gamma(M_H)$ used here are shown in the right plot of Fig. 2: one could appreciate the significant variation of these quantities when δ_c is also significantly varying with respect M_H, compared to the

constant values of the critical collapse during the radiation dominated epoch of the early Universe, indicated here with a black dashed line. A general trend is observed: for $M_H \lesssim 3M_\odot$, there is a tendency to generate heavier PBHs, while the opposite is found when $M_H \gtrsim 3M_\odot$. This can be seen in the left panel of Fig. 2, where orange (light blue) lines fall above (below) the dashed black line indicating the result for a radiation perfect fluid. The fitted values of $\mathcal{K}(M_H)$ and $\gamma(M_H)$ shown in the right panel of Fig. 2 aim to describe with enough accuracy this trend.

E. Curvature power spectrum

Our model is based on a parametrization of the curvature power spectrum, which we assume to have a nearly scale invariant shape of the form

$$\mathcal{P}_\zeta(k) = A\left(\frac{k}{k_{\min}}\right)^{n_s-1}\Theta(k - k_{\min})\Theta(k_{\max} - k), \quad (10)$$

where A defines the characteristic amplitude, n_s is the spectral tilt,[3] and $k_{\min}$ and $k_{\max}$ are the cutoff scales in momentum (k) space. This functional form generally describes broad spectra [98,99], whose consequent PBH mass distribution may be modulated by the QCD epoch. Notice that, due to the exponential dependence of the PBH abundance to the spectral amplitude, even mildly tilted spectra with $n_s \neq 1$ generate *narrow mass distributions* strongly peaked towards small (when blue with $n_s > 1$) or large (when red with $n_s < 1$) masses. See Sec. II F for more details.

In reality, sharp cutoffs in momentum space do not appear in physically motivated curvature power spectra, which are also typically constrained to obey maximum growth or decay rates as a function of the wave number [100–102]. However, due to the exponential dependence of the PBH abundance to the variance of the density contrast, we do not expect such a simplification to affect our result. Indeed, in Sec. VIII we shall show how the features of the parametrization (10) are naturally reproduced in a USR inflationary model.

Even though we restrict our parameter space to nearly scale invariant spectra, the variations of the spectral tilt would require considering potentially different shapes of collapsing overdensities (see e.g., [94]). Capturing this effect on the threshold and the other parameters of collapse would necessitate numerically simulating the PBH formation across the QCD epoch over a fine grid of variations beyond the scale invariant spectrum, which is computationally very demanding, and is left to future extensions of this work.

---

[3]We warn the reader that the spectral tilt n_s defined Eq. (10) specifically refers to the small (PBH) scales. The tilt observed at large (CMB) scales will be referred to as $n_s(k_*)$, where k_* is the CMB pivot scale, see discussion in Sec. VIII. We stress that, as we shall discuss, they are not *a priori* related to each other.

TABLE I. Hyperparameters of the PBH model and their prior ranges adopted in the inference analysis. The mass scales M_S and M_L are intended as expressed in units of the solar mass $M_\odot$.

Model	PBH			
λ	$\log_{10} f_{PBH}$	n_s	$\log_{10} M_S$	$\log_{10} M_L$
Prior	$[-6, 0]$	$[0, 1.5]$	$[-2.5, \log_{10} M_L]$	$[\log_{10} M_S, 4]$

We also assume the absence of primordial non-Gaussianities of the curvature perturbations (see e.g., Refs. [103–106]) while we fully account for the unavoidable intrinsic non-Gaussianities induced by the nonlinear relation between the curvature perturbation and the energy density contrast [107,108].

Notice that, for fixed spectrum shape parameters $[n_s, k_{\min}, k_{\max}]$, the overall PBH abundance f_{PBH} (to be defined later on) is degenerate with the amplitude A. Also, the minimum and maximum scales at which the power spectrum is cut correspond to characteristic horizon mass scales $M_S \equiv M_H(k_{\max})$ and $M_L \equiv M_H(k_{\min})$. In other words, M_S and M_L are respectively the smallest and largest horizon masses bracketing the PBH formation epoch. Therefore, we equivalently choose to adopt the following hyperparameters describing the PBH model as

$$\lambda_{PBH} = [\log_{10} f_{PBH}, n_s, \log_{10} M_S, \log_{10} M_L], \quad (11)$$

where, if not explicitly indicated, the mass scales M_S and M_L are intended as expressed in units of the solar mass $M_\odot$. In Table I, we summarize the choice of priors of the PBH model later adopted in the GWTC-3 Bayesian inference analyses.

F. Computation of the mass distribution

In this section we report the computation of the PBH mass distribution starting from the primordial power spectrum defined in Eq. (10). We shall follow the derivation reported in Ref. [108], to which we refer for more details.

Looking at (8) it has already been observed that this equation can be written in terms of a Gaussian component linearly related to the curvature perturbation $\delta_l \equiv -2\Phi r_m \zeta'(r_m)$ as

$$\delta_m = \left(\delta_l - \frac{1}{4\Phi}\delta_l^2\right). \quad (12)$$

The probability density function of the linear component of the smoothed energy density contrast δ_l is Gaussian, and thus can be written as

$$P(\delta_l) = \frac{1}{\sqrt{2\pi\sigma_0^2}}\exp\left(-\frac{\delta_l^2}{2\sigma_0^2}\right). \quad (13)$$

The variance σ_0^2 and the first momentum of the distribution σ_1^2 are

$$\sigma_i^2(r_m) = \frac{4}{9}\Phi^2 \int_0^\infty \frac{dk}{k}(kr_m)^4 \tilde{W}^2(k, r_m)T^2(k, r_m)k^{2i}\mathcal{P}_\zeta(k),$$

(14)

where $i = 0, 1$; $\tilde{W}(k, r_m)$ is the Fourier transform of the top-hat smoothing function,

$$\tilde{W}(k, r_m) = 3\left[\frac{\sin(kr_m) - kr_m\cos(kr_m)}{(kr_m)^3}\right], \quad (15)$$

and $T(k, r_m)$ is the linear transfer function

$$T(k, r_m) = 3\left[\frac{\sin(kr_m/\sqrt{3}) - kr_m\cos(kr_m/\sqrt{3})/\sqrt{3}}{(kr_m/\sqrt{3})^3}\right].$$

(16)

In the following, we are going to identify the smoothing scale r_m with the corresponding horizon mass M_H (fixed by the horizon crossing condition $aHr_m = 1$ [91]) using the relation with power spectral modes,

$$r_m k \equiv \kappa = 4.49, \quad (17)$$

found for a broad (and nearly scale invariant) power spectrum [94]. This relation is strictly valid for a shape parameter $\alpha = 3$ [94], consistently with the approximations described above.

By consequence, one finds that the horizon mass M_H is related to power spectral modes through

$$M_H \simeq 17 M_\odot \left(\frac{g_*}{10.75}\right)^{-1/6}\left(\frac{k/\kappa}{\text{pc}^{-1}}\right)^{-2}, \quad (18)$$

where g_* is the number of degrees of freedom of relativistic particles. We reiterate here for clarity that Eq. (18) relates the horizon mass M_H to the epoch of horizon crossing of the peak of the compaction function (of size r_m) produced by the single mode k. This differs from the horizon mass corresponding to the crossing time of modes k themselves and we point the attention of the reader to the relating coefficient κ that has been frequently (but incorrectly) omitted in the past.

In principle, the transfer function defined in Eq. (16) is derived using linear perturbation theory in a radiation dominated Universe ($w = 1/3$). While $T(k, r_m)$ is modified by varying the EOS, and this would lead to a modified evolution of subhorizon modes, the presence of a window function already efficiently smooths curvature perturbations with $kr_m \gg 1$ and the impact of a softer EOS should be small. As discussed in Sec. II G, we will capture both $T(k, r_m)$ and $\tilde{W}(k, r_m)$ with an effective smoothing

function, neglecting further modifications of $T(k, r_m)$ from a time-dependent w around the QCD epoch.

The threshold for PBH formation can be translated into a critical amplitude of the linear component $\delta_{c,l\pm}$ by inverting Eq. (12) as

$$\delta_{c,l\pm} = 2\Phi\left(1 \pm \sqrt{1 - \frac{\delta_c}{\Phi}}\right). \quad (19)$$

In the computation of the mass distribution we only include values of δ_l falling in the range

$$\delta_{c,l-} < \delta_l < 2\Phi \equiv \delta_l^+, \quad (20)$$

corresponding to type-I PBH formation [91], and neglect the contribution from PBHs formed in the second branch whose contribution is exponentially suppressed.

The number density of sufficiently high peaks can be computed adopting the theory of random Gaussian fields [109], which gives

$$\mathcal{N} = \frac{\sigma_1^3}{4\pi^2\sigma_0^3}\nu^3\exp\left(-\frac{\nu^2}{2}\right), \quad (21)$$

where we introduced the rescaled peak height $\nu \equiv \delta_l/\sigma_0$. The mass fraction for each peak of given height ν which collapses to form a PBH can be expressed by evaluating

$$\beta_\nu = \frac{m_{\text{PBH}}(\nu)}{M_H}\mathcal{N}(\nu)\theta(\nu - \nu_c), \quad (22)$$

where the Heaviside step function θ implements the threshold for collapse.

The total energy fraction of the Universe composed by PBHs formed at a given time (equivalently identified with a single horizon mass M_H) is given by integrating the relevant range of ν between $\nu_{c-} \equiv \delta_{c,l-}/\sigma_0$ and $\nu_+ = \delta_l^+/\sigma_0$ [using Eqs. (19) and (20)], which can be written as

$$\beta(M_H) = \int_{\nu_{c-}}^{\nu_+} d\nu \frac{\mathcal{K}}{3\pi}\left(\nu\sigma_0 - \frac{1}{4\Phi}(\nu\sigma_0)^2 - \delta_c\right)^\gamma$$
$$\times \left(\frac{\sigma_1}{aH\sigma_0}\right)^3 \nu^3\exp\left(-\frac{\nu^2}{2}\right). \quad (23)$$

The term $1/aH$ is fixed by the horizon crossing condition $aH = 1/r_m = k_H/\kappa$ [91]. Finally, the entire energy fraction composed by PBHs after formation is found by integrating over all relevant epochs (corresponding to the time span when modes within $k_{\text{min}} < k < k_{\text{max}}$ cross the Hubble horizon) as

$$\Omega_{\text{PBH}} = \int_{M_s}^{M_L} d\ln M_H \left(\frac{M_{\text{eq}}}{M_H}\right)^{1/2}\beta(M_H), \quad (24)$$

where $M_{eq} = 2.8 \times 10^{17} M_{\odot}$ is the horizon mass at the time of matter-radiation equality [110]. The corresponding total PBH abundance is then simply defined as

$$f_{PBH} \equiv \frac{\Omega_{PBH}}{\Omega_{DM}}, \tag{25}$$

where $\Omega_{DM} = 0.265$.

The mass function $\psi(m_{PBH})$ is defined as the fraction of PBHs with mass in the infinitesimal interval $(m_{PBH}, m_{PBH} + dm_{PBH})$. This can be obtained by differentiating Ω_{PBH} with respect to the PBH mass as

$$\psi(m_{PBH}) = \frac{1}{\Omega_{PBH}} \frac{d\Omega_{PBH}}{dm_{PBH}}. \tag{26}$$

Our definition of the mass distribution implies unit normalization under integration as

$$\int dm_{PBH} \psi(m_{PBH}) = 1, \tag{27}$$

so that $\psi(m_{PBH})$ has the dimensions of [1/mass]. Notice that an alternative definition of the mass distribution may be given in terms of logarithmic mass intervals. This is found by computing

$$f(m_{PBH}) \equiv \frac{1}{\Omega_{DM}} \frac{d\Omega_{PBH}}{d\ln m_{PBH}} = m_{PBH} f_{PBH} \psi(m_{PBH}), \tag{28}$$

yielding a dimensionless function. This alternative quantity will be useful when comparing the mass distribution resulting from our analysis with PBH constraints [21], see Sec. VII.

In order to compute the full mass distribution, it is convenient to invert the relation between horizon and PBH mass through the critical collapse relation (9), focusing only on the type-I branch, as

$$\delta_l = 2\Phi(1 - \sqrt{\Lambda}), \tag{29}$$

where

$$\Lambda = 1 - \frac{\delta_c}{\Phi} - \frac{1}{\Phi} \left(\frac{m_{PBH}}{\mathcal{K} M_H} \right)^{1/\gamma}. \tag{30}$$

At this point, using Eq. (29), we can change the variable of integration in Eq. (23) and write

$$\psi(m_{PBH}) = \frac{8}{3\pi \Omega_{PBH} m_{PBH}} \int_{M_s}^{M_L} \frac{dM_H}{M_H} \left(\frac{M_{eq}}{M_H} \right)^{1/2} \left(\frac{\sigma_1}{aH\sigma_0} \right)^3 \frac{\Phi^3 \mathcal{K}}{\gamma \sigma_0^4} \left(\frac{m_{PBH}}{\mathcal{K} M_H} \right)^{\frac{1+\gamma}{\gamma}} \frac{(1 - \sqrt{\Lambda})^3}{\Lambda^{1/2}} \exp\left[-\frac{2\Phi^2}{\sigma_0^2} (1 - \sqrt{\Lambda})^2 \right], \tag{31}$$

and the integration range of M_H is subject to the condition $\Lambda > 0$ (because we require $\delta > \delta_c$). The quantities $\mathcal{K}(M_H)$, $\gamma(M_H)$, $\Phi(M_H)$, $\delta_c(M_H)$, and $\sigma_i(M_H)$ are left within the integration over the horizon mass scale, as they all explicitly depend on M_H when thermal effects are included. In the low mass limit, i.e., $m_{PBH} \ll M_s$, one can find that the mass distribution (31) scales as

$$\psi(m_{PBH}) \propto \left(\frac{m_{PBH}}{\mathcal{K} M_s} \right)^{1/\gamma}, \tag{32}$$

which gives the characteristic tail $\psi(m_{PBH}) \propto m_{PBH}^{2.8}$ if one assumes the energy density of the Universe behaving as a relativistic fluid with $w = 1/3$, which gives $\gamma \approx 0.36$ [10].

In Fig. 3 we show the mass distribution generated by the collapse of a single mode k_H. Depending on the exact moment of the cosmological horizon crossing, which fixes exactly at which epoch across the QCD era the collapse takes place, the consequent mass distribution deviates from the one obtained when the Universe is radiation dominated. In particular, we observe differences in the low mass tail and in the location of the peak of $\psi(m_{PBH})$. Modes collapsing before (after) $M_H \approx M_{\odot}$ tend to generate a mass distribution peaked at larger (smaller) values compared to
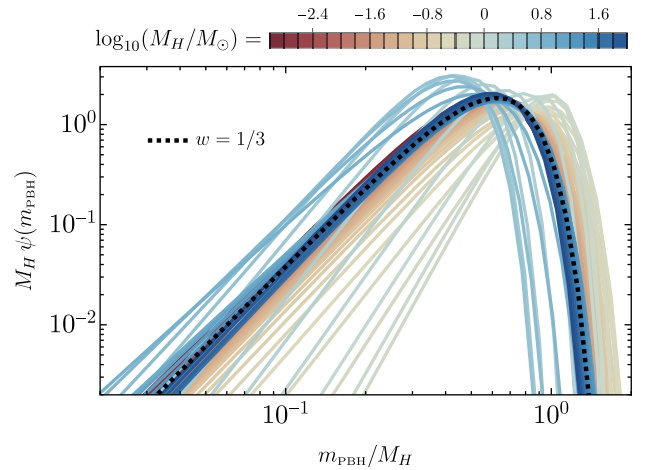

FIG. 3. Mass distribution resulting from the collapse of a single spectral scale k_H crossing the horizon at various M_H. The low mass tail is dictated by the critical collapse scaling γ, see Eq. (32). The dashed black line denotes the result assuming a Universe with the radiation EOS $w = 1/3$. In that case, the peak of the mass function for given M_H contribution would sit at $m_{PBH}/M_H = 0.602$.

the reference result of a radiation dominated medium. This generates a pile-up effect around the solar mass, which additionally contributes to enhance the QCD peak induced by the reduced value of the threshold around $M_H \approx M_\odot$.

To summarize, with $\mathcal{K}(M_H)$, $\gamma(M_H)$, $\Phi(M_H)$, and $\delta_c(M_H)$ computed from the simulations in Ref. [16] and shown in Figs. 1 and 2, alongside $\sigma_i(M_H)$ from Eq. (14), the algorithm presented above can be applied to compute the PBH mass function and the corresponding total abundance in terms of the parameters of the cosmological power spectrum.

In Fig. 4, we show a few representative examples of such a mass distribution, obtained by fixing the hyperparameters λ_{PBH} of the curvature power spectrum. In particular, we focus the attention on the role of the tilt n_s. In case $n_s = 1$, the spectrum is sizable at modes collapsing during the QCD epoch and a bump around the solar mass is obtained [15], on top of what is expected from a scale invariant spectrum $\psi(m_{\mathrm{PBH}}) \approx m_{\mathrm{PBH}}^{-3/2}$ [99]. On the other hand, already for slightly red spectra (with $n_s < 1$), the mass distribution becomes independent of the UV spectrum cutoff k_{max} (i.e., of M_S) and increasingly tilted towards larger masses, up to the point where the QCD enhancement becomes irrelevant, due to the slightly reduced power at the QCD scale, compensating the reduced value of threshold δ_c with respect $w = 1/3$. For intermediate values of n_s (e.g., $n_s \approx 0.7$), a doubly peaked mass distribution can be realized, where the location of the light peak is fixed by the QCD epoch and the heavy one is instead controlled by M_L.
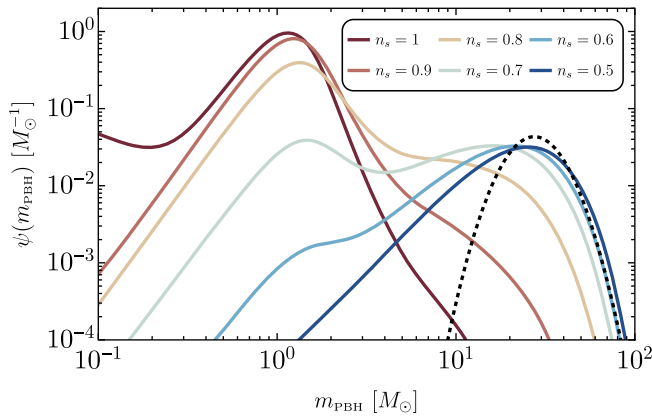

FIG. 4. The mass function obtained with a few choices of the curvature power spectrum compatible with the posterior distribution inferred by the analysis presented in Sec. VI (see Table V). This plot assumes $f_{\mathrm{PBH}} = 10^{-3}$, the minimum horizon mass to be smaller than $M_S \lesssim 10^{-2.5} M_\odot$, the largest mass in the spectrum $M_L = 10^{2.8} M_\odot$ and a variable tilt n_s. (Only for an exactly scale invariant spectrum $n_s = 1$, does the mass function depend on M_S, in which case we fix $M_S = 10^{-2.5} M_\odot$.) The black dashed line reports the log-normal mass distribution found as the best fit in the analysis of Ref. [45]. Overall, the *ab initio* distribution shaped by the QCD phase transition has larger support for PBHs with $m_{\mathrm{PBH}} \lesssim 10 M_\odot$ compared to the log-normal parametrization.

We can compare these examples with the best-fit log-normal mass distribution obtained in the analysis of Ref. [45] (black dashed line in Fig. 4). As one can see, red tilted spectra may produce similar mass distributions peaked at around $m_{\mathrm{PBH}} \approx 30 M_\odot$, for which the QCD softening of the EOS plays no role. The critical collapse, however, generates an asymmetry in the mass distributions that can only be taken into account by introducing additional parameters controlling the skewness of the distribution, as pointed out in Ref. [111].

G. The semianalytical mass distribution

The computation of the integral (31), which should be performed on a sufficiently dense grid of values of m_{PBH} for each choice of the PBH hyperparameters λ_{PBH}, may be rather time consuming, because it requires computing numerically the integrals (14) at each M_H.

In order to simplify the description of the PBH abundance and speed up the hierarchical Bayesian analysis, we absorb the effect of both the window function and linear transfer function, which are cutting subhorizon curvature modes, in a single Gaussian window function of the form

$$\hat{W}(k, \hat{R}) = \exp\left[-(k\hat{R})^2/4\right], \qquad (33)$$

where the smoothing scale $\hat{R}$ is fitted appropriately. In particular, $\hat{R}$ have been adjusted to match the average smoothing between r_m and $r_m/\sqrt{3}$ through the factor $\hat{R} = s r_m$ with

$$s = \frac{(1 + 1/\sqrt{3})}{2} \simeq 0.78. \qquad (34)$$

We checked that this approximation, solely intended to speed up the computation of the mass distribution when running the Monte Carlo Markov chain (MCMC) analysis, does not introduce any appreciable modification to the mass distribution.

Within this simplifying assumption, one can solve Eq. (14) analytically,

$$\sigma_0^2 = \frac{4}{9} \Phi^2 2^{(n_s+1)/2} A s^{-4} (k_{\mathrm{min}} r_m s)^{1-n_s}$$
$$\times \left[\Gamma\left(\frac{n_s+3}{2}, \frac{(k_{\mathrm{min}} r_m s)^2}{2}\right) - \Gamma\left(\frac{n_s+3}{2}, \frac{(k_{\mathrm{max}} r_m s)^2}{2}\right) \right], \qquad (35)$$

while $\sigma_1 = k_{\mathrm{min}} \sigma_0 (n_s \rightarrow n_s + 2)$, and $\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} \mathrm{d}t$ is the incomplete Gamma function. The variance can be expressed in terms of the model hyperparameters by setting
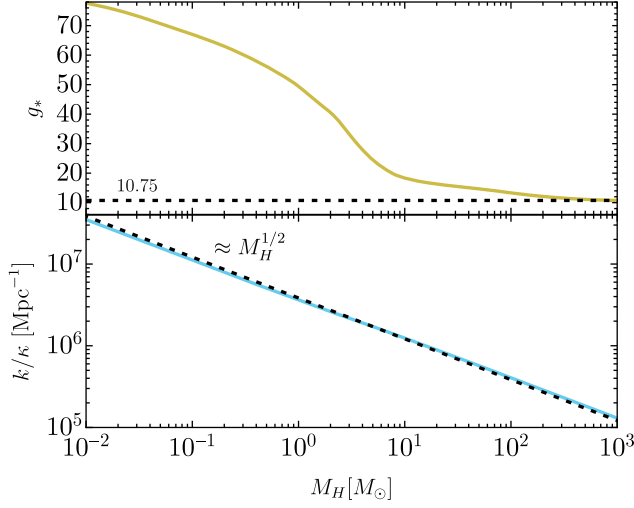
FIG. 5. The top plot shows how the effective degrees of freedom g_* varies as a function of M_H. The bottom plot shows the impact g_* has on the relation between spectral modes and the horizon mass at horizon crossing time.

$$k_{\min} r_m = \kappa \sqrt{M_H/M_{\rm L}},$$
$$k_{\max} r_m = \kappa \sqrt{M_H/M_{\rm S}}, \qquad (36)$$

while the amplitude A is fixed with a bisection method to select the desired PBH abundance $f_{\rm PBH}$ (within subpercent accuracy on the latter).

In Eq. (36) we implicitly fixed the number of the effective degrees of freedom g_* appearing in Eq. (18), and in Fig. 5 (top panel) we show how g_* varies as a function of the temperature of the Universe, which is tracked by M_H in our formalism. This induces a small deviation from the scaling reported in Eq. (36) (see bottom panel of Fig. 5), which we neglect to make the variance fully analytical with the aim of speeding up the computations. Therefore, in the following we will fix $g_* = 25$, i.e., the value at the central region of our interest. By fixing g_*, we neglect a small running of n_s effectively induced by the change of degrees of freedom when computing the mass distribution.

One final simplification one may attempt is to neglect the critical collapse and remove the integration over the horizon mass scale M_H. This, however, cannot be done consistently as the width of the QCD modulation around the solar mass is *narrower* than the one induced by the critical collapse. As it can be seen in Fig. 4, the critical mass distribution has a crucial role in shaping the peak of the mass distribution around the solar mass. This can be deduced by realizing that the mass distributions obtained with different n_s have the same scaling $\approx m_{\rm PBH}^{1/\gamma}$ below the QCD peak, induced by the critical collapse of the mode k corresponding to the minimum of the threshold δ_c, while they are basically insensitive to the variance σ_0 at $M_H < M_\odot$).

III. THE PBH MERGER RATE

The standard PBH formation mechanism we consider assumes PBHs are generated from the collapse of sizable Gaussian cosmological perturbations in the radiation dominated epoch of the early Universe [5–8]. In this scenario, PBHs are predicted to be characterized by small natal spins [112,113], and are not clustered at high redshift [98,114–118]. Furthermore, the PBH merger rate at low redshift is dominated by binaries that gravitationally decouple from the Hubble flow before the matter-radiation equality [119,120]. We compute the differential volumetric PBH merger rate density following Refs. [31,34,37,118] as

$$\frac{d\mathcal{R}_{\rm PBH}}{dm_1 dm_2} = \frac{1.6 \times 10^6}{{\rm Gpc}^3 \, {\rm yr}} f_{\rm PBH}^{\frac{53}{37}} \left(\frac{t(z)}{t_0}\right)^{-\frac{34}{37}} \eta^{-\frac{34}{37}} \left(\frac{M}{M_\odot}\right)^{-\frac{32}{37}}$$
$$\times S(M, f_{\rm PBH}, \psi, z)\psi(m_1)\psi(m_2), \qquad (37)$$

where $M = m_1 + m_2$, $\eta = m_1 m_2/M^2$, and t_0 is the current age of the Universe.

The suppression factor $S < 1$ accounts for environmental effects in both the early- and late-time Universe. We can separately define each contribution as

$$S \equiv S_{\rm early}(M, f_{\rm PBH}, \psi) \times S_{\rm late}(f_{\rm PBH}, z). \qquad (38)$$

An analytic expression for S can be found in Ref. [40], which we report here for completeness. In the early Universe, suppression results as a consequence of inter-actions between PBH binaries and both the surrounding dark matter inhomogeneities, as well as neighboring PBHs at high redshift [27,29,31,121]. This factor takes the form[4]

$$S_{\rm early} \approx 1.42 \left[\frac{\langle m^2 \rangle/\langle m \rangle^2}{\bar{N}(y) + C} + \frac{\sigma_{\rm M}^2}{f_{\rm PBH}^2}\right]^{-21/74} \exp\left[-\bar{N}(y)\right],$$
$$(39)$$

with

$$\bar{N}(y) \equiv \frac{M}{\langle m \rangle}\left(\frac{f_{\rm PBH}}{f_{\rm PBH} + \sigma_{\rm M}}\right), \qquad (40)$$

and the rescaled variance of matter density perturbations takes the value $\sigma_{\rm M} \simeq 0.004$. In Eq. (39), the constant factor C is defined as [see Eq. (A.5) of Ref. [40]]

---

[4]The suppression factor in Eq. (39) was tested against N-body simulations in Ref. [31], also assuming a wide (but log-normal) mass distribution. While, in this work, we adopt a different mass distribution, derived from first principles, its width in the stellar mass range is compatible with the one tested in Ref. [31], supporting our adoption of Eq. (39).

$$C = \frac{f_{\mathrm{PBH}}^2}{\sigma_{\mathrm{M}}^2} \frac{\langle m^2 \rangle}{\langle m \rangle^2} \left\{ \left[\frac{\Gamma(29/37)}{\sqrt{\pi}} U\left(\frac{21}{74}, \frac{1}{2}, \frac{5 f_{\mathrm{PBH}}^2}{6 \sigma_{\mathrm{M}}^2} \right) \right]^{-\frac{74}{21}} - 1 \right\}^{-1},$$

(41)

where $\Gamma(x)$ is the Euler Gamma function and $U(a, b, z)$ denotes the confluent hypergeometric function. We warn the reader that we are adopting a different notation for the mass distribution compared to the one used in Ref. [40], which here is normalized such that $\int \mathrm{d}m \psi(m) = 1$. With this choice, the mass average reads

$$\langle m^n \rangle = \int m^n \psi(m) \mathrm{d}m.$$

(42)

In the late Universe, multiple encounters with other PBHs that populate small clusters formed from the initial Poisson conditions lead to a thermalization of the eccentricity distribution, which enhances the merger time and effectively reduces the late-time universe merger rate [65,122–125]. By accounting for the fraction of binaries which avoids dense enough clusters and are not disrupted, one can write down this additional suppression factor as [34,40,118,126]

$$S_{\mathrm{late}}(x) \approx \min\left[1, 9.6 \times 10^{-3} x^{-0.65} \exp\left(0.03 \ln^2 x \right) \right],$$

(43)

where we introduced the variable $x \equiv (t(z)/t_0)^{0.44} f_{\mathrm{PBH}}$. Notice also that, for $f_{\mathrm{PBH}} \lesssim 0.003$, one always finds $S_{\mathrm{late}} \simeq 1$, i.e., the suppression of the merger rate due to disruption inside PBH clusters is negligible. This is also supported by the results obtained through cosmological N-body simulations finding that PBHs are essentially isolated when their abundance is small enough [117].

It is important to mention that the late-time suppression factor was only computed for a sufficiently narrow mass distribution [34,40,118]. So far, a full computation considering wide distributions was not performed in the literature. However, we do not expect this extension to modify significantly the formulation used here as S_{late} is found to be only mildly dependent on the mass scale (see e.g., [127]). Furthermore, we generically expect $S_{\mathrm{late}} \approx 1$ for the PBH abundance inferred *a posteriori* by our analysis. In the computation of the merger rate, we are also neglecting the contribution from binaries that can form dynamically within PBH clusters from either capture or three-body interactions. This is justified because, in this scenario and for the small values of f_{PBH} we obtain, the contribution of those channels to the total merger rate is subdominant relative to the early universe binaries [128].

Finally, the natal distribution of PBH masses and spins (the latter being initially negligible [112,113], see also [129]) can be modified if PBHs undergo an efficient accretion phase during cosmic history [37,130,131]. For a given accretion model, the peculiar accretion-driven and redshift-dependent mass-spin distribution can be used to add extra information in the inference [17,45,132] and also impact the merger rate [37,130,131]. However, while certain features of PBH accretion are robust and should be model independent, there remain large uncertainties in the mass (and, especially, spin) accretion. Thus, in order to remain agnostic and conservative, here we neglect PBH accretion and do not include spin information in the merger events. In practice, in the inference we shall only use the dependence of the merger rate on the individual masses and redshift, conservatively limiting the information that can be inferred from single merger events [17,132].

Let us conclude by stressing that, while in this work we compute the mass distribution from first principles across the QCD era as described in Sec. II, thus going beyond the parametrization often used in the literature, certain characteristics of the PBH model are general and arise from the form of the PBH merger rate in Eq. (37). These features are the monotonic merger rate evolution with redshift, $\mathcal{R}_{\mathrm{PBH}} \approx (t/t_0)^{-34/37}$, a scaling of the merger rate with the PBH abundance,

$$\mathcal{R}_{\mathrm{PBH}} \propto \begin{cases} f_{\mathrm{PBH}}^{2/3} & \text{for } f_{\mathrm{PBH}} \gtrsim 10^{-3}, \\ f_{\mathrm{PBH}}^2 & \text{for } f_{\mathrm{PBH}} \lesssim 10^{-3}, \end{cases}$$

(44)

a lack of preference towards symmetric mass ratios enforced by the term $\eta^{-34/47}$, and an exponential suppression of heavy (i.e., $M \gg \langle m \rangle$) mergers due to the suppression factor (39).

IV. LVK PHENOMENOLOGICAL MODELS

In the following section, we will compare and mix our *ab initio* PBH channel with phenomenological models used by the LVK collaboration to fit the BH and NS binary events in the GWTC-3 catalog. This approach is very conservative, because we choose to confront the PBH scenario with the best working model specifically tailored to describe the coarse-grained properties of the observed merger population. As such, this approach is not meant to be used to *search for* a subpopulation of PBHs in the data, but rather to place an *upper bound* on the PBH abundance compatible with the data and to assess whether certain events are more likely ascribed to a putative primordial channel. As we will discuss in the Conclusions, one natural extension of this analysis would entail considering *ab initio* astrophysical models, as attempted in Ref. [45].

According to the LVK prescription, compact objects with masses below $3 M_\odot$ are labeled as NSs, whereas heavier objects are labeled as BHs. Two different mass distributions are used to describe mergers of these families, as discussed below. However, at variance with the LVK analysis, we shall adopt a more agnostic approach and allow for the light events to be BH binaries (of primordial origin), with the exception of GW170817 [133] for which sufficient evidence for the interpretation as a NS binary was

gathered with the observation of an electromagnetic counterpart compatible with a NS merger [134].

In the standard scenario, PBH mergers at low redshift are due to binaries that had gravitationally decoupled from the Hubble flow before the matter-radiation equality [119,120], i.e., much before the first stars were born. Thus, "mixed" binaries formed by an isolated PBH and either an astrophysical-origin BH or a NS can be assembled only through dynamical capture, e.g., in dense clusters. The probability of forming these binaries is very low [41,135] and we shall neglect such possibility. In other words, we shall assume that all primordial binaries are formed by two PBHs and that all astrophysical-origin binaries are formed by astrophysical BHs and/or NSs.

A. Astrophysical BH binaries

We describe the merger rate of astrophysical BH binaries with the reference population model called POWER LAW + PEAK [136] adopted by the recent LVK population analyses (see e.g., Ref. [59]). Henceforth we shall refer to this as the "astrophysical" BH (ABH) population, although it should be kept in mind that the model is phenomenological and not based on *ab initio* astrophysical simulations. The ABH model assumes that the distribution of primary binary BH mass m_1 is described by a mixture of a power-law model,

$$P_{\mathrm{ABH}}(m_1|\lambda, m_{\min}, m_{\max}) \propto m_1^{-\alpha} \tag{45}$$

and a Gaussian peak,

$$N_{\mathrm{ABH}}(m_1|\mu_m, \sigma_m, m_{\min}, m_{\max}) \propto \exp\left[-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2}\right], \tag{46}$$

normalized to unity across the range $m_{\min} \le m_1 \le m_{\max}$. The mixing fraction between the two components is dictated by λ_{peak} as

$$p_{\mathrm{ABH}}^{m_1}(m_1) = (1 - \lambda_{\mathrm{peak}})P_{\mathrm{ABH}}(m_1) + \lambda_{\mathrm{peak}}N_{\mathrm{ABH}}(m_1). \tag{47}$$

We describe the distribution of mass ratio via a power law as

$$p_{\mathrm{ABH}}^{m_2}(q|m_1, \beta) \propto q^\beta, \tag{48}$$

constrained within the range $m_{\min}/m_1 \le q \le 1$. For simplicity, we do not introduce the term $S_{\mathrm{ABH}}(m|\delta_m)$ adopted in LVK analyses to smooth the sharp cutoff below $m_{\min}$. Finally, the evolution of the merger rate at high redshift follows

$$p_{\mathrm{ABH}}^z(z|\kappa) \propto (1 + z)^\kappa. \tag{49}$$

The number of events can be found by integrating the merger rate density with the additional factor of $dV_c/dz/(1 + z)$. Since observations are limited to small redshift, the merger rate evolution is still rather poorly constrained. In order to simplify the analysis, we fix the power-law evolution of the astrophysical phenomenological model to the best-fit value of $\kappa = 2.9$ [59]. To summarize, we write the differential merger rate density of ABH as

$$\frac{\mathrm{d}\mathcal{R}_{\mathrm{ABH}}}{\mathrm{d}m_1 \mathrm{d}m_2} = \mathcal{R}_{\mathrm{ABH}}^0 p_{\mathrm{ABH}}^z(z)p_{\mathrm{ABH}}^{m_1}(m_1)p_{\mathrm{ABH}}^{m_2}(m_2|m_1), \tag{50}$$

and the hyperparameters of the ABH model are

$$\lambda_{\mathrm{ABH}} = [\log_{10}\mathcal{R}_{\mathrm{ABH}}^0, \lambda_{\mathrm{peak}}, \alpha, \beta, m_{\min}, m_{\max}, \mu_m, \sigma_m], \tag{51}$$

where we introduced the quantity $\mathcal{R}_{\mathrm{ABH}}^0 \equiv \mathcal{R}_{\mathrm{ABH}}(z = 0)$ controlling the present-day ABH merger rate density. In Table II we report the prior ranges for the ABH model parameters adopted in the following Bayesian analysis.

While evidence of additional features on top of the POWER LAW + PEAK coarse-grained model was found by the LVK collaboration [59] (see also Refs. [132,137–142]) we do not expect our results—especially the upper bound on f_{PBH}—to be affected by potential systematic effects in our choice of benchmark mass model.

B. Binaries involving NSs

Following the LVK population analysis, we model the distribution of NSs as an underlying Gaussian mass distribution that is common to all NSs, with random pairing into compact binaries. For mixed NSBH mergers, the BH mass distribution is fixed to be uniform between $[3–60]M_\odot$. The joint mass distribution takes the form

$$p_{\mathrm{NS}}(m_1, m_2) = \begin{cases} N_{\mathrm{NS}}(m_1)N_{\mathrm{NS}}(m_2), \\ U(m_1, [3M_\odot, 60M_\odot])N_{\mathrm{NS}}(m_2), \end{cases} \tag{52}$$

TABLE II. Population hyperparameters λ for the ABH and NS models considered in this work, along with their prior distributions. We refer to a uniform distribution between two values $\theta_{\min}$ and $\theta_{\max}$ as $[\theta_{\min}, \theta_{\max}]$. Rates ($\mathcal{R}^0$) are reported in units of [yr^{-1} Gpc^{-3}] while (m, μ, σ) are written in units of $[M_\odot]$.

Model	ABH								NS			
λ	$\log_{10}\mathcal{R}_{\mathrm{ABH}}^0$	α	β	$m_{\min}$	$m_{\max}$	λ_{peak}	μ_m	σ_m	$\log_{10}\mathcal{R}_{\mathrm{NS}}^0$	$m_{\min}^{\mathrm{NS}}$	$m_{\max}^{\mathrm{NS}}$	μ_m^{NS}	σ_m^{NS}
Prior	[−3, 3]	[0,5]	[0,7]	[3,10]	[30,100]	[0,1]	[20,50]	[1,10]	[−1, 5]	[1,1.5]	[1.5,3]	[1,3]	[0.01,2]

for NS and mixed NSBH binaries, respectively, where the Gaussian peak is defined as

$$N_{\mathrm{NS}}(m|\mu_m^{\mathrm{NS}}, \sigma_m^{\mathrm{NS}}, m_{\min}^{\mathrm{NS}}, m_{\max}^{\mathrm{NS}}) \propto \exp\left[-\frac{(m-\mu_m^{\mathrm{NS}})^2}{2(\sigma_m^{\mathrm{NS}})^2}\right], \quad (53)$$

normalized to unity across the range $m_{\min}^{\mathrm{NS}} \leq m \leq m_{\max}^{\mathrm{NS}}$. We assume the redshift evolution of the merger rate for this channel follows the same behavior of the ABH model, namely $\mathcal{R}_{\mathrm{NS}}(z) \approx \mathcal{R}_{\mathrm{NS}}^0 (1+z)^{2.9}$. This evolution is, however, practically irrelevant, as light mergers are currently observable only at $z \approx 0$. Finally, we can write the differential merger rate as

$$\frac{\mathrm{d}\mathcal{R}_{\mathrm{NS}}}{\mathrm{d}m_1 \mathrm{d}m_2} = \mathcal{R}_{\mathrm{NS}}^0 p_{\mathrm{NS}}^z(z) p_{\mathrm{NS}}(m_1, m_2), \quad (54)$$

and the hyperparameters of the NS model are (see also Table II)

$$\lambda_{\mathrm{NS}} = [\log_{10}\mathcal{R}_{\mathrm{NS}}^0, m_{\min}^{\mathrm{NS}}, m_{\max}^{\mathrm{NS}}, \mu_m^{\mathrm{NS}}, \sigma_m^{\mathrm{NS}}]. \quad (55)$$

One may also consider splitting the rate of BNS and NSBH binaries, thus introducing an additional parameter in Eq. (55). However, due to the small number of detections with at least one component lighter than $3M_\odot$, merger rate densities remain affected by large uncertainties in the light sector of the catalog [59], and both contributions are broadly compatible with each other.

V. ANALYSIS SETUP

In this section we summarize the statistical framework we use to perform the analysis and model comparison (see e.g., [143,144]), alongside our event selection within the GWTC-3 dataset [24].

A. Hierarchical Bayesian inference

The aim of the hierarchical Bayesian inference is to produce posterior distributions for the hyperparameters of a model $\mathcal{M}$ which is assumed to explain the GW dataset, alongside the corresponding evidence $Z_\mathcal{M}$ allowing for statistical model comparisons. The LVK collaboration's Gravitational Wave Open Science Center [145,146] releases the output of the parameter estimation performed on each GW signal as a collection of posterior distributions for the parameters describing the properties of each individual merger. We denote this as *event posteriors* $p(\boldsymbol{\theta}|d_i)$, where $\boldsymbol{\theta}$ indicates the binary *event parameters*. The index i runs over all the detected GW events while, in our analysis, we restrict the set of intrinsic binary parameters to $\boldsymbol{\theta} = (m_1, m_2, z)$.

We compute the number of GW events produced in a given model within the observation time as

$$N(\lambda) \equiv \int \mathrm{d}\boldsymbol{\theta} N_{\mathrm{pop}}(\boldsymbol{\theta}|\lambda) = T_{\mathrm{obs}} R(\lambda) \int \mathrm{d}\boldsymbol{\theta} p_{\mathrm{pop}}(\boldsymbol{\theta}|\lambda), \quad (56)$$

where $R(\lambda)$ is the intrinsic merger rate, $p_{\mathrm{pop}}(\boldsymbol{\theta}|\lambda)$ is the *population likelihood*, corresponding to the distribution of event parameters for the model $\mathcal{M}$ characterized by hyperparameters λ, and T_{obs} is the duration of the various LVK observing runs.

One can account for the selection effects induced by the finite sensitivity of the detectors by introducing the *observable* number of events,

$$N_{\mathrm{det}}(\lambda) \equiv \alpha(\lambda) N(\lambda), \quad (57)$$

where the selection bias parametrized by $\alpha(\lambda) \leq 1$ will be discussed in the next subsection.

Given a vector of hyperparameters λ (or *population parameters*) describing the model $\mathcal{M}$, the posterior distribution inferred from the data is

$$\frac{p(\lambda|\boldsymbol{d})}{\pi(\lambda)} \propto e^{-N_{\mathrm{det}}(\lambda)} N(\lambda)^{N_{\mathrm{obs}}} \prod_{i=1}^{N_{\mathrm{obs}}} \int \mathrm{d}\boldsymbol{\theta}_i \frac{p(\boldsymbol{\theta}_i|\boldsymbol{d}) p_{\mathrm{pop}}(\boldsymbol{\theta}_i|\lambda)}{\pi(\boldsymbol{\theta}_i)}, \quad (58)$$

where the prefactor introduces the standard terms describing the statistics of an inhomogeneous Poisson process (see e.g., Refs. [143,147–149] for detailed derivations), $\pi(\lambda)$ is the prior distribution assumed for the model hyperparameters, and $\pi(\boldsymbol{\theta}_i)$ is the prior distribution over the intrinsic parameters adopted by the LVK collaboration when performing the parameter estimation for each individual event. The factor $\pi(\boldsymbol{\theta}_i)$ in the denominator removes the dependence of the analysis on the priors adopted by the LVK collaboration to perform parameter estimation, which was shown to potentially affect the interpretation of individual events [150,151] (see also Ref. [152] for this analysis with PBH informed priors).

In order to speed up the evaluation of Eq. (58), the integral is performed using importance sampling, i.e., by computing the expectation value of the prior-reweighted population likelihood as a discrete sum over the samples of the event posteriors. In practice, this can be equivalently written as

$$\frac{p(\lambda|\boldsymbol{d})}{\pi(\lambda)} \propto e^{-N(\lambda)\alpha(\lambda)} \prod_{i=1}^{N_{\mathrm{obs}}} \frac{1}{\mathcal{S}_i} \sum_{j=1}^{\mathcal{S}_i} \frac{N_{\mathrm{pop}}(^j\boldsymbol{\theta}_i|\lambda)}{\pi(^j\boldsymbol{\theta}_i)}, \quad (59)$$

where j labels the jth sample of the ith event, and $\mathcal{S}_i$ identifies the length of the ith posterior. We sample Eq. (59) using the MCMC package `emcee` [153].

Given a model $\mathcal{M}$, the evidence $Z_\mathcal{M}$ is defined as the marginal population likelihood. This is found by performing the integral of the population posterior,

$$Z_{\mathcal{M}} \equiv \int d\lambda\, p(\lambda|d). \qquad (60)$$

We compute the evidence for each model from the posterior data following Ref. [154]. One can then compare different models by computing the so-called Bayes factors, defined as

$$\mathcal{B}_{\mathcal{M}_2}^{\mathcal{M}_1} \equiv \frac{Z_{\mathcal{M}_1}}{Z_{\mathcal{M}_2}}. \qquad (61)$$

According to Jeffreys' scale criterion [155], a Bayes factor larger than $(10, 10^{1.5}, 10^2)$ would imply a strong, very strong, or decisive evidence in favor of model $\mathcal{M}_1$ with respect to model $\mathcal{M}_2$ given the available dataset.

B. Selection bias

One of the most time consuming tasks when evaluating the likelihood function in Eq. (59) is the computation of the selection bias $\alpha(\lambda)$, quantifying the fraction of observable events in model $\mathcal{M}$ characterized by the hyperparameters λ. Following recent work (see e.g., Ref. [156]), we estimate the selection bias by computing the signal to noise ratio (SNR) for LIGO Hanford, LIGO Livingston and Virgo operating at midhighlatelow sensitivity [157] while adopting the IMRPhenomPv2 waveform approximant [158,159] built in the PyCBC package [160]. The network SNR threshold for detection is set by requiring the quadrature sum of the SNRs from the three detectors to be above $\rho_{th} = 10$, a value which is consistent with the false-alarm-rate threshold used as a detection criterion for events in LVK searches [157].

Analogously to what is done in the LVK analyses, we speed up the computation of the observable number of events [Eq. (57)] by building an injection which covers all the parameter space reached by the models we consider (which is larger than the injection released by the LVK collaboration). We select successfully found injections (i.e., SNR $> \rho_{th}$) and reweight to the population with hyperparameters λ as

$$\alpha(\lambda) = \frac{1}{N_{inj}} \sum_{j=1}^{N_{found}} \frac{p_{pop}(\theta_j|\lambda)}{p_{inj}(\theta_j)}. \qquad (62)$$

In the previous step, we introduced N_{found} as the number of recovered events, N_{inj} as the total number of injections (including those that are not observable with low SNR) and $p_{inj}(\theta)$ as the reference distribution from which injections were built. In particular, the injected distribution of masses follows $p_{inj}(m_1) \propto m_1^{-2.35}$ for $0.1 M_\odot \leq m_1 \leq 500 M_\odot$ and $p_{inj}(q|m_1) \propto q^2$, $p_{inj}(z) \propto (1+z)^{2-1} dV_c/dz$, and again we neglect the binary spins. In order to efficiently cover the wide mass range, we split the injection in two parts with primary mass below and above $5 M_\odot$. We analyze events for

the latter region injecting a population up to redshift $z \leq 2$. The light events, given the much smaller detection horizon, are injected with a redshift distribution extending up to redshift $z \leq 0.2$.

When computing the expected number of events during the future O4/O5 observing runs, we adopt the same framework presented here but with updated LIGO and Virgo future sensitivity curves from Ref. [161].

C. The GWTC-3 dataset

Out of the ≈90 GW detection candidates found by the first three LVK observing runs, here we use the same subset of confident detections selected for the GWTC-3 population analysis in Ref. [59]. Following this choice, the GWTC-3 dataset contains 69 binary BH events and seven potential NS-involving binaries (which are characterized by at least one object with mass below $3M_\odot$).

It is particularly important to include light events in our analysis due to the potential PBH contribution to light binary components, in particular in the solar-mass range mostly affected by the QCD phase. This implies, in particular, that we do include the light events GW170817, GW190425, GW190426_152155, GW190814, GW190917_114630, GW200105_162426, GW200115_042309, out of which only the first one is confidently regarded as a NS binary due to the observation of the electromagnetic counterpart [134]. We do not consider the additional candidate events found by independent searches performed outside the LVK collaboration (e.g., [162,163]), and leave such a task for future work.

We adopt the Overall_posterior samples provided in Ref. [164] for the 11 considered events from the GWTC-1 catalog, the PrecessingSpinIMRHM posteriors provided in Refs. [165,166] for events in the GWTC-2 and GWTC-2.1 catalogs, respectively, while we adopt the C01:Mixed samples for the O3b events reported in the GWTC-3 dataset [167].

VI. CONSTRAINTS ON PBHs FROM GWTC-3

In this section we report the results of the Bayesian inference analyses of GWTC-3 data, assuming either the astrophysical phenomenological models or PBHs (or a mixture of both) are generating mergers of binary BHs and NSs. While it was already shown that PBHs alone are not able to explain all the features observed in the recent GW catalogs [38,40,42,45] assuming a log-normal mass distribution, repeating such simplified analysis is useful to confirm this conclusion remains valid also in a first-principle model including the effects of the QCD phase transition.

A. Single-population inference

We start by discussing the inference on each model separately, focusing on the two subsets of the total events
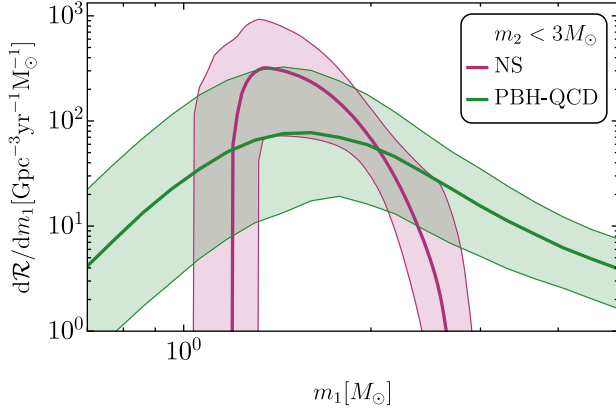
FIG. 6. NS and PBH merger rate distributions as a function of primary mass as inferred from a subset of the GWTC-3 catalog ($m_2 < 3M_\odot$) assuming either the NS phenomenological model or the *ab initio* PBH model with QCD effects and including GW190814.

divided by the condition $m_2 \lesssim 3M_\odot$ (which are defined by the LVK collaboration as events containing NS components or not).

Focusing on the light set of GWTC-3 events, in Fig. 6 we show the inferred merger rate distribution, where at least one of the binary components has mass smaller than $3M_\odot$. We either assume the NS phenomenological model or the PBH model. In the latter case, we do not include the binary NS event GW170817. We report the corresponding posterior distributions in Appendix A.

As previously discussed, the probability of a binary formed by only one PBH is very low, so we neglect this possibility. This implies that, for a highly asymmetric binary with $m_2 < 3M_\odot$ and $m_1 \gg 3M_\odot$ (like, e.g., GW190814 [168]), if the secondary is identified as a PBH then also the primary should be. This is not the case for the astrophysical channels, where the secondary is naturally identified as a NS

and the primary as an ABH. This difference explains why the best-fit PBH merger rate distribution in Fig. 6 has support at larger masses compared to the NS case, although they both peak when $1 \lesssim m_1/M_\odot \lesssim 2$, driven by the events with $m_{1,2} < 3M_\odot$ commonly identified as NS binaries. As a consequence, one falsifiable prediction that follows from the interpretation of GW190814 as a PBH binary is the generation of events filling the lower mass gap potentially existing in the ABH sector between the heaviest allowed NS mass (see e.g., Ref. [169] for a review) and the lightest BH observed [170–172] (see also [173]). Furthermore, due to the features of the *ab initio* PBH mass distribution previously discussed, the PBH merger rate is broader than in the NS case and inevitably has a non-negligible support also in the subsolar range induced by the critical collapse tail. As we shall see, this is a general feature of the model that allows making predictions on subsolar mergers in the PBH scenario.

In Fig. 7, we show the merger rate distribution inferred using only the heavy GWTC-3 events, where both binary components have mass larger than $3M_\odot$ and are therefore identified as BHs. We assume a single binary BH population, either described by the phenomenological ABH model (red) or by our *ab initio* PBH model (green), or also by a phenomenological PBH model using a log-normal mass distribution (blue) often used in the literature (see, e.g., [38,40,42,45]) and shown here for comparison. More details on these are given in Appendix A. Interestingly, in this case we observe that the two PBH models yield fairly similar distributions both for the primary mass (left panel) and mass ratio (right panel). This is because the effects of the power spectrum and QCD phase are largely washed out by the absence of detections with masses below $\approx 6M_\odot$ in this subset of events and some universal properties of the merger rate in Eq. (37), which make the final result largely independent of the details of the two specific parametrization of the PBH mass distribution.
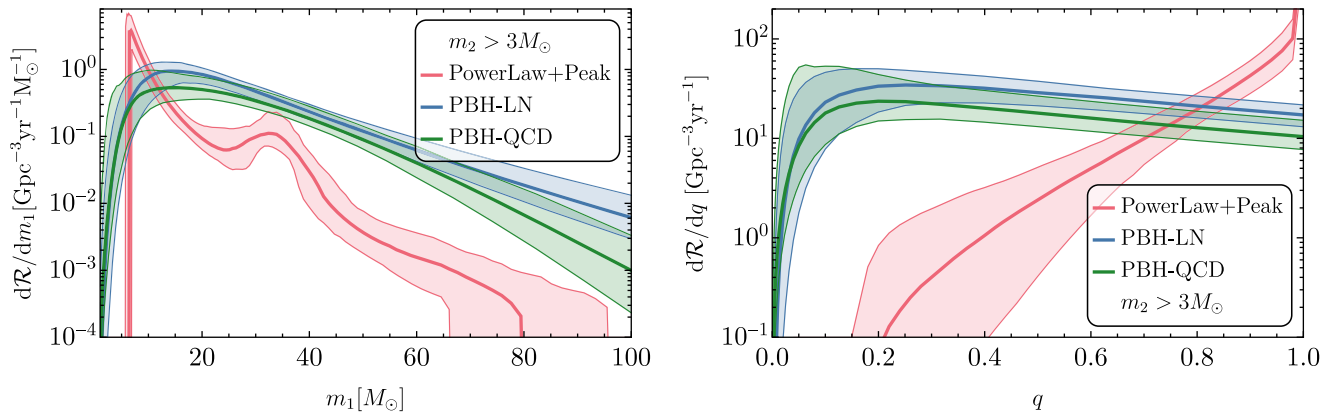

FIG. 7. BH merger rates as a function of the primary mass (left panel) or of the mass ratio (right panel) as inferred from a subset of the GWTC-3 catalog ($m_2 > 3M_\odot$) and assuming only a single binary BH population, either described by the ABH phenomenological model (red), the PBH model assuming a log-normal mass distribution (blue), or the *ab initio* PBH model with a mass distribution fixed by the curvature spectrum in Eq. (10) and modulated by the QCD phase (green).
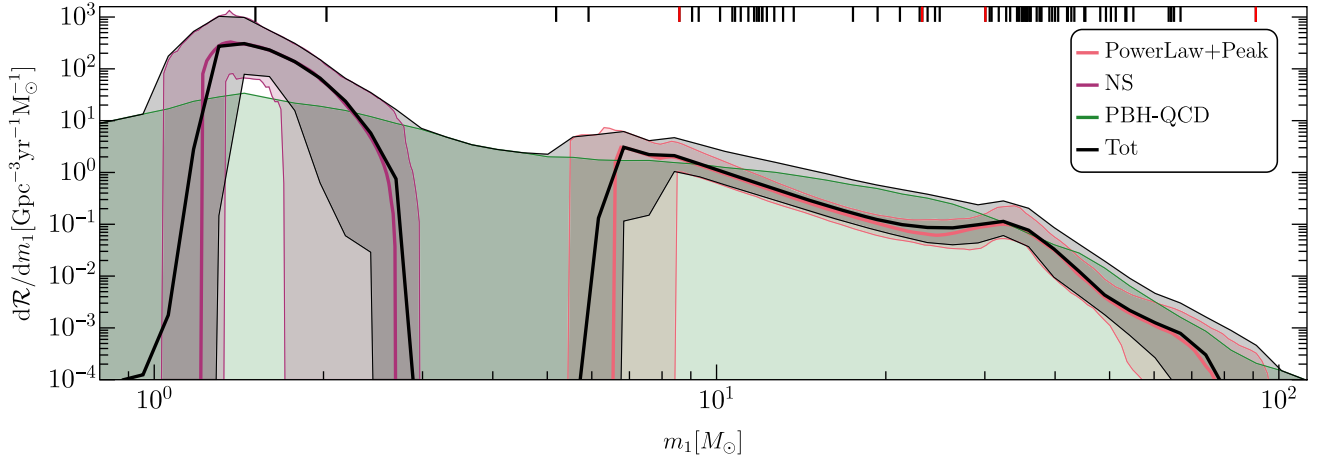
FIG. 8.   Merger rate distribution as a function of the primary mass for the GWTC-3 population inference and including contribution from three channels: ABH phenomenological model, NS phenomenological model, and *ab initio* PBH model modulated by the QCD phase. Events in the lower mass gap (e.g., GW190814) are more naturally interpreted as PBHs rather than being included in the ABH or NS phenomenological channels. This is where the black line mostly deviates from the median distribution of the NS/BBH astrophysical channels. Each black tick at the top of the frame indicates the median values for the primary mass of each GWTC-3 event. In red, we highlight those with non-negligible probability (i.e., $>5\%$) of being of primordial origin in our analysis, see Table III.

Figure 7 confirms previous results (e.g., [38,40,42,45]) finding that the PBH merger rate distribution is markedly different from the ABH one, in particular it lacks a double peak in the mass distribution, it predicts a larger merger rate at high masses, and (in the absence of accretion [17,37,45,130–132]) it predicts a broader merger-rate distribution as a function of the mass ratio which does not favor $q = 1$. The Bayes factors strongly disfavor the interpretation of the totality of the events as coming from the PBH channel alone. In particular, we find $\log_{10}\mathcal{B}_{\mathrm{PBH,LN}}^{\mathrm{ABH}} = 18.5$ and $\log_{10}\mathcal{B}_{\mathrm{PBH,QCD}}^{\mathrm{ABH}} = 17.0$. The value obtained for a log-normal mass distribution is consistent with what was estimated in previous analysis [38,42], but scaled considering the larger statistical sample available with the newest GWTC-3 catalog.

B. Multipopulation inference

Let us now move to population inferences assuming multiple channels. The corresponding posterior corner plots are presented in Appendix A.

In Fig. 8, we show the (differential) merger rate distribution as a function of the primary mass for the entire GWTC-3 catalog allowing contributions from three channels: the LVK ABH and NS phenomenological models, and the *ab initio* PBH model originating from the curvature power spectrum and modulated by the QCD phase transition. The most striking feature of this plot is the fact that the PBH distribution can cover the entire mass range, from subsolar ($m_1 \lesssim M_\odot$) to intermediate mass ($m_1 \gtrsim 100 M_\odot$), with a large support also in the lower mass gap ($3 \lesssim m_1/M_\odot \lesssim 5$) which is instead avoided by the ABH and NS distributions. Note that this last

property is nontrivial, since the ABH and NS models are phenomenological and not informed by astrophysical priors, so *a priori* there is no constraint preventing the best-fit ABH and NS distributions from having support in the lower mass gap.

The support of the PBH distribution at high masses is due both to heavy events potentially interpretable as PBHs and also (as in the previous case of single-population analyses) to the fact that if the secondary is interpreted as a PBH then automatically also the heavier primary is primordial, so the low-mass ($\leq 3 m_\odot$) and high-mass ranges ($\geq 3 m_\odot$) are intertwined. This is not the case for the ABH/NS models, since they allow for mixed BH-NS binaries following independent distributions.

Furthermore, the PBH merger rate distribution shown in Fig. 8 has significant support in the subsolar range. This contribution is only bounded from above by the non-observation of subsolar events in the GWTC-3 catalog. As previously discussed, this interesting property is due to the inevitable broadness of the PBH mass function below the QCD peak induced by the critical collapse. Since the PBH distribution has support in the solar mass range, as it provides a competitive explanation for GW190814 and a marginal contribution to the otherwise NS binaries (although the PBH merger rate is 1–2 orders of magnitude smaller than for the NS distribution), then in our *ab initio* PBH model it is inevitable to have support for subsolar mergers. This is not the case for the NS phenomenological model, whose mass distribution abruptly drops in the subsolar range.

Finally, note that while the ABH and NS distributions shown in Fig. 8 have an upper and lower value given by their corresponding 90% credible interval, the PBH

TABLE III. GWTC-3 events with highest PBH likelihood listed in chronological order. The two groups refer to $m_2 > 3M_\odot$ (top) or $m_2 < 3M_\odot$ (bottom). We also report the measured masses of each event.

GW event	PBH probability [%]	$m_1[M_\odot]$	$m_2[M_\odot]$
GW151012	1.2	$23.2^{+14.9}_{-5.5}$	$13.6^{+4.1}_{-4.8}$
GW190412	25.4	$30.1^{+4.7}_{-5.1}$	$8.3^{+1.6}_{-0.9}$
GW190512_180714	1.6	$23.3^{+5.3}_{-5.8}$	$12.6^{+3.6}_{-2.5}$
GW190519_153544	1.5	$66.0^{+10.7}_{-12.0}$	$40.5^{+11.0}_{-11.1}$
GW190521	7.2	$95.3^{+28.7}_{-18.9}$	$69.0^{+22.7}_{-23.1}$
GW190602_175927	2.7	$69.1^{+15.7}_{-13.0}$	$47.8^{+14.3}_{-17.4}$
GW190701_203306	1.4	$53.9^{+11.8}_{-8.0}$	$40.8^{+8.7}_{-12.0}$
GW190706_222641	1.3	$67.0^{+14.6}_{-16.2}$	$38.2^{+14.6}_{-13.3}$
GW190828_065509	2.8	$24.1^{+7.0}_{-7.2}$	$10.2^{+3.6}_{-2.1}$
GW190924_021846	40.3	$8.9^{+7.0}_{-2.0}$	$5.0^{+1.4}_{-1.9}$
GW191109_010717	2.9	65^{+11}_{-11}	47^{+15}_{-13}
GW191129_134029	1.2	$10.7^{+4.1}_{-2.1}$	$6.7^{+1.5}_{-1.7}$
GW190425	2.8	$2.0^{+0.6}_{-0.3}$	$1.4^{+0.3}_{-0.3}$
GW190426_152155	1.2	$5.7^{+3.9}_{-2.3}$	$1.5^{+0.8}_{-0.5}$
GW190814	29.1	$23.2^{+1.1}_{-1.0}$	$2.59^{+0.08}_{-0.09}$
GW190917_114630	3.0	$9.3^{+3.4}_{-4.4}$	$2.1^{+1.5}_{-0.5}$
GW200105_162426	3.6	$8.9^{+1.2}_{-1.5}$	$1.9^{+0.3}_{-0.2}$
GW200115_042309	1.2	$5.9^{+2.0}_{-2.5}$	$1.44^{+0.85}_{-0.29}$

distribution has no lower value since the posterior of f_{PBH} is also compatible with zero (see Appendix A). This property is natural in our analysis, since the ABH and NS distributions are phenomenological models built to reproduce most of the features of the data.[5] Therefore, as previously remarked, our analysis is not meant to search for a PBH subpopulation but rather to place an upper limit on the PBH abundance compatible with the data (see next section).

Nonetheless, it is interesting that there exist events with a significant likelihood to be interpreted as PBH binaries by our inference, as shown in Table III. In general, the most interesting events are those being either in the light or heavy portions of the catalog, close to either mass gaps, or being characterized by a small mass ratio. While many

---

[5]For example, the phenomenological ABH distribution can accommodate the upper mass-gap event GW190521 [174], even if it is challenging to explain the latter in standard astrophysical scenarios due to the pulsational pair supernova instability preventing the formation of binaries with masses above the (uncertain) limit $\approx 50M_\odot$ [175–187]. One possibility widely investigated in the literature is the interpretation of such an event as a second generation merger in globular clusters or galactic nuclei [188–194], even though it may be challenging to explain the observed rate of this event (see also [45,58]).

events have $\mathcal{O}(\%)$ probability, for GW190924_021846, GW190814, GW190412, and GW190521 the probability is approximately 40%, 29%, 25%, and 7% respectively. We stress that we are comparing an *ab initio* PBH model with phenomenological LVK fits tailored to match current data without any astrophysical input. In particular, the LVK fits do not enforce any mass gap in the ABH/NS distribution, so it is possible that events like GW190814 and GW190924_021846 (with masses $m_2 \approx 2.7M_\odot$ and $m_2 \approx 5M_\odot$, which respectively lie squarely in the lower-mass gap and on its upper end) are well fitted by the ABH or NS phenomenological models. Thus, it is interesting and *a priori* not granted that precisely these events have a sizable probability to be interpreted as primordial. This is due to the fact that they nevertheless lie in a relatively scarcely populated mass range, so the phenomenological distributions should stretch significantly to accommodate them, possibly reducing their ability to fit the many other heavier events in the catalog. Overall, these results may indicate that such events, regardless of their primordial interpretation, may not fit consistently within the population described by the LVK reference model and may belong to distinct populations of NS and BH binaries.

It is also interesting that the light events ($m_{1,2} \lesssim 3M_\odot$) that are interpreted as standard NS binaries by the LVK analysis (e.g., GW190425) have only $\mathcal{O}(\%)$ likelihood to be interpreted as PBHs. This is due to the fact that, even if the PBH distribution modulated by the QCD phase peaks at $m_1 \approx M_\odot$, its magnitude is anyway much smaller than the inferred value of the NS distribution. This is most likely due to the combination of the critical collapse tail (which does not allow for a sharp drop of the mass function below the solar mass) and the constraint from the absence of subsolar detections in GWTC-3.

To conclude this section, we report the Bayes factors comparing the ABH + NS model to the one which includes a PBH subpopulation, found to be

$$\log_{10}\mathcal{B}^{ABH+NS+PBH,QCD}_{ABH+NS} = 0.9, \qquad (63)$$

showing a marginal evidence in favor of a contribution from a PBH channel. This interpretation implicitly includes the downplaying effect of a larger set of parameters introduced in the model when a PBH subpopulation is allowed. Indeed, the ratio between the best-fit likelihood of the two models is

$$\log_{10}\left(\frac{L^*_{ABH+NS+PBH,QCD}}{L^*_{ABH+NS}}\right) = 1.4. \qquad (64)$$

Therefore, the PBH subpopulation improves the fit to the data but not to a sufficient level that would make their absence strongly disfavored.

Overall, this analysis suggests the presence of more features in the GWTC-3 data than what is captured by the

LVK NS and BBH phenomenological models. Even in the most conservative setting, we found that a PBH subpopulation may capture some of these features, even when the nonobservation of a subsolar merger population is taken into account. We now proceed to discuss some interesting implications of our results for future detections and constraints on PBHs and early universe models of inflation.

VII. IMPLICATIONS FOR FUTURE GW EXPERIMENTS AND PBH MODELS

In this section we discuss some implications of our results for the upcoming LVK observation runs and for the PBH scenario.

A. Predicted rate of subsolar mergers and mass-gap events in future LVK searches

As previously discussed, a general property of the *ab initio* PBH model is to predict a significant merger rate in the subsolar range and in the lower mass gap, due to the broadness of the PBH mass function. Thus, once fixing the best-fit PBH abundance distribution through the Bayesian inference, it is possible to make *falsifiable* predictions about the expected numbers of events in the subsolar mass range and in the lower mass gap, assuming some of the GWTC-3 events already detected are interpreted as a primordial binary.

In Table IV, we show these predictions, assuming GW190814 is primordial (29% likelihood in our analysis). Assuming a primordial origin for GW190924_021846 (40% likelihood in our analysis) provides similar predictions.[6] First of all, the first row in Table IV shows that the interpretation of at least GW190814 as a PBH binary implies the current catalog may include a fraction between 1% and 29% of PBH mergers. On the other hand, the number of expected subsolar mergers within the O1-O2-O3 observation runs is below unity, consistently with the absence of observations in that mass range.

Due to the much improved sensitivity of future observation runs, we notice that O4 and O5 would be bound to detect many PBH events, as expected. However, unless some of these events have smoking-gun features [17], it would be hard to distinguish them from ordinary astrophysical channels. Therefore, a more interesting prediction of Table IV is the number of subsolar and mass-gap events detectable in O4 and O5. In particular, in O5 there could be as many as ≈8 subsolar events per year (but the 90% confidence interval is also compatible with zero

---

[6]In the following we shall mostly assume that GW190814 is a primordial binary, even though the PBH likelihood of GW190924_021846 is higher. Besides the fact that the two assumptions would provide similar results, GW190814 is more challenging to fit within standard astrophysical scenarios and the mass of its secondary [195] motivates exploring other explanations for this event.

TABLE IV. Assuming GW190814 had primordial origin, this table reports the 90% confidence interval (C.I.) for the number of detected PBH events within GWTC-3, and predicted events (per year) with O4 and O5 sensitivity. We also indicate forecasted detections within the subsolar ($m_2 < M_\odot$, SS), lower mass gap (m_1 or $m_2 \in [2.5, 5]M_\odot$, LMG), and upper mass-gap ($m_1 > 50M_\odot$, UMG) ranges.

	$N_{\rm PBH}^{\rm det}$	$N_{\rm PBH}^{\rm det}({\rm SS})$	$N_{\rm PBH}^{\rm det}({\rm LMG})$	$N_{\rm PBH}^{\rm det}({\rm UMG})$
O1–O3	[0.8, 22.4]	[0.0, 0.6]	[0.1, 2.3]	[0.0, 6.1]
O4	[1.9, 43.7]	[0.0, 1.3]	[0.3, 13.0]	[0.0, 13.1]
O5	[10.3, 216.7]	[0.0, 8.6]	[0.8, 25.2]	[0.0, 47.3]

events). More interestingly, if GW190814 is assumed to be primordial then O5 should detect one to a few dozen events per year in the lower mass gap (and up to ≈50 upper mass-gap events), which might be more difficult to interpret in astrophysical scenarios.

While detecting a subsolar merger would be a unique smoking gun for PBHs (or would anyway call for new physics beyond the standard astrophysical formation scenario [196–203]), the lower mass gap [204] could be populated also by second-generation mergers formed in dense stellar clusters, whose rates in this mass range are particularly uncertain. A way to distinguish second-generation BH mergers from PBH mergers is by measuring the binary spins, since in the former case the spin is expected to be non-negligible [204,205], at variance with the latter case [17,132].

We conclude this section by speculating that the existence of a lower mass-gap population of PBHs may be compatible with the OGLE-2011-BLG-0462 low mass BH microlensing observation [206,207], whose x-ray luminosity is consistent with the small radiative efficiency expected for a BH and disfavors a NS interpretation [208], see also Ref. [209].

B. PBH constraints

The posterior distribution describing the parameters of the PBH population (see Appendix A) can be used to set an upper bound on the PBH abundance in the solar mass range.

In Fig. 9 we show the posterior predictive distribution for the PBH mass function $f(m_{\rm PBH})$ in a logarithmic scale obtained from the GWTC-3 inference, together with existing constraints in this mass range (see, e.g., [21] for a recent review). In the mass range of interest for our discussion, the most relevant constraints come from CMB anisotropies produced by accreting PBHs in the early universe [20,210]. Other constraints come from comparing the *late time* emission of electromagnetic signals from interstellar gas accretion onto PBHs with observations of galactic radio and x-ray isolated sources (XRay) [211,212] and x-ray binaries (XRayB) [213], x-ray and radio backgrounds (XRR) [214],
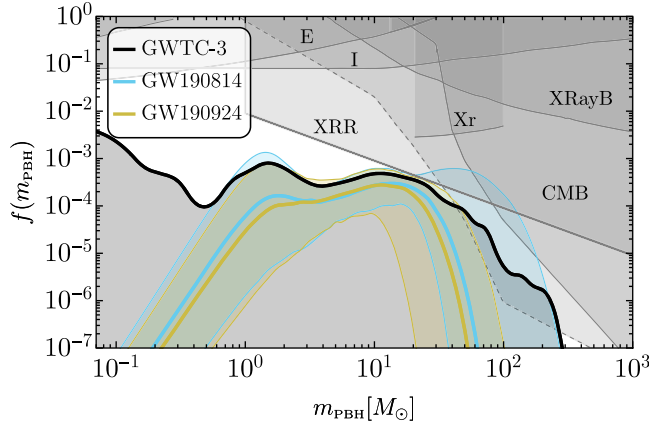
FIG. 9. Constraints on the PBH mass distribution derived in this work and compared to existing ones [21]. The black curve shows the upper bound (90% C.I.) for the mass distribution obtained from the GWTC-3 inference. The cyan (yellow) band shows the posterior distribution assuming GW190814 (GW190924) is interpreted as a PBH binary.

FIG. 10. Upper bound on the PBH abundance as a function of average mass $\langle m_{\rm PBH} \rangle$ derived at 90% C.I. from the GWTC-3 dataset assuming a power-law parametrization of the primordial power spectrum and (conservatively) a dominant contribution of astrophysical mergers in the LVK band.

lensing searches of massive compact halo objects (MACHOs) towards the Large Magellanic Clouds (EROS, E) [215], fast transient events near critical curves of massive galaxy clusters (ICARUS,I) [216], and observations of stars in the Galactic bulge by the Optical Gravitational Lensing Experiment (OGLE,O) [217]. The deflection of light by PBHs in the density spike likely existing around the M87 supermassive black hole combined with EHT measurements give rise to additional constraints [218], which are not shown as their conservative version would fall behind the region already excluded by CMB. Consistently with the assumptions made in the previous sections, here we also do not account for the potential impact of PBH accretion that may shift CMB constraint to higher masses [131].

The black area in Fig. 9 corresponds to the (90% confidence level) upper bound on the PBH mass distribution $f(m_{\rm PBH})$ derived from the GWTC-3 multipopulation inference and hence extends up to $f(m_{\rm PBH}) \to 0$. The cyan region is instead the posterior distribution assuming that the lower mass-gap event GW190814 is a primordial binary, which forces $f(m_{\rm PBH})$ to be nonzero and therefore bounded from below. A similar bound can be obtained by assuming that GW190924 is primordial (yellow curves). First of all, we note that the allowed region for the PBH model is not excluded by other constraints not based on GW events. Only a minor overlap between the cyan band and the CMB constraints is observed, which is not however sufficient to constrain the scenario. Finally, the mass distribution is allowed to gain a higher contribution going towards masses well below $\approx M_\odot$ due to the reduced sensitivity of LVK deep in the subsolar mass range.

We can also translate the constraint on the mass distribution on the overall value of the abundance defined from Eq. (28) as
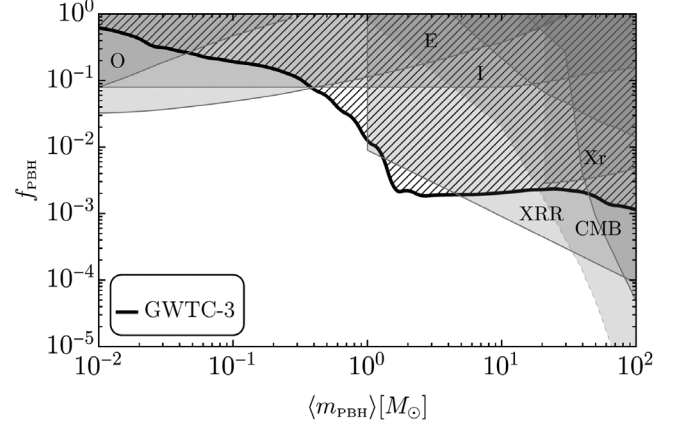
$$ f_{\rm PBH} \equiv \int {\rm d}\ln m_{\rm PBH} f(m_{\rm PBH}). \tag{65} $$

While most of the posterior of $f_{\rm PBH}$ is constrained to be much smaller than unity, see Fig. 20, there is also a small support for a tail reaching $f_{\rm PBH} = 1$. This tail is correlated with blue spectra (i.e., large n_s) giving larger support to light masses, and small $M_{\rm S}$. This means that values of the PBH abundances of order unity can only be reached for light PBH populations where the LVK sensitivity sufficiently degrades. We can better visualize this result by computing the maximum $f_{\rm PBH}$ at 90% C.I. as a function of the average PBH mass $\langle m_{\rm PBH} \rangle$. This upper bound represents the maximum value of the fraction of the dark matter which can be explained by a PBH population derived assuming a power spectrum of the form (10) and an average mass $\langle m_{\rm PBH} \rangle$, when also an ABH population of mergers is allowed to efficiently explain the majority of mergers in the GWTC-3 dataset. This bound is shown in Fig. 10, showing a marked plateau around $f_{\rm PBH} \approx 2 \times 10^{-3}$, consistently with previous approximated studies [29,34,39,40,45], which drastically degrades at masses below $M_\odot$, eventually hitting other non-GW-based constraints. This also confirms that LVK observations set the most stringent constraints in the mass range $\langle m_{\rm PBH} \rangle \in [0.3, 50] M_\odot$.[7]

Furthermore, as already shown in Table III, we see that the best-fit PBH model allows for a certain number of events (GW190412, GW190924_021846, GW190814, GW190521) to have a primordial origin with probability

---

[7]While our constraint is derived assuming nearly Gaussian perturbations and a consequent initial Poisson spatial distribution of PBHs, it was recently shown that even assuming a (more exotic) clustered initial condition does not allow to evade constraints preventing stellar mass PBHs from being a dominant component of the dark matter [219].

TABLE V. Posterior 90% C.I. for PBH population parameters assuming GW190814 is primordial (similar results are found by assuming that GW190924 is primordial).

Parameter	All	GW190814	GW190924
$\log_{10} A$	$-1.9^{+0.4}_{-0.6}$	$-1.93^{+0.10}_{-0.05}$	$-1.9^{+0.1}_{-0.1}$
n_s	$0.68^{+0.66}_{-0.61}$	$0.68^{+0.18}_{-0.40}$	$0.64^{+0.29}_{-0.56}$
$\log_{10}(k_{\min}/\mathrm{Mpc}^{-1})$	$6.0^{+1.6}_{-0.6}$	$5.9^{+0.2}_{-0.4}$	$6.0^{+0.3}_{-0.2}$
$\log_{10}(k_{\max}/\mathrm{Mpc}^{-1})$	$7.8^{+0.6}_{-0.9}$	$8.1^{+0.3}_{-0.9}$	$8.0^{+0.4}_{-1.2}$
$\log_{10} f_{\mathrm{PBH}}$	$-3.4^{+2.2}_{-2.3}$	$-3.1^{+0.5}_{-0.4}$	$-3.2^{+0.3}_{-0.5}$
$\log_{10}(M_S/M_\odot)$	$-1.2^{+1.8}_{-1.2}$	$-1.6^{+1.7}_{-0.7}$	$-1.6^{+2.5}_{-0.9}$
$\log_{10}(M_L/M_\odot)$	$2.4^{+1.3}_{-3.2}$	$2.6^{+0.7}_{-0.3}$	$2.5^{+0.5}_{-0.5}$

respectively about (25%, 40%, 29%, 7%). This means, in contrast with the analysis of Ref. [69], that the absence of subsolar events during the past LVK runs (which is automatically included in our analysis) does not exclude the possibility that *some* of the detected events have a primordial origin. This is due to the fact that in our *ab initio* model we allow the tilt n_s to vary and its inferred value is given in Table V (along with the posteriors of the other PBH population hyperparameters), while the same parameter was fixed to $n_s \approx 0.95$ (very close to its value at the much larger, and uncorrelated, CMB scales) in Ref. [69] (following the choice made in Refs. [64,65]).

Although error bars on n_s are large, the population inference systematically selects a *redder* tilted curvature power spectrum which reduces the abundance in the (sub) solar mass range and erases the dependence to the high scale $k_{\max}$ (i.e., low mass M_S) cutoff, which is compatible with the left boundary of its prior range $M_S = 10^{-2.5} M_\odot$. This is needed in order to counteract the QCD enhancement at the solar mass and reduce the hierarchy in mass distribution between the solar mass and $\mathcal{O}(\text{tens})$ of solar masses (where GW190814 and other events get support). The PBH abundance is found to depend strongly on the tilt, so even a change by $\sim 10\%$ can change the abundance significantly. In particular, a smaller value of n_s makes the QCD peak less pronounced and the slope at higher masses less steep, resulting in observable rates in the $\mathcal{O}(10 M_\odot)$ range even in the absence of subsolar events.[8] Finally, the contribution to the heavier portion of the catalog depends instead on the scale where the power spectrum grows from the CMB values ($k_{\min}$ or M_L). This is an inevitable ingredient in PBH models, as we shall discuss in the next section.

---

[8]We note that current constraints in the subsolar mass [70] rely on assuming a given PBH mass distribution, which is not the one induced by the QCD phase transition considered here. We stress that our analysis automatically accounts for possible subsolar events and the absence thereof in GWTC-3.

VIII. PBHs FROM INFLATIONARY DYNAMICS: A DATA-DRIVEN MODEL

We now come to the theoretical interpretation of our data-driven results. The goal we set in this section is simple but ambitious: we aim to construct a model of PBH formation that gives an abundance distribution compatible with the allowed region shown in Fig. 9. This question will be addressed in Sec. VIII A. Even more ambitiously, we may ask whether such PBH distribution could comprise the entirety of the dark matter observed in the universe. This question will be addressed in Sec. VIII B.

Before entering into the details, let us illustrate the general strategy. At first sight, the answer to the last question is a resounding no—after all the very same constraint extracted in Fig. 10 limits the maximum abundance of PBHs to be far below order-one values. It is well known, in fact, that the only mass range, consistent with observational bounds, in which dark matter could entirely consist of PBHs is $10^{-16} \lesssim m_{\mathrm{PBH}}/M_\odot \lesssim 10^{-12}$ [21] (dubbed the asteroid mass range in the following), that is for PBHs way lighter than the solar-mass range covered by LVK data. However, a first exception to this apparent incompatibility was pointed out in Ref. [220], where it was shown that a broad curvature power spectrum in the form of a double-step Heaviside theta function could potentially give birth to a population of PBHs with a mass distribution covering vastly different scales (see also Ref. [73]).

In light of this result, one could be tempted to interpret a putative mass distribution compatible with the constraint in Fig. 9 as the proverbial tip of the iceberg, which is just the final part of a much wider mass distribution possibly reaching order-one abundances at values of m_{PBH} compatible with the asteroid mass range. The idea of Ref. [220]—originally thought in reference to the stochastic signal of GWs generated, as a second-order effect, by the large scalar perturbations that form PBHs—was recently explored in much more detail in Ref. [72]. Two of the results of this paper are worth emphasizing. First, it was shown how to engineer consistent (that is, compatible with CMB observations and the end of inflation) inflationary dynamics which give rise, starting from a handful of physically meaningful parameters, to a curvature power spectrum compatible with the toy model given in Ref. [220]; second, and most importantly, it was shown that, once the relevant parameters that control the background dynamics have been identified, it is relatively simple to understand what are the conditions that are needed in order to generate a PBH mass distribution that gives an order-one abundance of dark matter in the asteroid mass range and, *simultaneously*, a detectable fraction of solar-mass merger events. In this section, we will add one more piece of information to this picture, quantitatively assessing and reinforcing its robust observational consequences in the solar-mass range.

As in Ref. [72], our starting point for the following discussion is the analytical *ansatz*:

$$\eta(N) = \frac{1}{2}\left\{\left[\eta_{\rm I} - \eta_{\rm II} + (\eta_{\rm II} - \eta_{\rm I})\tanh\left(\frac{N - N_{\rm I}}{\delta N_{\rm I}}\right)\right]\right.$$
$$+ \left[\eta_{\rm II} + \eta_{\rm III} + (\eta_{\rm III} - \eta_{\rm II})\tanh\left(\frac{N - N_{\rm II}}{\delta N_{\rm II}}\right)\right]$$
$$\left. + \left[\eta_{\rm IV} - \eta_{\rm III} + (\eta_{\rm IV} - \eta_{\rm III})\tanh\left(\frac{N - N_{\rm III}}{\delta N_{\rm III}}\right)\right]\right\} \tag{66}$$

that describes the time evolution of the Hubble parameter $\eta \equiv -\ddot{H}/2H\dot{H}$, where $\dot{H} = dH/dt$ is the cosmic-time derivative of the Hubble rate H and N, defined by $dN/dt = H$, is the number of e-folds.

A. Solar-mass PBHs from inflationary dynamics

Consider first the limit $\eta_{\rm III} = \eta_{\rm IV}$ in Eq. (66). The last line vanishes, and we are left with the expression

$$\eta(N) = \frac{1}{2}\left\{\left[\eta_{\rm I} - \eta_{\rm II} + (\eta_{\rm II} - \eta_{\rm I})\tanh\left(\frac{N - N_{\rm I}}{\delta N_{\rm I}}\right)\right]\right.$$
$$\left. + \left[\eta_{\rm II} + \eta_{\rm III} + (\eta_{\rm III} - \eta_{\rm II})\tanh\left(\frac{N - N_{\rm II}}{\delta N_{\rm II}}\right)\right]\right\}. \tag{67}$$

The meaning of the free parameters entering in Eq. (66) becomes manifest by looking at Fig. 11. The right-side y axis of this figure shows Eq. (67) as a function of the number of e-fold N (upper-side x axis) or, equivalently, the comoving wave number k (lower-side x axis); these two quantities are indeed related by the horizon-crossing condition $k = a(N)H(N)$ that we normalize in such a way that $N = 0$ corresponds to the crossing time of the CMB pivot scale $k_\star = 0.05$ Mpc^{-1}. Once the time evolution of η is given, it is immediately possible to obtain the time evolution of the Hubble parameter $\epsilon \equiv -\dot{H}/H^2$ by solving the differential equation $\eta = \epsilon - 1/2 d \log \epsilon/dN$ with initial condition $\epsilon_{\rm I}$ at $N = 0$; at this stage, therefore, the background inflationary dynamics is completely specified. The curvature power spectrum can be now obtained by solving numerically the Mukhanov-Sasaki equation (cf. Ref. [72] for technical details). In Fig. 11 we superimpose the curvature power spectrum $\mathcal{P}_\zeta(k)$ that corresponds to the time evolution of η shown in the same figure.

The most important part of the dynamics is the presence of a phase of USR during which we have $\eta \gtrsim 3/2$ in our parametrization, inducing an exponential growth of a specific set of modes. Such USR phase takes place in the e-fold time interval $N_{\rm I} \lesssim N \lesssim N_{\rm II}$.

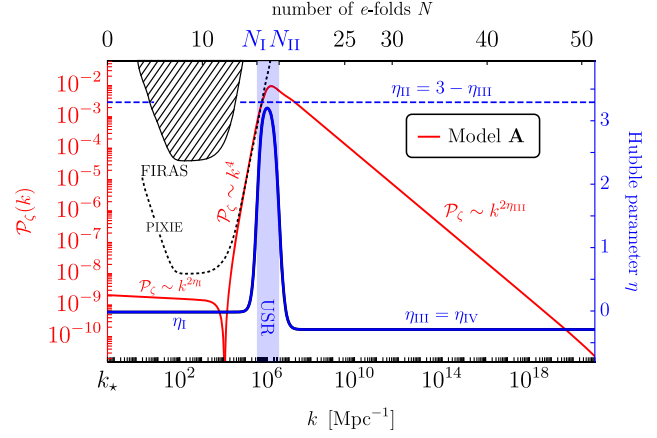Curvature perturbations that cross the horizon well before the USR phase are not affected by the latter, and

FIG. 11. Curvature power spectrum (red, left-side y axis) as a function of the comoving wave number k (lower-side x axis); on the upper-side x axis, we indicate the number of e-folds N according to the horizon-crossing condition $k = a(N)H(N)$, normalized at $N = 0$ for the pivot scale $k_\star = 0.05$ Mpc^{-1}. We superimpose (blue, right-side y axis) the time evolution of the Hubble parameter η. The figure refers to the explicit realization of our model A in Table VI. The vertical region shaded in blue indicates the USR phase ($N_{\rm I} \leq N \leq N_{\rm II}$). The meshed region shows the FIRAS constraint from CMB spectral distortions computed for the steepest growth power spectrum, $\mathcal{P}_\zeta \sim k^4$, cf. Ref. [100]. The dotted black line illustrates the projected constraints from a future PIXIE-like spectral distortion experiment [55,221].

contribute to the power spectrum according to the usual slow-roll approximation. This part of the power spectrum follows the scaling $\mathcal{P}_\zeta \sim k^{2\eta_{\rm I}}$ (cf. Fig. 11) and the numerical values of $\eta_{\rm I}$ and $\epsilon_{\rm I}$ are chosen in such a way to fit CMB data at the pivot scale.

Curvature perturbations that cross the horizon right before the USR phase are those that are mostly affected by the latter. These modes are exponentially enhanced and give rise to a steep growth of the power spectrum that in our model follows the scaling $\mathcal{P}_\zeta \sim k^4$. Because of this growth, the curvature power spectrum experiences a parametric change (with respect to the preceding slow-roll value) of the order $\Delta\mathcal{P}_\zeta \sim e^{2\eta_{\rm II}\Delta N_{\rm USR}}$, with $\Delta N_{\rm USR} \equiv N_{\rm II} - N_{\rm I}$. To fix ideas, in order to get a 7 orders-of-magnitude enhancement of the power spectrum (that would bring the typical slow-roll amplitude $\mathcal{P}_\zeta \sim 10^{-9}$ up to $\mathcal{P}_\zeta \sim 10^{-2}$) one needs $\eta_{\rm II}\Delta N_{\rm USR} \sim 8$. In other words, the combination of parameters $\eta_{\rm II}\Delta N_{\rm USR}$ controls the height that the curvature power spectrum reaches as a consequence of the USR phase.

Curvature perturbations that cross the horizon well after the USR phase during the e-fold time interval $N \gtrsim N_{\rm II}$ give rise to the final part of the power spectrum with scaling $\mathcal{P}_\zeta \sim k^{2\eta_{\rm III}}$. In this part of the dynamics the slow-roll approximation is again applicable. The value of $\eta_{\rm III}$ is negative, and this is crucial for inflation to end. We fix the numerical value of $\eta_{\rm III}$ by imposing a total number

TABLE VI. Free parameters of our models together with their numerical benchmark values. We define $\Delta N_{\rm USR} \equiv N_{\rm II} - N_{\rm I}$ and $\Delta N_{\rm plateau} \equiv N_{\rm III} - N_{\rm II}$. Consistently with Planck data [225], at the CMB pivot scale $k_\star = 0.05$ Mpc^{-1}, all models give $n_s(k_\ast) = 0.965$ and $A_s = 2.1 \times 10^{-9}$ for, respectively, spectral index and amplitude of the curvature power spectrum, and a tensor-to-scalar ratio $r = 0.005$. Only models C and D produce PBHs that can account for the entirety of the dark matter without violating existing constraints (see also Appendix B for more details).

Model parameters in Eq. (66)	$\epsilon_{\rm I}, \eta_{\rm I}$	$N_{\rm ref}$	$N_{\rm I}$	$\eta_{\rm II}, \Delta N_{\rm USR}$	$\eta_{\rm III}$	$\Delta N_{\rm plateau}$	$N_{\rm IV}, \eta_{\rm IV}$	$\delta N_{\rm I}$	$\delta N_{\rm II}$	$\delta N_{\rm III}$	$f_{\rm PBH}=1$
Model A	$\epsilon_{\rm I} = 3.125 \times 10^{-4}$ $\eta_{\rm I} = -1.68 \times 10^{-2}$	0	15.75	$\eta_{\rm II} = 3 - \eta_{\rm III}$ $\Delta N_{\rm USR} = 2.342$	-0.292	✗	$N_{\rm IV} = 55$ $\eta_{\rm IV} = \eta_{\rm III}$	0.50	0.59	✗	✗
Model B	$\epsilon_{\rm I} = 3.125 \times 10^{-4}$ $\eta_{\rm I} = -1.68 \times 10^{-2}$	0	15.5	$\eta_{\rm II} = 3.17$ $\Delta N_{\rm USR} = 2.47$	-0.294	✗	$N_{\rm IV} = 55$ $\eta_{\rm IV} = \eta_{\rm III}$	0.50	0.59	✗	✗
Model C	$\epsilon_{\rm I} = 3.125 \times 10^{-4}$ $\eta_{\rm I} = -1.68 \times 10^{-2}$	0	15.75	$\eta_{\rm II} = 3.197$ $\Delta N_{\rm USR} = 2.44$	0	17.56	$N_{\rm IV} = 55$ $\eta_{\rm IV} = -0.576$	0.50	0.68	0.50	✓
Model D	$\epsilon_{\rm I} = 3.125 \times 10^{-4}$ $\eta_{\rm I} = -1.68 \times 10^{-2}$	0	15.5	$\eta_{\rm II} = 3.11$ $\Delta N_{\rm USR} = 2.44$	0.012	17.56	$N_{\rm IV} = 55$ $\eta_{\rm IV} = -0.567$	0.50	0.50	0.50	✓

of 55 inflationary e-folds; different choices of this benchmark value would slightly affect the parameters of our model while not significantly changing the dynamics.

Finally, the parameters $\delta N_{\rm I}$ and $\delta N_{\rm II}$ control the width of the tanh transitions between different values of η at, respectively, $N_{\rm I}$ and $N_{\rm II}$. The limit $\delta N_{\rm I,II} \to 0$ corresponds to a step transition.

The above discussion captures the most evident features of the curvature power spectrum and explains the formation of the peak shown in Fig. 11. The precise form of the power spectrum at the tip of the peak is shaped by curvature modes that cross the horizon during and immediately after the USR phase. This aspect is truly crucial for our analysis since this is the part of the power spectrum that should be compared with the model in Eq. (10). On the theory side, the key aspect is the possibility to establish the so-called Wands duality [222] between the USR phase that takes place during the e-fold time interval $N_{\rm I} \leq N \leq N_{\rm II}$ and the subsequent phase $N > N_{\rm II}$. In short, the Wands duality is the statement that phases with η and $3 - \eta$ give rise to the same spectral slope in the curvature power spectrum (cf. also Refs. [223,224]). In our model, this implies that if we set $\eta_{\rm II} = 3 - \eta_{\rm III}$ we expect that the form of the power spectrum right after the tip of the peak will take the same power-law form $\mathcal{P}_\zeta \sim k^{2\eta_{\rm III}}$ that, as discussed before, characterizes the last part of the dynamics. In model A, therefore, we enforce the condition $\eta_{\rm II} = 3 - \eta_{\rm III}$, see Table VI.

In the left panel of Fig. 12 we enlarge the peak of the power spectrum. The solid black line corresponds to model A. The numerical solution of the Mukhanov-Sasaki equation confirms our analytical intuition: right after the peak, the curvature power spectrum can be well approximated by a power law with spectral index that, in the notation of Eq. (10), takes the approximate value $n_s = 1 + 2\eta_{\rm III}$. Since $\eta_{\rm III}$ is negative and $|\eta_{\rm III}| < \mathcal{O}(1)$ (otherwise the inflaton will

roll too fast towards the end of inflation) it is natural to expect a red tilted power spectrum, consistently with our previous analysis using GW data. In the explicit realization given by model A, we find $n_s \approx 0.45$, consistent with our population inference (see Table III). In the left panel of Fig. 12 we also show (region shaded in cyan) the 90% C.I. posterior distribution assuming GW190814 is interpreted as a PBH binary. This is the same region shown in Fig. 9 but recomputed in terms of the parameters of the power spectrum. The comparison shows that the USR dynamics in model A gives a good agreement with the data-driven results derived in Sec. VI B. En route, we note that, after mapping the inference to the parameters of the power spectrum in Eq. (10), the posterior distribution of $P_\zeta(k)$ is extremely well constrained at around $k = 3 \times 10^6$ Mpc^{-1}, whereas the error bars become larger towards the two cutoff scales in momentum. Intuitively, this is expected. We remind that this posterior is based on the assumption that GW190814 is a primordial binary, which forces $f(m_{\rm PBH})$ to be nonzero (and pretty well determined in particular for $m_{\rm PBH} \simeq 20 M_\odot$ that is the primary mass of GW190814). Since the abundance has an exponential dependence on the amplitude A of the power spectrum, the latter cannot change too much around the corresponding wave numbers. Additionally, the remaining parameters entering in the spectrum (10) correlate in such a way to respect the stringent bound around 3×10^6 Mpc^{-1} while broadening the permitted regions at both sides.

In the right panel of Fig. 12 we show the mass distribution $f(m_{\rm PBH})$ computed according to the formalism set[9] in

---

[9]It should be noted that we now fully compute the variances in Eq. (14) numerically without relying on the analytical approximation in Eq. (35). The agreement between both approaches confirms the validity of the approximations adopted to perform the MCMC Bayesian analysis.
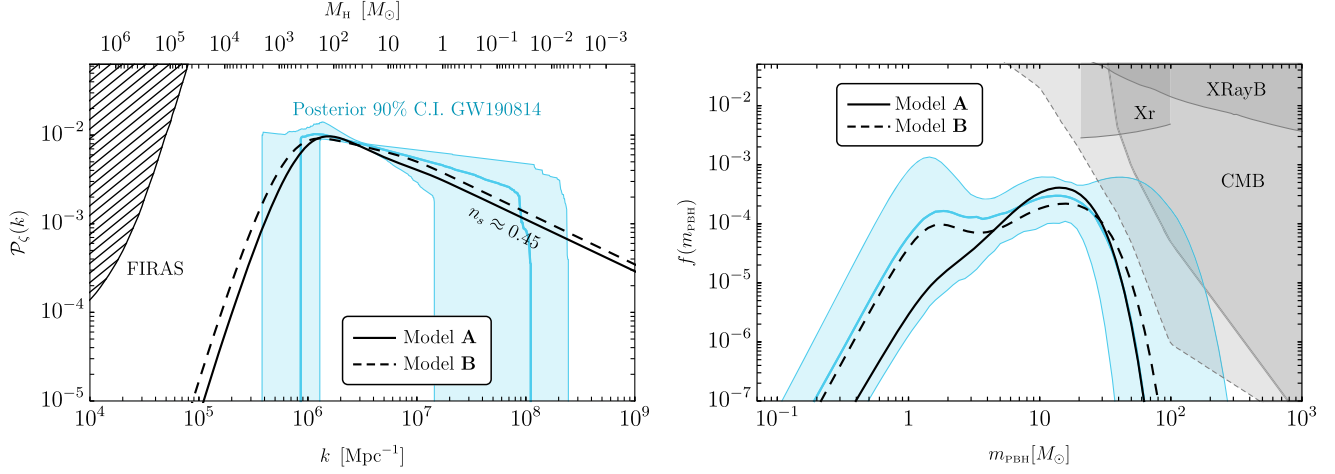
FIG. 12.   Left panel: enlargement of Fig. 11 near the peak of the spectrum. On the top x axis we consider, instead of comoving wave numbers k, the horizon mass according to Eq. (18). In addition to model A we also show the power spectrum that corresponds to model B in Table VI (the latter is not shown in Fig. 11 since the small differences between models A and B can be only appreciated near the peak). The cyan shaded region corresponds to the 90% C.I. posterior distribution obtained in Sec. VI B assuming GW190814 is interpreted as a PBH binary (cf. Table V). Right panel: mass distribution $f(m_{PBH})$ for both models A and B compared with the 90% C.I. posterior distribution assuming GW190814 is interpreted as a PBH binary (cf. Fig. 9).

Sec. II. We compare the distribution given by model A with the 90% C.I. posterior distribution assuming GW190814 is interpreted as a PBH binary. As expected, the model is consistent with the region bracketed by the confidence interval. Since we have $n_s \approx 0.45$, the model tends to underproduce PBHs in the subsolar mass range with respect to the median value.

It is important to stress that model A relies on the condition $\eta_{II} = 3 - \eta_{III}$. If we break the Wands duality, it is no longer guaranteed that, after the tip of the peak, the curvature power spectrum will be described by a single power law. On the contrary, we expect that curvature modes that cross the horizon during and immediately after the USR phase will give to the power spectrum a slightly different scaling compared to the one that characterizes the subsequent phase, $\mathcal{P}_\zeta \sim k^{2\eta_{III}}$. To better illustrate this point, we consider model B in Table VI. In this model, the Wands duality is broken, $\eta_{II} \neq 3 - \eta_{III}$. In Fig. 12 model B is represented by the black dashed line. As expected, we see that model B gives a curvature power spectrum that, right after the peak, is characterized by a broken power law with two slightly different spectral indices. The second one is fixed by the last part of the dynamics and always given by $n_s = 1 + 2\eta_{III}$. The first one, on the contrary, can be tuned to match more accurately the central value $n_s \approx 0.68$ given in Table V and, therefore, it would enhance the number of PBHs in the subsolar mass range. As shown in the right panel of Fig. 12, this is exactly what model B was designed for and it matches the data-driven distribution much more closely than model A.

At this point of the analysis, we are already in the position to draw a number of relevant conclusions. The formation of PBHs is a rare event that requires some finely tuned underlying dynamics. This statement seems to be true whatever formation mechanism one decides to consider and, in our analysis, we focused on the presence of a phase of USR during inflation. Once we are willing to accept the presence of this tuned dynamics, the point that we would like to stress is that the latter naturally comes with a number of features that fully justify the simplified approach taken in our numerical analysis.

 (i) First, we note that the cutoff k_{min} (equivalently, M_L) arises naturally as a consequence of the sharp enhancement of the power spectrum (with respect to CMB values) that is essential for the generation of a sizable abundance of PBHs; in our explicit realization, such enhancement is provided by the presence of the USR phase.
 (ii) Second, curvature modes that cross the horizon during and after the USR phase shape the form of the power spectrum for $k > k_{min}$. In the context of the parametrization given in Eq. (67), and imposing the Wands duality condition $\eta_{II} = 3 - \eta_{III}$, it is possible to get a red tilted power-law functional dependence with $n_s = 1 + 2\eta_{III}$ (cf. model A in Table VI and Fig. 12).
 (iii) More in general, if we drop the condition $\eta_{II} = 3 - \eta_{III}$, the power spectrum is better approximated by a broken power law with two spectral indices (cf. model B in Table VI and Fig. 12).
 (iv) Finally, as already noticed, a red tilted spectrum as that suggested by GW data makes the PBH mass distribution practically insensitive to the cutoff k_{max} (equivalently, M_S). This aspect is well illustrated by our model since the explicit USR dynamics that we consider does not really give any specific value for

$k_{\max}$; on the contrary, we find that the power spectrum just decreases as $\mathcal{P}_\zeta \sim k^{2\eta_{\mathrm{III}}}$ following the last part of the dynamics that ends inflation.

Overall, our analysis shows that it is possible to devise USR inflationary models that produce the curvature power spectrum in Eq. (10) assumed as the chief starting ingredient of our GW data-driven population inference. In practice, instead of parametrizing the spectrum as in Eq. (10) one could directly start by parametrizing the evolution of the Hubble parameter η [e.g., Eq. (67)] or the potential and couplings of the inflaton field(s), and directly run the inference on the values of the inflationary model.

B. Solar-mass PBHs *and* dark matter from inflationary dynamics

We now move to consider the second question raised in the introductory part of this section: Is it possible to make the presence of a PBH subpopulation that explains a fraction of GWTC-3 events compatible with the assumption that the entirety of dark matter observed in the universe consists of PBHs?

Answering this question requires devising a realization of inflationary dynamics tuned in such a way that the logarithmic integral of the mass distribution gives unity, cf. Eqs. (28) and (65). Since in the solar mass range the fraction of dark matter in the form of PBHs is constrained to be at most $\mathcal{O}(10^{-3})$, the integral must be dominated by the peak in the asteroid mass range (for further details see Appendix B).

1. Power spectrum with a plateau: How to bridge PBH populations with widely different mass

We consider the full evolution given by Eq. (66). Compared to the situation discussed in Sec. VIII A, we now have $\eta_{\mathrm{III}} \neq \eta_{\mathrm{IV}}$ and one additional tanh transition at e-fold time N_{III}. In the left panel of Fig. 13 we show (blue, right-side y axis) the evolution of η dictated by Eq. (66) in the explicit realization given by model C in Table VI.

We superimpose the (red, left-side y axis) the curvature power spectrum that corresponds to such dynamics.

As discussed in Ref. [72], we impose the condition $\eta_{\mathrm{III}} = 0$. This condition generates a wide plateau in the power spectrum, raised in amplitude with respect to CMB values because of the preceding USR phase. The subsequent transition at N_{III} from $\eta_{\mathrm{III}} = 0$ to $\eta_{\mathrm{IV}} < 0$ is necessary to end inflation.

The presence of the plateau in the power spectrum provides the concrete possibility to have a mass distribution of PBHs that covers many orders of magnitude.

What is actually crucial for our analysis is the precise form of the power spectrum at the two edges of the aforementioned plateau. The left-side edge is shaped by curvature modes that cross the horizon during and immediately after the USR phase while the right-side edge is shaped by curvature modes that cross the horizon during and immediately after the transition at $N \simeq N_{\mathrm{III}}$.

At the left-side edge of the plateau, the power spectrum is characterized by a bumplike feature (cf. Ref. [72] for a detailed discussion about its formation). This bump provides the link with our numerical analysis. To make this point more transparent, in the right panel of Fig. 13 we enlarge the bumplike feature at the left-side edge of the plateau.

It is instructive to compare the curvature power spectrum with the 90% C.I. posterior derived from our numerical analysis assuming GW190814 is a PBH binary. From this comparison we see that the curvature power spectrum features a cutoff at small wavelengths. In full analogy with the previous case (cf. Fig. 12, left), this cutoff is naturally generated by the sharp transition ($\mathcal{P}_\zeta \sim k^4$) that, because of the USR phase, brings the power spectrum from CMB values up to the typical amplitudes $\mathcal{O}(10^{-2})$ that are needed to generate PBHs. After the initial $\sim k^4$ growth, the power spectrum decreases before it settles to the constant value of the plateau. This decreasing part of the bump plays the role of the red-tilted power spectrum found in our numerical analysis. This is evident from the comparison shown in the
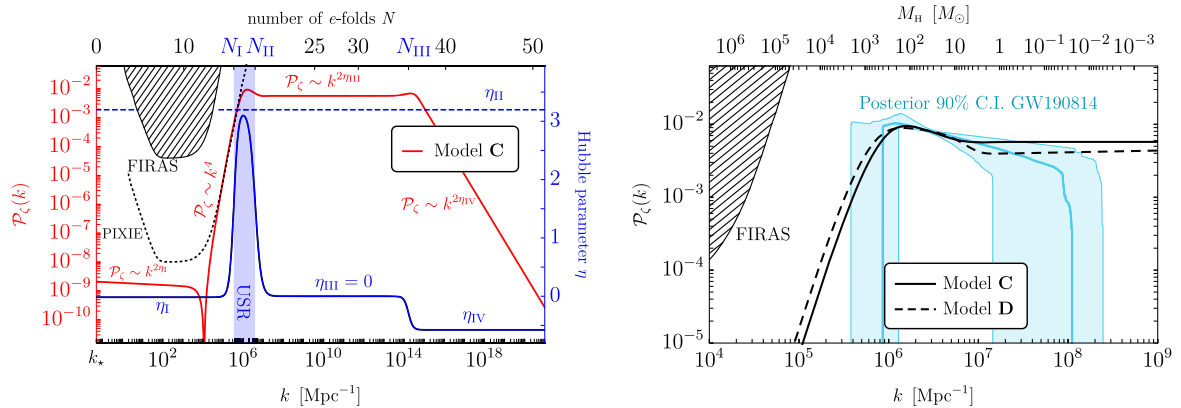

FIG. 13. Left panel: same as in Fig. 11 but for model C in Table VI. Right panel: we enlarge the left-side edge of the plateau (same as in the left panel of Fig. 12). In addition to model C we also show (dashed black line) the curvature power spectrum of model D in Table VI.

right panel of Fig. 13 between the posterior distribution and the power spectrum of model C. It should be noted that the power spectrum of model C is constrained to match the power-law behavior of the posterior distribution, in particular in the interval of comoving wave number where the latter is almost precisely nailed down by the numerical analysis (at about $m_{PBH} \approx 20 M_\odot$, i.e., the primary mass of GW190814). On the contrary, away from this k-interval deviations are possible. This is consistent with the fact that the numerical analysis is practically insensitive to k_{max}. In the explicit realization of our model discussed in Sec. VIII A, this freedom was exploited to directly connect the power spectrum to the last part of the dynamics that ends inflation (cf. Fig. 11). In the present scenario, we exploit the same freedom to connect the bump to the subsequent plateau.

At the right-side edge of the plateau, the curvature power spectrum is characterized by a second bumplike feature (cf. Ref. [72] for a detailed discussion about its formation). As in Ref. [72], we will exploit the bump at the left-side edge of the plateau for the generation of a solar-mass population of PBHs while the bump at the right-side edge of the plateau will be responsible for the generation of much lighter PBHs in the asteroid mass range. To this end, Eq. (18), together with the approximate horizon crossing condition $N = \log(k/k_\star)$, gives a good intuition about how to choose the values of N_I and N_{III}. We compute the full PBH mass distribution following the formalism introduced in Sec. II F. Intuitively, the PBH abundance roughly scales as $\approx \exp\left[-1/\mathcal{P}_\zeta(k)\right]$ and in our model it will be dominated by the two bumps of $\mathcal{P}_\zeta(k)$ where the latter takes its largest values. In addition, the abundance of heavier PBHs will be further boosted by the

effect of the QCD phase transition. We expect, therefore, two peaks in the mass distribution of PBHs, one in the solar mass range and the other in the asteroid mass range. In between the two peaks, we expect the typical redshift-induced scaling $m_{PBH}^{-1/2}$ associated to scale invariant power spectra (because of the plateau in between the two bumps).

The above expectations are confirmed by the numerical result shown in Fig. 14. Model C corresponds to the mass distribution given by the solid black line. In the same figure, we also plot a fourth realization of our model, dubbed model D in Table VI, that, contrary to the previous case, is characterized by $\eta_{III} \neq 0$; the corresponding mass distribution is given by the dot-dashed black line and features, as expected, a violation of the scaling $m_{PBH}^{-1/2}$ in between the two peaks. The rationale behind the different choice of η_{III} that distinguishes model C from model D is discussed in details in Appendix B.

Let us summarize here our findings:

(i) As discussed in Ref. [72], it is possible to tailor an USR dynamics that gives a population of asteroid-mass PBHs consistent with the abundance of dark matter observed in the present-day universe and, at the same time, a subpopulation of solar-mass PBHs. Remarkably, what we have shown with our analysis is that this subpopulation of solar-mass PBHs has the right features to explain a fraction of GWTC-3 events.

(ii) As in Sec. VIII A, the cutoff k_{min} (equivalently, M_L) arises naturally as a consequence of the sharp enhancement of the power spectrum (with respect to CMB values) that is essential for the generation of a sizable abundance of PBHs.
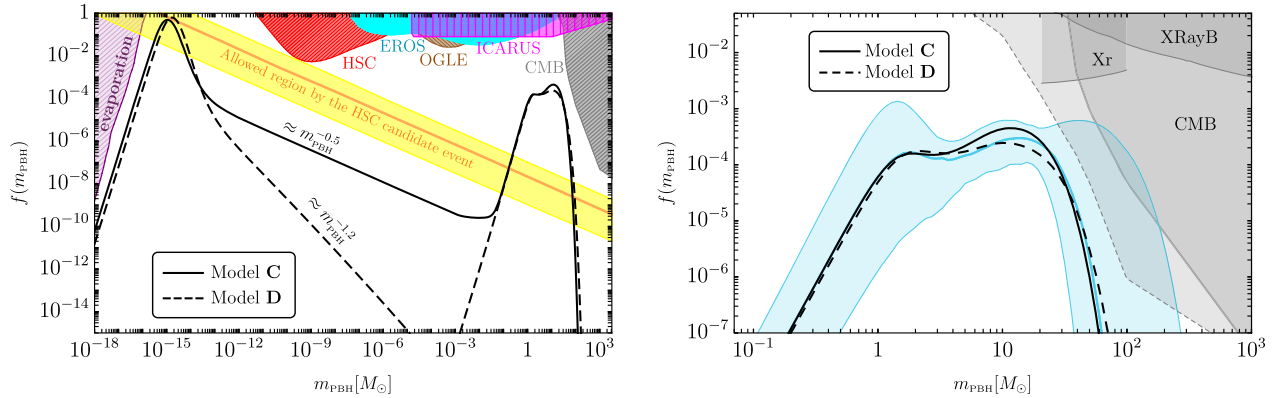

FIG. 14. Left panel: PBH mass distribution as a function of the PBH mass. We show the curves that correspond to model C and model D in Table VI. We show the following constraints (see Ref. [55] for a review and /bradkav/PBHbounds). Envelope of evaporation constraints (see also [226–228]): EDGES [229], CMB [230], INTEGRAL [231,232], 511 keV [233], Voyager [234], EGRB [235]; microlensing constraints from the Hyper-Supreme Cam (HSC) [236]; microlensing constraints from EROS [237]; microlensing constraints from OGLE [217]; Icarus microlensing event [216]; constraints from modification of the CMB spectrum due to accreting PBHs [20]. The yellow band corresponds to the allowed region for a PBH mass function $\propto m_{PBH}^{-0.5}$ consistent with the HSC microlensing candidate event [236] (see also [238]). Right panel: same as the left panel but we enlarge the solar-mass region. To guide the eye, we add the posterior distribution assuming GW190814 is interpreted as a PBH binary (cyan band, same as in Fig. 9).

(iii) The bulk of the PBH distribution in the solar-mass range is given by the bump at the left-side edge of the plateau in the curvature power spectrum (cf. the right panels of Figs. 13 and 14). The form of this bump is shaped by curvature modes that cross the horizon during and immediately after the USR phase. Despite its simplicity, Eq. (10) captures well the form of the bump. In particular, the red tilt $n_s < 1$ is absolutely crucial since it models the transition between the k^4 growth of the power spectrum and the subsequent plateau.

(iv) The PBH mass distribution in the solar-mass range is practically insensitive to the cutoff $k_{\max}$ (equivalently, M_s). We exploit such freedom to connect the part of the power spectrum that matches the ansatz in Eq. (10) with the plateau that in our model bridges solar- to asteroid-mass PBHs.

Let us mention that, while in this draft we focused on a model of USR inflation, the key ingredient is the peculiar shape of the power spectrum. Thus, we expect similar results would hold for any early universe model that can produce a similar curvature power spectrum.

C. Reconstructed inflaton potential

Once the Hubble parameters ϵ and η are determined, one can derive the inflationary potential by computing [72]

$$V(N) = V(N_{\rm ref}) \exp\left\{-2\int_{N_{\rm ref}}^{N} dN' \left[\frac{\epsilon(3-\eta)}{3-\epsilon}\right]\right\}, \quad (68)$$

$$\phi(N) = \phi(N_{\rm ref}) - \int_{N_{\rm ref}}^{N} dN' \sqrt{2\epsilon}. \quad (69)$$

The combination of $V(N)$ and $\phi(N)$ allows reconstructing the profile of the inflationary potential $V(\phi)$ in field space. Equation (68) highlights the advantages of our approach, based on parametrizing the inflationary dynamics in terms of the Hubble parameters, as in Eq. (66). As both ϵ and η enter in the exponent of Eq. (68), their determination is free from the fine-tuning necessary when working directly on a parametrization of the potential.

In Fig. 15 we show the reconstructed inflationary potentials in the case of model A and model C, cf. Ref. [72] for details. In both cases, we denote as $\phi_{\rm ref}$ the field value at which we fit CMB observables [and define $V_{\rm ref} = V(\phi_{\rm ref})$]. Both models exhibit the presence of a transition region that corresponds to the USR phase. The blue band limits, in field space, the e-fold time interval $N_{\rm II} < N < N_{\rm I}$. It is interesting to notice that the inflaton velocity is drastically reduced by the USR phase enhancing the power spectrum. As a consequence, generating the extended plateau in Fig. 13 (which is absent in models A and B) only requires the inflaton to remain in the second slow-roll configuration for a very short displacement in field space (still related to numerous e-foldings). Therefore,
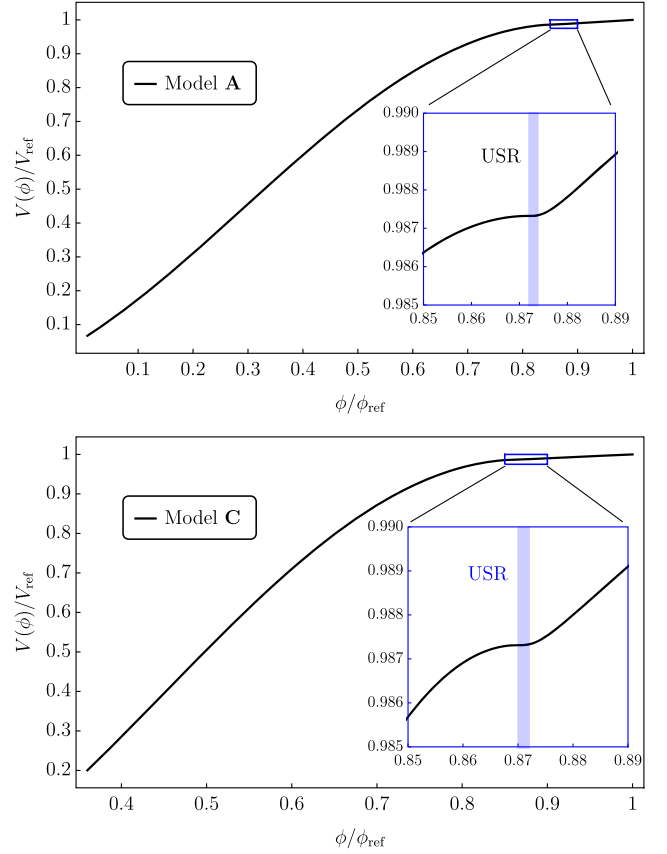

FIG. 15. Reconstructed inflationary potentials computed by means of the approach discussed in Ref. [72] in the case of model A (top panel) and model C (bottom panel). In the inset plots we enlarge the USR region.

small modifications to the inflaton potential are needed in order to generate the various models, as shown in Fig. 15.

We now move to describe the phenomenological consequences of our model as far as stochastic GWs are concerned.

D. Predictions: Stochastic GWs from PBHs

Once the free parameters of our model have been fixed by the condition $f_{\rm PBH} = \mathcal{O}(1)$ and the consistency with the posterior spectrum of curvature perturbation compatible with the primordial interpretation of GW190814, we are in the position to compute the *predicted* signal of scalar-induced stochastic GWs [239–247] and the SGWB produced by PBH mergers [28,46,248].

As shown in Refs. [72,220], the scalar-induced GW signal in the case of a very broad power spectrum, like the ones we are considering in models C and D, covers the wide range of frequencies $10^{-9} \lesssim f/{\rm Hz} \lesssim 1$, and its amplitude is both compatible with the putative signal recently reported by the NANOGrav collaboration [249] at about $f = \mathcal{O}(10^{-9})$ Hz (also independently supported by other
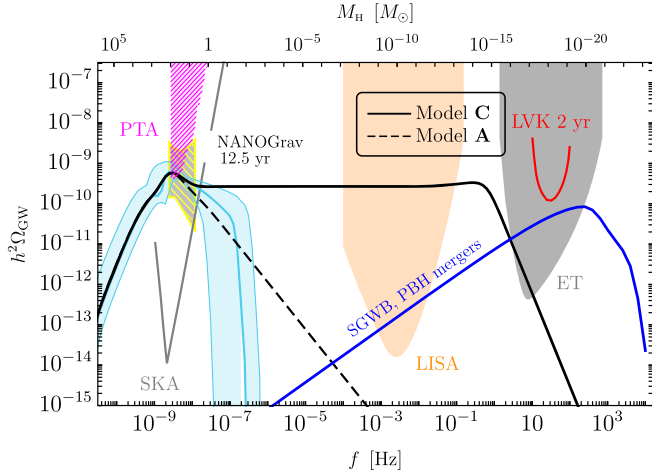
FIG. 16. Fraction of the energy density in GWs relative to the critical energy density of the Universe as a function of the frequency. We show the power-law integrated sensitivity curves of future ground- and space-based GW experiments (two-years observation with LVK at design sensitivity, the Einstein Telescope and LISA, cf. Ref. [258]) as well as previous Parkes Pulsar Timing Array (PTA) constraint [259], NANOGrav putative band [249] and SKA projected sensitivity [260]. We plot the signals predicted by our model in the two realizations A and C proposed in Table VI.

Pulsar Timing Array experiments [250–252])[10] and detectable by future space-based GW interferometers like LISA [254] (in the interval $10^{-4} \lesssim f$ [Hz] $\lesssim 10^{-1}$). We confirm this expectation in the case of our models in Fig. 16. The scalar-induced signal of GWs is proportional to $\mathcal{P}_\zeta^2(k)$ [see Eq. (71) below] and, therefore, it inherits its shape. The bump at the left-side edge of the plateau falls precisely inside the contour favored, at the 2-σ level, by the putative signal reported by NANOGrav (see also [255–257]); this is interesting because it means that our dynamics may predict a peculiar frequency dependence that could be tested by future pulsar timing array measurements.

For completeness, we also show the stochastic signal of scalar-induced GWs computed using the posterior distribution, expressed in terms of the parameters of the power spectrum in Eq. (10), assuming GW190814 is a PBH binary (cyan region). Interestingly, we note that the part of the signal that is most constrained by the numerical analysis falls precisely in the frequency band of the NANOGrav region.

It should be noted that, as in Refs. [72,220], we compute the scalar-induced GW signal in Fig. 16 assuming a radiation-dominated universe while a more accurate

---

computation should include the effect of the quark-hadron phase transition through the change of the number of effective degrees of freedom, the EOS parameter w, and the sound speed c_s.

To comment more quantitatively on this point, let us write the amplitude of induced GW spectral density measured today in the form [247]

$$\Omega_{\text{GW},0}h^2 = 0.39\Omega_{r,0}h^2 \underbrace{\left[\frac{g_*(T_H)}{106.75}\right]\left[\frac{g_{*,s}(T_H)}{106.75}\right]^{-\frac{4}{3}}}_{\equiv c_g(T_H)}\Omega_{\text{GW},H}, \quad (70)$$

with

$$\Omega_{\text{GW},H} = \left(\frac{k}{k_H}\right)^{-2b}\int_0^\infty dv \int_{|1-v|}^{1+v} du\,\mathcal{T}(u,v)\mathcal{P}_\zeta(ku)\mathcal{P}_\zeta(kv),$$
$$(71)$$

where $b \equiv (1-3w)/(1+3w)$, $\Omega_{r,0}$ is the density fraction of radiation, $g_*(T)$ and $g_{*,s}(T)$ the temperature-dependent effective degrees of freedom for energy density and entropy density, $\mathcal{T}(u,v)$ the transfer function that fully depends on the universe EOS; the subscript $_H$ stands for the time when induced GWs of given wave number k are sufficiently inside the cosmological horizon to be treated as a radiation fluid in an expanding universe.

There are two effects induced by the thermal history of the universe across the QCD era. First, $c_g(T_H)$ is constant and equal to unity only for perturbation modes reentering the Hubble horizon deep in the radiation epoch; as the left-side edge of the curvature power spectrum reenter the Hubble horizon at around the quark-hadron phase transition, the reduction of g_* and $g_{*,s}$ induces a modulation of the SGWB spectrum. In Fig. 17 we show the evolution of $c_g(T)$ trading its temperature dependence for the dependence on the horizon mass M_H (top x axis) as [15]

$$M_H = 1.5 \times 10^5 M_\odot \left[\frac{g_*(T)}{10.75}\right]^{-1/2}\left(\frac{T}{\text{MeV}}\right)^{-2} \quad (72)$$

and the comoving wave number k (bottom x axis) using Eq. (18). To guide the eye, we superimpose the frequency range (translated into a wave number interval by means of $f = k/2\pi$) favored by the putative NANOGrav signal. We conclude that modeling the temperature dependence of the factor c_g enhances the GW signal in the low-frequency part of the spectrum relevant for the comparison with pulsar timing array data (about a factor 2 but with some frequency dependence). The second physical effect is induced by the dependence of $\mathcal{T}(u,v)$ on both w and c_s in Eq. (71).

Both effects have been discussed in Ref. [261] (see also Ref. [262]) specifically addressing the thermal history induced by the QCD phase transition. For a scale-invariant power spectrum, it turns out that the evolution of w and c_s
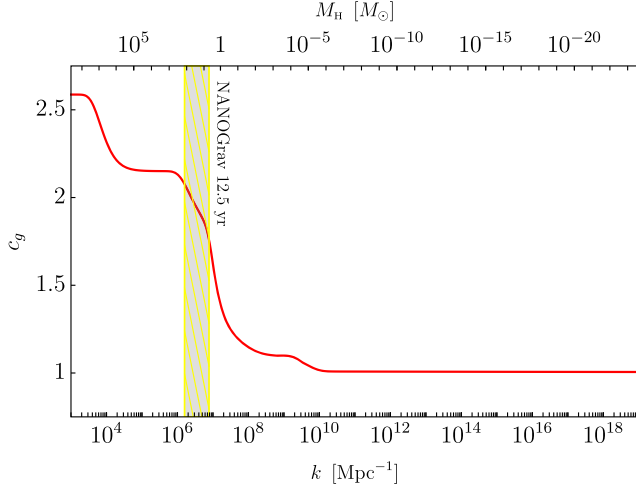
---

[10]The band compatible with recent NANOGrav observations is in partial tension with previously derived PTA constraints. According to the NANOGrav collaboration [253], the improved priors for the intrinsic pulsar red noise used in the most recent analysis relaxes previous bounds.

FIG. 17. Evolution of the factor c_g defined in Eq. (70) as a function of the horizon mass M_H (top x axis) and the comoving wave number k (bottom x axis). The yellow region shaded in gray marks the frequency range (converted into $k = 2\pi f$) $2.4 \times 10^{-9} \leq f$ [Hz] $\leq 1.2 \times 10^{-8}$ favored by the NANOGrav putative signal.

only induces a subleading modification with respect to the effect of the changing effective degrees of freedom. However, in view of future tests of the putative NANOGrav signal, it would be certainly important to include, following Ref. [261], the full effect of the QCD phase transition in the computation of the spectrum of induced GWs, and nail down more precisely the frequency dependence of the signal that our inflationary dynamics predicts in the range relevant for pulsar timing array measurements. We leave this investigation for future work.

In Fig. 16 we also show the SGWB produced by the population of mergers in the solar mass range, again under the assumption of GW190814 being a primordial binary and adopting the best-fit values from Table V. We do not show the astrophysical contribution as it strongly depends on the rate evolution above the peak expected around redshift $z \approx 2$, following the star formation rate [263]. This can, therefore, be regarded as a lower bound on such a background from mergers in the stellar mass range.

We compute the spectrum at frequency ν by integrating the PBH merger rate across the cosmological history as

$$\Omega_{\mathrm{GW}}(f) = \frac{f}{\rho_0} \int_0^{f_3/f - 1} dz \frac{\mathcal{R}_{\mathrm{PBH}}}{(1+z)H(z)} \frac{dE_{\mathrm{GW}}(f_s)}{df_s}, \quad (73)$$

in terms of the redshifted source frequency $f_s = f(1+z)$, the present energy density $\rho_0 = 3H_0^2/8\pi$, the Hubble constant H_0, and the energy spectrum of GWs denoted $dE_{\mathrm{GW}}/d\nu_s$. Notice that Eq. (73) implicitly requires an integration over $m_{1,2}$. Finally, f_3 controls the maximum redshift beyond which mergers cannot contribute to a given

spectral frequency f and it is determined by the effective cutoff of the spectrum (see Appendix A of Ref. [46] and references therein for more details).

It may be possible to distinguish the contribution to the SGWB coming from either PBH or ABH/NS mergers thanks to their predicted different merger rate evolution. The SGWB results from the integrated contribution of the merger history [264–268], and PBHs are characterized by an extended rate growth reaching much before star formation. Therefore, given the same detection rate of resolved binaries at low redshift, a PBH contribution produces a larger SGWB. Correlating rates of individual detections and the SGWB amplitude may allow to set a lower bound on the primordial contribution at future third-generation experiments [46] (see also [269,270]).

The peak frequency of GWs emitted from BH mergers is close to innermost stable circular orbit frequency, $f_{\mathrm{ISCO}} \simeq 4.4 \times 10^3$ Hz$(M_\odot/(m_1 + m_2))$. As the solar-mass and intermediate-mass PBH population is bounded to be below $\mathcal{O}(10^2)M_\odot$ by CMB accretion constraints (cf. Fig. 9) and eventually by FIRAS/PIXIE data, the SGWB cannot get sizable contributions at frequencies smaller than $\mathcal{O}(10)$ Hz, if not from the $\approx f^{2/3}$ tail produced by the inspiral phase. Therefore, the contribution to the SGWB from PBHs with masses smaller than $\mathcal{O}(10^2)M_\odot$[11] falling in the LISA band cannot overcome the one induced at second order by the formation of an asteroid mass population of PBHs explaining the dark matter [273,274].[12] Finally, we neglect the second peak potentially generated by the asteroidal mass PBHs which would fall at much higher frequencies, of interest for UHF-GW experiments [275] (see in particular Ref. [127] and references therein).

IX. CONCLUSIONS AND OUTLOOK

We have performed the first Bayesian PBH population inference on GW data directly using *ab initio* curvature power spectrum parameters and including the effect of the modified threshold due to the QCD EOS [16]. We critically confronted this state-of-the-art PBH model with LVK phenomenological population models that describe the GWTC-3 catalog both in the NS and in the BH mass range.

We found that the upper bound on the PBH abundance is consistent with previous analyses ($f_{\mathrm{PBH}} \lesssim 10^{-3}$) and it is stronger than other constraints in this mass range. Nonetheless, we also found marginal evidence for extra information in the data on top of the LVK phenomenological distributions, which may be captured by a

---

[11]One also expects GWs signals in the LISA band from mergers of supermassive BH binaries [271,272], which we do not quantify in Fig. 16. As the majority of those mergers will be resolved, and subtracted, they would marginally contaminate SGWB searches.

[12]These constraints were neglected in Refs. [66,67].

primordial subpopulation of binaries. Indeed, a PBH subpopulation can explain a fraction of GWTC-3 events, in particular binaries with light (such as the lower mass-gap event GW190814) or heavy (e.g., GW190521) components. Interestingly, the light events that are assigned the highest PBH likelihood by our inference happen also to be those which are more challenging to accommodate within standard astrophysical scenarios.

Intriguingly, our *ab initio* PBH distribution allows us to make some falsifiable predictions: if some of the GWTC-3 events are primordial (in particular the lower mass-gap events GW190814, which is assigned ≈29% probability of being primordial by our inference), then the merger rates in the subsolar mass range and in the lower mass gap are high enough to be detectable by future LVK runs. In particular, the absence of subsolar mergers in O5 would automatically exclude the primordial origin of the light events within GWTC-3.

Our work is just a first attempt to use *ab initio* PBH models in GW population inference, and we hope it will be extended in several ways. Most importantly, one should perform multipopulation Bayesian inference by mixing our PBH model with astrophysical models for BH and/or NS binaries, similarly to what was recently done in Ref. [45].

It is also possible to improve the PBH modeling, in particular by considering a different parametrized curvature spectrum (e.g. peaked Gaussian bump etc. [111,224]), primordial non-Gaussianities [103–106,276,277], accretion effects [37,130,131], and spin information in the inference [132]. Eventually, extending the numerical simulations of Ref. [16] used here for a sufficient set of shapes of the collapsing overdensities would allow capturing the threshold and mass dependence on deviations from the nearly scale invariant spectra (i.e., n_s), allowing us to include a full dependence of parameters of collapse on each specific spectral mode (or M_H in our formalism) beyond the effect of the QCD EOS.

On the theory side, building on the reverse engineering approach recently devised in Ref. [72], we have mapped the GW data-driven curvature power spectrum into an USR inflationary model. We remark that the reverse engineering approach used in this analysis goes beyond the mere parametrization of the dynamics given in Eq. (66) since it allows to numerically reconstruct the inflationary potential, which we showed in Fig. 15 in the case of models A and C. It is legitimate to ask whether it would be possible to move directly to the analysis based on some scalar potential and skip the reverse engineering approach of Ref. [72]. The answer is certainly positive; however, working directly at the level of the potential may not give the same control on the shape of the power spectrum compared to the reverse engineering approach, thus making the analysis much more difficult and way less transparent. Ultimately, it may be possible to run the population inference directly on the

fundamental coupling constants of a given inflationary model and investigate the possible quantum field theory origin of the reconstructed potential. We leave these tasks for future work.

We also confirmed a remarkable feature of this approach [72], namely that a single USR model can consistently accommodate a double-peaked PBH mass function. The dominant peak occurs in the asteroid-mass range and it is responsible for explaining the totality of the dark matter in small PBHs, while the second (subleading) peak is produced by the enhancement beyond the effect of the QCD phase transition and provides a detectable PBH merger rate in the band of current and future GW detectors.

An important by-product of our analysis is that the inferred value of the (red) tilt of the spectrum makes the above scenario fully compatible with the absence of subsolar mergers in GWTC-3, although it also predicts that subsolar mergers and more lower-mass-gap events can be detectable in the future.

Finally, we showed that other falsifiable predictions of the designed curvature power spectrum are: (i) a detectable scalar-induced SGWB signal compatible with the NANOGrav putative measurement and detectable by future PTA observations and by LISA; and (ii) a SGWB produced by PBH mergers which will be detectable by the Einstein Telescope. In both cases, an urgent extension of our work is to properly account for the full richness of the *ab initio* PBH model (including the effects of the QCD phase transition) in shaping the frequency dependence of these SGWB signals. These advancements are required in order to fully exploit the constraining power of GW data soon to be available.

ACKNOWLEDGMENTS

Note added.—Recently, we became aware of Ref. [279], which independently explores the role of the QCD phase transition in shaping the PBH mass function and PBH merger rate distribution. Reference [279] provides a detailed numerical description of the threshold for PBH formation, which looks to be in reasonable agreement with the results of Ref. [16] used in our analysis. However, the computation of the mass distribution in [279] does not include the dependence of the critical collapse and density variance on the EOS, leading to an approximate mass function which does not take into account the additional pileup of PBHs around the solar mass produced by the critical collapse. More importantly, they restrict the discussion to a nearly scale invariant shape of the enhanced spectrum at small scales with tilt $n_s = 0.965$–0.975, and do not compare the PBH model to GWTC-3 data through a Bayesian analysis. The values of $f_{\rm PBH}$ considered in Ref. [279] are larger than ours by 2–3 orders of magnitude and compatible with $f_{\rm PBH} \approx 1$. This discrepancy can be attributed to various differences, which we list in the following: (i) a different suppression factor in the PBH merger rate formula is used in Ref. [279]. While Ref. [279] used the analytical treatment of Refs. [57,66], we adopt the results of Refs. [31,34,118] (informed by N-body simulations), which give smaller values of $f_{\rm PBH}$, compatible with the analyses of Refs. [29,38–40,42,45] and also not excluded by other non-GW constraints. (ii) Reference [279] restricts the

parameter space to masses above $m_i > M_\odot$ and mass ratio larger than $q \gtrsim 0.1$ when computing[13]

$$\frac{{\rm d}R_{\rm PBH}}{{\rm d}m_1} \equiv \int {\rm d}m_2 \times 2\theta(m_1 - m_2) \times \frac{{\rm d}R_{\rm PBH}}{{\rm d}m_1 {\rm d}m_2}. \quad (74)$$

This leads to a drastic reduction of the differential rate, which however would only be justified if LVK were unable to detect such neglected events. Using their approximated mass distribution and their choice of merger rate formula with $f_{\rm PBH} = 1$, we find LVK would have observed around $N_{\rm det}(m_i < M_\odot) \approx 20$ mergers with at least one subsolar component during GWTC-3. With our modeling of the mass distribution, while fixing $n_s = 0.97$ and $f_{\rm PBH} = 1$, we obtain $N_{\rm det}(m_i < M_\odot) \simeq 36$. This means that the absence of subsolar detections in the various LVK runs is incompatible with such mass function and large values of the abundance, as already pointed out in Ref. [69]. This constraint is automatically included in our MCMC analysis. (iii) Finally, our inclusion of a dominant contribution from ABH mergers to the GWTC-3 catalog only has a minor impact on the constraint on $f_{\rm PBH}$ (as shown by our single population analyses).

APPENDIX A: POSTERIOR DISTRIBUTIONS

Here we present the posterior distributions resulting from the various Bayesian inferences. While the main phenomenological consequences of these posteriors are described in the main text, here we report a few interesting insights on these distributions, also highlighting relevant correlations between parameters. We note that, in order to simplify the notation, we report mass scales in solar mass units $[M_\odot]$ and rate densities in units of $[{\rm Gpc}^{-3}\,{\rm yr}^{-1}]$.

We start with the phenomenological NS or ABH channels. Focusing first on light events with $m_2 \leq 3M_\odot$ (i.e., left panel of Fig. 18), which is based on fitting only seven detections, we observe that the posterior selects narrow mass distributions centered around the solar mass scale, while the mass cutoffs are poorly constrained beyond the basic requirement of encompassing all the events in this mass range. This can be observed by noticing that $m_{\rm min}^{\rm NS}$ and $m_{\rm max}^{\rm NS}$ flatten reaching the lower, or upper, boundaries of their respective prior range. The latter, in particular, is forced to be above $m_{\rm max}^{\rm NS} \approx 2.5M_\odot$ to include the secondary mass of GW190814.

The posterior distribution for the ABH model (i.e., right panel of Fig. 18) is instead better constrained, due to the larger number of detections (i.e., 69 in GWTC-3). Similar conclusions as for the NS case can be drawn on the

---

[13]Also, Ref. [279] neglected the factor 2 when deriving the merger rate as a function of primary (heavier) mass due to the two possible mass orderings in Eq. (37), which allows for both $m_1 < m_2$ and $m_2 < m_1$ by construction [31].
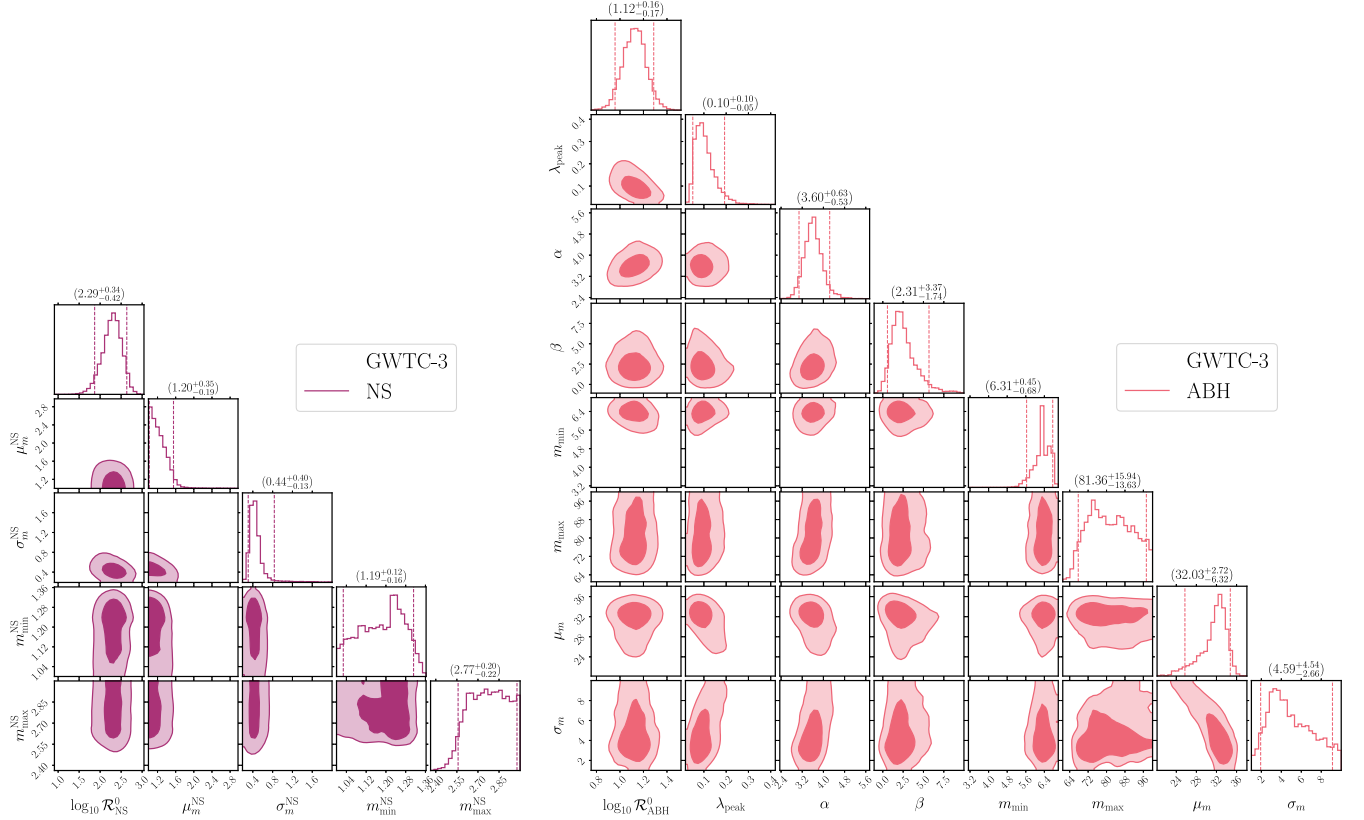
FIG. 18.   Posterior distribution for the hyperparameters characterizing the LVK NS (left) and ABH (right) phenomenological models, obtained within a single-population inference of the GWTC-3 catalog.

minimum and maximum scales bracketing the ABH population $m_{\min}$ and $m_{\max}$. In particular, the latter is bounded to be above $\approx 70 M_\odot$ to capture the mass-gap event GW190521, whose primary mass is measured to be $m_1 = 95.3^{+28.7}_{-18.9} M_\odot$ [59] (see Table III). Also, a distinct anticorrelation is observed between the central scale and the width of the Gaussian peak, accounting for a small fraction of the intrinsic population of mergers. This is most probably enforced by the requirement of not overproducing mergers in the heavy tail (i.e., $m_1 > \mu_m$) of the Gaussian contribution.

These general features are consistent with the results of the LVK population analysis reported in Ref. [59]. While the overall posterior is fully compatible with LVK findings, slight deviations are observed, most probably introduced by the omission of spin information in our inference, the absence of subdominant smoothing terms enforced at the tails of the mass distribution of the LVK model, and the adoption of a selection bias solely based on SNR computations, see Sec. V B (instead of the one based on the LVK injection campaign). The latter choice, which is customarily adopted in the recent literature (see e.g. [45,156]), is required in our setting, as our analysis necessitates consistently computing the selection bias also in the subsolar mass range, which is not captured by the LVK injection campaign [280] and which is crucial to enforce the

constraint on the PBH population from the absence of subsolar detections in GWTC-3. We do not expect these approximations to impact our results.

In Fig. 19 we report the analogous posteriors obtained assuming the PBH population alone explains the population of mergers. In the left panel, we report the result assuming an *ab initio* mass distribution of PBHs derived from the curvature spectrum in Eq. (10) and the effect of the QCD epoch. The gray (cyan) color indicates the result of the inference on the light (heavy) events. The much larger uncertainties observed in the gray posterior is due to the aforementioned smaller sample of events with $m_2 < 3 M_\odot$. Strikingly, both analyses provide similar best-fit values for the hyperparameters λ_{PBH}, apart from M_S which is unbounded from below—and allows for the presence of the QCD induced bump just above $M_\odot$—in the first case, while it is constrained to be $\log_{10}(M_S/M_\odot) = 0.95^{+0.11}_{-0.19}$ in the second case. The correlation between n_s and M_L can be explained by noticing that the smaller values of the tilt, corresponding to *redder* spectra, enhance $\psi(m_{\mathrm{PBH}})$ at high masses, and the high mass cutoff M_L needs to adjust to reduce the prominence of heavy mergers. Both analyses constrain the abundance to be much smaller than unity in this mass range, namely $\log_{10} f_{\mathrm{PBH}} = -2.70^{+0.32}_{-0.25}$ and $\log_{10} f_{\mathrm{PBH}} = -2.78^{+0.08}_{-0.07}$, respectively.
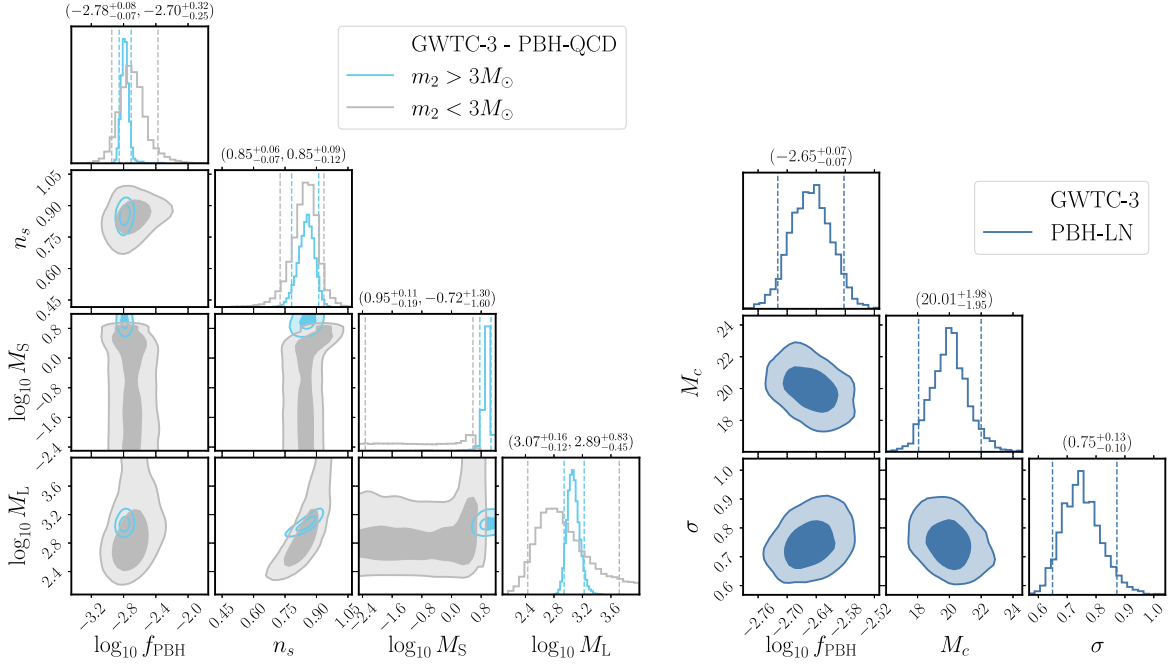
FIG. 19. Same as Fig. 18 but for the PBH model assuming a mass distribution obtained from first principles with the effect of QCD softening of the EOS (left) or a log-normal mass function (right). Different colors correspond to separate fits of the light (heavy) portion of the catalog, with events characterized by $m_2 < 3M_\odot$ ($m_2 > 3M_\odot$) (canonically referred to events containing NS or not by the LVK analysis).

In order to fully compare the constraint on the PBH abundance obtained with this single population analysis of GWTC-3 with previous literature, we also repeat the inference assuming the PBH population is described by a log-normal mass distribution of the form (e.g., [281])

$$\psi(m_{\rm PBH}) = \frac{1}{m_{\rm PBH}\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\log^2(m_{\rm PBH}/M_c)}{2\sigma^2}\right], \quad \text{(A1)}$$

where M_c is the central mass scale and σ the width. The right panel of Fig. 19 shows that the best-fit values of such a scenario are consistent with results previously derived in the literature [29,31,37–40] under analogous assumptions but with older datasets. In particular, the mass distribution is found to be broad and peaked at $\approx 20 M_\odot$. However, as discussed in the main text, such shape would overproduce mergers in the heavy portion of the catalog and give rise to a flat distribution of mass ratio, in sharp contrast with what was observed in the data; see the detailed discussion in Sec. VI A. Finally, assuming a log-normal mass distribution gives a slightly less stringent, but statistically compatible, bound on the PBH abundance, which is found to be $\log_{10} f_{\rm PBH} = -2.65^{+0.07}_{-0.07}$.

Let us conclude this Appendix by discussing the result of the mixed population inference. The corresponding posterior distribution is shown in Fig. 20. In this case, the dataset includes all 76 detections in GWTC-3, and allows for ABH, NS and PBH mergers (with a QCD induced mass distribution) to contribute to the population of mergers.

First, we notice that the ABH and NS models are mostly uncorrelated with each other, as can be observed by focusing the bottom-left 8×5 box. This is because they explain different sets of events, and a crosstalk between them would only be introduced by a dominant contribution from PBH mergers that, instead, can cover both mass ranges. Second, bimodal distributions are observed in various mass cutoffs. In particular, secondary peaks appear in the distributions of $m_{\rm min}$, $m_{\rm max}$, and $m_{\rm max}^{\rm NS}$ when extreme events, potentially outliers of the astrophysical populations, such as GW190814, GW190924_021846 and GW190521, are explained by the PBH channel, respectively. In the portions of the posterior where PBHs are *necessary* to explain the various special events, the PBH abundance $f_{\rm PBH}$ is found to be bounded from below and takes values around $f_{\rm PBH} \approx 10^{-3}$ (see also Table V). We also observe that the posterior shows small support for the simultaneous interpretation of GW190814 and GW190521 as PBH mergers, see the $(m_{\rm max}, m_{\rm max}^{\rm NS})$ [or (6,17)] panel of the posterior.

Finally, we observe that the hyperparameters of the PBH population are all characterized by a pronounced peak, corresponding to the high likelihood regions where PBHs contribute to the observations and improve the fit (see discussion in Sec. VI B). However, $f_{\rm PBH}$ is not bounded from below and has a tail reaching the left boundary $f_{\rm PBH} = 10^{-6}$, where PBH contribution is negligible. This also implies that the remaining parameters have posterior distributions with broad tails filling the prior volume, with the small mass scale subject to the condition $M_{\rm S} < M_{\rm L}$.
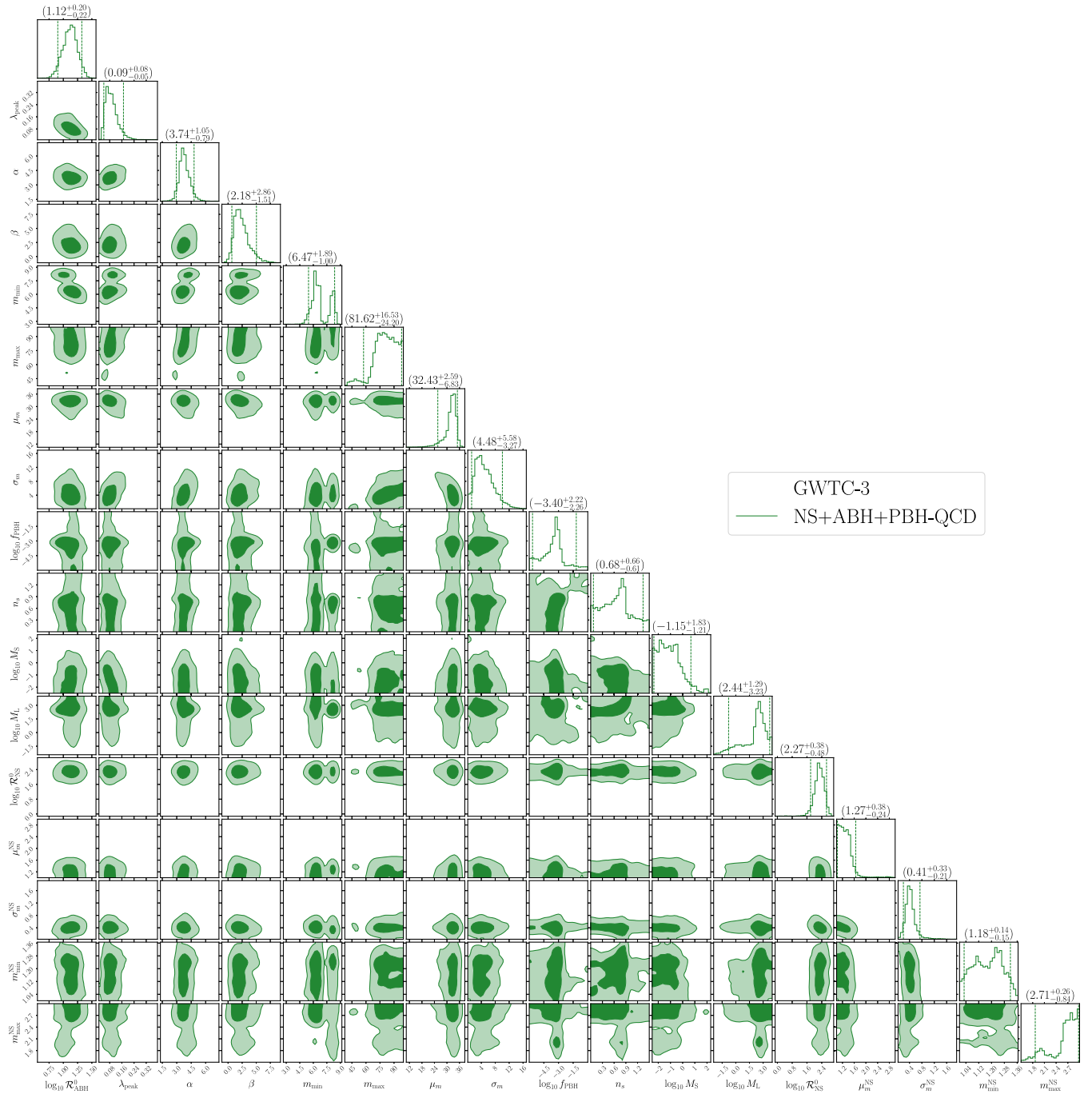
FIG. 20. Posterior distribution for the hyperparameters in the mixed NS + ABH + PBH analysis where the primordial channels assume the *ab initio* mass distribution shaped by the QCD epoch.

A correlation between n_s and M_L is found also in this case, due to the requirement of not overproducing heavy mergers. We conclude by highlighting that the tail at large values of f_{PBH} reaching unity is strongly correlated with high values of $n_s > 1$ (blue tilts) and small M_S. This is because one can evade the constraint from LVK measurements only with light enough populations strongly peaked at light mass scales below $\mathcal{O}(10^{-1})M_\odot$, where the LVK sensitivity strongly deteriorates. This is reflected in the bound on f_{PBH} as a function of the average mass $\langle m_{PBH} \rangle$ shown in Fig. 10.

APPENDIX B: PBHS WITH WIDELY DIFFERENT MASS, A TECHNICAL INSIGHT

This is a slightly technical appendix providing key details useful to fully understand the rationale behind the numerical values of the parameters chosen in Table VI.

The two realizations of our model discussed in Sec. VIII B 1 (models C and D) are tuned to give the totality of dark matter in the form of PBHs. Concretely, this means that the logarithmic integral of the mass distribution gives unity, cf. Eqs. (28) and (65). Since in the solar mass range the fraction of dark matter in the form of PBHs is constrained to be at most $\mathcal{O}(10^{-3})$, the integral is dominated by the peak in the asteroid mass range. In turn, this implies that the value of the curvature power spectrum in correspondence of the bump at the right side of the plateau should be high enough to get the desired order-one abundance of PBHs.

This is a nontrivial task to accomplish. The reason is that one should be careful to enhance the amplitude of the power spectrum at the right-side end of the plateau without also altering too much the amplitude of the left-side edge since, otherwise, the risk is to overproduce solar mass PBHs which are incompatible with LVK merger rates.

We envisage four possible ways to tackle this problem (see also Fig. 21):

(a) First, we consider the case in which we take $\eta_{\mathrm{III}} = 0$ and tune the value of $\eta_{\mathrm{II}} \Delta N_{\mathrm{USR}}$ appropriately to get $f_{\mathrm{PBH}} = \mathcal{O}(1)$. Furthermore, we fix the widths of the three transitions to the benchmark value $\delta N_{\mathrm{I}} = \delta N_{\mathrm{II}} = \delta N_{\mathrm{III}} = 0.5$.

The above tuning of $\eta_{\mathrm{II}} \Delta N_{\mathrm{USR}}$ basically corresponds to a rigid shift of the whole power spectrum towards larger values. Consequently, a larger abundance of asteroid mass PBHs will unavoidably enhance also the abundance of solar mass PBHs. Numerically, we find that (in this initial setup with $\delta N_{\mathrm{I}} = \delta N_{\mathrm{II}} = \delta N_{\mathrm{III}} = 0.5$) it is not possible to make the totality of dark matter in the form of PBHs without violating the constraints in the solar mass range (that is, without exceeding the allowed region in Fig. 9).
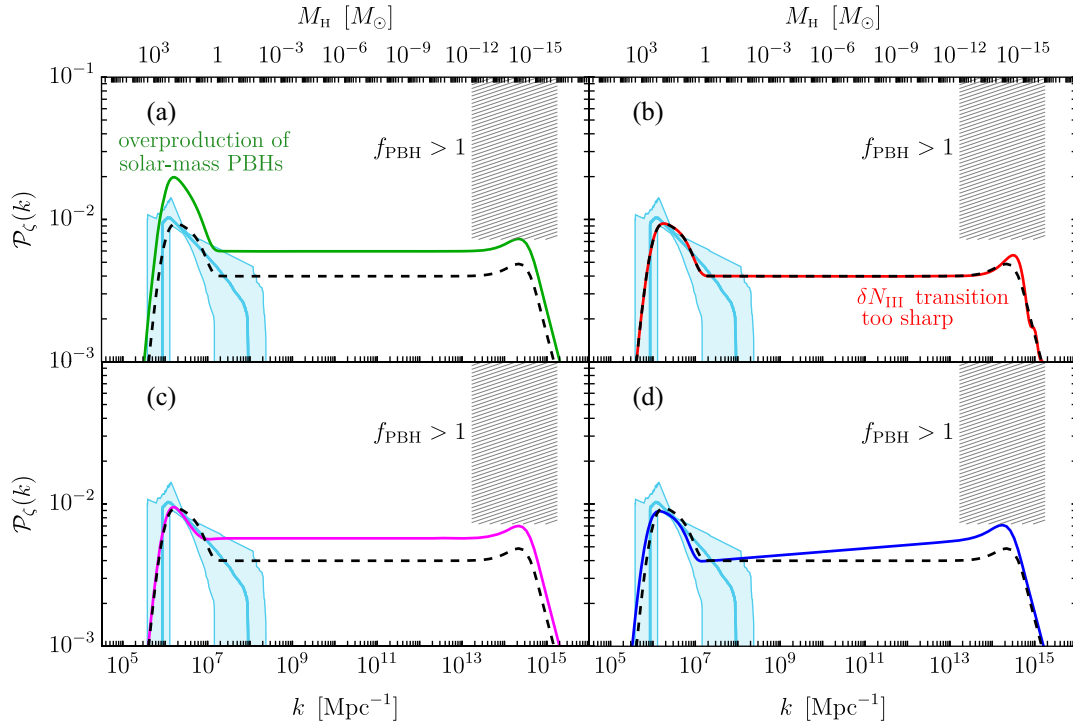

FIG. 21. Four examples for the enhancement of the abundance of asteroid mass PBHs while also confronting with the constraints in the solar-mass range. In all four panels, the dashed black line represents the curvature power spectrum (we enlarge the very top part) that corresponds to model C in Table VI but taking $\delta N_{\mathrm{I}} = \delta N_{\mathrm{II}} = \delta N_{\mathrm{III}} = 0.5$ and with $\Delta N_{\mathrm{USR}} \eta_{\mathrm{II}}$ tuned in order to make the abundance of solar mass PBHs compatible with Fig. 9. In this realization, the model gives $f_{\mathrm{PBH}} \ll 1$. The task, therefore, is to enhance the height of the bump at the right-side edge of the plateau [that is, enhance the abundance of asteroid mass PBHs in order to get $f_{\mathrm{PBH}} = \mathcal{O}(1)$] without altering the bump at the left-side edge of the plateau. (a) We take larger η_{II}; this change shifts the whole power spectrum towards larger values, and overproduces solar mass PBHs. (b) We take smaller δN_{III}; the tanh transition at N_{III} becomes too sharp, and this fact introduces additional non-Gaussianity in the computation of the PBH abundance. (c) We take larger η_{II} and larger δN_{II}; as in (a), increasing η_{II} shifts the whole power spectrum towards larger values. However, δN_{II} controls the amplitude of the bump at the left-side edge (without altering the rest of the spectrum). Increasing δN_{II} has the consequence of decreasing the amplitude of the left-side bump without altering the amplitude of the right-side one. As shown in the figure, from the combinations of these two effects one gets the desired enhancement in the amplitude of the right-side bump while the amplitude of the left-side bump is kept at the level of the dashed line. This is model C. (d) We take η_{II} slightly different from zero and positive. This introduces a tilt in the plateau that enhances the amplitude of the bump at the right-side edge without changing the left-side one. This is model D.

(b) To fix this problem, a possible way out is to keep $\eta_{\mathrm{II}}\Delta N_{\mathrm{USR}}$ fixed to some value that is compatible with constraints in the solar mass range and change the shape of the bump at the right side of the plateau. As discussed in Ref. [72], this is possible by tuning the value of δN_{III} (smaller values of δN_{III} make the bump more pronounced). However, we find that, in order to boost the abundance of asteroid mass PBHs to values $f_{\mathrm{PBH}} = \mathcal{O}(1)$, we need $\delta N_{\mathrm{III}} \ll 1$. The drawback is that such a sharp transition typically generates sizable non-Gaussianities that may threaten the validity of our computation of the abundance [106,282,283]. For this reason, we discard this possibility (in addition, it is unclear whether very sharp transitions in the evolution of η are realizable in concrete models).

(c) The third possibility is the one we adopted in model C. As in *(a)*, we take $\eta_{\mathrm{III}} = 0$ and tune the value of $\eta_{\mathrm{II}}\Delta N_{\mathrm{USR}}$ appropriately to get $f_{\mathrm{PBH}} = \mathcal{O}(1)$; as discussed, we end up with an overabundance of solar mass PBHs. However, as noticed in Ref. [72], the value of δN_{II} controls the height of the bump at the left-side edge of the plateau. In particular, increasing the value of δN_{II} decreases the amplitude of the bump. It is, therefore, sufficient to consider a slightly larger value of δN_{II} to smooth out the abundance of solar mass PBHs and get a perfect fit of $f_{\mathrm{PBH}} = \mathcal{O}(1)$.

(d) Finally, the fourth possibility is the one we adopted in model D. We keep $\eta_{\mathrm{II}}\Delta N_{\mathrm{USR}}$ fixed to some value that is compatible with the constraints in the solar mass range (in particular, compatible with the posterior in

Fig. 9). If we now take η_{III} nonzero and positive, the power spectrum will scale as $\mathcal{P}_\zeta \sim k^{2\eta_{\mathrm{III}}}$ in the region between the two bumps, and this will enhance the height of the bump at the right-side edge of the plateau without affecting the one at the left-side edge. Numerically, we find that values of η_{III} as small as few $\times 10^{-2}$ are enough to get the desired enhancement that gives $f_{\mathrm{PBH}} = \mathcal{O}(1)$.

A bonus possibility is to move the asteroid mass peak towards smaller masses in order to exploit the enhancement of the abundance due to the redshift factor $M_H^{-1/2}$ in Eq. (24). In our model this means taking larger values of N_{III}. However, we find that one quickly clashes with the constraint given by Hawking evaporation. For clarity's sake, we illustrate the four possibilities (a)–(d) in Fig. 21, see caption for details.

As a technical remark, we would like to emphasise the power of the parametrization in Eq. (66). As clear from the above discussion, the free parameters that enter in the evolution of η have a clear connection with the shape of the curvature power spectrum and, therefore, it turns out to be extremely simple to manipulate the dynamics and carve out the desired form of $\mathcal{P}_\zeta(k)$.

Since in model D η_{III} is not exactly zero, we expect, as anticipated, a violation of the redshift-induced scaling $m_{\mathrm{PBH}}^{-1/2}$ associated to scale invariant power spectra. Numerically, we find the power-law scaling $m_{\mathrm{PBH}}^{-1.2}$, cf. Fig. 14. To the contrary, in model C we have $\eta_{\mathrm{III}} = 0$. In this model, therefore, the scaling $m_{\mathrm{PBH}}^{-1/2}$ is recovered, as confirmed in Fig. 14.

[1] Y. B. Zel'dovich and I. D. Novikov, Sov. Astron. AJ (Engl. Transl.), **10**, 602 (1967).

[2] S. W. Hawking, Nature (London) **248**, 30 (1974).

[3] G. F. Chapline, Nature (London) **253**, 251 (1975).

[4] B. J. Carr, Astrophys. J. **201**, 1 (1975).

[5] P. Ivanov, P. Naselsky, and I. Novikov, Phys. Rev. D **50**, 7173 (1994).

[6] J. Garcia-Bellido, A. D. Linde, and D. Wands, Phys. Rev. D **54**, 6040 (1996).

[7] P. Ivanov, Phys. Rev. D **57**, 7145 (1998).

[8] S. Blinnikov, A. Dolgov, N. K. Porayko, and K. Postnov, J. Cosmol. Astropart. Phys. 11 (2016) 036.

[9] M. Shibata and M. Sasaki, Phys. Rev. D **60**, 084002 (1999).

[10] J. C. Niemeyer and K. Jedamzik, Phys. Rev. Lett. **80**, 5481 (1998).

[11] K. Jedamzik and J. C. Niemeyer, Phys. Rev. D **59**, 124014 (1999).

[12] I. Musco, J. C. Miller, and L. Rezzolla, Classical Quantum Gravity **22**, 1405 (2005).

[13] I. Musco, J. C. Miller, and A. G. Polnarev, Classical Quantum Gravity **26**, 235001 (2009).

[14] I. Musco and J. C. Miller, Classical Quantum Gravity **30**, 145009 (2013).

[15] C. T. Byrnes, M. Hindmarsh, S. Young, and M. R. S. Hawkins, J. Cosmol. Astropart. Phys. 08 (2018) 041.

[16] I. Musco, S. Young, and K. Jedamzik (to be published).

[17] G. Franciolini, R. Cotesta, N. Loutrel, E. Berti, P. Pani, and A. Riotto, Phys. Rev. D **105**, 063510 (2022).

[18] M. Volonteri, Astron. Astrophys. Rev. **18**, 279 (2010).

[19] S. Clesse and J. García-Bellido, Phys. Rev. D **92**, 023524 (2015).

[20] P. D. Serpico, V. Poulin, D. Inman, and K. Kohri, Phys. Rev. Res. **2**, 023204 (2020).

[21] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, Rep. Prog. Phys. **84**, 116902 (2021).

[22] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. X **9**, 031040 (2019).

[23] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. X **11**, 021053 (2021).

[24] R. Abbott *et al.* (LIGO-Virgo-KAGRA Collaborations), arXiv:2111.03606.

[25] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli, and A. G. Riess, Phys. Rev. Lett. **116**, 201301 (2016).

[26] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, Phys. Rev. Lett. **117**, 061101 (2016); **121**, 059901(E) (2018).

[27] Y. N. Eroshenko, J. Phys. Conf. Ser. **1051**, 012010 (2018).

[28] S. Wang, Y.-F. Wang, Q.-G. Huang, and T. G. F. Li, Phys. Rev. Lett. **120**, 191102 (2018).

[29] Y. Ali-Haïmoud, E. D. Kovetz, and M. Kamionkowski, Phys. Rev. D **96**, 123523 (2017).

[30] Z.-C. Chen and Q.-G. Huang, Astrophys. J. **864**, 61 (2018).

[31] M. Raidal, C. Spethmann, V. Vaskonen, and H. Veermäe, J. Cosmol. Astropart. Phys. 02 (2019) 018.

[32] L. Liu, Z.-K. Guo, and R.-G. Cai, Eur. Phys. J. C **79**, 717 (2019).

[33] G. Hütsi, M. Raidal, and H. Veermäe, Phys. Rev. D **100**, 083016 (2019).

[34] V. Vaskonen and H. Veermäe, Phys. Rev. D **101**, 043015 (2020).

[35] A. D. Gow, C. T. Byrnes, A. Hall, and J. A. Peacock, J. Cosmol. Astropart. Phys. 01 (2020) 031.

[36] Y. Wu, Phys. Rev. D **101**, 083008 (2020).

[37] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, J. Cosmol. Astropart. Phys. 06 (2020) 044.

[38] A. Hall, A. D. Gow, and C. T. Byrnes, Phys. Rev. D **102**, 123524 (2020).

[39] K. W. K. Wong, G. Franciolini, V. De Luca, V. Baibhav, E. Berti, P. Pani, and A. Riotto, Phys. Rev. D **103**, 023026 (2021).

[40] G. Hütsi, M. Raidal, V. Vaskonen, and H. Veermäe, J. Cosmol. Astropart. Phys. 03 (2021) 068.

[41] K. Kritos, V. De Luca, G. Franciolini, A. Kehagias, and A. Riotto, J. Cosmol. Astropart. Phys. 05 (2021) 039.

[42] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, J. Cosmol. Astropart. Phys. 05 (2021) 003.

[43] H. Deng, J. Cosmol. Astropart. Phys. 04 (2021) 058.

[44] R. Kimura, T. Suyama, M. Yamaguchi, and Y.-L. Zhang, J. Cosmol. Astropart. Phys. 04 (2021) 031.

[45] G. Franciolini, V. Baibhav, V. De Luca, K. K. Y. Ng, K. W. K. Wong, E. Berti, P. Pani, A. Riotto, and S. Vitale, Phys. Rev. D **105**, 083526 (2022).

[46] S. S. Bavera, G. Franciolini, G. Cusin, A. Riotto, M. Zevin, and T. Fragos, Astron. Astrophys. **660**, A26 (2022).

[47] L. Liu, X.-Y. Yang, Z.-K. Guo, and R.-G. Cai, arXiv:2112.05473.

[48] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, J. Cosmol. Astropart. Phys. 11 (2021) 039.

[49] O. Pujolas, V. Vaskonen, and H. Veermäe, Phys. Rev. D **104**, 083521 (2021).

[50] K. K. Y. Ng, S. Chen, B. Goncharov, U. Dupletsa, S. Borhanian, M. Branchesi, J. Harms, M. Maggiore, B. S. Sathyaprakash, and S. Vitale, Astrophys. J. Lett. **931**, L12 (2022).

[51] K. K. Y. Ng, G. Franciolini, E. Berti, P. Pani, A. Riotto, and S. Vitale, Astrophys. J. Lett. **933**, L41 (2022).

[52] M. Martinelli, F. Scarcella, N. B. Hogg, B. J. Kavanagh, D. Gaggero, and P. Fleury, J. Cosmol. Astropart. Phys. 08 (2022) 006.

[53] P. S. Cole, A. Coogan, B. J. Kavanagh, and G. Bertone, arXiv:2207.07576.

[54] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, Classical Quantum Gravity **35**, 063001 (2018).

[55] A. M. Green and B. J. Kavanagh, J. Phys. G **48**, 043001 (2021).

[56] G. Franciolini, Primordial black holes: From theory to gravitational wave observations, Ph.D. thesis, University of Geneva, 2021, https://archive-ouverte.unige.ch/unige:156136.

[57] S. Clesse and J. Garcia-Bellido, Phys. Dark Universe **38**, 101111 (2022).

[58] V. De Luca, V. Desjacques, G. Franciolini, P. Pani, and A. Riotto, Phys. Rev. Lett. **126**, 051101 (2021).

[59] LIGO Scientific, Virgo, and KAGRA Scientific Collaborations, arXiv:2111.03634.

[60] M. Maggiore *et al.*, J. Cosmol. Astropart. Phys. 03 (2020) 050.

[61] V. Kalogera *et al.*, arXiv:2111.06990.

[62] K. Jedamzik, Phys. Rep. **307**, 155 (1998).

[63] B. Carr, S. Clesse, and J. García-Bellido, Mon. Not. R. Astron. Soc. **501**, 1426 (2021).

[64] B. Carr, S. Clesse, J. García-Bellido, and F. Kühnel, Phys. Dark Universe **31**, 100755 (2021).

[65] K. Jedamzik, Phys. Rev. Lett. **126**, 051302 (2021).

[66] E. Bagui and S. Clesse, Phys. Dark Universe **38**, 101115 (2022).

[67] M. Braglia, J. Garcia-Bellido, and S. Kuroyanagi, J. Cosmol. Astropart. Phys. 12 (2021) 012.

[68] M. Braglia, J. Garcia-Bellido, and S. Kuroyanagi, arXiv:2201.13414.

[69] J. I. Juan, P. Serpico, and G. F. Abellán, J. Cosmol. Astropart. Phys. 07 (2022) 009.

[70] A. H. Nitz and Y.-F. Wang, Phys. Rev. D **106**, 023024 (2022).

[71] Z.-C. Chen, C. Yuan, and Q.-G. Huang, Phys. Lett. B **829**, 137040 (2022).

[72] G. Franciolini and A. Urbano, arXiv:2207.10056.

[73] K. Inomata, M. Kawasaki, K. Mukaida, Y. Tada, and T. T. Yanagida, Phys. Rev. D **95**, 123510 (2017).

[74] J. Garcia-Bellido and E. Ruiz Morales, Phys. Dark Universe **18**, 47 (2017).

[75] G. Ballesteros and M. Taoso, Phys. Rev. D **97**, 023501 (2018).

[76] M. P. Hertzberg and M. Yamada, Phys. Rev. D **97**, 083509 (2018).

[77] K. Kannike, L. Marzola, M. Raidal, and H. Veermäe, J. Cosmol. Astropart. Phys. 09 (2017) 020.

[78] I. Dalianis, A. Kehagias, and G. Tringas, J. Cosmol. Astropart. Phys. 01 (2019) 037.

[79] K. Inomata, M. Kawasaki, K. Mukaida, and T. T. Yanagida, Phys. Rev. D **97**, 043514 (2018).

[80] D. Y. Cheong, S. M. Lee, and S. C. Park, J. Cosmol. Astropart. Phys. 01 (2021) 032.

[81] G. Ballesteros, J. Rey, M. Taoso, and A. Urbano, J. Cosmol. Astropart. Phys. 07 (2020) 025.

[82] L. Iacconi, H. Assadullahi, M. Fasiello, and D. Wands, J. Cosmol. Astropart. Phys. 06 (2022) 007.

[83] S. Kawai and J. Kim, Phys. Rev. D **104,** 083545 (2021).

[84] M. Hindmarsh and O. Philipsen, Phys. Rev. D **71,** 087302 (2005).

[85] S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, and K. K. Szabo, Phys. Lett. B **730,** 99 (2014).

[86] A. Helou, I. Musco, and J. C. Miller, Classical Quantum Gravity **34,** 135012 (2017).

[87] D. S. Salopek and J. R. Bond, Phys. Rev. D **42,** 3936 (1990).

[88] A. G. Polnarev and I. Musco, Classical Quantum Gravity **24,** 1405 (2007).

[89] T. Harada, C.-M. Yoo, T. Nakama, and Y. Koga, Phys. Rev. D **91,** 084057 (2015).

[90] C.-M. Yoo, T. Harada, S. Hirano, and K. Kohri, Prog. Theor. Exp. Phys. **2021,** 013E02 (2021).

[91] I. Musco, Phys. Rev. D **100,** 123524 (2019).

[92] T. Papanikolaou, Phys. Rev. D **105,** 124055 (2022).

[93] T. Harada, C.-M. Yoo, and K. Kohri, Phys. Rev. D **88,** 084051 (2013); **89,** 029903(E) (2014).

[94] I. Musco, V. De Luca, G. Franciolini, and A. Riotto, Phys. Rev. D **103,** 063538 (2021).

[95] M. W. Choptuik, Phys. Rev. Lett. **70,** 9 (1993).

[96] C. Gundlach, Living Rev. Relativity **2,** 4 (1999).

[97] A. M. Green and A. R. Liddle, Phys. Rev. D **60,** 063509 (1999).

[98] A. Moradinezhad Dizgah, G. Franciolini, and A. Riotto, J. Cosmol. Astropart. Phys. 11 (2019) 001.

[99] V. De Luca, G. Franciolini, and A. Riotto, Phys. Lett. B **807,** 135550 (2020).

[100] C. T. Byrnes, P. S. Cole, and S. P. Patil, J. Cosmol. Astropart. Phys. 06 (2019) 028.

[101] A. Kalaja, N. Bellomo, N. Bartolo, D. Bertacca, S. Matarrese, I. Musco, A. Raccanelli, and L. Verde, J. Cosmol. Astropart. Phys. 10 (2019) 031.

[102] P. S. Cole, A. D. Gow, C. T. Byrnes, and S. P. Patil, arXiv:2204.07573.

[103] G. Franciolini, A. Kehagias, S. Matarrese, and A. Riotto, J. Cosmol. Astropart. Phys. 03 (2018) 016.

[104] V. Atal, J. Garriga, and A. Marcos-Caballero, J. Cosmol. Astropart. Phys. 09 (2019) 073.

[105] V. De Luca, G. Franciolini, and A. Riotto, Phys. Rev. D **104,** 063526 (2021).

[106] M. Taoso and A. Urbano, J. Cosmol. Astropart. Phys. 08 (2021) 016.

[107] V. De Luca, G. Franciolini, A. Kehagias, M. Peloso, A. Riotto, and C. Ünal, J. Cosmol. Astropart. Phys. 07 (2019) 048.

[108] S. Young, I. Musco, and C. T. Byrnes, J. Cosmol. Astropart. Phys. 11 (2019) 012.

[109] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304,** 15 (1986).

[110] T. Nakama, J. Silk, and M. Kamionkowski, Phys. Rev. D **95,** 043511 (2017).

[111] A. D. Gow, C. T. Byrnes, and A. Hall, Phys. Rev. D **105,** 023503 (2022).

[112] V. De Luca, V. Desjacques, G. Franciolini, A. Malhotra, and A. Riotto, J. Cosmol. Astropart. Phys. 05 (2019) 018.

[113] M. Mirbabayi, A. Gruzinov, and J. Noreña, J. Cosmol. Astropart. Phys. 03 (2020) 017.

[114] Y. Ali-Haïmoud, Phys. Rev. Lett. **121,** 081304 (2018).

[115] V. Desjacques and A. Riotto, Phys. Rev. D **98,** 123533 (2018).

[116] G. Ballesteros, P. D. Serpico, and M. Taoso, J. Cosmol. Astropart. Phys. 10 (2018) 043.

[117] D. Inman and Y. Ali-Haïmoud, Phys. Rev. D **100,** 083528 (2019).

[118] V. De Luca, V. Desjacques, G. Franciolini, and A. Riotto, J. Cosmol. Astropart. Phys. 11 (2020) 028.

[119] T. Nakamura, M. Sasaki, T. Tanaka, and K. S. Thorne, Astrophys. J. Lett. **487,** L139 (1997).

[120] K. Ioka, T. Chiba, T. Tanaka, and T. Nakamura, Phys. Rev. D **58,** 063003 (1998).

[121] L. Liu, Z.-K. Guo, and R.-G. Cai, Phys. Rev. D **99,** 063523 (2019).

[122] K. Jedamzik, J. Cosmol. Astropart. Phys. 09 (2020) 022.

[123] S. Young and A. S. Hamers, J. Cosmol. Astropart. Phys. 10 (2020) 036.

[124] M. Trashorras, J. García-Bellido, and S. Nesseris, Universe **7,** 18 (2021).

[125] M. Tkachev, S. Pilipenko, and G. Yepes, Mon. Not. R. Astron. Soc. **499,** 4854 (2020).

[126] https://agenda.infn.it/event/23799/contributions/125718/attachments/78986/102370/rome110221.pdf (2021).

[127] G. Franciolini, A. Maharana, and F. Muia, Phys. Rev. D **106,** 103520 (2022).

[128] G. Franciolini, K. Kritos, E. Berti, and J. Silk, Phys. Rev. D **106,** 083529 (2022).

[129] Y. Koga, T. Harada, Y. Tada, S. Yokoyama, and C.-M. Yoo, Astrophys. J. **939,** 65 (2022).

[130] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, J. Cosmol. Astropart. Phys. 04 (2020) 052.

[131] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, Phys. Rev. D **102,** 043505 (2020).

[132] G. Franciolini and P. Pani, Phys. Rev. D **105,** 123024 (2022).

[133] B. P. Abbott et al. (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **119,** 161101 (2017).

[134] B. P. Abbott et al. (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, AS-KAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT Collaborations), Astrophys. J. Lett. **848,** L12 (2017).

[135] M. Sasaki, V. Takhistov, V. Vardanyan, and Y.-l. Zhang, Astrophys. J. **931,** 2 (2022).

[136] C. Talbot and E. Thrane, Astrophys. J. **856,** 173 (2018).

[137] T. A. Callister, C.-J. Haster, K. K. Y. Ng, S. Vitale, and W. M. Farr, Astrophys. J. Lett. **922,** L5 (2021).

[138] V. Tiwari and S. Fairhurst, Astrophys. J. Lett. **913,** L19 (2021).

[139] B. Edelman, Z. Doctor, J. Godfrey, and B. Farr, Astrophys. J. **924,** 101 (2022).

[140] V. Tiwari, Astrophys. J. **928,** 155 (2022).

[141] Y.-J. Li, Y.-Z. Wang, S.-P. Tang, Q. Yuan, Y.-Z. Fan, and D.-M. Wei, Astrophys. J. Lett. **933,** L14 (2022).

[142] S. Biscoveanu, T. A. Callister, C.-J. Haster, K. K. Y. Ng, S. Vitale, and W. M. Farr, Astrophys. J. Lett. **932,** L19 (2022).

[143] I. Mandel, W. M. Farr, and J. R. Gair, Mon. Not. R. Astron. Soc. **486,** 1086 (2019).

[144] S. Vitale, D. Gerosa, W. M. Farr, and S. R. Taylor, *Handbook of Gravitational Wave Astronomy* (Springer, New York, 2020).

[145] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, J. Phys. Conf. Ser. **610,** 012021 (2015).

[146] https://www.gw-openscience.org.

[147] T. J. Loredo, Am. Inst. Phys. Conf. Ser. **735,** 195 (2004).

[148] S. R. Taylor and D. Gerosa, Phys. Rev. D **98,** 083017 (2018).

[149] E. Thrane and C. Talbot, Pub. Astron. Soc. Aust. **36,** e010 (2019); **37,** e036(E) (2020).

[150] S. Vitale, D. Gerosa, C.-J. Haster, K. Chatziioannou, and A. Zimmerman, Phys. Rev. Lett. **119,** 251103 (2017).

[151] M. Zevin, C. P. L. Berry, S. Coughlin, K. Chatziioannou, and S. Vitale, Astrophys. J. Lett. **899,** L17 (2020).

[152] S. Bhagwat, V. De Luca, G. Franciolini, P. Pani, and A. Riotto, J. Cosmol. Astropart. Phys. 01 (2021) 037.

[153] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, Publ. Astron. Soc. Pac. **125,** 306 (2013).

[154] M. A. Newton and A. E. Raftery, J. R. Stat. Soc. Ser. B **56,** 3 (1994).

[155] H. Jeffreys, *The Theory of Probability (3rd ed.)* (Oxford, England, 1998).

[156] M. Zevin, S. S. Bavera, C. P. L. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow, Astrophys. J. **910,** 152 (2021).

[157] B. P. Abbott *et al.* (KAGRA, LIGO Scientific, Virgo Collaborations), Living Rev. Relativity **21,** 3 (2018).

[158] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett. **113,** 151101 (2014).

[159] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, Phys. Rev. D **93,** 044007 (2016).

[160] A. Nitz *et al.*, gwastro/pycbc: Pycbc release v1.14.4 (2019).

[161] https://dcc.ligo.org/LIGO-T2000012/public (2022).

[162] A. H. Nitz, S. Kumar, Y.-F. Wang, S. Kastha, S. Wu, M. Schäfer, R. Dhurkunde, and C. D. Capano, arXiv:2112.06878.

[163] S. Olsen, T. Venumadhav, J. Mushkin, J. Roulet, B. Zackay, and M. Zaldarriaga, Phys. Rev. D **106,** 043009 (2022).

[164] https://dcc.ligo.org/LIGO-P1800370/public (2018).

[165] https://dcc.ligo.org/LIGO-P2000223/public (2020).

[166] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run—Parameter estimation data release (2021), https://inspirehep.net/literature/1897574.

[167] R. Abbott *et al.* (LIGO Scientific, Virgo, KAGRA Collaborations), GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run—Parameter estimation data release (2021), https://inspirehep.net/literature?sort=mostrecent&size=25&page=1&q=find%20eprint%202111.03606&ui-citation-summary=true.

[168] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. Lett. **896,** L44 (2020).

[169] K. Chatziioannou, Gen. Relativ. Gravit. **52,** 109 (2020).

[170] C. D. Bailyn, R. K. Jain, P. Coppi, and J. A. Orosz, Astrophys. J. **499,** 367 (1998).

[171] F. Özel, D. Psaltis, R. Narayan, and J. E. McClintock, Astrophys. J. **725,** 1918 (2010).

[172] W. M. Farr, N. Sravan, A. Cantrell, L. Kreidberg, C. D. Bailyn, I. Mandel, and V. Kalogera, Astrophys. J. **741,** 103 (2011).

[173] A. M. Farah, M. Fishbach, R. Essick, D. E. Holz, and S. Galaudage, Astrophys. J. **931,** 108 (2022).

[174] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **125,** 101102 (2020).

[175] Z. Barkat, G. Rakavy, and N. Sack, Phys. Rev. Lett. **18,** 379 (1967).

[176] A. Heger and S. E. Woosley, Astrophys. J. **567,** 532 (2002).

[177] S. E. Woosley, S. Blinnikov, and A. Heger, Nature (London) **450,** 390 (2007).

[178] K. Belczynski *et al.*, Astron. Astrophys. **594,** A97 (2016).

[179] S. E. Woosley, Astrophys. J. **836,** 244 (2017).

[180] S. Stevenson, M. Sampson, J. Powell, A. Vigna-Gómez, C. J. Neijssel, D. Szécsi, and I. Mandel, Astrophys. J. **882,** 121 (2019).

[181] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham, Astrophys. J. **887,** 53 (2019).

[182] M. Renzo, R. J. Farmer, S. Justham, S. E. de Mink, Y. Götberg, and P. Marchant, Mon. Not. R. Astron. Soc. **493,** 4333 (2020).

[183] M. Mapelli, M. Spera, E. Montanari, M. Limongi, A. Chieffi, N. Giacobbo, A. Bressan, and Y. Bouffanais, Astrophys. J. **888,** 76 (2020).

[184] D. Croon, S. D. McDermott, and J. Sakstein, Phys. Rev. D **102,** 115024 (2020).

[185] P. Marchant and T. Moriya, Astron. Astrophys. **640,** L18 (2020).

[186] J. Ziegler and K. Freese, Phys. Rev. D **104,** 043015 (2021).

[187] K. Belczynski, Astrophys. J. Lett. **905,** L15 (2020).

[188] M. Fishbach, D. E. Holz, and B. Farr, Astrophys. J. Lett. **840,** L24 (2017).

[189] D. Gerosa and E. Berti, Phys. Rev. D **100,** 041301 (2019).

[190] C. L. Rodriguez, M. Zevin, P. Amaro-Seoane, S. Chatterjee, K. Kremer, F. A. Rasio, and C. S. Ye, Phys. Rev. D **100,** 043027 (2019).

[191] V. Baibhav, D. Gerosa, E. Berti, K. W. K. Wong, T. Helfer, and M. Mould, Phys. Rev. D **102,** 043002 (2020).

[192] C. Kimball, C. Talbot, C. P. L. Berry, M. Carney, M. Zevin, E. Thrane, and V. Kalogera, Astrophys. J. **900,** 177 (2020).

[193] J. Samsing and K. Hotokezaka, Astrophys. J. **923,** 126 (2021).

[194] M. Mapelli, F. Santoliquido, Y. Bouffanais, M. A. Sedda, M. C. Artale, and A. Ballone, Symmetry **13,** 1678 (2021).

[195] R. Abbott et al. (LIGO Scientific, Virgo Collaborations), Astrophys. J. Lett. **896,** L44 (2020).

[196] S. Shandera, D. Jeong, and H. S. G. Gebhardt, Phys. Rev. Lett. **120,** 241102 (2018).

[197] V. Cardoso and P. Pani, Living Rev. Relativity **22,** 4 (2019).

[198] H.-K. Guo, K. Sinha, and C. Sun, J. Cosmol. Astropart. Phys. 09 (2019) 032.

[199] J. Bramante, T. Linden, and Y.-D. Tsai, Phys. Rev. D **97,** 055016 (2018).

[200] V. Takhistov, G. M. Fuller, and A. Kusenko, Phys. Rev. Lett. **126,** 071101 (2021).

[201] B. Dasgupta, R. Laha, and A. Ray, Phys. Rev. Lett. **126,** 141105 (2021).

[202] P. Giffin, J. Lloyd, S. D. McDermott, and S. Profumo, Phys. Rev. D **105,** 123030 (2022).

[203] S. Barsanti, V. De Luca, A. Maselli, and P. Pani, Phys. Rev. Lett. **128,** 111104 (2022).

[204] A. Gupta, D. Gerosa, K. G. Arun, E. Berti, W. M. Farr, and B. S. Sathyaprakash, Phys. Rev. D **101,** 103036 (2020).

[205] F. Hofmann, E. Barausse, and L. Rezzolla, Astrophys. J. Lett. **825,** L19 (2016).

[206] C. Y. Lam et al., Astrophys. J. Lett. **933,** L23 (2022).

[207] K. C. Sahu et al. (OGLE, MOA, PLANET, microFUN, MiNDSTEp Consortium, RoboNet Collaborations), Astrophys. J. **933,** 83 (2022).

[208] S. Mereghetti, L. Sidoli, G. Ponti, and A. Treves, Astrophys. J. **934,** 62 (2022).

[209] M. Abramowicz, M. Bejger, A. Udalski, and M. Wielgus, Astrophys. J. Lett. **935,** L28 (2022).

[210] Y. Ali-Haïmoud and M. Kamionkowski, Phys. Rev. D **95,** 043534 (2017).

[211] D. Gaggero, G. Bertone, F. Calore, R. M. T. Connors, M. Lovell, S. Markoff, and E. Storm, Phys. Rev. Lett. **118,** 241101 (2017).

[212] J. Manshanden, D. Gaggero, G. Bertone, R. M. Connors, and M. Ricotti, J. Cosmol. Astropart. Phys. 06 (2019) 026.

[213] Y. Inoue and A. Kusenko, J. Cosmol. Astropart. Phys. 10 (2017) 034.

[214] F. Ziparo, S. Gallerani, A. Ferrara, and F. Vito, Mon. Not. R. Astron. Soc. **517,** 1086 (2022).

[215] R. A. Allsman et al. (Macho Collaboration), Astrophys. J. Lett. **550,** L169 (2001).

[216] M. Oguri, J. M. Diego, N. Kaiser, P. L. Kelly, and T. Broadhurst, Phys. Rev. D **97,** 023518 (2018).

[217] H. Niikura, M. Takada, S. Yokoyama, T. Sumi, and S. Masaki, Phys. Rev. D **99,** 083503 (2019).

[218] J. Silk and L. Stodolsky, Phys. Rev. D **105,** 063506 (2022).

[219] V. De Luca, G. Franciolini, A. Riotto, and H. Veermäe, Phys. Rev. Lett. **129,** 191302 (2022).

[220] V. De Luca, G. Franciolini, and A. Riotto, Phys. Rev. Lett. **126,** 041303 (2021).

[221] J. Chluba et al., Exp. Astron. **51,** 1515 (2021).

[222] D. Wands, Phys. Rev. D **60,** 023507 (1999).

[223] M. Biagetti, G. Franciolini, A. Kehagias, and A. Riotto, J. Cosmol. Astropart. Phys. 07 (2018) 032.

[224] A. Karam, N. Koivunen, E. Tomberg, V. Vaskonen, and H. Veermäe, arXiv:2205.13540.

[225] P. A. R. Ade et al. (Planck Collaboration), Astron. Astrophys. **594,** A13 (2016).

[226] A. K. Saha and R. Laha, Phys. Rev. D **105,** 103026 (2022).

[227] R. Laha, Phys. Rev. Lett. **123,** 251101 (2019).

[228] A. Ray, R. Laha, J. B. Muñoz, and R. Caputo, Phys. Rev. D **104,** 023516 (2021).

[229] S. Mittal, A. Ray, G. Kulkarni, and B. Dasgupta, J. Cosmol. Astropart. Phys. 03 (2022) 030.

[230] S. Clark, B. Dutta, Y. Gao, L. E. Strigari, and S. Watson, Phys. Rev. D **95,** 083006 (2017).

[231] R. Laha, J. B. Muñoz, and T. R. Slatyer, Phys. Rev. D **101,** 123514 (2020).

[232] J. Berteaud, F. Calore, J. Iguaz, P. D. Serpico, and T. Siegert, Phys. Rev. D **106,** 023030 (2022).

[233] W. DeRocco and P. W. Graham, Phys. Rev. Lett. **123,** 251102 (2019).

[234] M. Boudaud and M. Cirelli, Phys. Rev. Lett. **122,** 041104 (2019).

[235] B. J. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, Phys. Rev. D **81,** 104019 (2010).

[236] H. Niikura et al., Nat. Astron. **3,** 524 (2019).

[237] P. Tisserand et al. (EROS-2 Collaboration), Astron. Astrophys. **469,** 387 (2007).

[238] S. Sugiyama, V. Takhistov, E. Vitagliano, A. Kusenko, M. Sasaki, and M. Takada, Phys. Lett. B **814,** 136097 (2021).

[239] K. Tomita, Prog. Theor. Phys. **54,** 730 (1975).

[240] S. Matarrese, O. Pantano, and D. Saez, Phys. Rev. Lett. **72,** 320 (1994).

[241] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nucl. Phys. **B667,** 119 (2003).

[242] S. Mollerach, D. Harari, and S. Matarrese, Phys. Rev. D **69,** 063002 (2004).

[243] K. N. Ananda, C. Clarkson, and D. Wands, Phys. Rev. D **75,** 123518 (2007).

[244] D. Baumann, P. J. Steinhardt, K. Takahashi, and K. Ichiki, Phys. Rev. D **76,** 084019 (2007).

[245] J.-Z. Zhou, X. Zhang, Q.-H. Zhu, and Z. Chang, J. Cosmol. Astropart. Phys. 05 (2022) 013.

[246] X. Zhang, J.-Z. Zhou, and Z. Chang, Eur. Phys. J. C **82,** 781 (2022).

[247] G. Domènech, Universe **7,** 398 (2021).

[248] S. Wang, T. Terada, and K. Kohri, Phys. Rev. D **99,** 103531 (2019); **101,** 069901(E) (2020).

[249] Z. Arzoumanian et al. (NANOGrav Collaboration), Astrophys. J. Lett. **905,** L34 (2020).

[250] B. Goncharov et al., Astrophys. J. Lett. **917,** L19 (2021).

[251] S. Chen et al., Mon. Not. R. Astron. Soc. **508,** 4970 (2021).

[252] J. Antoniadis et al., Mon. Not. R. Astron. Soc. **510,** 4873 (2022).

[253] Z. Arzoumanian et al. (NANOGrav Collaboration), Astrophys. J. Lett. **905,** L34 (2020).

[254] P. Auclair et al. (LISA Cosmology Working Group), arXiv:2204.05434.

[255] V. Vaskonen and H. Veermäe, Phys. Rev. Lett. **126,** 051303 (2021).

[256] K. Kohri and T. Terada, Phys. Lett. B **813**, 136040 (2021).

[257] A. Ashoorioon, K. Rezazadeh, and A. Rostami, Phys. Lett. B **835**, 137542 (2022).

[258] S. S. Bavera, T. Fragos, Y. Qin, E. Zapartas, C. J. Neijssel, I. Mandel, A. Batta, S. M. Gaebel, C. Kimball, and S. Stevenson, Astron. Astrophys. **635**, A97 (2020).

[259] R. M. Shannon *et al.*, Science **349**, 1522 (2015).

[260] G. Janssen *et al.*, Proc. Sci., AASKA14 (**2015**) 037 [arXiv:1501.00127].

[261] K. T. Abe, Y. Tada, and I. Ueda, J. Cosmol. Astropart. Phys. 06 (2021) 048.

[262] K. Saikawa and S. Shirai, J. Cosmol. Astropart. Phys. 05 (2018) 035.

[263] P. Madau and M. Dickinson, Annu. Rev. Astron. Astrophys. **52**, 415 (2014).

[264] B. P. Abbott *et al.*, Phys. Rev. Lett. **116**, 131102 (2016).

[265] B. P. Abbott *et al.*, Phys. Rev. Lett. **120**, 091101 (2018).

[266] J. A. de Freitas Pacheco, arXiv:2001.09663.

[267] C. Périgois, C. Belczynski, T. Bulik, and T. Regimbau, Phys. Rev. D **103**, 043002 (2021).

[268] C. Périgois, F. Santoliquido, Y. Bouffanais, U. N. Di Carlo, N. Giacobbo, S. Rastello, M. Mapelli, and T. Regimbau, Phys. Rev. D **105**, 103032 (2022).

[269] Z.-C. Chen and Q.-G. Huang, J. Cosmol. Astropart. Phys. 08 (2020) 039.

[270] S. Mukherjee and J. Silk, Mon. Not. R. Astron. Soc. **506**, 3977 (2021).

[271] A. Sesana, M. Volonteri, and F. Haardt, Mon. Not. R. Astron. Soc. **377**, 1711 (2007).

[272] S. Banks, K. Lee, N. Azimi, K. Scarborough, N. Stefanov, I. Periwal, N. Chen, C. DeGraf, and T. Di Matteo, Mon. Not. R. Astron. Soc. **512**, 6007 (2022).

[273] N. Bartolo, V. De Luca, G. Franciolini, A. Lewis, M. Peloso, and A. Riotto, Phys. Rev. Lett. **122**, 211301 (2019).

[274] N. Bartolo, V. De Luca, G. Franciolini, M. Peloso, D. Racco, and A. Riotto, Phys. Rev. D **99**, 103521 (2019).

[275] N. Aggarwal *et al.*, Living Rev. Relativity **24**, 4 (2021).

[276] M. Biagetti, V. De Luca, G. Franciolini, A. Kehagias, and A. Riotto, Phys. Lett. B **820**, 136602 (2021).

[277] G. Ferrante, G. Franciolini, A. J. Iovino, and A. Urbano, arXiv:2211.01728.

[278] LIGO Scientific Collaboration and Virgo Collaboration and KAGRA Collaboration, GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run—O3 search sensitivity estimates, 10.5281/zenodo.5546676.

[279] A. Escrivà, E. Bagui, and S. Clesse, arXiv:2209.06196.

[280] LIGO-Virgo-KAGRA Scientific Collaborations, GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run—O3 search sensitivity estimates (2021).

[281] A. S. Josan, A. M. Green, and K. A. Malik, Phys. Rev. D **79**, 103520 (2009).

[282] Y.-F. Cai, X. Chen, M. H. Namjoo, M. Sasaki, D.-G. Wang, and Z. Wang, J. Cosmol. Astropart. Phys. 05 (2018) 012.

[283] S. Passaglia, W. Hu, and H. Motohashi, Phys. Rev. D **99**, 043536 (2019).