Topical Review

Primordial black holes—perspectives in gravitational wave astronomy

Misao Sasaki¹, Teruaki Suyama², Takahiro Tanaka¹,³ and Shuichiro Yokoyama⁴,⁵

¹ Center for Gravitational Physics, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan
² Research Center for the Early Universe (RESCEU), Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan
³ Department of Physics, Kyoto University, Kyoto 606-8502, Japan
⁴ Department of Physics, Rikkyo University, Tokyo 171-8501, Japan
⁵ Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

E-mail: misao@yukawa.kyoto-u.ac.jp

Received 5 November 2016, revised 11 January 2018
Accepted for publication 15 January 2018
Published 8 February 2018

Abstract

This article reviews current understanding of primordial black holes (PBHs), with particular focus on those massive examples ($\gtrsim 10^{15}$ g) which remain at the present epoch, not having evaporated through Hawking radiation. With the detection of gravitational waves by LIGO, we have gained a completely novel observational tool to search for PBHs, complementary to those using electromagnetic waves. Taking the perspective that gravitational-wave astronomy will make significant progress in the coming decades, the purpose of this article is to give a comprehensive review covering a wide range of topics on PBHs. After discussing PBH formation, as well as several inflation models leading to PBH production, we summarize various existing and future observational constraints. We then present topics on formation of PBH binaries, gravitational waves from PBH binaries, and various observational tests of PBHs using gravitational waves.

Keywords: black holes, gravitational waves, primordial perturbations

(Some figures may appear in colour only in the online journal)

1. Introduction

The history of primordial black hole (PBH) study dates back to sixties, when Zeldovich and Novikov pointed out that BHs in the early Universe might grow catastrophically by accreting the surrounding radiation [1]. In 1971, Hawking proposed [2] that a highly overdense region
of inhomogeneities in the primordial Universe could directly undergo gravitational collapse to form BHs; this initiated the modern theory of the mechanism of PBH formation. In contrast to astrophysical processes (i.e. collapse of stars), which can only form BHs heavier than a particular mass (around three solar masses [3]), extremely strong gravitational forces inside the highly compressed radiation/matter existing in the early Universe allows formation of not only stellar/super-massive BHs but also small BHs that could in principle be as light as the Planck mass $\sim 10^{-5}$ g (see e.g. [4] and references therein). After the advent of inflationary cosmology, formation of PBHs and their properties—such as mass and abundance—had been studied in close connection with inflation models. Conversely, knowledge of observational information about PBHs provides important clues to build inflation models. In particular, it is worth mentioning that even the non-detection of PBHs gives us useful information about the early Universe [4].

Observational searches for PBHs have been conducted intensively and continuously over several decades. Depending on the mass, PBHs trigger different observational signals. PBHs lighter than a certain mass $M_c$, given by [5]

$$M_c \simeq \left(\frac{3 \hbar c^2 \alpha_0}{G^2} t_0\right)^{\frac{1}{3}} \sim 10^{15} \text{ g} \left(\frac{\alpha_0}{4 \times 10^{-4}}\right)^{\frac{1}{3}} \left(\frac{t_0}{13.8 \text{ Gyr}}\right)^{\frac{1}{3}},$$

have already evaporated by the cosmic age $t_0$ due to the Hawking radiation. Thus, PBHs lighter than $\simeq 10^{15}$ g do not exist in the present Universe. Nevertheless, they leave some traces by which we can investigate how many PBHs could have existed in the early Universe. For instance, PBHs in the mass range $10^9$–$10^{13}$ g change the abundance of light elements produced by the Big Bang nucleosynthesis, due to high energy particles emitted by the evaporating PBHs [6]. Comparison between the observed light elements and the theoretical prediction tightly constrains the abundance of such PBHs (see e.g. [7] and references therein).

PBHs heavier than $10^{15}$ g have not yet lost mass significantly by evaporation, and remain in the present Universe. They not only imprint observational traces in the early Universe (such as by accretion, and indirect effects by the primordial density perturbations that seed PBHs) but also produce various distinct signals at the present time, such as gravitational lensing, dynamical effects on baryonic matter, radiation emanating from the matter accreting into PBHs, etc. What physical process among them becomes the most prominent to show us the existence of PBHs depends on the PBH mass. For instance, gravitational lensing of background stars is the most powerful method to search for sub-solar PBHs (see section 3). Accretion, and dynamical effects on baryonic matter become more important for heavier PBHs. One of the important questions regarding non-evaporating PBHs is whether they comprise all dark matter or not. Thanks to the achievements of many different types of cosmological and astrophysical observation over decades, a stringent upper limit on PBH abundance has been obtained for a vast PBH mass range [8]. Currently, it appears that PBHs do not explain all dark matter, constituting at most a fraction thereof [9]. Reviewing the existing observational constraints on non-evaporating PBHs is one of the main purposes of this article.

The discovery at LIGO of the merger event (GW150914) of binary BHs [10] triggered a renewed interest of PBHs, especially in the stellar mass range. The unexpectedly large mass of the detected BHs (around $30 \ M_\odot$) brought us a new mystery about this component of the Universe. After the LIGO event, elucidating the origin of these BHs—and binary formation—has emerged as an important topic in cosmology and astrophysics (e.g. see [11]). Soon after LIGO’s announcement on the first detection of the BH merger, several research groups [12–14] independently pointed out that the inferred merger rate can be explained by the merger of PBHs without violating the trivial bound that the PBH abundance is equal to or less
than the total dark matter abundance. In [12, 13], binary formation via accidental encounters of PBHs in a dense environment, which works in the low-redshift Universe, has been considered, while a different mechanism of binary formation via the tidal perturbation caused by distant PBHs, which works in the radiation-dominated epoch in the early Universe and was originally proposed earlier in [15], has been investigated in [14]. These studies demonstrate that gravitational waves (GWs)—a brand new observable—provide a powerful and useful tool to probe parameters of PBHs (mass, abundance etc) in ways that had not been possible through electromagnetic waves alone. In other words, the roles of GWs are complementary to those of electromagnetic waves.

In addition to the PBH scenario, several astrophysical processes, such as the field binary scenario and dynamical formation in the dense stellar environment, have also been proposed to explain the observed properties of the binary black holes (for instance, see [11] and references therein). At the time of writing this article (autumn 2017), LIGO and Virgo have detected five BH merger events [10, 16–19]. For the moment, both the PBH scenario and the astrophysical scenarios are allowed as possible explanations of these events, because of the limited number of events detected.

We are at the dawn of the golden age of GW astronomy. In the future, ongoing experiments such as LIGO and Virgo (and soon KAGRA) will gain better sensitivity and be further upgraded, and new types of experiment such as the Einstein telescope [20], Cosmic Explorer [21], LISA [22] and DECIGO [23], each covering different frequency bands, will follow. Very likely, many more merger events will be detected, and more statistical information on BHs and BH binaries such as mass, spin, eccentricity, redshift, spatial inhomogeneities, etc will become available. This information will enable us to test the individual scenarios, and possibly to pin down the best one—or it is also conceivable that the truth is a mixture of multiple scenarios, e.g. some events due to PBHs and others due to astrophysical BHs. In order to reveal the true nature of BHs and their binaries, it is indispensable to go through three stages: to devise as many possible scenarios as we can (first); to understand each scenario theoretically, and derive distinct features predicted in each scenario (second); and then to test the predictions through real observations, and identify the answer (third). In addition to the main purpose as mentioned above, another purpose of this article is to review the PBH scenario in connection with the LIGO events, and various proposals to test this scenario by using the future GW observations—i.e. the first and the second stages of the above classification.

The organization of this article is as follows. In the next section, we will review the basics of the PBH formation and related topics. Inflation models leading to the production of PBHs are also introduced. In section 3, we cover various observational constraints on the abundance of non-evaporating PBHs obtained via electromagnetic observation. Expected future constraints are also briefly presented. We do not put much focus on smaller PBHs that have already evaporated—but briefly mention the constraints—because they are less relevant to GWs. Regarding the constraints on such small PBHs, [7] provides a detailed analysis. In section 4, we review the PBH scenario as an explanation of the LIGO events and future prospects of its observational testing via GWs. The final section is a summary.

Throughout this article, we use a natural unit scheme in which $c = \hbar = 1$.

2. Formation of PBHs

To date, several mechanisms of PBH formation in the early Universe have been proposed. As examples, [24, 25] recently discussed the possibility of PBH formation by domain walls, and [24, 26] also proposed a scenario in which PBHs are formed by vacuum bubbles which
nucleate during inflation. There have also been several works about PBH formation from cosmic string loops [27–29]. However, the most frequently studied PBH formation scenario must be the gravitational collapse of overdense regions in the early Universe. Here, we briefly review understanding of the latter PBH formation process, and also the inflationary models which could produce such an overdense region.

2.1. Basis of primordial black holes formation in the early Universe

2.1.1. PBH formation based on the simple physical picture. In the early radiation-dominated Universe, a highly overdense region would gravitationally collapse into a black hole, directly. Such a black hole formed in the early Universe is called primordial black hole (PBH). Details of PBH formation from overdense regions in the early Universe have been extensively investigated numerically and analytically [30–36]. Here, let us give a brief review of a rough sketch of PBH formation which unifies the traditional view based on density contrast [37, 38] with the more recent view based on curvature perturbation [31, 39].

First, in the early Universe after inflation, the background spacetime can be well-described by the spatially-flat Friedmann–Lemaître–Robertson–Walker (FLRW) metric (homogeneous and isotropic space):

\[ ds^2 = -dt^2 + a(t)^2 \delta_{ij} dx^i dx^j, \]  

where \( a(t) \) is a scale factor. From the Einstein equation, we can derive a background Friedmann equation as

\[ \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \bar{\rho}(t), \]  

where a dot denotes the derivative in terms of \( t \) and \( \bar{\rho} \) is the background energy density.

Against this background, we consider a locally perturbed region that would eventually collapse to a black hole. Such a region would be a very rare region in the space. Hence, it may be approximated by a spherically symmetric region of positive curvature. Since the comoving size of such a region is initially much larger than the Hubble horizon size, one may apply the separate universe approach or the leading order spatial gradient expansion to it—that is, we may assume a metric of the form

\[ ds^2 = -dt^2 + a(t)^2 e^{2\psi(r)} \delta_{ij} dx^i dx^j, \]  

where \( \psi > 0 \) and is assumed to be monotonically decreasing to zero as \( r \to \infty \). It is known that the above form agrees with the metric on comoving slices on superhorizon scales, where \( \psi \) corresponds to the nonlinear version of the conserved comoving curvature perturbation—that is, \( \psi = \mathcal{R}_c \) [40]. The above metric can be cast into the more familiar form of a locally closed universe with the metric

\[ ds^2 = -dt^2 + a(t)^2 \left[ \frac{dR^2}{1-K(R)R^2} + R^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right], \]  

where the coordinates \( r \) and \( R \) are related to each other by \( R = r e^{\psi(r)} \), and \( K \) is given by

\[ K = -\frac{\psi''(r)}{r} + \frac{r \psi'(r)}{e^{2\psi(r)}}. \]  

We note that the 3-curvature of the \( t = \text{const.} \) hypersurface is given by
\[ R^{(3)} = -\frac{e^{-2\psi}}{3a^2} \delta^i_j \left[ 2\partial_i \partial_j \psi + \partial_i \psi \partial_j \psi \right] = \frac{K}{a^2} \left( 1 + \frac{d \ln K(R)}{3d \ln R} \right). \]  

(7)

Ignoring the spatial derivative of \( K \) in the spirit of leading order gradient expansion, the time–time component of the Einstein equations (the Hamiltonian constraint) gives

\[ H^2 + \frac{K(r)}{a^2} = \frac{8\pi G}{3} \rho, \]  

(8)

where \( H = \dot{a}/a \). This is equivalent to the Friedmann equation, except for a small inhomogeneity induced by the curvature term. One could regard this as the Hamiltonian constraint on the comoving hypersurface, or that on the uniform Hubble hypersurface on which the expansion rate is spatially homogeneous and isotropic.

The above equation naturally leads us to define the density contrast on the comoving hypersurface by

\[ \Delta := \frac{\rho \cdot \rho \bar{\rho}}{\bar{\rho}} = \frac{3K}{8\pi G \rho a^2} = \frac{K}{H^2 a^2}. \]  

(9)

From the fact that \( \bar{\rho}(t) \propto a^{-4} \) during the radiation-dominated universe, this is vanishingly small initially—consistent with the picture that it is the curvature perturbation that induces the density perturbation.

As the universe evolves, \( \Delta \) grows to become of order unity. If we were to ignore the spatial dependence of \( K \), the universe with \( K > 0 \) would eventually stop expanding, and recollapse. This happens when \( 3K/a^2 = 8\pi G \rho \), i.e. when the comoving scale of this positively curved region becomes of the order of the Hubble horizon scale, at which our separate universe approximation precisely breaks down. Also, the equivalence between the comoving and uniform Hubble slices no longer holds. Nevertheless, we may expect that equation (8) will still be used in obtaining a qualitatively acceptable criterion for black hole formation, which has actually been shown to be valid in fully nonlinear numerical studies.

Since \( \Delta = 1 \) is the time when the universe would stop expanding if it were homogeneous and isotropic, let us assume this epoch to be the time of black hole formation, \( t = t_c \). Since a perturbation on scales smaller than the Jeans length cannot collapse, we set this to happen at \( c_s^2 k^2 / a^2 = H^2 \) or \( k^2 / a^2 = 3H^2 \) for \( c_s^2 = 1/3 \). Specifically, we have

\[ 1 = \Delta(t_c) = \frac{K}{k^2 H^2 a^2} = \frac{K}{c_s^2 k^2}. \]  

(10)

This implies we should identify \( K \) with \( c_s^2 k^2 \). It is then straightforward to find the criterion for black hole formation: that the comoving slice density contrast at the time when the scale of interest re-enters the Hubble horizon is greater than \( \Delta_c = c_s^2 \).

\[ \Delta(t_k) = \frac{K}{H^2(t_k) a^2(t_k)} = \frac{c_s^2 k^2}{H^2(t_k) a^2(t_k)} \geq \Delta_c = c_s^2 = \frac{1}{3}, \]  

(11)

where \( t_k \) is the time at which \( k/a = H \).

Within the scope of the present level of approximation, it is not meaningful to distinguish between the Jeans length, \( R = c_s H^{-1} \), and the Hubble horizon.Crudely speaking, the mass of the PBH formed is equal to the horizon mass at the time of formation.

6 Originally, in [37] there was an upper bound on \( \Delta(t_k) \) in order to avoid the formation of the separate closed universe. However, recently [41] pointed out that this is not the case. Actually, from the point of view of the curvature perturbation on comoving slices, it is by definition the fact that \( \Delta \) can never exceed unity on superhorizon scales.
In the above, we have presented a basis for the formation of PBHs out of the primordial perturbation, based on the simple physical picture. Although such analysis captures the essence of PBH formation, it is also important to clarify the impact of the various effects that have been ignored in the above discussion, which are addressed below.

2.1.2. Precise value of the threshold. The equation (11) only tells us that PBH formation occurs when the density perturbation becomes comparable to 1/3. A precise value of the threshold for PBH formation has been extensively investigated both numerically and analytically [30–36]. For instance, [36] has derived a new analytic formula for the threshold as

\[ \delta_{UH} = \sin^2(\pi \sqrt{w/(1 + 3w)}/(1 + 3w)) \]

for the uniform overdense profile surrounded by the underdense layer, where \( \delta_{UH} \) is the amplitude of the density perturbation at the horizon crossing time in the uniform Hubble slice, and \( w \) is the equation of state of the dominant component in the Universe at formation. Although the analytic formula given above shows good agreement with the results of the numerical simulations, numerical simulations also demonstrate that there is no unique value of the threshold. Different density/curvature perturbation profiles collapse to BHs above different thresholds, which is quite natural from the physical point of view. In terms of the comoving density perturbation, the spread of the threshold was found to be 0.3–0.66. Interestingly, the value 1/3, which had been obtained in the crude approximation, lies in this range. Notice that the spread changes if we use perturbation variables defined in a different time slicing.

2.1.3. Effects of long wavelength modes. The physical picture presented above shows that PBHs can form when the size of the overdense region becomes equal to the Hubble horizon. Because of causality, any additional large-scale perturbations longer than the Hubble horizon at the time of the PBH formation must not affect PBH formation. However, this is not explicitly visible for some perturbation variables.

For example, if we use the curvature variable \( \psi \) on the uniform Hubble slicing, which is identical to the comoving curvature perturbation \( \mathcal{R} \) on super-horizon scales (and is also equal to the curvature perturbation on uniform energy density slicing \( \zeta \)), the threshold depends on how much the longer wavelength modes are. The reason for this is that the curvature variable is not a local quantity, but one the determination of which requires information about the distant distribution of matter [42]. Although one can in principle use the curvature variable (or any other perturbation variable, as long as it is well-defined) to compute any observables such as the abundance of PBHs, appropriate prescription is needed to obtain the correct results if the perturbation variable under consideration is not local.

In [39], the formation criterion is discussed in terms of the spatial 3-curvature and the density contrast, both of which naturally suppress the super-Hubble modes. Actually, the picture we presented in the above is based on this approach. As clear from equation (7), it is the spatial derivative of \( \psi \) that determines the 3-curvature, and when normalized with the background expansion term \( H^2 \) this is approximately equal to the density contrast on the comoving hypersurface \( \Delta \), as defined in equation (9). Use of these quantities is especially convenient when one studies the formation of PBHs out of inhomogeneities which are randomly distributed, and whose overdensity regions are non-spherical.

2.1.4. Near critical collapse. It is known that when the overdensity \( \delta \) is very close to the threshold, there is a simple scaling law for the mass of the black holes formed [43–45]:

\[ M_{BH} = C(\delta - \delta_{th})^\gamma. \] (12)
The mass function of the PBHs has been calculated, based on the above scaling law, in [30, 46]. The results show that the typical mass of PBHs is still about the horizon mass evaluated at the time of formation. Yet, since the critical collapse produces PBHs of smaller mass, the mass function has a power-law tail for the smaller mass range. This could be important when one tries to adopt precise observational constraints for the abundance of PBHs with extended mass spectra to constrain inflation models [8, 47].

2.1.5. Non-sphericity of the overdense region. In the discussion above, the criterion for PBH formation is considered for an isolated spherically symmetric perturbation where the universe approaches a perfectly homogeneous and isotropic spatially flat universe in the large scale limit. In reality, any overdense region due to random density perturbation is, generically, not completely spherically symmetric. Yet, compared to studies of PBH formation out of spherically symmetric overdensity, studies on how non-sphericity affects PBH formation are in short supply. In [48], a non-spherical overdense region was approximated as an ellipsoid, and it was required as the collapse condition that a spherical region enclosed by the shortest axis of the ellipsoid should satisfy the collapse criterion for spherical overdensity. Then, combining the above condition with Carr’s argument [37]—which relates the density contrast at the time of horizon crossing to the mass of the overdense region—it was shown that the threshold density contrast of the ellipsoid over density $\delta_{ec}$ is given by

$$\frac{\delta_{ec}}{\delta_c} \approx 1 + 3e,$$

(13)

where $e$ is the ellipticity of the ellipsoid, and $\delta_c$ is the threshold for the spherical overdensity. This shows that density contrast larger than the spherical one is needed for a non-spherical overdense region to turn into a PBH. Although the numerical value on the right-hand side of equation (13) is obtained by the crude approximation in [48], this result is qualitatively natural, since more overdensity is needed than in the spherical case to pull the overdense region along the longest axis by gravity. Thus, for primordial density perturbations obeying the probability distribution function for which higher-sigma peaks are more suppressed—which are realized in most inflation models—the overdense regions for the PBH formation are nearly spherically symmetric. A more quantitative argument is possible if we further assume the Gaussianity for the density perturbations. In this case, a typical value of the ellipticity for the peak amplitude $\delta$ much greater than the square root of the variance $\sigma (\delta \gg \sigma)$ is given by

$$e \sim \frac{3\sigma}{\sqrt{10\pi\delta}}.$$

(14)

Thus, if PBHs are formed from $n$-sigma peaks, the degree of non-sphericity of the overdense regions involved is limited by $O(n^{-1})$ [48–50].

2.1.6. PBH formation during a matter-dominated-like era. Finally, let us mention the case of PBH formation during the matter-dominated era. Of course, primordial black holes may also form in the matter-dominated era, and this has also been discussed in the literature; it has been emphasized that taking account of deviations from spherical configurations is essential [51–53]. Besides, in reality the matter cannot be exactly dust with $P = 0$, but behaves either as a fluid with small pressure or as a collisionless fluid. These effects should be very important, but they have not been fully explored yet. The massive scalar field is known to behave like non-relativistic matter when it oscillates around the potential minimum. Thus, for models of inflation in which the inflaton or another scalar field oscillates after inflation and dominates the universe, PBH formation during that era may become important [54–56].
2.2. Clustering of PBHs

Let us consider a simplified situation where primordial perturbations consist of two components, short and long wavelength modes, and PBHs are formed out of the short wavelength modes (see figure 1). We denote by $k_S$ and $k_L$ the representative comoving wavenumbers of the short and long wavelength modes, respectively ($k_S \gg k_L$). At the time of the horizon crossing $k_S = aH$, the long wavelength modes are still super-Hubble, and such modes are absorbed into the homogeneous component on the Hubble scale $aH^{-1}$.

If there are no correlations between the short and long wavelength modes, which is the case for perturbations obeying Gaussian statistics, spatial distribution of PBHs at the time of formation should not trace the large scale perturbations. In fact, adopting the comoving density contrast as the perturbation variable and the peak-background picture, it has been shown [57] that the effect of large scale perturbations is significantly suppressed. In other words, PBHs are little clustered initially on the comoving scale $aH^{-1}$. Their spatial distribution is uniform, and the number density fluctuates according to the Poisson statistics. Notice that even if the PBHs are not clustered initially, they can cluster later due to the gravitational potential created by the long modes, just like the formation of dark matter halos in the $\Lambda$CDM scenario.

If, on the other hand, the short modes correlate with the long modes, it can happen that the amplitudes of the short modes are modulated by the long modes. One example is the local-type non-Gaussianity for which the curvature variable $\psi$ on super-Hubble scales can be written as

$$\psi(\vec{x}) = \psi_g(\vec{x}) + f_{NL} \psi_g^2(\vec{x}),$$

(15)

where $\psi_g(\vec{x})$ is a Gaussian variable containing both short and long modes, and $f_{NL}$ is the so-called nonlinearity parameter [60], which is dimensionless and constant. This type of perturbation arises when the initially isocurvature modes are converted to the adiabatic modes. Decomposing $\psi_g$ into the short mode $\psi_{g,S}$ and the long mode $\psi_{g,L}$ by

$$\psi_g = \psi_{g,S} + \psi_{g,L},$$

(16)

we find that $\psi$ can be written as

$$\psi = \psi_S + \psi_L,$$

(17)

It has been claimed in [58], and later in [13, 59], that PBHs are clustered initially even for Gaussian density perturbations. As far as the authors are concerned, there is not a broad consensus on this possibility.

The definition of $f_{NL}$ in equation (15) is different from the standard one by a factor $3/5$. We have omitted this unessential factor to make equations simple.
where $\psi_S$ and $\psi_L$ are defined by

$$
\psi_S = (1 + 2f_{NL}\psi_L + f_{NL}\psi_S)\psi_S, \quad \psi_L = \psi_S + f_{NL}\psi_S^2.
$$

By definition, $|\partial_i\psi_S| \sim k_S \psi_S$, $|\partial_i\psi_L| \sim k_L \psi_L$ hold. Then, ignoring the terms suppressed by a factor $k_L$, the 3-curvature given by equation (7) can be written as

$$
R^{(3)} \approx -e^{-2\psi_S} 3^2 \delta^{ij} (2\partial_i\partial_j\psi_S + \partial_i\psi_S \partial_j\psi_S),
$$

where $a_L = e^{\psi_L} a$ is the local scale factor. For $|\psi_S| = O(1)$, $R^{(3)}$ given above represents space curvature on the comoving scale $k_S^{-1}$. Since $\psi_S$ contains the long mode $\psi_L$, the magnitude of the short mode 3-curvature modulates over the comoving scale $k_S^{-1}$. As a result, PBH number density also modulates on the comoving scale $k_L^{-1}$, and PBHs are clustered on the large scale $k_L^{-1}$. For more detail, see [57, 61].

2.3. Abundance of PBHs

In order to investigate the abundance of PBHs formed, let us introduce a parameter which represents the mass fraction (the energy density fraction) of PBHs at the formation as $\beta$, which can be defined as

$$
\beta := \frac{\rho_{\text{PBH}}}{\rho_{\text{tot}}} \bigg|_{\text{at formation}} = \left(\frac{H_0}{H_{\text{form}}}\right)^2 \left(\frac{a_{\text{form}}}{a_0}\right)^{-3} \Omega_{\text{CDM}} f_{\text{PBH}},
$$

where $H := \dot{a}/a$ is a Hubble parameter, $f_{\text{PBH}}$ and $\Omega_{\text{CDM}}$ are respectively a fraction of PBHs against the total dark matter component and a density parameter of the matter component at present, and ‘form’ and ‘0’, respectively, denote the values evaluated at the formation and the present time. As we have mentioned in the previous subsection, the mass of PBHs formed in the radiation-dominated era can approximately be evaluated to be equal to the horizon mass, $M_H := (4\pi/3)\rho H^{-3}$ with $\rho$ being the total energy density of the Universe), at the formation, and hence we have

$$
M_{\text{PBH}} = \gamma M_H \bigg|_{\text{at formation}} = \gamma \frac{4\pi}{3} \rho_{\text{form}} H_{\text{form}}^{-3} = \gamma \frac{4\pi}{3} \frac{3H^3}{8\pi G} H_{\text{form}}^3
$$

$$
= \gamma \frac{1}{2G} H_{\text{form}}^{-1}.
$$

Here, we introduce a correction factor, $\gamma$, which can be evaluated as $\gamma \simeq 0.2$ in a simple analytic calculation [37]. By using the above relation between the mass of PBHs and the Hubble parameter at the formation, mass fraction of PBHs, $\beta$, can be written as (see e.g. [7])

$$
\beta \simeq 3.7 \times 10^{-9} \left(\frac{\gamma}{0.2}\right)^{-1/2} \left(\frac{g_*_{\text{form}}}{10.75}\right)^{1/4} \left(\frac{M_{\text{PBH}}}{M_\odot}\right)^{1/2} f_{\text{PBH}},
$$

where $g_*$ is a number of relativistic degree of freedom. Thus, for each mass of PBHs, the observational constraint on $f_{\text{PBH}}$ can be interpreted as that on $\beta$.

As we have shown in the previous subsection, during the radiation-dominated era, PBHs are basically considered to be formed when a sufficiently overdense region, corresponding to the density fluctuations with a sufficiently large amplitude at a certain scale, enters the Hubble horizon. Once the probability distribution function of the density fluctuations is given, $\beta$ can
be regarded as the probability that the density contrast is larger than the threshold for PBH formation, and we can evaluate the mass fraction $\beta$ as

$$\beta = \gamma \int_{\delta_{th}}^{1} P(\delta) \, d\delta, \quad (23)$$

where $\delta_{th}$ is the threshold for PBH formation. For the Gaussian distribution function, $\beta$ is approximately given by

$$\beta(M_{PBH}) = \gamma \int_{\delta_{th}}^{1} \frac{d\delta}{\sqrt{2\pi} \sigma_{MPBH}} \exp \left[ -\frac{\delta^2}{2\sigma_{MPBH}^2} \right], \quad (24)$$

where $\sigma_{M_{PBH}}$ is the variance of the density fluctuations on the mass scale $M_{PBH}$, and $\nu_{th} := \delta_{th}/\sigma_{M_{PBH}}$. The variance $\sigma_{M_{PBH}}$ is estimated as

$$\sigma_{M_{PBH}}^2 = \int d\ln k P_{\delta}(k) W^2(kR) = \int d\ln k W^2(kR) \frac{16}{81} (kR)^4 P_{\zeta}(k), \quad (25)$$

where $P_{\delta}$ and $P_{\zeta}$ are, respectively, the power spectra of the primordial density fluctuations and the curvature perturbations on comoving slicing and $W(kR)$ is a window function smoothing over the comoving scale $R \simeq 1/(a_{form}H_{form}) = 2GM_{PBH}/a_{form} \gamma^{-1}$. Here, we assume $\sigma_{M_{PBH}} \ll 1$. Then, from the above expression, we can interpret the constraint on $\beta$ as that on $\nu_{th}$, which corresponds to the amplitude of the density fluctuations. Furthermore, this would give a hint for constructing the successful inflationary models with PBH formation. As an example, if the constraint on $f_{PBH}$ for $30 M_{\odot}$ would be obtained as $f_{PBH} < 10^{-3}$, which is equivalent to $\beta < 3.6 \times 10^{-11}$, it could be interpreted as $\nu_{th} \gtrsim 6.27$—that is, $\sigma_{M_{PBH}} \lesssim 0.08$ with $\delta_{th} = 0.5$. (However, see [62] for a caveat of the translation to the upper limit on the amplitude of the primordial density fluctuations.)

For a primordial power spectrum with a broad peak, it should be useful to give an expression for the mass function, which represents the fraction of PBHs with mass in $(M, M + d\ln M)$. It can be defined as

$$\frac{df_{PBH}(M)}{d \ln M} \, d\ln M \approx \nu(M)^2 \left| \frac{d \ln \nu(M)}{d \ln M} \right| \times f_{PBH}(M) \, d\ln M, \quad (26)$$

with

$$f_{PBH}(M) = 2.7 \times 10^8 \left( \frac{\gamma}{0.2} \right)^{1/2} \left( \frac{g_{* \text{form}}}{10.75} \right)^{-1/4} \left( \frac{M}{M_{\odot}} \right)^{-1/2} \beta(M), \quad (27)$$

and

$$\beta(M) \approx \gamma \sqrt{2\pi} \nu(M) \exp \left[ -\frac{\nu(M)^2}{2} \right]. \quad (28)$$

Harada et al [53] found that for the PBH formation in the matter-dominated era the production probability of PBHs, $\beta$, is approximately given by $\beta \approx 0.05556 \sigma_{\text{hor}}^2$ for $\sigma_{\text{hor}} \ll 1$. Here, $\sigma_{\text{hor}}$ is a variance of the density fluctuations at the horizon re-entry.
11

Here, \( \nu(M) := \delta_\text{th}/\sigma_M \), and we have assumed the Gaussian distribution function for the primordial density fluctuations and \( \nu(M) \gg 1 \).

There are several works about the effect of the non-Gaussianity in the estimation of \( \beta \) [63–70]. In principle, the non-Gaussian feature of the statistics of the primordial curvature perturbations should depend strongly on the generation mechanism of the primordial fluctuations—that is, the inflation models—and we should evaluate \( \beta \) for each exact form of the probability distribution function \( P(\delta) \). However, it might be a hard task to evaluate the exact form of the probability distribution function for each inflationary model, and hence it should be convenient to give an approximate formulation to take into account the non-Gaussian effect. One such formulation can be made by considering the cumulant expansion of the probability distribution functions [66]. In the standard single slow-roll inflation, the higher order cumulants are expected to be suppressed by the slow-roll parameter. However, as shown in the next subsection, for efficient PBH formation we need to consider non-standard inflationary models with the violation of the slow-roll condition. Thus, the effect of the non-Gaussianity can be relevant for some inflationary models [68].

2.4. Generating seeds of PBH from inflation

Here, we focus on the mechanism by which the PBH formation is induced by the primordial density fluctuations, and we give a review of inflationary mechanism which could generate the seeds of PBHs—that is, primordial curvature perturbations with large amplitudes at a certain scale.

In preparation, let us give a relation between the mass of PBHs and the comoving wavenumber of the perturbations. Basically, as we have discussed in the previous subsection, a PBH can form when an overdense region enters the Hubble horizon. If such an overdense region is sourced from the primordial curvature perturbations, the size of the overdense region can be characterized by the comoving wavenumber of the primordial perturbations, \( k \). Thus, we can obtain the relation between the Hubble scale at the PBH formation and the comoving wavenumber of the sourced primordial curvature perturbations as

\[
\frac{aH}{|at\,\text{formation}|} = k. \tag{29}
\]

During the radiation-dominated era, we have \( a \propto H^{-1/2} \) and hence the relation between the comoving wavenumber \( k \) and the Hubble parameter at the formation as \( H_{\text{form}} \propto k^2 \). Substituting the relation between \( k \) and \( H_{\text{form}} \) into the above expression for \( M_{\text{PBH}} \), we can obtain the relation between PBH mass and the comoving wavenumber as [71]

\[
M_{\text{PBH}}(k) \simeq 30 M_\odot \left( \frac{\gamma}{0.2} \right) \left( \frac{g_{\ast,\text{form}}}{10.75} \right)^{-1/6} \left( \frac{k}{2.9 \times 10^5 \text{ Mpc}^{-1}} \right)^{-2}. \tag{30}
\]

From this equation, we find that a hierarchy between the observable scales by cosmic microwave background (CMB) observations and \( 30 M_\odot \) PBHs is given by \(^{10}\)

\[
N_{\text{CMB} - 30 M_\odot} := \ln \frac{k_{30 M_\odot \text{PBHs}}}{k_{\text{CMB}}} = \ln \frac{2.9 \times 10^5 \text{ Mpc}^{-1}}{0.002 \text{ Mpc}^{-1}} \sim 20. \tag{31}
\]

On the other hand, the number of \( e \)-folds measured from the time when the present horizon scale exits the Hubble horizon during inflation to the end of inflation is required to be typically

\(^{10}\) Here, we use a pivot scale in Planck 2015 as a typical observable scale by CMB observations, \( k_{\text{CMB}} \).
Thus, based on the above discussion, we find that, if we want to form \(30 M_\odot\)-PBHs from the primordial density perturbations, we need to consider some mechanism which can amplify the perturbations during the inflationary phase\(^{11}\) (see figure 2).

### 2.4.1 Single-field inflation

First, let us consider the possibility of generating \(30 M_\odot\) PBHs in the standard slow-roll inflationary scenario. For the single-field standard slow-roll inflation, the power spectrum of the primordial curvature perturbations on comoving slicing is given by

\[
\mathcal{P}_R(k) \sim |R_c(k)|^2 \sim \left(\frac{H^2}{\dot{\phi}}\right)^2_{aH=k},
\]

where \(\dot{\phi}\) is the time derivative of the inflaton field and the subscript \(aH=k\) represents the value at a time when the scale of interest exits the Hubble horizon during inflation. For the power-law ansatz,

\[
\mathcal{P}_R(k) : = A R_c \left( k \right) \frac{k^{-n_s-1}}{k_*},
\]

Equation (32) gives a power-law index as\(^{12}\)

\[
n_s - 1 = -6 \epsilon + 2 \eta,
\]

where \(\epsilon\) and \(\eta\) are slow-roll parameters which can be characterized in terms of the potential of the inflaton field as

\[
\epsilon : = \frac{1}{16 \pi G} \left( \frac{V_\phi}{V} \right)^2, \quad \eta : = \frac{1}{8 \pi G} \frac{V_{\phi\phi}}{V},
\]

where \(V_\phi : = dV(\phi)/d\phi\) and \(V_{\phi\phi} : = d^2V(\phi)/d\phi^2\). Thus, positive large \(\eta\) can realize a blue-tilted power spectrum, that is, larger amplitude for smaller scales (larger comoving wavenumber).

\(^{11}\) References [72–76] discussed PBH formation during the QCD phase transition era, with the assumption that the phase transition is first order. However, currently the QCD phase transition has been shown to be not first order but a crossover, and hence efficient PBH formation during the QCD epoch might be difficult.

\(^{12}\) This expression is derived based on the slow-roll equations:

\[
H^2 \simeq \frac{8 \pi G}{3} V(\phi), \quad 3H \dot{\phi} \simeq -V_\phi.
\]
Current CMB observations have indicated that $P_R$ is about $10^{-9}$ (often called COBE normalization) over CMB observable scales, and the PBH formation is effective enough to be observationally interesting for $P_R = \mathcal{O}(10^{-2} - 10^{-1})$. Thus, if we want to realize such a large amplitude at $k = k_{30\,M_{\odot}\text{PBHs}} = 2.9 \times 10^5 \text{Mpc}^{-1}$ with a blue-tilted power spectrum, consistent with the COBE normalization, we need

$$n_s - 1 = \frac{\ln(P_{R,(k_{30\,M_{\odot}\text{PBHs}}})/P_{R,(k_{CMB})})}{\ln(k_{30\,M_{\odot}\text{PBHs}}/k_{CMB})} \simeq 0.85.$$  (36)

For successful inflation, we need to require $\epsilon, \eta \ll 1$. Hence, the above value for the spectral index seems to be large, and is not so easy to realize in the standard inflationary models. Furthermore, we have a strong constraint on the spectral index obtained from CMB observations, e.g. from the Planck observation, we have \[77\]

$$n_s = 0.968 \pm 0.006 \text{ at } k_\ast = 0.05 \text{ Mpc}^{-1}.$$  (37)

Thus, for PBH formation, even if we can realize the blue-tilted power spectrum, the single power-law model should be in conflict with the CMB observations.

**Running mass inflation model.** One possibility to realize a relatively large blue-tilted inflationary model consistent with the Planck result is the large running mass inflation model \[78–80\]\(^\text{13}\). In this model, the time dependence of $\eta$ during inflation could be large, that is, the scale-dependence of $n_s$ could also become large (see e.g. \[85, 86\])\(^\text{14}\). Thus, when $\eta$ remains small, so as to be consistent with the Planck result on CMB scales, but $\eta$ takes a positive larger value on smaller scales, we can realize the large amplitude of primordial curvature perturbations on an appropriately small scale which could be seeds of PBHs. For such kinds of model, the primordial power spectrum is not a simple power-law (33), but one that includes the scale-dependence of the spectral index perturbatively:

$$P_{R,(k)} := A_{R,}\left(\frac{k}{k_\ast}\right)^{n_s - 1 + \frac{\alpha_s}{2} \ln(k/k_\ast) + \frac{\beta_s}{3} \ln^2(k/k_\ast) + \cdots},$$  (38)

where $\alpha_s$ and $\beta_s$ are, respectively, called ‘running of spectral index’ and ‘running of running’ parameters. In the standard slow-roll inflation models, $\alpha_s$ and $\beta_s$ are strongly suppressed by slow-roll parameters. However, in the large running mass inflation model, these parameters would be relatively large. For the above parameterization, in order to produce $30\,M_{\odot}$ PBHs with $n_s = 0.968$ on CMB scales, we need to take $\alpha_s$ to be about 0.1, as shown by the red line in figure 3. However, assuming a negligibly small running of running parameter $\beta_s$, for such a large $\alpha_s$, PBHs with smaller masses (corresponding to larger comoving wavenumbers) would be overproduced. To avoid the overproduction of smaller PBHs, we need to have a cutoff in the primordial power spectrum at an appropriate scale. Based on the parameterization given as (38), such a cutoff can be realized by taking into account non-negligible running of running $\beta_s$. As shown in figure 3, by employing appropriately tuned positive $\alpha_s$ and negative $\beta_s$, one can realize a broader peak in the power spectrum, which predicts the broad mass spectrum of PBHs.

\(^{13}\) Here, we focus on inflation models which could produce PBHs with $O(1 - 10)\,M_{\odot}$. Formation of smaller PBHs in a running mass inflation model has also been discussed in \[81–84\].

\(^{14}\) Note that, in order to realize the end of inflation, this type of model needs a hybrid-inflation-type mechanism with a waterfall field other than the inflaton. In this sense, the running mass inflation model is not strictly a single field model \[86\].
The recent Planck result gives constraints not only on the spectral index $n_s$, but also on the running of the spectral index $\alpha_s$ as \cite{87}

$$\alpha_s = -0.003 \pm 0.007 \text{ at } k^* = 0.05 \text{ Mpc}^{-1}. \tag{39}$$

Thus, by comparing this observational constraint with the value for $\alpha_s$ shown in figure 3, we find that it is difficult to construct a viable inflationary model which can produce $30 M_\odot$ PBHs based on the perturbative scale-dependent spectral index (38).

**Inflection inflation model.** \cite{88} recently proposed a single field model which can produce a primordial power spectrum with a peak. They consider the inflaton’s potential with an inflection (plateau) point where the inflaton temporarily slows down during the inflationary phase (see figure 4). PBH formation in such an inflationary scenario with a ‘plateau’ in the scalar potential was discussed in \cite{89}. Ivanov et al \cite{89} calculated the spectrum of the adiabatic perturbations for the schematic representation of a scalar potential featuring two breaks and a flat plateau between these breaks\footnote{Bullock and Primack \cite{63} also studied PBH formation in several toy models.}. The reason why such a model can produce a peak in the primordial spectrum can easily be understood as follows. Let us recall the expression of the primordial power spectrum given by (32). This expression can be rewritten in terms of the slow-roll parameter as

$$P_R(k) = \left(\frac{8\pi G H^2}{\epsilon}\right)_{aH=k}. \tag{40}$$

From this expression, one can find that when the inflaton temporarily slows down, the slow-roll parameter $\epsilon$ becomes more suppressed, and the power spectrum has a peak at the scales which exit the Hubble horizon during the slow-down phase. However, if the plateau is completely flat, a problem appears: that the inflaton may stay too long at the plateau and the inflationary phase continue eternity—so-called eternal inflation. To avoid this problem, in \cite{88}, a ‘near’-inflection point has been introduced in the inflaton potential. Such kinds of model in the context of $O(10) M_\odot$ PBH formation have also been discussed in \cite{90–93}\footnote{Smaller PBHs formation in the context of an inflection point inflation is discussed in \cite{91, 94–96}.}.

In fact, as pointed out in \cite{93}, the standard slow-roll conditions might be generally violated near the inflection point. In the standard slow-roll approximation, we approximate $\dot{\phi} \simeq -V_\phi/(3H)$ where $V_\phi := dV/d\phi$ and $\phi$ is an inflaton field, and this means that in the

---

Figure 3. Primordial power spectrum with large running of spectral index and running of running parameters.
equation of motion of the inflaton we can neglect the acceleration of the inflaton, $\ddot{\phi}$, term. However, if $V_\phi$ becomes too small as around the inflection point, in the equation of motion $3H\dot{\phi}$ term would become balanced with $\ddot{\phi}$ term, that is, $|\ddot{\phi}| \approx |3H\dot{\phi}| (\gg |V_\phi|)$. Thus, conditions of the above discussion based on the slow-roll parameters would be violated. In fact, based on the detailed calculation, in such an inflection-point inflation, the amplification of the primordial density fluctuations can be realized, but $P_R \simeq 10^{-4}$ at most [88, 93].

This type of model tends to predict a broader peak, $\Delta \log k \gtrsim O(10)$, in the primordial power spectrum, and hence the mass spectrum of PBHs formed would also be broader [88].

**Single-field chaotic new inflation model.** Yokoyama–Saito [65, 66] proposed the possibility of PBH formation in the context of double inflation even for the single field case, called chaotic new inflation. In this scenario, the inflaton potential is basically given by so-called Coleman–Weinberg or double-well type potentials, which have been studied in the context of the new inflation scenario (figure 5).

In the standard new inflation scenario, inflation is caused by an inflaton which starts slow-roll from the vicinity of the origin, which is an unstable local maximum. On the other hand, if the initial amplitude of the scalar field is Planck scale largely exceeding the global minimum, it causes chaotic inflation. For potentials of this kind, if we tune the model parameters, two inflationary phases could be realized. As a first phase of inflation, chaotic inflation occurs with a large field value of the inflaton. After the phase of chaotic inflation, the inflaton oscillates...
around the origin, as well as in the standard chaotic inflationary model. However, in this model, the potential has a feature which can realize new inflation around the origin. Thus, after the oscillation, with some tuning of the model parameters, new inflation occurs as a second phase of the inflation.

In this scenario, the amplification of the primordial curvature perturbations would be expected to be efficient for the modes which exit the horizon at the transition era from the first phase to the second, where the slow-roll conditions are temporarily violated. As in the case of the inflection point inflation model, around the local maximum the slow-roll conditions are temporarily violated, and then the enhancement of the curvature perturbations occurs. This model would predict a relatively sharp peak in the power spectrum of the curvature perturbations, $\Delta \log k \lesssim O(10)$, and the PBH mass spectrum would become even narrower [66].

2.4.2. Multi-scalar inflation. Next, let us consider multi-scalar inflation models which can produce seeds of PBHs, i.e. the primordial curvature perturbations with large amplitudes.

Double inflation model. As we have discussed in the single-field inflation case, if we have a phase where the inflaton temporarily slows down during inflation, corresponding to a very flat region in the inflationary potential, we can realize a peak in the primordial power spectrum at an appropriate scale. The important point in the above models realizing the enhancement of the primordial curvature perturbations to form PBHs during the inflationary era is temporal violation of the standard slow-roll conditions. In the case of multi-scalar models, features of this kind could easily be realized in the inflation dynamics [97, 98]. Garcia-Bellido et al [99] investigated the possibility of PBH production in the context of hybrid inflation, where inflationary dynamics depends on two scalar fields.

In the original types of hybrid inflation, one scalar field which possesses a false vacuum energy is assumed to be massive, and inflation was driven by this false vacuum energy. Another field slow-rolls down the flat potential, and when it reaches a critical value, the massive field becomes tachyonic and inflation abruptly ends. In general, we can also consider the case where both fields existing in hybrid inflation are not massive\(^{17}\). In such a case, we can realize two stages of inflation (double inflation) and a phase of transition (see figure 6). Basically, in the transition phase, fields roll down the very flat region in the potential and hence the primordial curvature perturbations which exit the Hubble horizon during this phase may be amplified, as we have discussed in the inflection model.

Lots of models have now been proposed for PBH formation in the context of double inflation [59]. References [71, 101, 102] discuss the construction of such double inflation models in supergravity, named smooth hybrid new inflation.

In this type of scenario, the expected power spectrum of the curvature perturbations strongly depends on the details of the model. As an example, as shown in [71, 103, 104], the expected power spectrum has two peaks, one of which is very sharp and the other much broader.

Curvaton model. Another possibility to generate the primordial curvature perturbations with large amplitude at a certain scale can be considered in the context of the so-called curvaton scenario. In the simple curvaton scenario, we have two scalar fields: one (the inflaton) is responsible for the accelerating expansion of the Universe, that is, inflation; and the other (the curvaton) responsible for generating the primordial curvature perturbations. As we have mentioned, one of the difficulties of PBH formation in the single field inflation is that we

\(^{17}\) Non-Gaussianity of the super-Hubble perturbations in this case was shown to be very tiny [100].
should realize the large amplitude of fluctuations at a small scale for PBH formation and the COBE normalization on the CMB scales for the primordial curvature perturbations, simultaneously. In some extension of the simple curvaton scenario, the primordial curvature perturbations could be generated both from the inflaton and the curvaton fluctuations. Thus, if we can construct a model in which the primordial curvature perturbations on the CMB scales are generated from the inflaton fluctuations, while those on small scales are generated from the curvaton fluctuations, we can realize a successful inflationary model for PBH formation consistent with the CMB observations [38, 105, 106].

Basically, in this curvaton scenario, the power spectrum of the curvaton fluctuations (primordial isocurvature perturbations) has a cut-off on a large scale and a scale-invariant character on smaller scales. During the radiation-dominated era after reheating, the curvaton field starts to oscillate, and behaves as non-relativistic matter. Thus, the energy density of the curvaton field gradually gets to contribute to the total energy density of the Universe, meaning that the curvaton fluctuations, which are initially isocurvature perturbations, are gradually transformed into adiabatic curvature perturbations—i.e. the adiabatic curvature perturbations should realize the large amplitude of fluctuations at a small scale for PBH formation and the COBE normalization on the CMB scales for the primordial curvature perturbations, simultaneously. In some extension of the simple curvaton scenario, the primordial curvature perturbations could be generated both from the inflaton and the curvaton fluctuations. Thus, if we can construct a model in which the primordial curvature perturbations on the CMB scales are generated from the inflaton fluctuations, while those on small scales are generated from the curvaton fluctuations, we can realize a successful inflationary model for PBH formation consistent with the CMB observations [38, 105, 106].

Basically, in this curvaton scenario, the power spectrum of the curvaton fluctuations (primordial isocurvature perturbations) has a cut-off on a large scale and a scale-invariant character on smaller scales. During the radiation-dominated era after reheating, the curvaton field starts to oscillate, and behaves as non-relativistic matter. Thus, the energy density of the curvaton field gradually gets to contribute to the total energy density of the Universe, meaning that the curvaton fluctuations, which are initially isocurvature perturbations, are gradually transformed into adiabatic curvature perturbations—i.e. the adiabatic curvature perturbations should realize the large amplitude of fluctuations at a small scale for PBH formation and the COBE normalization on the CMB scales for the primordial curvature perturbations, simultaneously. In some extension of the simple curvaton scenario, the primordial curvature perturbations could be generated both from the inflaton and the curvaton fluctuations. Thus, if we can construct a model in which the primordial curvature perturbations on the CMB scales are generated from the inflaton fluctuations, while those on small scales are generated from the curvaton fluctuations, we can realize a successful inflationary model for PBH formation consistent with the CMB observations [38, 105, 106].
grow in time. In the standard curvaton scenario, the curvaton decays into radiation at a certain time. After the curvaton decays, the adiabatic curvature perturbations become constant in time. As shown in figure 7, in this scenario, both components of the primordial curvature perturbation spectra at the horizon re-entry are smooth. Thus, the mass spectrum of the formed PBHs would be broad. In [108], it was shown that if there are multiple curvaton fields, they can temporarily enhance the curvature perturbations on all super-Hubble scales, and PBHs could be copiously produced as those modes re-enter the Hubble horizon, until the enhancement terminates. There are also several models in which the primordial curvature perturbations were enhanced by non-trivial interactions between the inflaton and some field, e.g. a gauge field [110].

3. Observational constraints on non-evaporating PBHs ($\gtrsim 10^{15} \text{ g}$)

In this section, we review how the upper limit on the fraction of PBHs in dark matter $f_{\text{PBH}} \equiv \Omega_{\text{PBH}}/\Omega_{\text{DM}}$ can be obtained for various mass ranges of non-evaporating PBHs ($M_{\text{PBH}} \gtrsim 10^{15}$ g). Broadly speaking, constraints are divided into direct and indirect ones. The direct constraints are derived by investigating the observational effects that PBHs directly trigger by their gravitational potential, and are thus independent of the mechanisms of PBH formation. The direct constraints can be further classified into four categories by the manner of PBH effect: gravitational lensing, dynamical effect, accretion, and growth of large scale structures. All these topics are covered in this section. Indirect constraints are those that can be obtained by observational effects that are not caused by the PBHs but something else which is deeply connected to PBHs. Although such constraints cannot be applied to all the possible PBH scenarios, they are powerful in excluding some of them. Several known indirect constraints are also reviewed in the latter part of this section.

The existing constraints will be tightened (or may be replaced by the detection of PBHs) in the future both by the improvement of the apparatuses similar to the existing ones and the launch of the qualitatively new observations such as 21 cm lines. Future constraints are also briefly touched at the end of this section.

All the constraints discussed in this section are based on electromagnetic signals. The new observable, namely gravitational waves, in the context of PBHs is the topic of the next section. It is important to keep in mind that all the constraints in this section are derived under the assumption that the PBH mass function is monochromatic, which is valid when the width of the mass function is sufficiently narrow. The case for a broad mass spectrum requires separate analysis, and will be briefly addressed in the last part of this section.

Before ending the introduction of this section, we briefly mention the constraints on PBHs lighter than $\sim 10^{15}$ g that have already evaporated, or are in the final stage of evaporation, via Hawking radiation [7]. Although those PBHs, depending on their mass, do not exist or are fading at the present epoch, high-energy particles emitted from PBHs leave certain signals from which we can place the upper limit on the PBH abundance. These include production of the lightest supersymmetric particles (if they exist) ($10^4 \text{ g} < M_{\text{PBH}} < 10^9 \text{ g}$), entropy production in the early universe ($10^6 \text{ g} < M_{\text{PBH}} < 10^8 \text{ g}$), change of the abundance of the light elements produced by the big bang nucleosynthesis ($10^9 \text{ g} < M_{\text{PBH}} < 10^{13} \text{ g}$), extragalactic photon background ($10^{14} \text{ g} < M_{\text{PBH}} < 10^{15} \text{ g}$), and damping of the CMB temperature anisotropies on small scales by modifying the cosmic ionization history ($10^{13} \text{ g} < M_{\text{PBH}} < 10^{14} \text{ g}$).

19 The temporal enhancement of the curvature perturbations and PBH formation were also studied in the case where the non-inflaton field has a Galileon-type kinetic interaction [109].
Topical Review

By comparing these effects with observations, upper limits on the PBH fraction $\beta$ defined by equation (20) for various PBH mass (assuming monochromatic mass function) can be obtained (for comprehensive study on this topic, see [7, 111]). Except for the constraints from entropy production and primordial helium abundance, these limits are severe in the sense that they allow only a tiny fraction of PBHs in dark matter at any cosmic time before the PBHs evaporate. Assuming Gaussianity of the primordial perturbation, these limits can be converted to the upper limit on the power spectrum of the curvature perturbation $R$ as $P_R(k) \lesssim 10^{-2}$ for $10^9 < k/\text{pc}^{-1} < 10^{14}$ [111]. Although the upper limit on $\beta$ varies by many orders of magnitude over the PBH mass range corresponding to the above range of $k$, the upper limit in terms of $P_R$ is insensitive to $k$, which is traced back to the fact that $\beta$ depends on the $P_R$ exponentially (see section 2.3).

3.1. Gravitational lensing

Gravitational lensing is a very powerful method to constrain/detect PBHs. The outstanding virtue of gravitational lensing is that the individual lensing signal is solely based on the gravitational physics, and does not suffer from the uncertainties that exist in studying electromagnetic signals resulting from the interaction between PBHs and surrounding matter.

If PBHs are present in the Universe, they cause gravitational lensing on background objects such as stars. Under the thin-lens approximation, in which the deflection of light occurs at a point on the lens plane (which is a very good approximation in the astrophysical situations we are considering), the lens equation can be written as (figure 8)

$$\theta D_S = D_S \beta + D_{LS} \alpha, \quad \alpha = \frac{4GM_{BH}}{D_l \theta}. \quad (41)$$

In terms of the distance on the lens plane $r = D_l \theta$, the lens equation becomes

$$r^2 - r_0 r - R_E^2 = 0, \quad (42)$$

where $r_0 = D_L \beta$ and $R_E = \sqrt{\frac{4GM_{BH}}{D_S} \left| \frac{\partial^2 \theta}{\partial z^2} \right|}$ is the Einstein radius. Thus, the positions of the lensed images are given by

![Figure 8. Trajectory of light ray bent by the lens object (BH). $\alpha$ is the deflection angle.](image-url)
There are thus two lensed images, at \( r_1 \) and \( r_2 \). The only exception is when the source, lens object and observer are on the same line, i.e. \( \beta = 1 \). In this case, the image on the lens plane becomes a circle (the so-called Einstein Ring) with its radius given by \( R_E \). Lensing effects become significant when \( r_0 \lesssim R_E \) and typical angular separation of the two images is

\[
\Delta \sim \frac{R_E}{D_L} = \sqrt{\frac{4GM_{BH}}{D_S}} \frac{1-x}{x} \approx 0.3 \text{ mas} \left( \frac{M_{BH}}{10 \, M_\odot} \right)^{1/2} \left( \frac{D_S}{100 \, \text{kpc}} \right)^{-1/2} \sqrt{\frac{1-x}{x}},
\]

where we have introduced the parametrization as \( D_L = D_S x \) (\( 0 < x < 1 \)).

### 3.1.1. Microlensing

Gravitational microlensing refers to gravitational lensing events in which the angular separation of the images lensed by a compact object is so small that the individual images cannot be resolved by observation [112]. For instance, the angular separation given by equation (44) is much smaller than the angular resolution of the existing (or past) microlensing experiments such as the MACHO Project and EROS collaboration, which is about 0.6 arcseconds. In a gravitational microlensing event, what can be observed is only the superposition of two images which is brighter than the original source. The magnification factor \( A \), normalized to unity in the absence of the microlensing, is given by

\[
A = \frac{u^2 + 2}{u\sqrt{u^2 + 4}}, \quad u = \frac{r_0}{R_E};
\]

at \( r_0 = R_E, A = 1.34 \). When the lens object is moving relative to the line of sight connecting the source and the observer, \( u \) becomes time-dependent, and the magnification varies in time. Assuming a constant tangential velocity \( v_{BH} \), we can write \( u \) as

\[
u^2 = \frac{v^2 t^2 + b^2}{R_E^2} = \frac{v^2 t^2}{R_E^2} + u_{\text{min}}^2,
\]

where \( b \) is the impact parameter of the source in the lens plane, and the origin of \( t \) is chosen so that the source image in the lens plane comes closest to the lens object at \( t = 0 \). Magnification curves for various values of \( u_{\text{min}} \) are given in figure 9. As we can see, the magnification is
maximum at $t = 0$ and symmetric about it. The time scale for the rise and fall-off of the magnification is given by

$$T = \frac{R_E}{v} = \sqrt{\frac{4G M_{BH} D_S x(1-x)}{v^2}} \approx 2 \text{ yr} \sqrt{x(1-x)} \left( \frac{M_{BH}}{10 M_\odot} \right)^{1/2} \left( \frac{D_S}{100 \text{ kpc}} \right)^{1/2} \left( \frac{v}{200 \text{ km s}^{-1}} \right)^{-1}. \quad (47)$$

We find that the time scale is longer for larger lens mass and greater distance from the source.

A useful quantity which is used in the context of the microlensing is the optical depth $\tau$ [113, 114]. This quantity measures how likely a background source is micro-lensed by the compact objects with magnification factor greater than 1.34$^{20}$. In other words, the optical depth, when it is much smaller than unity, is the probability that the light from the source passes inside the circle defined by the Einstein radius on the lens plane. Assuming that all the lens objects have the same mass, the optical depth can be written as

$$\tau = \int_0^{D_L} dr \, \rho_{BH}(r) \pi R_E^2(r) = 4\pi G \int_0^{D_L} dD_L \, \rho_{BH}(D_L) \frac{D_L(D_S - D_L)}{D_S}, \quad (48)$$

where $\rho_{BH} = M_{BH}/\rho_{BH}$ is the energy density of PBHs. Interestingly, the last expression of the optical depth shows that it does not depend explicitly on the mass of PBHs. What determines the optical depth is how much of the total mass of dark matter PBHs constitute.

In 1986, Paczyński pointed out that it would be possible to test the hypothesis that some fraction of dark matter is in the form of compact objects by monitoring a few million stars in the Magellanic Clouds (MCs) [115]. Light from any star in the MCs passes through the dark matter halo encompassing the Milky Way Galaxy before it reaches the observer on the Earth. If the dark matter halo entirely (or partially) consists of compact objects, they produce microlensing events on the source stars in MC. According to the estimate in [115], the optical depth for the microlensing event is given by

$$\tau \sim 10^{-6} f_{PBH}. \quad (49)$$

If $f_{PBH} \sim 1$, about one star among a few million stars must produce a microlensing event. Conversely, non-observation of microlensing events places an upper limit on PBH abundance. The time scale of the magnification is $\sim 2$ yr for $M_{BH} = 10 M_\odot$ and $\sim 400$ min for $M_{BH} = 10^{-6} M_\odot$ (see equation (47)). Thus, a few years of observation of the stars in the MCs enables us to place a constraint on the number of PBHs with up to $M_{BH} \sim 10 M_\odot$.

It is important to mention that there is a lower limit on the PBH mass below which the finite-size effect of the source changes the magnification by the microlensing from that in the point lens case. The finite-size effect becomes relevant when the angular size of the source star projected on the lens plane becomes comparable to or larger than the angular size of the impact parameter $r_0$. The peak magnification becomes smaller or larger than that in the case in the point lens case depending on the magnitude relation between the projected source size and $r_0$ [116]. In particular, when $r_0$ is much smaller than the source size, the maximum of $A$ becomes

$$A_{\text{max}} = \frac{\sqrt{4 + r^2}}{r} \approx \frac{xR_{\text{star}}}{R_E}. \quad (50)$$

Thus, $A_{\text{max}}$ approaches unity for $r \gtrsim 1$. In terms of the lens mass, the condition $r \gtrsim 1$ can be written as

$^{20}$The actual threshold for $A$ for the detection of the microlensing event depends on the detection methods in individual experiments. The value $A = 1.34$ is just for convenience.
\[ M_{\text{PBH}} \lesssim \frac{D_L R_{\text{star}}^3}{4GD_S^2} \sim 3 \times 10^{-9} M_\odot \left( \frac{R_{\text{star}}}{R_\odot} \right)^2 \left( \frac{D_L}{10 \text{ kpc}} \right) \left( \frac{D_S}{100 \text{ kpc}} \right)^{-2}. \]  

(51)

If the lens mass satisfies this inequality, magnification is significantly reduced from that in the point lens case.

Following the intriguing proposal by Paczyński, several groups around the world carried out year-scale observations of millions of stars in the MCs. Detection of the microlensing events were first reported both by the MACHO Project [117], in which one event suggesting a compact object with \( \sim 0.1 M_\odot \) was detected and the EROS (Expérience pour la Recherche d’Objets Sombres) collaboration [118] in which two events suggesting compact objects between a few \( \times 10^{-2} \) and \( 1 M_\odot \) were detected. Continuing monitoring the stars in the Large MC (LMC), the MACHO Project finally reported the detection of 13–17 microlensing events and the lensing optical depth as [119]

\[ \tau_{\text{LMC}} = 1.2^{+0.4}_{-0.3} \times 10^{-7}, \]  

(52)

and suggested about 20% contribution of the compact objects in a mass range \( 0.1 \sim 1 M_\odot \) to the dark matter halo mass. The EROS collaboration [120], based on the 6.7 yr observations of stars in the LMC and Small MC (SMC), later reported that only one candidate microlensing event of the star in SMC was detected and the optical depth as

\[ \tau_{\text{LMC}} < 0.36 \times 10^{-7}, \quad 0.085 \times 10^{-7} < \tau_{\text{SMC}} < 8.0 \times 10^{-7}. \]  

(53)

Based on the earlier estimation that the optical depth for microlensing by objects in the SMC is \( \tau_{\text{SMC}} \sim 0.4 \times 10^{-7} \) [121], it was suggested that the detection of one event is consistent with the expectations of self-lensing by objects in the SMC. According to the EROS collaboration [120], the severe constraint on \( \tau_{\text{LMC}} \) that appears to be in contradiction with that measured by the MACHO Project could possibly be explained by the contamination of variable stars, the self-lensing in the inner parts of the LMC, blending effects (reconstructed fluxes receive contributions from more than one star), and the possibility that the fields observed by the MACHO Project lie just behind a clumpy region of the compact objects that does not cover the EROS-fields. The constraint found by the EROS collaboration is shown in figure 10.

After the EROS collaboration, another experiment called OGLE (Optical Gravitational Lensing Experiment) also reported the detection of two candidate microlensing events in the LMC [122] and three events in the SMC [123] and obtained

\[ \tau_{\text{LMC}} = (0.16 \pm 0.12) \times 10^{-7}, \quad \tau_{\text{SMC}} = (1.30 \pm 1.01) \times 10^{-7}. \]  

(54)

While the four events are consistent with contributions from the self-lensing in the Galactic disk and SMC, detailed analysis [124] shows that there is a possibility that one event (OGLE-SMC-02) is a binary BH lens in the dark matter halo, although the self-lensing scenario also remains as a possible explanation. The upper limit on the PBH fraction by OGLE is shown in figure 10. However, it is important to keep in mind that this constraint is derived under the assumption that all OGLE events are due to the self-lensing/background signals.

In [125], it was proposed that the data obtained by the Kepler satellite could be used to obtain a new upper limit on the abundance of PBHs in the mass range \( 5 \times 10^{-10} M_\odot \sim 10^{-4} M_\odot \), which is much smaller than the stellar mass covered by the previous microlensing experiments mentioned above. Although both the number of stars that are monitored (\( \sim 0.1 \) millions) and the distance to the stars (\( \sim 1 \text{ kpc} \)) are much smaller than the microlensing experiments targeting the MCs, advantages due to the high sensitivity of the photometry by the Kepler mission and the large size of the microlensing cross section can compensate for the above
disadvantages. It was reported in [126] that after removing the background events, no microlensing candidates have been found in the two years of Kepler data. By the non-detection of the microlensing events, an upper limit on the PBH fraction was obtained [126]. The resultant constraint is shown in figure 10. The actual constraint is much weaker than the theoretical estimate derived in [125]. According to [126], this is due to the optimistic assumption made in [125], which underestimates flare events of short duration.

Recently, a very stringent upper limit on the PBH fraction for the mass range $\left(10^{-13} M_\odot, 10^{-5} M_\odot\right)$ has been obtained by the observations of more than ten million stars in M31 by the Subaru hyper suprime-cam (HSC) [127]. A powerful and unique feature of the HSC is its high cadence: observations of stars every 2 min (90 s exposure plus about 35 s for readout) over about 7 h (196 exposures in total). This is why the HSC is sensitive to PBH masses much smaller than one solar mass (see equation (47)). After eliminating the fake events, one microlensing candidate was found. No convincing conclusion has been obtained as to the nature of the microlensing event. However, depending on whether the microlensing candidate is attributed to PBH or not, it is possible to derive an upper limit on PBH abundance, which is shown in figure 10. The value so obtained provides the strongest constraint for a wide mass range that fully covers the mass range probed by the Kepler mission.

A microlensing constraint on $f_{\text{PBH}}$ has been obtained in [128, 129] in a manner which is uniquely different from the ones discussed above. Observations by Hubble Space Telescope found a fast transient which turned out to be a magnified star at $z = 1.49$ [130] (estimated magnification is $A > 2000$, which is much bigger than 1—?). The star being at the vicinity of the caustic of the galaxy cluster called MACS J1149.6 + 2223 located at $z = 0.544$ (caustic is the place in the source plane where the magnification diverges) suggests that the magnification is caused by caustic crossing. However, detailed analysis showed that the observed magnification is not large enough to be explained solely by the caustic generated by the galaxy cluster, for which the magnification should be as large as $A \sim 10^6$ (the magnification does not diverge in reality, due to the finite size of the source). A natural explanation is contribution from a point mass lens in the same galaxy cluster. The caustic is known to be significantly distorted even by the presence of tiny point masses. The distortion of the caustic results in the reduction of magnification compared to the case where point masses are absent. The more point masses there are, the greater the reduction of magnification. The observed magnification curve requires a certain number of point masses. This does not mean that this observation yields both upper and lower limits on $f_{\text{PBH}}$, since there are stars responsible for the intra-cluster light (ICL) that produce the same effect on the caustic as the PBHs do. Within the uncertainty of the abundance of the ICL stars, it was shown that the ICL stars can explain the observed fast transient. As a result, only the upper limit on $f_{\text{PBH}}$ has been obtained. The constraint, which was given in [129], is shown in figure 10.

Quasar microlensing is also a useful method to detect/constrain PBHs [131]. Some distant quasars are observed as pairs due to gravitational lensing by intervening galaxies. Such quasar pairs can be additionally microlensed by PBHs (or any other types of compact object) that reside between the quasars and the Earth. Contrary to microlensing surveys such as the MACHO project, which look for the characteristic time variation of the source brightness, a salient aspect of the quasar microlensing is that we need only a single-epoch flux measurement for each quasar pair. The point of this method is that continuum component and the narrow emission lines of the photo spectrum for each quasar image of the pair show different sensitivity to the microlensing. Emission lines are generally thought to have originated from a region larger than the region for the continuum component, and not to be affected by the microlensing, while the continuum component is microlensed by PBHs. This enables extracting the signal of microlensing by comparing the flux ratio of the emission lines with...
that of the continuum component from the photo spectrum taken at a single epoch. In [132], the above program has been applied to the microlensing data of 24 gravitationally lensed quasars (see [133] for earlier study) and a tight upper limit on \( f_{\text{PBH}} \) was obtained for stellar-mass PBHs. However, since the upper limit on \( f_{\text{PBH}} \) was not given explicitly, we do not show it in figure 10.

### 3.1.2. Millilensing

Supermassive PBHs (\( \gtrsim 10^6 \, M_\odot \)) are also a target of search by the gravitational lensing. Supermassive compact objects (not necessarily PBHs) filling the Universe produce two lensed images of the source at the cosmological distance with mas scale angular separation (hence ‘millilensing’). Press and Gunn proposed that such double images can be resolved by the VLBI maps [134], and estimated the probability of such events. By using the VLBI observations of 48 compact radio sources, the constraint on the PBH fraction \( f_{\text{PBH}} < 0.2 \) (95 CL) for \( 10^7 < M_{\text{PBH}} / M_\odot < 10^9 \) was obtained under the assumption of the Einstein–de Sitter Universe in [135]. Later, the constraint was refined in [136], in which VLBI maps of 300 compact radio sources were analyzed, and no evidence of millilensing was found. Consequently, the tighter upper limit on \( f_{\text{PBH}} \) shown in figure 10 was obtained for the wider PBH mass range.

### 3.1.3. Femtolensing

PBHs with much smaller mass, such as \( 10^{-16} \, M_\odot \lesssim M_{\text{PBH}} \lesssim 10^{-13} \, M_\odot \), yet in the non-evaporating regime, can still be probed by gravitational lensing in a very unique manner. For the gravitational lensing by such tiny lens masses, the wave effects of light become important. When the wavelength of light from the source star becomes comparable to or longer than the Schwarzschild radius of the lens object, the diffraction considerably changes the amplification from that in the geometric optics approximation [137–139]. In the quasi-geometrical optics approximation, the magnification factor is given by [140]

\[
A = \frac{u^2 + 2 + 2 \sin(E_\gamma T_{12})}{u \sqrt{u^2 + 4}},
\]

(55)

where \( E_\gamma \) is the energy of gamma-ray photons, and \( T_{12} \) is the difference in arrival time between the two images,

\[
T_{12} = 4GM_{\text{PBH}} \left[ \frac{1}{2} u \sqrt{u^2 + 4} + \ln \left( \frac{\sqrt{u^2 + 4} + u}{\sqrt{u^2 + 4} - u} \right) \right].
\]

(56)

This shows that the arrival time difference is of the order of the Schwarzschild radius of the lens object for the typical value of \( u \) (\( \sim 1 \)) for which the signal of the gravitational lensing is strong. Effect of diffraction appears as a sinusoidal oscillation part in the magnification factor. This arises since the magnification factor is the square of the summation (not the summation of the square) of the lensed electro-magnetic waves corresponding to the individual images and the phase of the two waves differs by the arrival time difference multiplied by the photon energy. In other words, the sinusoidal part represents the interference of two images. In the geometrical optics limit, the oscillations become extremely fine, and disappear when \( A \) is averaged over the energy bin \( \Delta E_\gamma \sim 1 / T_{12} \). As a result, the standard formula (45) is recovered.

Gould pointed out that if small PBHs in the mass range \( 10^{-16} \, M_\odot \lesssim M_{\text{PBH}} \lesssim 10^{-13} \, M_\odot \) are distributed in the Universe, they should induce oscillatory features in the photon spectrum of the gamma-ray bursts (GRBs) that occur at cosmological distances [141]. This PBH mass range corresponds to the interval of \( E_\gamma^{-1} \), where \( 10 \, \text{keV} \lesssim E_\gamma \lesssim \text{MeV} \). The angular separation of the source at the cosmological distance lensed by such tiny PBHs is around the femto-as scale, thus the name ‘femtolensing’ [141]. In [142], such an oscillatory feature was searched
for in the GRB data obtained by the Fermi gamma-ray burst monitor (see [143] for the earlier constraint derived using the BATSE and Ulysses data). The GRB spectra were well fitted by the standard GRB models, and no evidence of femtolensing was found. This result excludes the possibility that these tiny PBHs constitute all dark matter [142]. The constraint graph is shown in figure 10.

3.2. Dynamical constraints

To a certain degree, PBHs affect any astrophysical system by gravitational interaction. By appropriately evaluating the impact of PBHs on astrophysical systems, and making comparison with observations, it is possible to put an upper limit on the PBH fraction for a wide range of PBH mass. To date, various astrophysical systems have been considered in this context.

3.2.1. Disruption of white dwarfs. White dwarfs are stars of about one solar mass and of size comparable to that of the Earth. Due to their compactness, their surface gravity is much stronger than that of main sequence stars. The strong self-gravity of white dwarfs is supported by the electron degeneracy pressure that is effective even at zero temperature [144]. White dwarfs are believed to be the progenitors of type Ia supernovae. When the mass of white dwarfs grows close to the Chandrasekhar limit, e.g. by accretion, nuclear fusion sets in and undergoes thermonuclear runaway.

It was pointed out in [145] that the passage of a PBH through a white dwarf can ignite the thermonuclear runaway that eventually makes a white dwarf well below the Chandrasekhar limit explode. The basic idea of [145] is as follows. When the PBH passes inside the white dwarf, all the nearby particles inside a thin tube along the trajectory of the PBH acquire energy due to the gravitational force from the PBH for a short time. Thus, the PBH loses kinetic energy, though by a the negligible amount compared with its initial kinetic energy. As a result, the temperature of the thin tube increases, and the nuclear fusion rate is significantly boosted—as the fusion rate typically has strong dependence on the temperature. If the time scale of the dissipation of the heat from the thin tube is shorter than the time scale on which the nuclear fusion occurs, then fusion occurs, releases energy and increases the temperature of the region outside the tube. Since the dissipation becomes less efficient for larger volumes, fusion is facilitated, ending up with an explosion.

Thus, if there were too many PBHs above a certain mass, the white dwarfs with corresponding mass could not exist in the present Universe. Conversely, we can constrain the abundance of the PBHs by using the observational confirmation of white dwarfs. We expect this argument to constrain a certain PBH mass range: light PBHs do not give sufficient heat to the white dwarfs, simply because of the weakness of their gravitational effect; and white dwarfs do not encounter many heavy PBHs, simply because of their rareness. The upper limit on the PBH fraction as a function of PBH mass is given in figure 10. It is also possible to derive constraints based on the requirement that the rate of supernovae induced by PBHs does not exceed the observed rate, but it is found that the derived constraint is less robust compared to that based on the observation of white dwarfs [145].

3.2.2. Disruption of neutron stars. Neutron stars are stars that are much more compact than white dwarfs. Their density is around the nuclear density, and the stars are largely supported by the degenerate pressure of neutrons. Neutron stars can also be used to constrain the abundance of PBHs, although in a slightly different manner from the case of the white dwarfs in section 3.2.1.
As is the case in white dwarfs, a PBH loses kinetic energy by the dynamical friction during its passage through a neutron star. The effect of PBHs on neutron stars was investigated in [146]. According to [146], the energy transferred from the PBH to the neutron star is given by

$$E_{\text{loss}} \approx 6.3 \times 10^{-12} \left( \frac{M_{\text{PBH}}}{10^{22} \, \text{g}} \right).$$

(57)

If this energy is greater than the initial kinetic energy of the PBH, then the PBH becomes gravitationally bound to the neutron star. After the first passage, the PBH subsequently undergoes orbital oscillations, passing through the neutron stars with every half oscillation period, and gradually loses all kinetic energy by transferring it to the neutron star. After a certain number of oscillations, the PBH finally becomes trapped inside the neutron star. The time scale for this to happen is given by [146]

$$t \sim 4 \times 10^4 \, \text{yr} \left( \frac{M_{\text{PBH}}}{10^{22} \, \text{g}} \right)^{-3/2}.$$  

(58)

Once the PBH lies inside the neutron star, it quickly accretes the nuclear matter and destroys the star [147]. Thus, neutron stars must not be exposed to frequent encounters with the PBHs, from which a certain upper limit on PBH abundance can be derived.

Equation (57) shows that the trapping of PBHs is effective for low-velocity PBHs. It was proposed in [146] that a strong constraint on the PBH abundance can be obtained from the observations of neutron stars in the cores of globular clusters, where the typical velocity dispersion of PBHs is \( \sim 10 \, \text{km s}^{-1} \) (see figure 10). However, it is important to keep in mind that the constraint relies on the assumption that there are PBHs as dark matter at the cores of globular clusters, and it is not known observationally whether dark matter exists in such regions. A much stronger constraint for the same phenomenon—PBH capture by neutron stars—has been obtained based on a different method for estimating the energy loss of the PBH in [148]. The subsequent papers [149, 150] preclude this possibility.

While the references above place an upper limit on \( f_{\text{PBH}} \), it was suggested in [151] that PBHs in the mass range \( 10^{-14} < M_{\text{PBH}}/M_{\odot} < 10^{-8} \) that account for a tiny fraction (a few percent) of dark matter can explain the observed r-process element abundances. According to the analysis in [151], millisecond pulsars in dense dark matter regions such as the Galactic center and dwarf spheroidal galaxies can efficiently swallow floating PBHs. As the eaten PBH swallows back the nuclear matter of the neutron star from the inside, the neutron star shrinks by losing mass to the PBH. Due to the conservation of angular momentum, the neutron star spins up as it shrinks. For millisecond pulsars, the rotational velocity at the equator exceeds the escape velocity at some point, and ejection of relatively cold neutron-rich material of 0.1–0.5 \( M_{\odot} \) results. r-process nucleosynthesis then occurs in the ejected material, which may explain the observed r-process abundance pattern (see also [152, 153] for follow-up studies.).

3.2.3. Disruption of wide halo binaries. Wide halo binaries are binaries of stars with wide separation (even \( O(1) \) pc is possible) residing in the Galactic halo. Because of their weak binding energy, they are vulnerable to disruption from encounters with PBHs. Thus, observational confirmation of wide halo binaries enables us to place an upper limit on PBH abundance.

Let us consider a situation where PBHs intermittently pass near the binary and transfer some energy to the binary. We denote the semi-major axis and mass of the binary by \( a \) and \( m_b \), respectively. When the transferred energy by the PBH encounter becomes greater than the binding energy of the binary, the binary is disrupted. If the binary is disrupted by a single
encounter, the disruption is in the catastrophic regime. On the other hand, if the binary is disrupted by the multiple encounters in which the individual encounters add energy smaller than the binding energy to the binary, the disruption is in the diffusion regime. According to [154], the time scales for the catastrophic and diffusive disruptions are given by

$$t_{d,\text{cat}} \simeq 0.07 \frac{m_b^{1/2}}{G^{1/2} \rho_{\text{PBH}} a^{1/2}}, \quad t_{d,\text{diff}} \simeq 0.002 \frac{V m_b}{G M_{\text{PBH}}/a},$$

(59)

where $V$ is the relative velocity between the binary and the PBH. Both time scales coincide for the PBH mass given by

$$M_{\text{PBH},c} \simeq 30 M_\odot \left( \frac{m_b}{2 M_\odot} \right)^{1/2} \left( \frac{a}{10^4 \, \text{AU}} \right)^{1/2} \frac{V}{200 \, \text{km s}^{-1}}.$$

(60)

Above this mass, the PBH is massive enough to disrupt by the single encounter, thus in the catastrophic regime. Below this mass, the binary is disrupted by the multiple encounters. Roughly speaking, the constraint on PBH abundance is obtained by requiring that the disruption time scale given above is longer than the age of the binary. A typical value of $t_{d,\text{cat}}$ for wide binaries is given by

$$t_{d,\text{cat}} \simeq 3 \, \text{Gyr} \left( \frac{m_b}{1 \, M_\odot} \right)^{1/2} f_{\text{PBH}}^{-1} \left( \frac{\rho_{\text{DM}}}{0.01 \, M_\odot/\text{pc}^3} \right)^{-1} \left( \frac{a}{0.1 \, \text{pc}} \right)^{-3/2},$$

(61)

which is shorter than the age of the binaries, which is of the order of the age of the Universe. This shows that wide binaries with $a \gtrsim 0.1$ pc can give a meaningful upper limit on $f_{\text{PBH}}$. Interestingly, the time scale for the catastrophic regime is independent of the PBH mass itself. Thus, in terms of the PBH fraction in total dark matter, the upper limit becomes independent of the PBH mass. On the other hand, for PBHs lighter than $M_{\text{PBH},c}$, the constraint becomes weaker for lighter PBHs.

Confirming the wide halo binaries is a challenging task. They are rare in the first place, and hard to distinguish from mere chance association. In [155], analyzing the wide binary candidates given in [156], the PBH fraction was constrained to be $\lesssim 0.2$ for PBH mass larger than a few $100 \, M_\odot$. Later, the authors of [157] found, based on their results of the radial velocity measurements of stars of four candidate wide binaries, that the second widest binary in [156] is spurious. However, the widest binary, whose separation is about 1.16 pc, was confirmed to be genuine. The derived constraint is weaker than that in [155] but still excludes PBHs as all dark matter for the mass range similar to [155] (see figure 10.).

Finally, it is worth mentioning that the time scales given by equation (59) are obtained under the impulse approximation, which is valid when the time scale on which the gravitational interaction between the PBH and the binary is effective is shorter than the dynamical time of the binary (i.e. orbital period) [154]. The former, which is equal to the time that PBH crosses the distance of the impact parameter with the velocity $V$, is given by

$$t = \frac{2 b_{\text{max}}}{V} \sim \frac{2}{V} \left( \frac{G M_{\text{PBH}}^2 a^3}{V^2 m_b} \right)^{1/4},$$

(62)

where $b_{\text{max}}$ is the maximum impact parameter for which the encounter is catastrophic. This time scale becomes equal to the orbital period $\sim 2\pi \sqrt{a^3/(G m_b)}$ for

$$M_{\text{PBH}} \simeq 9 \times 10^9 \, M_\odot.$$

(63)

The impulse approximation is justified for PBHs much lighter than this mass.
3.2.4. Disruption of globular clusters. The argument developed for the case of wide binaries can equally apply to globular clusters. The only difference between the two is that the globular clusters involve many more stars than the binaries. Plugging the typical values for the mass $m_{gc}$ and size of globular clusters $a_{gc}$ into equation (60), the PBH mass dividing the catastrophic regime and the diffusion regime becomes

$$M_{PBH,c} \simeq 10^5 M_\odot \left( \frac{m_{gc}}{10^5 M_\odot} \right)^{1/2} \left( \frac{a_{gc}}{10 \text{ pc}} \right)^{1/2} \frac{V}{200 \text{ km s}^{-1}}.$$  

(64)

The time scale of the catastrophic encounter becomes

$$t_{d,cat} \simeq 1 \text{ Gyr} \left( \frac{m_{gc}}{10^5 M_\odot} \right)^{1/2} f_{PBH}^{-1} \left( \frac{\rho_{DM}}{0.01 M_\odot/\text{pc}^3} \right) \left( \frac{a_{gc}}{10 \text{ pc}} \right)^{-3/2}.$$  

(65)

As is the case in the wide halo binaries, for $f_{PBH} \simeq 1$, this is again shorter than the age of the globular clusters, which is of the order of the age of the Universe. Thus, the observations of the existence of the globular clusters place a meaningful constraint on $f_{PBH}$. For more careful analysis, see [158].

3.2.5. Disruption of ultra-faint dwarf galaxies. Ultra-faint dwarf galaxies that are known to be present around the Galaxy and Andromeda Galaxy are also useful to place the upper limit on PBH abundance. If PBHs reside in the ultra-faint dwarf galaxies, stars inside intermittently interact with PBHs gravitationally. Two-body interaction tends to equalize the kinetic energies of the individual objects. This means that if PBHs are heavier than stars, stars on average acquire kinetic energy (dynamical heating). As a result, the interaction between stars and PBHs makes the stars move faster and spread wider and wider, which may contradict the observed distribution of stars.

Let us consider a star cluster inside an ultra-faint dwarf galaxy, and denote the cluster’s half-light radius by $r_h$. If the star cluster is immersed in a bath of PBHs, $r_h$ increases in time. Its evolution equation is given by [159]

$$\frac{dr_h}{dt} = \frac{4\sqrt{2} \pi G f_{PBH} M_{PBH} \ln \Lambda}{V} \left( \frac{M_{\text{cluster}}}{\rho_{DM} r_h^3} + 2 \beta r_h \right)^{-1},$$  

(66)

where $V$ is the typical PBH velocity relative to the cluster, $\rho_{DM}$ is the total dark matter density, $\ln \Lambda \approx 10$ is the Coulomb logarithm, and $\alpha \sim 0.4$, $\beta \sim 10$. $f_{PBH}$ is the fraction of PBHs among all dark matter. If the PBH is a subdominant component, then the rest of dark matter must be in some other form—for instance, unknown elementary particles. In such a case, dark matter other than the PBHs drags the stars gravitationally, and cools them. However, it was shown in [159] that the cooling effect is ineffective compared with the heating concerning the upper limit on $f_{PBH}$ by the considerations of the ultra-faint dwarf galaxies.

In [159], the upper limit on $f_{PBH}$ was obtained by investigating stellar distributions in ultra-faint dwarf galaxies and a star cluster in Eridanus II, a dwarf galaxy discovered as part of the dark energy survey [160, 161]. Eridanus II has a half-light radius of $\sim 300$ pc, and hosts a star cluster with half-light radius $r_h = 13$ pc. Because of the lack of our knowledge about the age of stellar clusters or dwarf galaxies, as well as their initial size, there are degrees of uncertainty in translating the observed stellar distributions to the upper limit on $f_{PBH}$. The constraints obtained with conservative assumptions are shown in figure 10. As is clear from the figure, the derived constraint is dependent on the adopted requirements. Within the range of uncertainties considered, PBHs heavier than $\sim 100 M_\odot$ as all the dark matter are excluded. In particular, the constraints in all cases are stronger than that obtained by the wide binaries.
discussed in section 3.2.3. More recently, by analyzing the Segue 1 dwarf galaxy, a similar constraint was derived in [162].

3.2.6. Dynamical friction on PBHs. If the Galactic halo is entirely or partially composed of massive PBHs, some of them must be in the region near the Galactic center. Such PBHs would receive strong dynamical friction from stars and the dark matter in the form of lighter PBHs or elementary particles, lose their kinetic energy, and spiral in to the center. If this infall time scale is shorter than the age of the Universe, accumulation of PBHs continues in the central region. As a result, the Galactic center would be dominated by a dense cluster of PBHs, or fewer but more massive BHs that result from the mergers of the accumulated PBHs. In either case, a highly concentrated region arises by the infall of the PBHs. Since there is an upper limit on the mass in the Galactic center, this limit can be translated into the fraction of PBHs in the Galactic halo for some PBH mass range.

A PBH moving in the surrounding matter loses its velocity at a rate given by [154]

\[
\frac{dV}{dt} = -\frac{4\pi G^2 M_{PBH}\rho_s \ln \Lambda}{V^2} \left( \text{erf}(X) - \frac{2X}{\sqrt{\pi}} e^{-X^2} \right), \quad X = \frac{V}{\sqrt{2}\sigma},
\]

where \(\sigma\) is the velocity dispersion of particles constituting the surrounding matter and \(\text{erf}\) is the error function. Now, suppose that PBHs move with the speed comparable to \(\sigma\) for which \(\text{erf}(X) - \frac{2X}{\sqrt{\pi}} e^{-X^2}\) is about 0.5. Then, we have

\[
\frac{dV}{dt} \approx -\frac{2\pi G^2 M_{PBH}\rho_s \ln \Lambda}{V^2}.
\]

This friction exerts torque on the PBHs, and the PBHs gradually lose angular momentum. Assuming that the orbit at each moment of time follows Kepler motion, the orbital velocity \(V\) at radius \(r\) (distance measured from the Galactic center) is given by \(V = \sqrt{GM_s(r)/r}\), where \(M_s(r)\) is the mass of the surrounding matter contained inside the radius \(r\). Using this relation, we can convert the time evolution of \(V\) into the decay rate of the orbit as

\[
\frac{1}{r} \frac{dr}{dt} = -\frac{1}{\tau_{\text{fall}}(r)}, \quad \tau_{\text{fall}}(r) = \left( 1 + \frac{4\pi r^3 \rho_s(r)}{M_s(r)} \right) \frac{1}{4\pi G^2 M_{PBH}\rho_s(r) \ln \Lambda} \left( \frac{GM_s(r)}{r} \right)^{3/2}.
\]

Thus, the PBH located at distance \(r\) spirals to the Galactic center on the time scale \(\tau_{\text{fall}}\). Near the Galactic center, the surrounding matter mainly consists of dark matter and stars in the bulge. Let us evaluate \(\tau_{\text{fall}}\) for the individual components respectively. For the stars in the bulge, we employ the Hernquist model as the density distribution [163]

\[
\rho_s = \frac{m_B}{2\pi} \frac{1}{r} \left( \frac{r + r_B}{r} \right)^{3},
\]

where \(r_B = 0.6\) kpc and \(M_B = 2.6 \times 10^{10} M_\odot\) is the total mass. The corresponding infall time is given by

\[
\tau_{\text{fall}} \approx 3\ \text{Gyr} \left( \frac{r}{r_B} + \frac{3}{1\ \text{kpc}} \right) \left( \frac{M_{BH}}{10^6 M_\odot} \right) \left( \frac{\ln \Lambda}{10} \right)^{-1}. \tag{71}
\]

For the dark matter halo, we employ the isothermal model as the density distribution,

\[
\rho_s = \frac{\sigma^2}{2\pi G r^2}, \tag{72}
\]
where $\sigma = 200 \text{ km s}^{-1}$. The corresponding infall time is given by

$$
\tau_{\text{fall}} \simeq 3 \text{ Gyr} \left( \frac{r}{1 \text{ kpc}} \right)^2 \left( \frac{M_{\text{BH}}}{10^6 M_\odot} \right)^{-1} \left( \frac{\ln \Lambda}{10} \right)^{-1}.
$$

These estimates show that PBHs as massive as $10^6 M_\odot$ within $\sim 1 \text{kpc}$ from the Galactic center fall into the center within the age of the Universe.

Let us now define the critical radius $r_{\text{df}}$ below which the PBHs spiral in to the Galactic center within the age of the Galaxy. Since the efficiency of the dynamical friction depends on the PBH mass, this radius is a function of $M_{\text{PBH}}$. Then, there would be a mass concentration in the galactic center by the present time by this amount,

$$
M_{\text{con}} = f_{\text{PBH}} \int_0^{r_{\text{df}}} 4\pi r^2 \rho_{\text{DM}}(r).
$$

In [158], the bound on $f_{\text{PBH}}$ was obtained by requiring that the mass concentration $M_{\text{con}}$ should not exceed the observational limit $\sim 3 \times 10^6 M_\odot$ on the central dark matter. According to their analysis, the result strongly constrains the PBH abundance for $M_{\text{PBH}} > 4 \times 10^4 M_\odot$, which is shown in figure 10. It is important to keep in mind that there are caveats in this constraint. Infalls of multiple PBHs into the central region may result in the ejection of PBHs by the slingshot mechanism, which would be efficient when the mergers of the PBHs in the central region by the gravitational radiation occur on time scales longer than that over which another PBH falls to the center. If ejections of PBHs happen, the mass growth of the central region is impeded. It has been demonstrated by $N$-body simulations that this mechanism is actually effective for BHs with $\lesssim 10^{6.5} M_\odot$ [164] when all the dark matter is assumed to be such BHs. Another potentially important effect is the rocket effect, which arises by the anisotropic emission of the gravitational radiation at the merger stage of the PBH binaries. According to [164], a recoil velocity $\gtrsim 1500 \text{ km s}^{-1}$ is needed to prevent runaway for $M_{\text{PBH}} = 10^7 M_\odot$. The newly formed BH will escape from the central region if the kicked velocity of the BH is significant.

### 3.2.7. Disk heating.

PBHs moving randomly in the Galactic halo have a chance to pass through the Galactic disk, and the stars in the disk are pulled by the gravitational attraction, acquiring velocity, every time a PBH passes nearby. Since the direction of the velocity that a star gains for an individual PBH passage is random, the time evolution of the velocity of the disk stars is described by a random walk. Thus, the variance of the star’s velocity increases in proportion to time. In other words, the disk stars become hotter and hotter over time. By requiring that the velocity increased by the PBHs does not exceed the observed velocity, we can constrain PBH abundance for some PBH mass range.

Under the influence of encounters with rapidly moving PBHs, the increase of variance of a star’s velocity during the time $t$ is given by [165]

$$
\sigma^2 \simeq \frac{8\pi G^2 f_{\text{PBH}} \rho_{\text{DM}} M_{\text{PBH}} \ln \Lambda}{V} \left( \frac{50 \text{ km s}^{-1}}{V} \right)^2 \left( \frac{M_{\text{PBH}}}{10^6 M_\odot} \right) \left( \frac{V}{200 \text{ km s}^{-1}} \right)^{-1} \left( \frac{t}{10 \text{ Gyr}} \right),
$$

when $V^2$ is much larger than the total velocity dispersion. Proper velocity of old stars in the solar neighborhood is around $50 \text{ km s}^{-1}$ [166]. Thus, the above equation shows that supermassive PBHs greater than a million solar masses cannot be the dominant component of dark matter. Furthermore, for PBHs heavier than $10^6 M_\odot$, we can derive a meaningful (i.e. $f_{\text{PBH}} < 1$) constraint on the PBH abundance as
\[ f_{\text{PBH}} < \left( \frac{M_{\text{PBH}}}{10^6 \, M_\odot} \right)^{-1}. \]  

(76)

For more detailed discussion, see [158].

Formulation of how the velocity variance of the stars evolves under the influence of massive objects (massive gas clouds were the main focus) was developed by Lacey [167]. In [165], the idea was proposed that the dark halos are composed of super-massive black holes (\( \sim 10^6 \, M_\odot \)), and they could explain the origin of the observed features suggesting the disk heating. However, recent precise measurements of the stars in the solar neighborhood disfavor this possibility. Introducing a free parameter \( \beta \) to parametrize a relation between the ages of the stars and their velocity dispersions as

\[ \sigma \propto t^\beta, \]  

(77)

the measurements show \( \beta \simeq 0.33 \) with uncertainty of \( \Delta \beta = 0.02-0.05 \) [166, 168]. The measured value does not coincide with the prediction \( \beta = 0.5 \) in the super-massive black hole scenario, and supports the spiral arms and the giant molecular clouds as the cause of disk heating [166].

3.3. Accretion constraints

Accretion of gas onto PBHs, and its impact on the constraint of PBH abundance has also been a subject of research [169]. Actually, previous studies show that the accretion constraint has potentially significant power in constraining PBHs for some mass range. However, we have to keep in mind that the physics is much more involved in the case of the accretion constraint compared to the lensing and dynamical constraints discussed in the previous sections. So far, it is impossible to derive the PBH constraint from the accretion from first-principles calculation, and all the derived constraints that will be presented later make some assumptions, and in some cases rely on observationally established empirical rules. In this section, we will review the accretion constraint on the PBH abundance based on two different processes: accretion effects on the CMB that arise in the early Universe and the electromagnetic waves from the accreted matter from the present PBHs.

3.3.1. Accretion effects on CMB. Baryonic gas around the PBH is attracted to the PBH by its gravity. As the gas falls into the central region, the gas is compressed, increasing its density and temperature. During the infall, the gas can be fully ionized, either by the internal collisions of gas particles or by the outgoing radiation. Near the black hole horizon, the gas temperature is enormous, and intense radiation from the ionized gas emanates outward. This radiation ionizes or heats the gas filling the Universe, and modifies the spectrum of the CMB photons from the Planckian distribution, the decoupling time of the CMB photons, and the ionization history. The latter two result in changing the power spectrum of the CMB temperature and the polarization anisotropies. In this way, PBHs leave non-standard features in the CMB observables. Non-detection of such features is translated into an upper limit on PBH abundance. Although the flow of the above logic is conceptually understandable, it is extremely difficult to predict how much the CMB observables are modified by the PBH accretion from first principles and self-consistent calculations, because of the complex nature of accretion. So far, various approximations and assumptions have been adopted in obtaining the PBH constraint from the accretion.
In early study [169], the standard Bondi accretion formula was adopted for the mass accretion rate, the efficiency of converting the accreted mass into the luminosity of the outgoing radiation was assumed to be independent of the cosmic time, and the spectrum of the emergent radiation was assumed to be flat and constant in time. In [170], the analysis was refined by taking into account the accumulation of dark matter, relative velocity between PBHs and the baryonic gas, and the coupling between CMB radiation and baryonic gas. In [171], a different relative velocity between PBHs and baryonic gas was used. However, the radiation efficiency was assumed to be a fixed value in those studies. In [172], the analysis in [170] was reexamined without adopting any *a priori* fixed value of the radiative efficiency, and it was found that the derived efficiency is much smaller than that assumed in [170]. As a result, the strong constraint on the PBH abundance derived in [170] was significantly weakened in [172]. In what follows, we briefly explain the main point of the analysis in [172].

A basic picture of the accretion investigated in [172] is as follows (see also figure 11). We start from considering a stationary system where a PBH is steadily swallowing the surrounding baryonic gas on the cosmological background. In other words, we assume that the accretion time scale is shorter than the Hubble time and the accretion follows adiabatically the cosmic expansion. It was shown in [170] that this picture is valid for $M_{\text{PBH}} \lesssim 3 \times 10^4 \, M_\odot$, and we focus on this mass range. Far from the PBH, the baryonic gas is little attracted by the PBH, and its density becomes equal to the cosmological density. Down to a certain radius (below the Bondi radius), the cooling of the baryonic gas by the background CMB is efficient, and the temperature of the gas remains the same as that of the CMB. Below this radius, cooling by the CMB becomes negligible, and the gas temperature increases as $r^{-1}$ due to adiabatic compression as the gas falls. When the temperature exceeds $\sim 10^4$ K, the collisional ionization starts to become important. During this phase, the increase of the internal energy gained by the infall is consumed as the ionization energy and the gas temperature remains constant. After the gas is completely ionized, the temperature again increases as $r^{-1}$ until the electrons become relativistic. The rate of temperature increase changes to $r^{-2/3}$ after the electrons become relativistic. Finally, the gas is swallowed by the PBH at the event horizon. Near the event horizon, the gas temperature is enormous, and the bremsstrahlung radiation produces intense outgoing radiation which eventually heats or ionizes the background gas. It is possible that the gas ionization is caused not by the collisions of gas particles but by the radiation emanating from the vicinity of the BH horizon. In [172], two extreme cases are considered, in which the complete ionization is achieved solely by either the collisional ionization or the photoionization.

Basic equations describing the accretion in the outer region where the ionization fraction coincides with that on the cosmological background ($\bar{x}_e$) are given by

$$4\pi r^2 \rho |v| = \dot{M} = \text{const.}, \quad (78)$$

$$\frac{d}{dr} \left( r^2 \rho \frac{dr}{d\rho} \right) = -\frac{GM}{r^2} - \frac{1}{3} \frac{d}{d\rho} \left( \frac{4}{3} \frac{\bar{x}_e \sigma_T \rho_{\text{CMB}}}{m_p} v \right), \quad (79)$$

$$\frac{d}{dr} \left( r^{2/3} \frac{T}{\rho^{2/3}} \right) = \frac{8}{3} \frac{\bar{x}_e \sigma_T \rho_{\text{CMB}}}{m_e (1 + \bar{x}_e)} \left( \frac{T_{\text{CMB}}}{T} - 1 \right), \quad (80)$$

where $v < 0$ is the radial gas velocity. The last terms in the second and the third equations represent the drag force and the cooling by the CMB respectively. From these equations, the mass accretion rate $\dot{M}$ is determined. Then, the continuity equation gives the electron number density at the vicinity of the event horizon as
\[
\dot{n}_e = \frac{\dot{M}}{4\pi m_e v^2 |\vec{v}|} = \frac{\dot{M}}{4\pi m_e v^2} \left( \frac{r}{r_g} \right)^{3/2},
\]

where we have used the free fall velocity \( v \approx \sqrt{g r / r} \). The gas temperature at the vicinity of the event horizon can also be determined by adiabatically extending the profile \( T(r) \) derived by solving the above basic equations down to the horizon. From the knowledge of the electron number density and the temperature, the luminosity \( L \) of the bremsstrahlung radiation is determined as

\[
L = \int 4\pi r^2 j \, dr, \quad j = n_e^2 \alpha \sigma T J(T/m_e),
\]

where \( J \) is a certain function fixed by the elementary process of the bremsstrahlung radiation. Then, the energy injection rate per volume of the radiation produced by the PBHs is given by

\[
\dot{\rho}_{\text{inj}} = f_{PBH} \frac{\rho_{DM}}{M_{PBH}} \langle L \rangle,
\]

where \( \langle L \rangle \) is the average of the luminosity over the PBH peculiar velocity relative to the baryonic gas. In [172], it was assumed that the luminosity in the case where the PBH moves relative to the baryonic gas was obtained by simply adding the relative velocity to the sound speed at infinity. The injected radiation interacts with the background gas via Compton scattering. Time evolution of the energy actually deposited to the gas is given by [172].
which is approximate, and derived under the assumption that most of the energy is injected near 0.1–10 MeV, corresponding to the characteristic temperature of the gas near the BH horizon. It was assumed that the energy deposited in this manner is distributed to the increase of the gas temperature, ionization fraction, and the excitation of the hydrogen atoms (to the first excited states)

$$\Delta \dot{T}_{\text{gas}} = \frac{2}{3 n_{\text{tot}}} \frac{1 + 2 \hat{x}_e}{3} \dot{\rho}_{\text{dep}}, \quad \Delta \dot{x}_e = \frac{1 - \hat{x}_e}{3 E_{\text{InH}}} \dot{\rho}_{\text{dep}}, \quad \Delta \dot{x}_2 = \frac{1 - \hat{x}_e}{3 E_{\text{InH}}} \dot{\rho}_{\text{dep}},$$

where $E_I = 13.6$ eV, $E_2 = 10.2$ eV. These equations describe how the temperature and the ionization history are modified by the accreting PBHs. For instance, increase of the ionization fraction enhances the CMB optical depth, which results in the damping of small scale CMB fluctuations and the enhancement of polarization power on large angular scales compared to the standard case. Also, the shift of the redshift of the last scattering changes the phase of the acoustic oscillations in the CMB spectrum. In [172], the modified CMB fluctuations were computed using the recombination and the Boltzmann codes [174, 175], and these were compared with the CMB TT, TE and EE power spectra provided by the Planck collaboration [176]. The resultant constraint on the PBH abundance is shown in figure 10.

Accretion of gas onto the PBHs in the early Universe (even before the CMB decoupling) also induces CMB spectral distortions, since the photons produced by the accreting PBHs are not completely thermalized. Photons generated in the redshift $5 \times 10^9 < z < 2 \times 10^9$ achieve kinetic equilibrium, and yield a non-vanishing chemical potential ($\mu$-distortion). Here, the upper and the lower limit on the redshift correspond to the temperature above which the photon-number changing process occurs efficiently, making the distribution Planckian, and the temperature below which the kinetic equilibrium is no longer maintained, respectively. Photons generated in the redshift $200 < z < 5 \times 10^9$ yield the Compton-$\gamma$ distortion. Here, the lower limit corresponds to the temperature at which the baryonic gas decouples from the CMB. In contrast to the meaningful constraint on $f_{\text{PBH}}$ by the temperature and the polarization...
anisotropies of CMB, it was also found in [172] that the induced amplitudes of both the $\mu$ and $\gamma$-distortions are too small to be relevant to present and future observations.

So far, the discussion has been based on the assumption that the accretion is spherical. The accretion is approximately spherical when the angular momentum of the gas at the Bondi radius is smaller than that determined by the Keplerian orbit at the innermost stable circular orbit. If this is not satisfied, the gas motion becomes Keplerian before it is absorbed by the PBH, and an accretion disk forms around the PBH. While the radiation in the case of the spherical accretion originates from the bremsstrahlung radiation, the radiation is produced more efficiently by the viscous heating in the presence of the accretion disk [144]. As a result, the radiative efficiency is typically higher for the accretion disk than the spherical accretion. In [177], it was claimed that the spherical assumption may be violated after recombination. In this reference, based on the analytical estimation, it was suggested that the accretion disk forms around the PBH. Performing the equivalent analysis for the CMB temperature and the polarization anisotropies described above for the case of the accretion disk, it was found that the corresponding constraint on $f_{\text{PBH}}$ becomes much stronger than that obtained in [172]. The constraints derived are shown in figure 10.

3.3.2. X-rays and radio from present-day PBHs. In the previous subsubsection, we have considered the gas accretion onto PBHs in the early Universe and its effects on the CMB. PBHs in the present Universe also accrete the surrounding gas if they are in a dense environment, and a constraint on PBH abundance can be obtained by comparison between observational data and the theoretical predictions of the electromagnetic waves from accreting PBHs.

In [178], radiation spectra originating from the gas accretion onto the PBHs of $0.5 M_\odot$ in the solar neighborhood was computed based on the assumption that the accretion is modeled by advection-dominated accretion flow (ADAF). This research was motivated by the microlensing observations that hint at PBHs of mass $\sim 0.5 M_\odot$ as dominant component of dark matter. It was found that the radiation spectra from the PBHs could be at the detectable level at the IR-optical band and is negligibly small at the x-ray band. This implies that a meaningful upper limit on PBH abundance could be obtained from the IR-optical band only. However, since the main interest in [178] was the detectability of the PBH signal due to the accretion, no explicit number for the upper limit on $f_{\text{PBH}}$ was given. Since there are contaminations from other astrophysical sources, for instance young stellar objects at the IR-optical band, it is important to keep in mind that it is crucial to do the consistency check over the multi-wavelength bands when one claims the positive detection of the PBHs from the accretion.

In [179], an upper limit on the PBH abundance for $10 < M_{\text{PBH}}/M_\odot < 100$ was obtained using observational data from the Very Large Array (VLA) radio and the Chandra x-ray. The flow of the argument is as follows. First, the PBH was assumed to accrete the surrounding gas with the mass accretion rate given by

$$\dot{M} = 4\pi \lambda (GM_{\text{PBH}})^2 \frac{\rho_{\text{gas}}}{(\frac{v_{\text{PBH}}^2}{c_s^2} + \frac{c_s^2}{c_s^2})^{3/2}},$$

where the value of $\lambda$, which measures the amount of mass accretion rate normalized by the Bondi accretion rate, was chosen to be $\lambda = 0.02$, which is consistent with x-ray observations of dwarf nova and the nuclear accretion in galaxies (see also [180]). The radiative efficiency, that converts the mass accretion rate to the bolometric luminosity from the accreting gas, was then assumed to be given by
\[ L_B = \eta \dot{M}, \quad \eta = 0.1 \left( \frac{\dot{M}}{0.01 \dot{M}_{\text{Edd}}} \right), \]  

(87)

where \( \dot{M}_{\text{Edd}} \) is the mass accretion rate corresponding to the Eddington limit. That the radiative efficiency is proportional to the mass accretion rate for \( \dot{M} < 0.01 \dot{M}_{\text{Edd}} \), which is the region of our current interest, is suggested by the observational data [181]. In order to obtain the x-ray luminosity out of the bolometric luminosity, the relation

\[ L_X = 0.3 L_B, \]  

(88)

was adopted. It is known that many accreting BHs accompany jets from which GHz radio waves are emitted. Thus, by making use of the so-called fundamental plane [182], which is an empirical relation among the mass of BH accompanying jets, x-ray luminosity \( L_X \), and radio luminosity \( L_R \), the radio luminosity was computed under the assumption that PBHs accreting gas produce jets. With this formalism to compute \( L_X \) and \( L_R \) from the given PBH mass (delta function mass spectrum), the Monte Carlo simulation has been performed by distributing the positions and velocities of PBHs according to the Navarro–Frenk–White (NFW) profile [183] and the Maxwell–Boltzmann distribution, respectively, and the expected detectable number of x-ray and radio sources from the PBHs have been evaluated for the Chandra [184] and the VLA observations [185]. It was found that both x-ray and radio observations result in a meaningful constraint \( f_{\text{PBH}} < 1 \). These constraints are shown in figure 10.

The PBH constraint due to present-day accretion has also been obtained in [186] by using the luminosity function of x-ray binaries derived in [187], in which 29 nearby star-forming galaxies in Chandra, Spizer, GALEX and 2MASS archives were analyzed. In [186], assuming the value \( \lambda = 1 \) in equation (86) and \( L_X = L_B \) for the x-ray luminosity, and introducing the dark matter disk in galaxies which has been suggested in simulations [188, 189] to the PBH distribution, luminosity function of the x-ray sources powered by PBHs was computed and compared with the observationally determined one given in [187]. It was found that the PBHs in the mass range of a few \( \sim 10^7 M_\odot \) are tightly constrained. This result is shown in figure 10.

As just described, there are theoretical uncertainties in modeling the accretion and the emitted radiation. However, these studies demonstrate that the accretion to the present-day PBHs has potential to provide a meaningful constraint on PBH abundance. Finally, before closing this section, we mention that searching for electromagnetic signals from isolated astrophysical black holes (i.e. non-PBH) has also been a target of active research (e.g. [178, 180, 190, 191]).

### 3.4. Large scale structure constraint

PBHs randomly distributed in space in the early Universe generate primordial density perturbations by their Poisson fluctuations on scales larger than the PBH mean distance, as first noted in [192]. The mean comoving distance of PBHs is given by

\[ \ell_{\text{mean}} = \left( \frac{M_{\text{PBH}}}{f_{\text{PBH/DM}}} \right)^{1/3} = 0.3 \text{ Mpc} \left( \frac{M_{\text{PBH}}}{10^6 M_\odot} \right)^{1/3} \left( \frac{f_{\text{PBH}}}{10^{-3}} \right)^{-1/3}. \]  

(89)

Denoting by \( n_{\text{PBH}} \) the comoving PBH number density, the number of PBHs in the comoving volume \( \lambda^3 \) fluctuates typically as \( \sim N_\Lambda^{1/2} \), where \( N_\Lambda = n_{\text{PBH}} \lambda^3 \) is the average PBH number in that volume. The fluctuations are isocurvature perturbations, since they are present on the hypersurface where the radiation energy density looks uniform. Thus, in the presence of the
PBHs, on top of the standard nearly scale-invariant adiabatic perturbations, the dark matter density contrast has the isocurvature perturbations whose variance is given by

$$\langle \delta_{\text{DM}}^2 \rangle = \left( \frac{\delta_{\text{PBH}}}{\rho_{\text{DM}}} \right)^2 = f_{\text{PBH}}^2 M_{\text{PBH}} \rho_{\text{DM}}^{-1} \lambda^{-3},$$

(90)
on the comoving scale $\lambda$. The scaling of the variance as $\propto \lambda^{-3}$ shows that the produced dark matter perturbations are more enhanced on smaller scales. Equivalently, in terms of the dimensionless power spectrum, the PBHs yield the contribution

$$P_{\delta_{\text{DM}}}(k) = \frac{k^3}{2\pi^2} f_{\text{PBH}}^2 M_{\text{PBH}} \rho_{\text{DM}},$$

(91)
for which the spectral index is 3, i.e. an extremely blue spectrum.

Based on this observation that the Poisson fluctuations in the PBH distribution enhance the dark matter perturbations on small scales, their impact on the Ly$\alpha$ forest observations was investigated in [193]. The basic idea is as follows. Spectra of distant quasars and galaxies show many absorption lines known as the Ly$\alpha$ forest. This arises due to the intervening neutral hydrogen in the intergalactic medium between the quasars and the Earth (typically $z = 2 \sim 5$), which absorbs photons from quasars by the Ly$\alpha$ transition ($n = 1$ to $n = 2$, where $n$ is the principal quantum number of the hydrogen atom). Although this epoch is after reionization, there is still a tiny fraction of neutral hydrogen in the intergalactic medium because of the balance between the photoionization by the surrounding UV radiation and the recombination, and this tiny amount of neutral hydrogen is sufficient to produce the Ly$\alpha$ forests. The wavenumber corresponding to this transition is $\lambda_{\text{Ly}\alpha} = 1216$ Å. When the photons are partially absorbed by the neutral hydrogen at redshift $z_c$, then the absorption line appears at $\lambda = \lambda_{\text{Ly}\alpha}(1 + z_c)$ in the observer frame. The optical depth of the Ly$\alpha$ transition to the direction of a quasar is given by (e.g. [194])

$$\tau_{\text{Ly}\alpha}(\lambda) \approx 1.3 \Delta_b(z_c) \left( \frac{x_{\text{HI}}(z_c)}{10^{-5}} \right) \left( \frac{1 + z_c}{4} \right)^{3/2},$$

(92)
where $\Delta_b = n_b/n_\text{b}$ is the baryon density normalized by the average value, and $x_{\text{HI}}$ is the fraction of the neutral hydrogen. This equation shows that the Ly$\alpha$ absorption is more efficient in the site where baryonic matter is denser. Because of this, many absorption lines in the spectra, viz. fine fluctuations of $\tau_{\text{Ly}\alpha}$ for the relevant range of $\lambda$, reflect the inhomogeneous distribution of baryonic matter. Given that baryon perturbations are affected gravitationally by dark matter perturbations, the statistical properties of the optical depth encode those of the dark matter perturbations. Conversely, observational analysis of the optical depth makes it possible to probe the matter perturbations on small scales down to $\sim$Mpc, where the minimum scale is determined by the thermal broading of the spectra of the Ly$\alpha$ forest. These scales are mildly nonlinear in the Ly$\alpha$ epoch.

In [193], superposing the dark matter isocurvature perturbation given by equation (91) (assuming $f_{\text{PBH}} = 1$) on the standard adiabatic perturbations, the resultant Ly$\alpha$ clouds were computed for various PBH masses by the use of hydro-dynamical cosmological simulations. The power spectrum $P_F(k)$, which is a two-point function of the Fourier transform of the transmitted flux $F(\lambda)$ defined by $F = e^{-\tau_{\text{Ly}\alpha}}$ was then evaluated and compared with the observational data presented in [195]. It was found that PBHs heavier than $\sim 10^4 M_{\odot}$ comprising all the dark matter produce too much power of the transmitted flux, and are inconsistent with the Ly$\alpha$ observations. Although the explicit upper limit on $f_{\text{PBH}}$ is not provided in [193], a simple analytical expression for the upper limit consistent with the results of [193]...
was given in [7]. The main point of [7] is that the matter perturbations sourced by equation (91) should not exceed $O(1)$, the typical amplitude of the matter perturbation, at the $\text{Ly} \alpha$ epoch, which yields

$$f_{\text{PBH}} < \left( \frac{M_{\text{PBH}}}{10^4 M_\odot} \right)^{-1}. \quad (93)$$

This constraint becomes meaningless when it implies that the PBH mean distance (89) is larger than the size of the $\text{Ly} \alpha$ clouds $\ell_{\text{Ly} \alpha}$. Thus, equation (93) makes sense when

$$f_{\text{PBH}} > 3 \times 10^{-4} \left( \frac{M_{\text{PBH}}}{10^4 M_\odot} \right) \left( \frac{\ell_{\text{Ly} \alpha}}{1 \text{ Mpc}} \right)^{-3}. \quad (94)$$

is satisfied. Equivalently, equation (93) can also be obtained by requiring that the power spectrum of equation (91) at the $\text{Ly} \alpha$ scale ($k \sim 1 \text{ Mpc}^{-1}$) is less than the value of the standard one $\sim 10^{-10}$ based on the consistency between the prediction of the standard adiabatic perturbations and the observed $\text{Ly} \alpha$ forests.

In [196], inspired by the first detection of GWs by LIGO, interesting suggestion was made that PBHs with mass $O(10) M_\odot$ constituting all dark matter could solve the potential tension of the observed near-IR cosmic infrared background (CIB) anisotropies. CIB is the accumulated emission of luminous objects throughout the history of the Universe (see e.g. [197]) and was detected for the first time by the COBE DIRBE instrument [198]. Intriguingly, measured anisotropies of CIB obtained by removing the foreground stars and galaxies are larger than expected, and cannot be explained only by the contribution from remaining faint galaxies [199]. This potential tension basically arises due to rareness of halos where luminous stars form in high redshift. Denoting by $f_*$ the efficiency with which baryons inside each halo are converted to luminous sources, the required value of $f_*$ to explain the observed amplitude of the CIB fluctuations is estimated as [196]

$$f_* \simeq 0.1 \left( \frac{f_{\text{Halo}}}{0.01} \right)^{-1} \left( \frac{z_{\text{eff}}}{10} \right), \quad (95)$$

where we have assumed that a mass-fraction $f_{\text{Halo}}$ of dark matter is contained in the halos that have luminous sources, and $z_{\text{eff}}$ is the effective average redshift. How to achieve such high value $f_* \simeq 0.1$ is a challenging task if $f_{\text{Halo}} \simeq 0.01$. Based on the idea explained above that the additional dark matter fluctuations sourced by the PBHs increase the minihalos in high redshift, $f_{\text{Halo}}$ was computed in the presence of PBHs, and was found to become much larger than that in the conventional $\Lambda$CDM case. Hence, in this scenario, the required value of $f_*$ is reduced to very modest one, which is the main point in [196]—that PBHs can explain the CIB fluctuations.

3.5. Indirect constraints

So far, we have considered the cosmological and astrophysical effects caused by PBHs, and discussed how PBH abundance can be constrained by them. These constraints are direct, in the sense that all the effects are directly triggered by the PBHs. A nice point of the direct constraints is that they do not resort to the formation mechanism of PBHs. Any model predicting the PBH formation must satisfy the direct constraints in order for it to be considered as a consistent model.

In addition to the direct constraints, there are indirect ones, which are the subjects of this section. As we have discussed in great detail in section 2, the most popular mechanism for
PBH formation is the direct gravitational collapse of primordial density perturbations, which can be naturally embedded in the framework of inflation in which the quantum fluctuations of the scalar fields are the ultimate origin of the primordial density perturbations. In this scenario, PBHs form only at high-σ (typically \( \sim 10\sigma \)) regions that are extremely rare. Although the other regions are not inhomogeneous enough to produce PBHs, they are still inhomogeneous enough to induce effects that are already excluded by or marginally consistent with observations, depending on the statistical properties of the primordial density perturbations. Such effects, which we are going to discuss in detail, are not produced by PBHs but by the density perturbations that seed the PBHs, and hence the resultant constraints on the PBHs are indirect. We have to keep in mind the underlying assumptions when trying to constrain a particular inflation model predicting PBHs by using indirect constraints.

3.5.1 Stochastic gravitational waves from the primordial density perturbations. The first effects that the primordial density perturbations seeding the PBHs causes are stochastic gravitational waves produced by the mode–mode coupling of the density perturbations. Density perturbations, which are classified as scalar-type, evolve independently of gravitational waves (GWs), which are classified as tensor-type, at the linear order in the perturbation (e.g. [200]). This independence no longer holds beyond the linear order, and the GWs are produced by the density perturbations at the second order by their mode–mode couplings\(^{21}\).

Assuming that primordial density perturbations existed on super-Hubble scales, the production of the GWs by the second-order effect happens predominantly at the time when the density perturbations re-enter the Hubble horizon [203, 204]. In other words, GWs are mainly generated at the same epoch as the PBH formation. Once produced, those GWs freely propagate in the subsequent epochs, and are still permeating the present Universe. Combining that the typical frequency of such GWs at the formation time is comparable to the Hubble horizon and that the horizon radius at that time is comparable to the size of the PBHs allows us to relate the PBH mass to the frequency of the GWs at present time as [205]

\[
 f_{\text{GW}} \simeq 1 \times 10^{-9} \text{ Hz} \left( \frac{M_{\text{PBH}}}{30 M_\odot} \right)^{-1/2}.
\]  

(96)

Thus, primordial density perturbations producing stellar-mass PBHs generate ultra-low frequency GWs in the nHz band. Quite interestingly, such low-frequency GWs are severely constrained by pulsar timing experiments.

Pulsars, which are rapidly rotating neutron stars emitting beams of radio waves, are useful to probe low-frequency GWs (e.g. [206]). In particular, millisecond pulsars, which rotate with a period of \( \mathcal{O}(\text{ms}) \), are observed to be significantly more stable than normal pulsars. The arrival times of each pulse from such millisecond pulsars have been measured accurately, and compared with the predictions of the pulsar timing model. Just as the expansion of space, i.e. time-varying space–space components of the metric, causes the cosmological redshift, the pulse frequency \( \nu = 1/T \), where \( T \) is the time interval between the arrival times of the successive pulses, is modulated if GWs are present. The modulation of the pulse frequency is given by [207]

\[
 \frac{\delta \nu}{\nu} = -H^j \left[ h_{ij}(t, \vec{x}_p) - h_{ij}(t - D, \vec{x}_p) \right],
\]  

(97)

\(^{21}\) Conversely, stochastic GWs also produce density perturbations due to mode–mode couplings. From the requirement that the resultant density perturbations do not lead to the overproduction of PBHs, an upper limit on the amplitude of stochastic GWs can be obtained [201, 202].
where $H^I_i$ is a geometrical factor depending on the propagation direction of the GWs relative to the direction of the pulsar, $h^I_i$ represent GWs, $D$ is the distance to the pulsar, and $\vec{x}_e$ and $\vec{x}_p$ are the positions of the Earth and the pulsar respectively. This induces the timing residuals of the pulse arrival time at time $t$ (with reference to time 0) as

$$R(t) = - \int_0^t \frac{\delta \nu}{\nu} dt.$$  \hspace{1cm} (98)

A few hundred millisecond pulsars have been observed. It is reasonable to think that the timing residuals caused by the GWs (98) for the pulsars are strongly correlated, since the frequency modulations due to the first term in equation (97) become almost identical. It was shown by Hellings and Downs [208] that the correlation function (normalized by the GW amplitude) of the timing residuals by the stochastic GWs between the pulsars separated by the sky angle $\theta$ is given by

$$c(\theta) = x \ln x - \frac{x}{6} + \frac{1}{3}, \quad x = \frac{1 - \cos \theta}{2}.$$  \hspace{1cm} (99)

By searching for this type of correlation in the measured data for many pulsar pairs, which is known as the pulsar timing array (PTA) experiment, GWs can be detected if they exist by amount more than the sensitivity the experiments can reach. Otherwise, an upper limit on the abundance of GWs is obtained.

Currently, there are three major PTA projects, the Parkes pulsar timing array (PPTA) [209], the North American nanohertz observatory for gravitational waves (NANOGrav) [210], and the European pulsar timing array (EPTA) [211], which aim to detect ultra-low frequency GWs in the nHz range. So far, no detection of stochastic GWs has been reported, and an upper limit on the GWs has thus been derived.

Now, let us return to the theoretical computations of the secondary GWs produced by the primordial density perturbations. First, let us write the metric representing the scalar-type perturbations and the sourced tensor-type perturbations on the flat FLRW spacetime [203, 204]:

$$ds^2 = a^2(\eta) \left[ - (1 + 2\Phi) d\eta^2 + (1 - 2\Psi) \left( \delta_{ij} + \frac{1}{2} h_{ij} \right) dx^i dx^j \right].$$  \hspace{1cm} (100)

where $\Phi$ and $\Psi$ are scalar-type perturbations (gravitational potential and curvature perturbation, respectively) and $h_{ij}$ are tensor-type perturbations satisfying the transverse and traceless conditions $\partial_i h^I_j = h^I_i = 0$, where $h^I_j = \delta^I_k h_{kj}$. In the absence of anisotropic stress, which is a good approximation [204], we have $\Phi = \Psi$. Here, we do not include $h_{ij}$ generated during or after inflation by some other mechanism, since they do not affect the GWs we are interested in this section. The evolution of $\Psi$ at the linear order in the radiation dominated era is given by

$$\Psi'' + \frac{4}{\eta} \Psi' - \frac{1}{3} \triangle \Psi = 0.$$  \hspace{1cm} (101)

The solution of this equation in the Fourier space with the requirement of vanishing decaying mode is given by

$$\Psi_k(\eta) = D_k(\eta) \Psi_k(0), \quad D_k(\eta) = \frac{3}{(k\eta)^2} \left[ \frac{\sqrt{3}}{k\eta} \sin \left( \frac{k\eta}{\sqrt{3}} \right) - \cos \left( \frac{k\eta}{\sqrt{3}} \right) \right],$$  \hspace{1cm} (102)

where the initial condition $\Psi_k(0)$ is fixed either by the assumed inflation model or by the assumed power spectrum of $\Psi$. The evolution of $h_{ij}$ is also obtained by expanding the Einstein equations. At the second-order in the scalar-type perturbations, we have [203, 204]
\[ h''_{ij} + 2\mathcal{H} h'_{ij} - \Delta h_{ij} = -4\hat{T}^{ij}_{rs}S_{rs}, \]

(103)

where \( \mathcal{H} = a'/a, \hat{T}^{ij}_{rs} \) is a projection operator that produces the transverse and traceless quantity, and the source term consisting of the products of scalar-type perturbations is given by

\[ S_{ij} = 2\Psi \dot{\Psi}_{,ij} - (\Psi_{,j} + \mathcal{H}^{-1} \dot{\Psi}_{,j}')(\Psi_{,i} + \mathcal{H}^{-1} \dot{\Psi}_{,i}'). \]

(104)

GWs can be Fourier-transformed as

\[ h_{ij}(\eta, \vec{x}) = \int \frac{d^3 k}{(2\pi)^{3/2}} e^{i\vec{k} \cdot \vec{x}} \sum_{l = +, \times} e_l^{ij}(\vec{k}) h_l(\eta, \vec{k}), \]

(105)

where \( e_l^{ij} \) are polarization tensors normalized by \( e^{+ ij} e^{+ ij} = e^{\times ij} e^{\times ij} = 1 \). Knowing the time evolution of \( S_{ij} \) by equation (102), the time evolution of \( h_l \) is determined as

\[ h_l(\eta, \vec{k}) = \frac{1}{a(\eta)} \int_0^\eta G_k(\eta, \eta') a(\eta') S_l(\eta', \vec{k}) d\eta', \]

(106)

where \( G_k(\eta, \eta') = \sin(k(\eta - \eta'))/k \) is the Green’s function of equation (103) and \( S_l(\eta, \vec{k}) \) is a Fourier transform of the right-hand side of equation (103) [104].

\[ S_l(\eta, \vec{k}) = \int \frac{d^3 q}{(2\pi)^{3/2}} 4\epsilon_l^{ij}(\vec{k}) q_i(q \cdot \vec{k} - \vec{q}, \eta) \Psi_{\vec{q}}(0) \Psi_{\vec{q} - \vec{k}}(0), \]

(107)

where

\[ f(\vec{k}, \vec{q}, \eta) = 2D_k(\eta) D_q(\eta) + \left( \frac{D_k'(\eta)}{\mathcal{H}} + \frac{D_q'(\eta)}{\mathcal{H}} \right) \left( D_k(\eta) + \frac{D_q(\eta)}{\mathcal{H}} \right). \]

(108)

The perturbations \( \Psi \) being stochastic, the GWs induced are also stochastic. A useful quantity parameterizing the amount of GWs is \( \Omega_{GW} \), the energy density of GWs per logarithmic frequency normalized by the critical density. Using the formula for the GW energy density [212]

\[ \rho_{GW} = \frac{1}{128\pi G} \langle h_0 \dot{h}_0 \rangle, \]

(109)

and the definition of the GW power spectrum

\[ \langle h_l(\vec{k}) h_l(\vec{q}) \rangle = \frac{2\pi^2}{k^3} P_h(k) \delta(\vec{k} + \vec{q}) \delta_{lj}, \]

(110)

we have

\[ \frac{\rho_{GW}}{\rho_{\Omega}} = \int d\ln f \Omega_{GW}(f), \quad \Omega_{GW}(f) = \frac{1}{24\mathcal{H}_0^2} k^2 P_h(k), \]

(111)

with \( k = 2\pi f \). Then, assuming the Gaussianity of the scalar-type perturbations, for which we can decompose the power spectrum of \( S_l \) into the product of the power spectrum of \( \Psi \), \( \Omega_{GW} \) can be written as
\[
\Omega_{GW}(f) = \frac{k^5}{3\pi H_0^2} \int_0^{\eta_0} d\eta' \int_0^{\eta_0'} d\eta'' G_k(\eta_0, \eta') G_k(\eta_0, \eta'') \frac{a(\eta') a(\eta'')}{a^2(\eta)} \times \int d^3q \frac{(e^{i(\vec{k} - \vec{q})})^2}{q^3 |\vec{k} - \vec{q}|^2} f(\vec{q}, \vec{q} - \eta', \eta') f(\vec{q}, \vec{q} - \eta, \eta'') P_\Psi(q) P_\Psi(|\vec{k} - \vec{q}|).
\]

(112)

Except for \( P_\Psi \), all the functions such as \( G_k \), \( a(\eta) \) and \( f \) are known. By using this equation, we can compute \( \Omega_{GW} \) from the power spectrum of the scalar-type perturbations \( P_\Psi \).

Then, assuming the delta-function type power spectrum for \( P_\Psi \) whose amplitude is large enough to produce a sizable amount of PBHs in the present Universe, it was claimed that the resultant \( \Omega_{GW} \) conflicts with the upper limit set by the pulsar timing experiments [205]. The analysis was then generalized in [213] to the case where \( P_\Psi \) has a silk hat type spectrum with a finite width. It was found that the predicted \( \Omega_{GW} \) becomes smaller for larger width (see also [214]). In [104, 215], computations of \( \Omega_{GW} \) for some concrete inflation models predicting the PBHs were performed, and comparisons with the data of the three PTA experiments (EPTA [216], PPTA [217], and NANOGrav [218]) were made (see also [110]). Although not all the regions in the parameter space of inflation models are excluded, it was shown that pulsar timing places stringent constraints on inflation models.

So far, the primordial density perturbations sourcing the stochastic GWs have been assumed to be Gaussian. In [219], it was suggested that the constraints mentioned above can weaken for non-Gaussian density perturbations, which may be typical when the amplitude of the density perturbation is significantly large. A basic observation behind this conclusion is that the variance of the density perturbation, which determines the magnitude of \( \Omega_{GW} \), can be smaller in the non-Gaussian perturbation than in the Gaussian case for a fixed PBH abundance. In [219], two different types of non-Gaussian density perturbation were considered. The first one is specified by the probability density function of the curvature perturbation smoothed over the Hubble horizon at the time of the PBH formation, which is given by

\[
P(\zeta) = \frac{1}{2^{3/2} \tilde{\sigma} \Gamma(1 + 1/p)} \exp \left[ -\left( \frac{\zeta}{\sqrt{2} \tilde{\sigma}} \right)^p \right],
\]

(113)

where \( p \) and \( \tilde{\sigma} \) are free parameters. When \( p = 2 \), the perturbations are Gaussian. As explained in section 2.3, the PBH fraction at the formation time is given by the probability that the perturbation exceeds the formation threshold \( \zeta_{th} \),

\[
\beta = \int_{\zeta_{th}}^{\infty} P(\zeta) d\zeta.
\]

(114)

For fixed \( \beta \), the variance of \( \zeta \) becomes smaller than the Gaussian case for \( p < 2 \) [220]. The second example is the local-type non-Gaussian perturbation, for which the smoothed curvature perturbation in the position space is written as

\[
\zeta = \zeta_G + \frac{3}{5} f_{NL} \zeta_G^2,
\]

(115)

where \( \zeta_G \) is Gaussian, and \( f_{NL} \) is a free parameter that parametrizes the significance of the non-Gaussian contribution. For positive \( f_{NL} \), the variance of \( \zeta \) is suppressed than the Gaussian case [221]. According to the analysis in [219], the PTA constraint can be safely evaded if \( p \lesssim O(1) \) or \( f_{NL} \gtrsim O(1) \) (the precise value depends on the theoretical uncertainties about
PBH formation). A Similar conclusion has been obtained in [222], where the curvature perturbation obeying the $\chi^2$ statistics, motivated by the rolling axion scenario [223], was considered.

3.5.2. CMB spectral distortions from the primordial density perturbations. The second indirect constraint can be obtained from the generation of the CMB spectral distortion of the primordial density perturbations [224, 225]. The basic argument is as follows.

The perturbations of photons and baryonic gas that re-entered the Hubble horizon prior to the CMB decoupling ($z \simeq 1100$) undergo acoustic oscillations due to the tight coupling between photons and baryons. These oscillations are eventually erased by the photon diffusion, i.e. imperfect coupling between photons and free electrons, known as Silk damping. Using the photon mean free path $\ell_{mp} = 1/(\sigma_T n_e)$, where $n_e$ is the number density of free electrons, and that the diffusion is described by random walks, the comoving wavenumber of a perturbation is related to the damping time of the perturbation by [226]

$$k = \frac{1}{1 + z} \sqrt{n_e \sigma_T H(z)} \simeq 4 \times 10^{-6} (1 + z)^{3/2} \text{Mpc}^{-1}. \quad (116)$$

This shows that the Silk damping occurs earlier for smaller scale perturbations. Through the damping, the oscillation energy of the perturbations is transferred to the background homogeneous plasma gas. As briefly discussed in section 3.3.1, if the Silk damping occurs before $z \approx 2 \times 10^6$, the photon-number changing interactions occur frequently, and the injected energy is completely consumed for the thermalization. As a result, the net effect is just a slight increase of the entropy per unit comoving volume. If the perturbations undergo Silk damping after that redshift but before $z \approx 5 \times 10^4$, only the kinetic equilibrium is achieved. As a result, the photons acquire chemical potential, and the distribution of photons becomes the Bose–Einstein distribution ($\mu$-type). Using the above equation (116), the interval of wavenumber of perturbations that yield the $\mu$-type distortion is $50 \lesssim k/\text{Mpc}^{-1} \lesssim 10^4$. If the perturbations dissipate after $z \approx 5 \times 10^4$, even the kinetic equilibrium is no longer reached, and the distribution of photons is characterized by the Compton-$y$ parameter ($y$-type). Thus, the perturbations of smaller comoving wavenumber than $50 \text{ Mpc}^{-1}$ generate $y$-type distortion in the CMB spectrum.

Now, let us consider for simplicity the delta-function power spectrum of the density perturbation at $k = k_*$, with total power $A$. To produce PBHs, $A = O(0.01 - 0.1)$ is required. Then, the $\mu$-type distortion produced is given by [227, 228]

$$\mu \simeq 2 \times A \left[ \exp \left( -\frac{k_*}{5400 \text{ Mpc}^{-1}} \right) - \exp \left( -\left[ \frac{k_*}{31.6 \text{ Mpc}^{-1}} \right]^2 \right) \right]. \quad (117)$$

So far, spectral distortion of the CMB has not been detected, and the strongest upper limit is placed by the COBE/FIRAS experiment as $\mu \lesssim 9 \times 10^{-5}$ [229]. Using this limit and the fiducial value $A = 0.02$, the level of $\mu$-distortion produced (117) is inconsistent with the COBE/FIRAS bound for $2 \lesssim k/\text{Mpc}^{-1} \lesssim 3 \times 10^4$. This interval is only logarithmically sensitive to the choice of $A$. In terms of the PBH mass, non-detection of the $\mu$-type distortion excludes PBHs in the mass range

$$2 \times 10^4 M_\odot \lesssim M_{\text{PBH}} \lesssim 2 \times 10^{13} M_\odot \quad (118)$$

if PBHs are produced from the direct collapse of the nearly Gaussian primordial density perturbations [228] (see also [230]).

In [220], it was pointed out that the above conclusion that the PBHs in the mass range (118) are excluded can in principle be circumvented if the primordial density perturbations...
are strongly non-Gaussian such that amplitude of the density perturbations in all regions other than the sites of PBH formation is too tiny to induce sizable $\mu$-distortion. Such a situation can be realized if only patches that later convert into PBHs experience different expansion history during inflation (see figure 3 of [220]).

3.5.3. Big-bang nucleosynthesis. From the last discussion in section 3.5.2, it may appear that it is impossible to constrain the amplitude of the primordial perturbation for $k \gtrsim 10^4 \text{Mpc}^{-1}$, translating to $M_{\text{PBH}} \lesssim 2 \times 10^4 \text{M}_\odot$, since the Silk damping of such small scale perturbations does not leave any spectral feature in the CMB. However, big-bang nucleosynthesis (BBN) has a potential to exclude smaller PBHs than the above mass.

Primordial density perturbations in the range $10^4 \lesssim k/\text{Mpc}^{-1} \lesssim 10^9$ dissipate by the Silk damping in the redshift range $2 \times 10^6 < z < 10^7$, which is after the BBN era but much before the decoupling time of the CMB (perturbations with $k \gtrsim 10^5 \text{Mpc}^{-1}$ dissipate prior to the neutrino decoupling due to neutrino diffusion). Injected energy by the Silk damping of these perturbations results in the increase of the temperature of the plasma gas. Because of the conservation of the baryon number, only the photon number density is increased by this process. In other words, the baryon-to-photon ratio $\eta$ defined by the ratio between the two number densities as $\eta \equiv n_b/n_\gamma$ decreases through Silk damping. This means that $\eta$ during the BBN era is bigger than that in the CMB era, i.e. $\eta_{\text{BBN}} > \eta_{\text{CMB}}$. The abundance of light elements produced by BBN is controlled by $\eta_{\text{BBN}}$, and the measurement of the light elements in the present Universe can determine $\eta_{\text{BBN}}$. Acoustic peaks in the CMB temperature power spectrum are controlled by $\eta_{\text{CMB}}$, and the measurement of the CMB anisotropies can determine $\eta_{\text{CMB}}$.

The difference between $\eta_{\text{BBN}}$ and $\eta_{\text{CMB}}$ depends on how much energy was injected by the Silk damping, i.e. the amplitude of the primordial density perturbations. Therefore, the determination of $\eta_{\text{CMB}}$ and $\eta_{\text{BBN}}$ by independent observations constrains the amplitude of the primordial density perturbations in the relevant length scale [231]. From observations, we have $\eta_{\text{BBN}} = (6.19 \pm 0.21) \times 10^{-10}$ and $\eta_{\text{CMB}} = (6.11 \pm 0.08) \times 10^{-10}$ [232]. Using these values, the upper limit on the total power $A$ for the delta-function type perturbation was obtained as [231]

$$A \lesssim 0.06.$$  \hfill (119)

In [233], a more stringent upper limit on $A$ was obtained by investigating the freezeout neutron fraction during BBN in the presence of the (adiabatic) primordial density perturbations (see also [234]). These density perturbations, which dissipate after BBN, produce local temperature fluctuations during BBN, since $\rho_s \propto T^4$. In local regions where the temperature is higher than the average, interactions persist longer, and the freezeout is delayed, resulting in a smaller amount of neutrons. On the other hand, more neutrons remain in colder regions. Because of two enhancement effects in hotter regions—that more baryons are available ($\delta n_b \propto \delta T$), and that physical volume is also enlarged by the metric perturbation compared to colder regions for fixed comoving volume—the averaged neutron fraction is biased toward the higher temperature region. As a result, the predicted helium abundance becomes smaller than the standard case. The derived upper limit on $A$ is [233]

$$A \lesssim 0.02.$$  \hfill (120)

Although this value does not immediately exclude the production of PBHs, this falls into the typical power required for PBHs. This suggests that reduction of errors in future measurements may find a PBH signal, or eventually rule out PBHs in the relevant mass range.
3.6. Future constraints

So far, all the constraints we have discussed are established constraints, in the sense that they are obtained from the existing observational data. Thus, as long as the data and theory translating to the constraint are correct, the derived constraint is real, and must be taken into account when one tests any early Universe model predicting PBH formation.

In the future, those constraints will be improved by new experimental apparatus similar to what is existing/existed but with better sensitivity. Furthermore, appearance of new technology will enable us to probe PBHs by using completely new methods. In this section, we briefly review several proposals to probe PBHs by future experiments.

3.6.1. Fast radio bursts. Fast radio bursts (FRBs) are radio transients that last only for $O(\text{ms})$ [235]. Since their discovery in 2007 [235], many FRBs have been detected. The measured values of dispersion measures suggest that FRBs occur at cosmological distances [236]. The origin of FRBs is not known yet.

In [237], it was pointed out that lensing of FRB signals by intervening PBHs can be used to probe PBHs with $\gtrsim 20 \, M_\odot$. The idea is as follows. Gravitational lensing causes appearance of two source images. Although the angular separation of the two images is too small to be resolved experimentally, the difference of arrival times of the images reaches the order of

$$T_{12} \simeq 1 \text{ ms } \left( \frac{M_{\text{PBH}}}{30 \, M_\odot} \right)$$

for $u = 1$ (see equation (56)). Thus, PBHs heavier than $\sim 10 \, M_\odot$ produce double bursts separated longer than the burst width, hence can in principle be resolved. Considering the two redshift distributions of FRBs, constant comoving number density and the one that follows the star formation rate, the lensing optical depth has been evaluated in [237]. It was found that planned experiment the Canadian Hydrogen Intensity Mapping Experiment (CHIME), which is expected to measure 700–15 000 FRBs per year [238], may detect tens of double bursts caused by PBHs if PBHs comprise all the dark matter and provide constraint $f_{\text{PBH}} \lesssim O(0.01)$ if no such events are detected.

3.6.2. Pulsar timing array (PTA) experiments. As explained in section 3.5.1, primordial density perturbations that generated PBHs from high-$\sigma$ peaks also produce the stochastic GWs by the mode–mode coupling. The present frequency of the GWs corresponding to the PBHs in the stellar mass range is in the nHz range, to which PTA experiments are sensitive. The Square Kilometer Array (SKA) is a planned giant radio telescope consisting of thousands of receptors with a total collecting area of about one square kilometer. With better sensitivity, SKA is expected to detect more pulsars, and improve the constraint on $\Omega_{\text{GW}}$ by $3 \sim 4$ orders of magnitude stronger than the current limit if no GWs are detected [239]. Such a constraint, if achieved, will further tighten the constraint in 3.5.1 much more severely.

While the above constraint is indirect in the sense that the constraint is derived by looking at the GWs from the primordial density perturbations that failed to turn into PBHs, pulsar timing also has potential to give a direct constraint on PBH abundance by measuring the effect of the Shapiro time delay caused by the PBHs intervening between pulsars and the Earth [240]. A PBH moving relative to the line of sight produces non-stationary change of the arrival times of pulses as

$$t(n) \approx nP_{\text{obs}} + \frac{n^2}{2}P_{\text{obs}}P'_{\text{obs}} + n^3P^3_{\text{obs}} \frac{4GM_{\text{PBH}}c^3}{3r_L^2},$$

(122)
for the $n$th pulse [240]. Here $P_{\text{obs}}, \dot{P}_{\text{obs}}, v_r, r_L$ are the observed pulse period, its time derivative, (constant) velocity of the PBH toward the line of sight, and the distance between the PBH and the line of sight, respectively. The observable effect of the PBH appears at $O(n^3)$ term. According to [240], non-detection of such effects in the long term observations ($O(10)$ years) of known pulsars or pulsars that would be newly detected by SKA can constrain PBHs as $f_{\text{PBH}} \lesssim 0.01–0.1$.

### 3.6.3. 21 cm

SKA is also expected to detect the cosmological 21 cm line from neutral hydrogen in the dark age before reionization. Observations of this 21 cm line have potential to constrain PBHs.

The 21 cm line is radiation emitted/absorbed by the transition between the two levels in the hydrogen 1s ground states. Energy of the spin singlet state is lower than that of the spin triplet states by $T^* = 5.9 \times 10^{-6} \text{eV} = 0.068 \text{mK}$. Relative number density between singlet ($n_0$) and triplet ($n_1$) hydrogen atoms is parametrized by the spin temperature $T_s$ as

$$\frac{n_1}{n_0} = 3 \exp \left( -\frac{T_s}{T^*} \right).$$  

(123)

There are three ingredients to determine the spin temperature as

$$T_s = \frac{T_{\text{CMB}} + y_\alpha T_\alpha + y_c T_K}{1 + y_\alpha + y_c},$$  

(124)

where the second and the third terms in the numerator represent the Lyman-$\alpha$ pumping, i.e. 21 cm transition through ambient Lyman-$\alpha$ photons as singlet/triplet $\rightarrow 2p \rightarrow$ singlet/triplet, and the collisions among hydrogen atoms, respectively [241]. The 21 cm intensity is normally expressed in terms of the brightness temperature $T_b$. With reference to the CMB brightness temperature, the differential temperature measured at the Earth is given by

$$\delta T_b = (1 + z)^{-1} (T_s - T_{\text{CMB}})(1 - e^{-\tau}),$$  

(125)

where $\tau$ is the optical depth for the 21 cm photons [241].

In [242], the effect of x-rays emitted by the accreting PBHs on the 21 cm fluctuations has been investigated. Radiation emitted from the accreting PBHs in the dark age ionizes and heats the surrounding neutral hydrogen gas. Ionization reduces the amount of neutral hydrogen, Lyman-$\alpha$ photons from PBHs affect the spin temperature through the $T_\alpha$ term, and the heating of neutral hydrogen by emanating radiation also affects the spin temperature by changing the kinetic temperature $T_K$. Assuming radiation intensity corresponding to 10% of the Eddington luminosity with a power-law x-ray spectrum, brightness temperature fluctuations was computed. It was found that with SKA-like experiments upper limit on PBH will be $\Omega_{\text{PBH}} = 10^{-5}(M_{\text{PBH}}/10^3 M_\odot)^{-0.2}$ at $z = 30$ and $\Omega_{\text{PBH}} = 10^{-7}(M_{\text{PBH}}/10^3 M_\odot)^{-0.2}$ at $z = 20$ for $10^2 < M_{\text{PBH}}/M_\odot < 10^5$.

In [243], based on the observation discussed in section 3.4 that the random distribution of PBHs adds an isocurvature component on top of the adiabatic perturbations, formation of minihalos sourced by PBHs and their effect on 21 cm fluctuations has been studied. Minihalos can change the 21 cm signals, since the gas temperature and the density of neutral hydrogen inside the minihalos differ from the background values. According to the analysis in [243], the PBHs with $M_{\text{PBH}} \gtrsim 10 M_\odot$ leave enhancement of the brightness temperature that is detectable for SKA for $f_{\text{PBH}} > 10^{-3} - 10^{-4}$.

### 3.7. Constraints for the extended PBH mass function
So far, we have assumed that the PBH mass function is monochromatic, i.e. all the PBHs have the same mass. Thus, the derived constraints are valid only when the PBH mass function is sufficiently narrow, and can no longer be trusted when the mass function is broad. In this subsection, we briefly mention how to generalize the constraint on the PBH abundance to the case of the extended PBH mass function (see [8, 9, 244] for more details).

A simple formalism for deriving the PBH constraint for the extended mass function was given in [9]. Let $\psi(M)$ be the PBH mass function, and consider an observable $A$ which PBHs produce or contribute. Only in this subsection, we define the normalization of the PBH mass function as

$$ f_{\text{PBH}} = \int \psi(M_{\text{PBH}}) dM_{\text{PBH}}. \quad (126) $$

The quantity $A$ depends on the details of experiments as well as the astrophysical phenomena one is interested in. For instance, in the microlensing experiment, $A$ is the expected number of the microlensing events detected by a particular experiment, and depends on the sensitivity of the detectors. Generally, $A$ is a functional of the mass function, and can be expressed as

$$ A[\psi(M)] = A_0 + \int dM \psi(M) K_1(M) + \int dM_1 dM_2 \psi(M_1) \psi(M_2) K_2(M_1, M_2) + \cdots, \quad (127) $$

where $\cdots$ are higher order terms in $\psi(M)$. Here $A_0$ represents any contribution other than from PBHs. For the microlensing experiments, $A[\psi(M)]$, the total expected number of the microlensing events during the observation period, contains only the $K_1(M)$-term as the PBH contribution, which is given by [244, 245]

$$ K_1(M) = E \int_0^{E} \frac{32 D_{\text{urb}}(t_e)}{u_c^2 M} \int_{0}^{t_e} \rho_{\text{dm}}(x) R^2_{Q}(x) e^{-Q(x)} dx, \quad Q(x) \equiv \frac{4R^2_{Q}(x) u^2_c}{R^2_{Q}(x)} . \quad (128) $$

where $E$ is the number of stars multiplied by the observation period, $u_T$ is the maximum $u$ (see equation (45)) below which the microlensing magnification becomes greater than the threshold, $v_c \approx 220 \text{ km s}^{-1}$ is the circular velocity of the Sun, and $\epsilon(t_e)$ is the probability that the detector detects microlensing events that last for the period $t_e$. As we will see in the next section, if $A$ is the merger event rate of the PBH binaries that are formed in the early Universe, the expansion (127) starts at the $K_1$ term if only the third BH is taken into account as the dominant source of the tidal force, and starts at even higher-order terms if more distant PBHs are also included. On the other hand, for PBH binaries that are formed in the low-redshift Universe by close encounters, $A$ is given by the $K_2$-term. These examples show that the order in $\psi$ at which the expansion of (127) starts varies for different observables.

Let us consider the simplest case, where $A$ is dominantly given by the $K_1$ term. Suppose the observations place an upper bound on $A$ as

$$ |A[\psi(M)]| \leq A_{\text{exp}}. \quad (129) $$

Then, for the monochromatic mass function $\psi(M) = f_{\text{PBH}}(M_{\text{PBH}}) \delta(M - M_{\text{PBH}})$, the above constraint becomes

$$ f_{\text{PBH}}(M_{\text{PBH}}) \leq A_{\text{exp}} / K_1(M_{\text{PBH}}) \equiv f_{\text{max}}(M_{\text{PBH}}), \quad (130) $$

where $f_{\text{max}}(M_{\text{PBH}})$ is the maximally allowed value of $f_{\text{PBH}}$ by the observation under consideration when all the PBHs have mass $M_{\text{PBH}}$. Replacing $K_1$ by $f_{\text{max}}$, we finally obtain

$$ \int \frac{\psi(M_{\text{PBH}})}{f_{\text{max}}(M_{\text{PBH}})} \leq 1. \quad (131) $$
This formula enables us to derive the constraint on the extended mass function, once we know the upper limit $f_{PBH}$ for the monochromatic mass function. This simple conversion does not hold in general when higher order terms contribute to $A$.

In [9], based on the above formalism, constraints on the PBH abundance for several types of the extended mass function were obtained. The results show that the constraints generally become stringent for the extended mass function compared to the case of the monochromatic mass function, due to the combination of the multiple observational limits for different PBH masses.

4. PBHs as sources of gravitational waves (GWs)

As we have discussed in detail in section 3, before the direct detection of GWs by LIGO, the search for PBHs in the Universe relied to a varying degree on electromagnetic waves. For instance, gravitational lensing uses the background electromagnetic sources such as stars and quasars, and the dynamical constraints are derived by the observations of stars which PBHs affect. So far, although there have been a couple of studies which attributed some unexplained observational signals to the PBHs, no observational searches for PBHs using electromagnetic waves have detected indisputable evidence for the existence of PBHs.

Direct detection of GWs by laser-interferometry is a completely novel method to search for PBHs that does not rely on electromagnetic waves. Soon after LIGO announced the first detection of GWs, caused by the merger of two BHs in a binary, in February 2016, several groups pointed out the possibility of the scenario that the observed BHs are PBHs [12–14],[196]. Thus, the exciting possibility has arisen that we might have discovered PBHs for the first time by the direct observation of GWs (not just constraint!).

Explaining the LIGO event by the PBHs is non-trivial in two respects. First, since the GWs are emitted from BH binaries, the formation mechanism of the PBH binary must be considered in order to test the PBH scenario with GW observations. Secondly, as discussed in section 3, there are existing constraints on the PBH abundance for the mass around the observed BH mass $\sim 30M_\odot$. It needs to be checked whether the PBH scenario does not conflict with those constraints. These issues, which we will address in detail later, are the first main topic in this section. It will turn out that these issues are cleared, and the PBH scenario can be considered as a reasonable candidate scenario for the observed BH merger event.

There are also astrophysical explanations for the observed heavy stellar-mass BHs (see [11] and references therein): the isolated field binary scenario, in which two stars in an isolated binary collapse to BHs in decreasing order of star mass; and the dynamical formation scenario, in which isolated BHs in a dense stellar environment form BH binaries at the core and eventually are ejected by three-body interactions. In both cases, studies [246–248] suggest that heavy BHs about $30M_\odot$ are born out of stars with low metallicity (at most half solar metallicity). Although there are many astrophysically uncertain factors in estimating the merger event rate, these scenarios are roughly consistent with the GW observations. Thus, our next task is to clarify how to discriminate these scenarios of binary BH formation, and to pin down the correct one.

GW astronomy has just began, and its future is bright. In the coming decade, the sensitivity of the existing GW experiments will greatly improve, and many merger events of the BH binaries will be detected. Looking at further future, new detectors with a larger arm length will be built, both on ground and in space, which can probe GWs in different frequency bands. By such experiments, we will obtain much information about BH binaries such as mass distribution, space distribution, redshift distribution and spin distribution. These information will...
definitely help us discriminate different scenarios of binary BH formation. How such information can be used for this purpose is the second topic in this section. Concerning this point, discussions about the astrophysical scenarios are beyond the scope of this monograph, and we will concentrate on the PBH scenario in what follows.

4.1. Formation of PBH binaries

Here, we will briefly review two distinct formation mechanisms for PBH binaries. It is important to remark here that the two mechanisms are not incompatible—i.e. not like the relationship between oil and water, but operate at different epochs in the cosmic history. Thus, what matters is to figure out which mechanism is more efficient to make PBH binaries that merge by present time.

4.1.1. PBH binary formation in the early Universe. The first mechanism we are going to discuss operates in the epoch when the Universe was dominated by radiation. This mechanism was proposed in [15] to investigate the detectability of binary mergers of MACHO PBHs (∼0.5 M⊙). In [15], it was assumed that \( f_{\text{PBH}} = 1 \), all the PBHs have the same mass, and PBHs are initially distributed randomly in space (Poisson distribution). In the following, we consider the monochromatic PBH mass function and uniform distribution of PBHs, and treat \( f_{\text{PBH}} \) as a free parameter.

Just after the PBHs were formed in the very early Universe, they were distributed sparsely—that is, the mean distance at that time was much longer than the Hubble horizon. Because of the rapid cosmic expansion, they were on the expansion flow and the mean distance grew in proportion to the scale factor \( a(t) \propto t^{1/2} \). Since the Hubble horizon grows as \( H^{-1}(t) \propto t^{1/2} \), the mean distance relative to the Hubble horizon decreases as the Universe expands. Denoting by \( \ell_{\text{PBH}}(z) \) the mean PBH distance at redshift \( z \), its length normalized by the Hubble horizon is given by

\[
H(z)\ell_{\text{PBH}}(z) = H(z)f_{\text{PBH}}^{-1/3}(1+z)^{-1} \simeq 6 \times 10^{-6} f_{\text{PBH}}^{-1/3} \left( \frac{1+z}{1+z_{\text{eq}}} \right) \left( \frac{M_{\text{PBH}}}{30 M_{\odot}} \right)^{1/3},
\]

for \( z > z_{\text{eq}} \), where \( z_{\text{eq}} \) is the redshift of the matter–radiation equality and \( n_{\text{PBH}} \) is the comoving (initial) PBH number density. Thus, for the stellar-mass PBHs, unless \( f_{\text{PBH}} \) is extremely tiny as \( f_{\text{PBH}} \lesssim 10^{-15} \), there is a period in the radiation-dominated epoch in which there are typically more than one PBHs in the Hubble horizon. The range of \( f_{\text{PBH}} \) of our interest is much larger than this value.

Let us focus on a PBH and the PBH closest to it, and let their comoving distance be \( x \). The physical separation at redshift \( z \) is then \( x/(1+z) \). The cosmic expansion acts as a force that pulls two PBHs away from each other. Two PBHs are also pulled by the gravitational force acting between them. The corresponding free-fall time becomes shorter than the Hubble time at a certain time during the radiation-dominated epoch if the comoving distance is shorter than

\[
x < x_{\text{max}} \equiv f_{\text{PBH}}^{1/3} \ell_{\text{PBH}}(z = 0).
\]

A PBH pair satisfying this condition decouples from the cosmic expansion and becomes gravitationally bound\(^{22}\). Conversely, a PBH pair with comoving distance longer than \( x_{\text{max}} \) never

\(^{22}\) This picture has been confirmed to be correct by numerically solving the Newtonian equations of motion [249, 250].
becomes gravitationally bound since a ratio of the free-fall time to the Hubble time remains constant in the subsequent matter dominated epoch. The decoupling redshift $z_{\text{dec}}$ is given by

$$1 + z_{\text{dec}} = (1 + z_{\text{eq}}) \left(\frac{x_{\text{max}}}{x}\right)^3.$$  \hfill (134)

As the two PBHs come closer, the surrounding PBHs, especially the nearest one, exert torques on the bound system. As a result, the two PBHs avoid a head-on collision, and typically form a highly eccentric binary (see figure 12). The major axis $a$ of the binary orbit is equal to $x/(1 + z_{\text{dec}})$. The angular momentum $J$ of the binary is estimated by multiplying the exerted torque from the nearest PBH by the free-fall time, and is given by

$$J \simeq t_{\text{ff}} GM_{\text{PBH}}^2 (1 + z_{\text{dec}})^2 x^2 y^3,$$ \hfill (135)

where $y$ is the comoving distance to the nearest PBH, and a factor of $O(1)$ has been ignored.

Assuming Keplerian motion after formation of the binary, the angular momentum is related to the eccentricity $e$ of the orbit as $J^2 = G\mu^2 Ma(1 - e^2)$, where $\mu$ and $M$ are the reduced mass and total mass respectively. We can convert these equations in terms of $x$ and $y$ as

$$a = \frac{\rho_c(1 + z_{\text{eq}})M_{\text{PBH}} x^4}{4\pi x^2 dx \frac{4\pi y^3 dy}{y_{\text{PBH}}^3}}, \quad e = \sqrt{1 - \left(\frac{x}{y}\right)^6}.$$ \hfill (136)

Because of the random distribution of PBHs, the probability that comoving distances are in the intervals $(x, x + dx)$ and $(y, y + dy)$ is given by

$$dP = \frac{4\pi x^2 dx \frac{4\pi y^3 dy}{y_{\text{PBH}}^3}}{\frac{4\pi y^3 dy}{n_{\text{PBH}}}} \exp \left(-\frac{4\pi y^3}{3n_{\text{PBH}}}ight) \Theta(y - x) \Theta(y_{\text{max}} - y).$$ \hfill (137)

Instead of dealing with this probability distribution, the simplified one as

$$dP = \frac{4\pi x^2 dx \frac{4\pi y^3 dy}{y_{\text{PBH}}^3}}{\frac{4\pi y^3 dy}{n_{\text{PBH}}}} \Theta(y - x) \Theta(y_{\text{max}} - y), \quad y_{\text{max}} = \left(\frac{4\pi}{3 n_{\text{PBH}}}ight)^{1/3},$$ \hfill (138)

was used in [15]. Notice that because of $y_{\text{max}}$, there is an upper limit on the eccentricity for fixed $x$ (and hence for fixed $a$) via

$$e_{\text{max}}^2 = 1 - \left(\frac{4\pi}{3 n_{\text{PBH}}}ight)^2 \left(\frac{(1 + z_{\text{eq}})M_{\text{PBH}}}{\rho_c(1 + z_{\text{eq}})M_{\text{PBH}}}ight)^2 a.$$ \hfill (139)

\textit{Figure 12.} A schematic picture of the formation of PBH binaries in the radiation-dominated epoch.

\textsuperscript{23} It is worth mentioning that the exact distribution function of the initial angular momentum, taking into account all the distant PBHs, was obtained in [250].
The schematic graph of $e = e_{\text{max}}$ as a function of $a$ is shown in figure 13 as a blue curve. Notice that there is also a maximum for $a$ as $a_{\text{max}} = x_{\text{max}}/(1 + z_{\text{eq}})$.

Knowing the probability distribution of $(x, y)$, we can translate it in terms of $(a, e)$ by using the formula (136) as

$$dP = \frac{4\pi^2}{3} n_{\text{PBH}} (1 + z_{\text{eq}})^{3/2} f_{\text{PBH}}^{3/2} a^{1/2} e (1 - e^2)^{-3/2} \, da \, de.$$  \hspace{1cm} (140)

This equation tells us how many PBH binaries with orbital parameters in $(a, a + da), (e, e + de)$ exist in the Universe at formation time. After PBH binaries were formed in the radiation-dominated epoch, each PBH binary continuously emit gravitational waves, and finally merge much later. The estimate of the merger rate will be discussed in the following subsection.

4.1.2. PBH binary formation in the present Universe. In addition to the PBH binary formation in the radiation-dominated epoch, PBHs can form binaries in the present Universe, which we review in this subsection.

Let us consider a situation where a PBH traveling in space accidentally has a near-miss with another PBH. These PBHs may be concentrated in a local region, such as inside a larger dark matter halo, or simply moving freely in space. For the moment, we do not make a particular assumption on how PBHs are distributed in the region of our interest. Figure 14 shows a schematic picture of the close encounter with impact parameter $b$ and periastron $r_p$. Near the periastron, the relative acceleration of the PBHs becomes maximal, and dominant emission of
Gravitational radiation occurs. If the amount of energy of the emitted GWs is greater than the kinetic energy of PBHs, then the PBHs cannot escape to infinity any more, and form a bound system. Since a direct head-on collision is probabilistically unlikely, a binary typically results. Let us investigate this problem more quantitatively [251]. We assume that the power of GWs is estimated by the unperturbed trajectory (without backreaction due to GW emission) in the Newtonian approximation. According to Peters [252], the time-averaged energy loss rate of the binary in the Keplerian orbit due to gravitational radiation is given by

\[
\left\langle \frac{dE}{dt} \right\rangle = -\frac{32}{5} \frac{G^4 m_1^3 m_2^2 (m_1 + m_2)}{a^3 (1 - e^2)^{3/2}} \left( 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right).
\]

Thus, the energy loss during one orbital period \(T\) becomes

\[
\Delta E = -T \left\langle \frac{dE}{dt} \right\rangle = \frac{85\pi \sqrt{G(m_1 + m_2)} G^3 m_1^2 m_2}{12 \sqrt{2} r_p^{5/2} (1 + e)^{7/2} (1 + e)} \left( 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right),
\]

where we have used Kepler’s third law and \(r_p = a(1 - e)\). We can approximate the trajectory of the close encounter by the ellipse with \(e = 1\), since the true trajectory is physically indistinguishable from the ellipse near the periastron where the dominant emission of GWs occurs. Then, the energy loss by the close encounter is obtained by plugging \(e = 1\) into the above equation:

\[
\Delta E = \frac{85\pi \sqrt{G(m_1 + m_2)} G^3 m_1^2 m_2}{12\sqrt{2} r_p^{1/2}}.
\]

If this energy is greater than the kinetic energy \(\mu v^2/2\), where \(\mu\) is the reduced mass and \(v\) is the relative velocity at large separation, then the PBHs form a binary. This imposes a condition on \(r_p\) as

\[
r_p < r_{p,max} = \frac{85\pi G^2 (m_1 + m_2)^{3/2} m_1 m_2}{6\sqrt{2} \mu v^2}.
\]

In the Newtonian approximation, the relation between \(b\) and \(r_p\) is given by

\[
b^2 (r_p) = r_p^2 + \frac{2GMr_p}{v^2}.
\]

The encounter with the impact parameter less than \(b(r_{p,max})\) yields a binary. In the limit of the strong gravitational focusing \((r_p \ll b)\), which we are interested in, the cross section for forming a binary becomes

\[
\sigma = \pi b^2 (r_{p,max}) \simeq \left( \frac{85\pi}{3} \right)^{2/7} \left( \frac{2GM_{PBH}}{v^{18/7}} \right)^2.
\]

Contrary to the PBH binaries that are formed in the radiation dominated epoch, the PBH binaries produced by the present mechanism merge in less than the age of the Universe [253].

4.2. Merger event rate of PBH binaries

Having explained two mechanisms for the formation of PBH binaries, we review the expected merger event rate of PBH binaries formed in each mechanism, separately.

4.2.1. PBH binaries formed in the early Universe. The following discussion is based on [14]. PBH binaries that are formed in the radiation-dominated epoch continuously emit gravitational
waves, gradually shrink, and finally merge. Since the initial orbital parameters of the binaries are stochastic, some binaries merge in the past, others at the present epoch, and yet others in the future. According to Peters [252], a binary consisting of point masses \( m_1 \) and \( m_2 \) with orbital parameter \((a, e)\) merges due to gravitational radiation after time \( t \) given by

\[
t = \frac{15}{304} \frac{a^4}{G^2 m_1 m_2 (m_1 + m_2)} \left[ (1 - e^2)^{1/2} \left( 1 + \frac{121}{304} e^2 \right)^{1/2} \right]^4 \int_0^{e'} \frac{e'^{3/2}}{(1 - e'^2)^{1/2}} \left( 1 + \frac{121}{304} e'^2 \right)^{e'/e} \, de'.
\]

The lower limit of the integral is set to 0 by the assumption that the binary is almost circular \((e' = 0)\) just before the binary merges. Because of the smallness of the tidal force from the outer PBHs compared to the gravitational force between the PBHs that form the binary, the orbital eccentricity at the binary formation time is typically close to unity. When \( e \approx 1 \), the above formula can be simplified as

\[
t = \frac{3}{85} \frac{1}{G^2 m_1 m_2 (m_1 + m_2)} (1 - e^2)^{7/2} a^4.
\]

This shows that a highly eccentric binary merges in shorter time by a factor \( \frac{268}{257} \approx 1.03 \) than the circular binary with the same semi-major axis, which simply reflects that the binary radiates GWs dominantly around the periastron. It is this factor that makes PBH binaries form in the radiation-dominated epoch efficiently merge within the age of the Universe, and yields a merger event rate that can even exceed that estimated by LIGO when PBHs constitute all the dark matter.

For simplicity, let us consider the case where all the PBHs have the same mass. In this case, equation (148) becomes

\[
\frac{t(a, e)}{t_c} = Q(1 - e^2)^{7/2} a^4,
\]

where \( Q = \frac{1}{170 G^3 M_{\text{PBH}}^2} \). From this equation, we can express \( a \) as a function of \( t \) and \( e \) as \( a = a(t, e) \). Then, we can rewrite the probability (140) as

\[
dP = \frac{4\pi^2}{3} \frac{1}{n_{\text{PBH}}} \left( 1 + z_{\text{eq}} \right)^{3/2} f_{\text{PBH}} a^{1/2} e(1 - e^2)^{-3/2} \frac{\partial a}{\partial t} \, dt \, de.
\]

Since the observations do not measure the initial eccentricity of the binaries, we integrate this probability over \( e \) and find

\[
dP = \frac{3}{58} \left( \frac{t}{T} \right)^{1/2} \left[ \frac{1}{1 - e_{\text{upper}}^2} \right]^{3/2} - 1 \frac{dt}{T},
\]

where \( e_{\text{upper}} \) is defined by

\[
e_{\text{upper}} = \begin{cases} \sqrt{1 - \left( \frac{t}{T} \right)^2}, & \text{for } t < t_c, \\ \sqrt{1 - \left( \frac{4 \pi f_{\text{PBH}}}{3} \right)^2 \left( \frac{t}{T} \right)^2}, & \text{for } t \geq t_c, \end{cases}
\]
and $t_c$ is defined by $t_c = T\left(\frac{4n_{\text{PBH}}}{3}\right)^{\frac{3}{4}}$. This gives the probability that a given PBH forms a binary and merges at a time in $(t, t + dt)$. Thus, the merger event rate $R$ per unit volume per unit time (at time $t$) is given by

$$R = n_{\text{PBH}} \frac{dP}{dt} = \frac{3n_{\text{PBH}}}{58} \left(\frac{t}{T}\right)^\frac{1}{2} \left[\frac{1}{(1 - e_{\text{upper}}^2)^\frac{3}{4}} - 1\right] \frac{1}{t}. \tag{153}$$

The red curve in figure 15 shows $R$ for $M_{\text{PBH}} = 30 M_\odot$, BH mass close to the first event (GW150914) detected by LIGO, and for $t = 14$ Gyr as a function of $f_{\text{PBH}}$. We find that the merger rate largely exceeds LIGO observation when $f_{\text{PBH}} \simeq 1$ and lies in the band estimated by LIGO for $f = 5 \times 10^{-4} \sim 2 \times 10^{-3}$. Thus, in the present mechanism of the PBH binary formation, the case that PBHs constitute only a fraction of dark matter is observationally relevant.

The knee at $f \simeq 7 \times 10^{-4}$ corresponds to the bifurcation point $t = t_c$. Physically, this bifurcation can be understood as follows. Since there is an upper limit on the distance between neighboring PBHs that form binaries, equation (149) tells that there is also an upper limit on the eccentricity for fixed $t$. Since the eccentricity is caused by the tidal force from the outer BH, the eccentricity becomes closer to unity, as the outer BH is more distantly separated. The probability of the location of the outer BH peaks at the mean separation of PBHs. If $f_{\text{PBH}}$ falls below about $7 \times 10^{-4}$, the eccentricity caused by the outer BH at the mean separation exceeds the mentioned upper limit on the eccentricity, and the outer BH must be located closer than the mean distance to cause the merger at time $t$. As a result, the merger rate is suppressed more by the volume factor than in the case where the outer BH is at the mean distance, which produces the knee.

One may wonder if the mergers of PBH binaries that have accumulated over the age of the Universe significantly modify the PBH mass function in the present Universe. Figure 16 shows the probability, which is obtained by integrating equation (153) over the cosmic time up to $t$, that any PBH undergoes a merger by the time $t$. From this figure, we find that even in the case with $f_{\text{PBH}} = 1$, the probability that a PBH has merged by the present time is about...
Thus, in this case the PBH mass function in the present universe is reduced by ∼0.01 at $M = M_{\text{PBH}}$, and has a little spike at $2M_{\text{PBH}}$. For smaller $f_{\text{PBH}}$, the time evolution of the mass function is more moderate.

From the accumulated merger probability $P$ given in figure 16, we can also estimate the accumulated merger rate per unit time as $\sim H_0^{-2} n_{\text{PBH}} P \sim 10^{-4} \text{ s}^{-1}$ for $M_{\text{PBH}} = 30 M_\odot, f_{\text{PBH}} = 10^{-3}$, and $t \simeq 10^{10} \text{ yr}$. Thus, the typical time interval between the successive merger events that occurred in the Hubble volume is much longer than the duration of a single merger event, which is classified as the ‘shot noise’ type according to [255].

We have assumed that PBHs are not initially clustered. The effect of clustering on the merger rate (153) was addressed in [256]. According to their result, the merger rate is enhanced by the clustering, which is a natural consequence since the clustering effectively increases the PBH comoving number density.

There are several effects that have been ignored in deriving the merger rate (153). These include tidal forces from outer PBHs other than the nearest one, subsequent capture of the outer BH by binaries already formed, initial peculiar velocities of PBHs, gravitational perturbation by surrounding non-PBH dark matter inhomogeneities, and subsequent accretion of dark matter and baryonic gas onto PBH binaries.

The first three effects have been investigated in [249], in which $f_{\text{PBH}} = 1$ was assumed. Because the tidal force is inversely proportional to the distance cubed and each outer PBH exerts torque in a random direction, only a few outer PBHs principally contribute to the tidal force. As a result, the merger event is reduced by at most 60%. The present mechanism of the formation of the bound system by the decoupling from the cosmic expansion works as long as the distance between PBHs is smaller than $x_{\text{max}}$ (see equation (133)). For $f_{\text{PBH}} = 1$, $x_{\text{max}}$ becomes the mean PBH distance. This means that the outer PBH, which generated angular momentum of the inner PBH binary, is commonly trapped by the binary later, resulting in the hierarchical triple system. As the three-body problem is difficult to analyze, the fate of such a triple system is hard to predict. In [249], such a case was left aside, and another case—that the distance to the nearest outer PBH is greater than $x_{\text{max}}$—was investigated. Restricting the PBH mergers to the latter case, it was found that the merger rate is reduced by about 40%, which is not a significant impact. If $f_{\text{PBH}} < 1$, which is the case relevant to LIGO observation, we expect that the probability of forming the triple system is more suppressed than the case of $f_{\text{PBH}} = 1$, since the fraction of PBHs satisfying the condition equation (133) is reduced by a factor $f_{\text{PBH}}$. Initial peculiar velocities, that are randomly directed for each PBH, naturally

![Figure 16. Probability that a PBH of $M_{\text{PBH}} = 30 M_\odot$ forms a binary and merges by cosmic time $t$ for four different values of $f_{\text{PBH}}$.](image)
yield angular momentum of the binary. If the angular momentum from the peculiar velocity is larger than that from the tidal force we discussed, the prediction of the merger rate will be modified. According to [249], the effect of the peculiar velocity is to reduce the merger rate by at most 30%, even if the initial peculiar velocity is equal to the speed of light.

In [257], the additional tidal force from the adiabatic perturbation of non-PBH dark matter (valid for $f_{\text{PBH}} < 1$) was taken into account, and its effect on the merger rate was investigated. Since we do not know the power of dark matter perturbations on relevant scales, smaller than those probed directly by the CMB observations, straightforward extrapolation of the Planck results was adopted to define the dark matter perturbations. It was found that the inclusion of this effect reduces the merger rate by at most a factor of 2. Effects of the tidal force from the nonlinear structures of dark matter on the PBH binaries have been estimated in [250], and were found to be insignificant as well.

The accretion of dark matter and baryonic gas onto PBH binaries was investigated in [258]. According to this study, the continuous accumulation of matter onto the binaries can rapidly decrease the binary radius by dynamical friction, and the merger may happen in the early Universe. More recent study in [250], based on the simple analytic calculation, suggests that the baryonic mass accumulated around the PBHs is likely overestimated in [258], and the baryonic effect is not significant enough to change the evolution of the PBH binaries. However, more detailed studies are needed to clarify the effectiveness of this scenario.

4.2.2. PBH binaries formed in the present Universe. The merger event rate of PBH binaries formed in the present Universe was estimated in [12]. The cross section of forming PBH binaries by the close encounter given by equation (146) shows that the binary formation is effective for low relative velocity. The encounters, which are accidental, are more frequent in high density than low density regions. These facts suggest that the PBH binary formation occurs efficiently inside low-mass dark halos, which are dense and have small virial velocity. The merger rate inside a halo with mass $M_h$ is given by

$$R_h(M_h) = \int_0^{R_{\text{vir}}} dr \frac{4\pi}{3} r^2 \left( \frac{\rho_{\text{PBH}}(r)}{M_{\text{PBH}}} \right)^2 \langle \sigma v_{\text{PBH}} \rangle,$$

(154)

where $\rho_{\text{PBH}}(r)$ is the density profile of PBHs inside the halo, and $\langle \sigma v_{\text{PBH}} \rangle$ denotes the average over relative velocity distribution with $\sigma$ given by equation (146). In [12], it was assumed that $\rho_{\text{PBH}}(r)$ and the velocity distribution are given by the Navarro–Frenk–White profile and the Maxwell–Boltzmann distribution respectively. The total merger rate per unit volume and unit time is then given by

$$\mathcal{R} = \int dM_M \frac{dn}{dM_h} R_h(M_h),$$

(155)

where $\frac{dn}{dM_h}$ is the halo mass function and $M_{\text{min}} \sim 400 M_\odot f_{\text{PBH}}^{-1}$ is the minimum mass of halos that have not yet evaporated by the present time. Three different mass functions, one obtained by the Press–Schechter formalism, one based on the simulations by Tinker et al [259], and another that has cutoff at small mass from Jenkins et al [260], were employed in computing $\mathcal{R}$. The result for $M_{\text{PBH}} = 30 M_\odot$ is given by

$$\mathcal{R} \approx 2\alpha f_{\text{PBH}}^2 \frac{\alpha}{10} Gpc^{-3}yr^{-1},$$

(156)

where $\alpha \approx 1, 0.6, 10^{-2}$ for Press–Schechter, Tinker et al and Jenkins et al respectively.

The blue curve in figure 15 shows $\mathcal{R}$ given by equation (156) with $\alpha = 1$. We find that the expected merger rate is consistent with the LIGO observation when $f_{\text{PBH}} \approx 1$. The fact that
there is a physically interesting region $f_{\text{PBH}} \leq 1$ consistent with the observations by LIGO makes the binary formation path by the present mechanism interesting. This formation path becomes important if the mergers of PBH binaries formed in the radiation dominated epoch are significantly reduced by some mechanism.

If PBHs cluster in the vicinity of the supermassive BHs at galactic centers, they can efficiently form binaries by the close encounter mechanism [261]. According to the study in [261], the expected merger rate exceeds that given by equation (156) for some possible range of the slope of the PBH density profile around the supermassive BHs. In such a case, PBHs in the galactic center may become the dominant channel to form binaries which emit GWs.

4.3. Constraint on the PBH abundance from the GW observations

In the previous subsection, we have seen that $30M_\odot$–$30M_\odot$ PBH binaries that are formed in the radiation-dominated era merge with frequency consistent with that estimated by the LIGO observations if $f_{\text{PBH}} \sim 10^{-3}$, and merge much more frequently if $f_{\text{PBH}} \sim 1$. This means we can use the GW observations to place an upper limit on $f_{\text{PBH}}$ irrespective of whether the observed mergers of BH binaries are attributed to PBHs or not. This was done in [250] for various PBH masses with the monochromatic mass function. The derived constraint is shown in figure 17. We see that the derived constraint $f_{\text{PBH}} \lesssim 10^{-3} - 10^{-2}$ excludes stellar mass PBHs as the dominant component of dark matter. This result demonstrates that GW observations provide a novel tool to probe/constrain PBHs independently of electromagnetic observations.

4.4. Distinguishing from other scenarios

In the previous subsections, we have reviewed the two mechanisms of PBH binary formation that work at different cosmic epochs, and discussed that the LIGO events can be explained by mergers of PBH binaries. Of course, the PBH scenario is not the only explanation of the LIGO events, and there are several astrophysical scenarios that have been proposed as the origin of the LIGO events. So far, due to the small number of events detected, those scenarios are allowed observationally. The next obvious task is to clarify how we can test the PBH scenario and discriminate it from the others by using future observations that will bring much more information. This is the main topic of this subsection.
4.4.1 Stochastic GW background from PBH binaries. At the time of writing this article (autumn 2017), LIGO has detected five GW events from BH–BH mergers. These GW events were close enough that they were heard as single events by the detectors. In addition to these loud events, there are other merger events that occur at more distant places. Although GWs from such mergers pass through the Earth more frequently than the louder ones, they are tiny, buried in noise, and may not be identified as single events. However, those events may be detected as a whole by taking correlations of the GW signals among different detectors and integrating it over some time. If such tiny GWs exist, after the time integration, the GW signal will emerge. Such GWs are referred to as stochastic GW background.

It is customary to represent the strength of stochastic GWs in terms of the GW energy density $\rho_{GW}$ per logarithmic frequency bin normalized by the critical density $\rho_c$ (for instance [262]):

$$\Omega_{GW}(f) = \frac{1}{\rho_c} \frac{d\rho_{GW}}{d\ln f}.$$  \hspace{1cm} (157)

For GWs that are emitted by the mergers of compact objects in binaries, $\Omega_{GW}$ can be written as [263]

$$\Omega_{GW}(f) = \frac{f}{\rho_c H_0} \int_0^{\infty} dz \frac{R(z)}{(1 + z) \sqrt{\Omega_m (1 + z)^3 + \Omega_\Lambda}} \frac{dE_{GW}(f')}{df'} \bigg|_{f' = (1+z)f},$$  \hspace{1cm} (158)

where $R(z)$ is the merger rate of the source we are interested in, and $dE_{GW}/df'$ is the spectral energy density of a source, which describes how much energy is released from the source in the form of GWs with frequency $f$. The spectral energy density can be determined by general relativity. Based on the phenomenological waveforms in the Fourier domain for the inspiral, merger and ringdown phases for non-spinning BH binaries [264], Zhu et al [265] converted it to the spectral energy density as
\[ dE_{GW} \left( \frac{df}{f} \right) = \frac{(G\pi)^{2/3} M_c^{5/3}}{3} \begin{cases} f^{-1/3}, & (f < f_1, \text{ inspiral phase}), \\ f_1^{-1}f_1^{5/3}, & (f_1 < f < f_2, \text{ merger phase}), \\ f_1^{-1}f_2^{-2/3} + f_2^{-4/3}, & (f_2 < f < f_3, \text{ ringdown phase}), \end{cases} \]

where \( M_c \) is the chirp mass, and \( f_1, f_2, f_3, \sigma \) are fitting parameters. Schematic shape of this function is shown in figure 18. Modified waveforms for BH–BH binaries, generalized to non-precessing spins, are given in Ajith et al [266].

For the evaluation of \( \Omega_{GW} \) from the mergers of PBH binaries, we use either equation (153) or (156) for \( R(z) \) in equation (158). For the PBH binaries formed in the radiation-dominated epoch, the expected \( \Omega_{GW} \) was computed for the first time in [267] for the MACHO mass range \( \sim 0.5 M_\odot \). In light of the detection of BH–BH binaries of \( \sim 30 M_\odot \) by LIGO, the computation of \( \Omega_{GW} \) for \( \sim 30 M_\odot \) was done in [268]. The analysis was generalized to the case where the PBH mass function is not monochromatic in [256]. For the PBH binaries formed in the present Universe, the expected \( \Omega_{GW} \) was computed in [269] (assuming monochromatic PBH mass function). Figure 19 shows curves of \( \Omega_{GW}(f) \) predicted in the aforementioned PBH scenarios: red curves for PBH binaries formed in the radiation-dominated epoch \( (M_{PBH} = 30 M_\odot, f_{PBH} = 10^{-2} \text{ and } f_{PBH} = 10^{-3}) \), and blue ones for PBH binaries formed in the present Universe \( (M_{PBH} = 30 M_\odot \text{ and } f_{PBH} = 1) \). As is anticipated from figure 15, PBH binaries formed in the radiation-dominated epoch produce larger stochastic GWs than those formed in the present Universe for the same \( f_{PBH} \). The orange colored band represents the contribution from astrophysical BH binaries within uncertainties. Black curves given in [263] represent sensitivities of LIGO at different observation stages.

Figure 19. Predicted \( \Omega_{GW} \) in the PBH scenarios. Two red curves obtained in [256] are the predicted \( \Omega_{GW} \) from PBH binaries that are formed in the radiation-dominated epoch for two different values of \( f_{PBH} \). Blue curve obtained in [269] is the predicted \( \Omega_{BH} \) from PBH binaries that are formed in the present Universe with \( f_{PBH} = 1 \). Orange band computed in [269] shows the expected \( \Omega_{GW} \) from the astrophysical BHs within uncertainties. Black curves given in [263] represent sensitivities of LIGO at different observation stages.
While different scenarios predict different curves of $\Omega_{GW}$ in figure 19, their shapes look similar. This is understandable, given that the shape is essentially determined by the shape of the spectral energy density $\frac{dE_{GW}}{df}$ and the source in different scenarios is physically identical: namely, BH–BH binaries. Since the physics that determines the spectral energy density is well understood (just general relativity), the qualitative shape of $\Omega_{GW}$-curves is robust. But this is a two-edged sword for those who try to test the PBH scenarios using stochastic GWs. When $\Omega_{GW}$ is measured, it becomes a big challenge to determine if the observed stochastic GWs originate from PBHs or from astrophysical BH binaries [268, 269]. Furthermore, if PBHs exist, the real $\Omega_{GW}$ should be a superposition of that from PBHs and that from astrophysical BHs. In order to extract the PBH signal from $\Omega_{GW}$, it is indispensable to reduce the theoretical uncertainties about the extent to which astrophysical BHs produce $\Omega_{GW}$. More studies are needed to figure out how far we can go.

Finally, we have to mention that $\Omega_{GW}$-curves in the PBH scenarios in figure 15 contain contributions from merger events that can be identified as a single event. How the individual events and the residual are decomposed depends on the detector’s sensitivity and the implementation of data analysis. Thus, we have to apply a similar operation to the predicted $\Omega_{GW}$ when comparing the PBH scenario with the real data.

4.4.2. Cosmic evolution of the merger rate. The second observable that can be potentially used to distinguish the PBH scenario from the astrophysical ones is the time evolution of the BH–BH merger rate. While PBHs exist from almost the beginning of the Universe, BHs resulting from the death of stars appear in the low-redshift Universe. Thus, the redshift evolution of the number density of PBHs differs from that of astrophysical BHs. From this simple fact, we expect that the redshift dependence of the merger rate of the BH–BH binaries should also exhibit a difference between the two scenarios.

Figure 20 shows the redshift evolution of the merger rate per unit source time and unit comoving volume ($\text{Gpc}^{-3}\text{yr}^{-1}$) for different scenarios. The black curve represents the merger rate of astrophysical BHs in the fiducial model mentioned in section 4.4.1. The orange band is the uncertainty of the model. These curves are taken from [269]. The merger rate has a peak in $z = 1–2$. This is caused by the assumption that the formation rate of BH binaries is proportional to the formation rate of stars below half solar metallicity [263]. Although there

Figure 20. Redshift dependence of the merger rate of BH–BH binaries per unit source time and unit comoving volume. Unit is $\text{Gpc}^{-3}\text{yr}^{-1}$. Black and orange curves, taken from [269], show the uncertainties of the expected merger rate from astrophysical BHs. Red curve is the merger rate from the PBH binaries formed in the radiation-dominated epoch with $f_{\text{PBH}} = 10^{-3}$. Blue curves, taken from [269], are the merger rate from PBH binaries formed in the low-redshift Universe.
is a large uncertainty, there is a clear tendency that the merger rate drops sharply beyond the redshift 1–2.

The red curve is the merger rate equation (153) of the PBH binaries formed in the radiation-dominated epoch for \( f_{\text{PBH}} = 10^{-3} \) and \( M_{\text{PBH}} = 30 \, M_\odot \). Interestingly, the merger rate increases continuously for higher redshift, even beyond \( z = 1 \sim 2 \). For the chosen parameters, the merger rate is comparable to the merger rate of stellar BHs at low redshift, but exceeds the astrophysical prediction at higher redshift, \( z \gtrsim 5 \). Thus, the increasing feature of the merger rate at higher redshift is the smoking gun of the PBH scenario, and searching for such events definitely constitutes one of the routes we have to tread in the future [270]. For the sake of completeness, we also show the merger rate of PBH binaries formed in the low redshift Universe for \( f_{\text{PBH}} = 1 \) and the chirp mass \( 30 M_\odot \) as blue curves, taken from [269]. The Press–Schechter formalism is used to obtain the halo mass function, and the thick and dashed curves correspond to different choices of the halo concentration models (for more details, see [269]). We find that, similarly to the PBH binaries formed in the early Universe, the merger rate in this case also increases monotonically for higher redshift. However, the predicted merger rate is significantly lower than that of the stellar origin up to \( z \sim 6 \)—and than the red curve at any redshift for the chosen parameter.

Detecting merger events at high redshift \( z \gg 1 \) is challenging. Figure 21 shows the maximum redshift that the advanced LIGO with the design sensitivity can detect merger events of binaries with signal-to-noise ratio 8 as a function of the (source-frame) chirp mass, which is constructed from [271]. We find that the advanced LIGO at the design sensitivity can detect merger events of BH–BH binaries up to \( z \sim 1.5 \) for \( M_c = 30 M_\odot \). This shows that the LIGO may not be powerful enough to distinguish the PBH scenario from the astrophysical scenarios in the context of looking into the high redshift merger events.

Beyond LIGO, there are several proposed GW detectors both on ground and in space. These include the Einstein Telescope, Cosmic Explorer, eLISA and DECIGO. These detectors will be able to detect the GWs coming from much more distant places than LIGO can. For instance, it was shown in [270] that pre-DECIGO (DECihertz laser Interferometer Gravitational wave Observatory), which consists of three spacecraft cruising around the Earth in a triangle of side length 100 km, and is planned to be launched in the late 2020s, can detect merger events of \( 30 M_\odot \) BHs up to \( z \sim 10 \) with a signal-to-noise ratio of 8. At such high redshift, merger events of astrophysical origin are rare, and we can perform a clear test of the PBH scenario.
4.4.3. Mass distribution. Masses of the individual BHs before the merger are 
\((m_1, m_2) = (35, 30)\) for GW150914, 
\((m_1, m_2) = (14, 8)\) for GW151226, 
\((m_1, m_2) = (31, 19)\) for GW170104, 
\((m_1, m_2) = (12, 7)\) for GW170608 and 
\((m_1, m_2) = (30, 25)\) for GW170814 
(in units of solar mass). Obviously, there is some spread in the mass distribution. It is natural 
to think that the event rate distribution in the 2D mass plane should reflect to a certain degree 
the formation mechanism of the BH binaries, and its statistical nature can be used to discrimi-
nate different formation scenarios. Although merger events that have been discovered are cur-
cently countable by hand, it is almost certain that many more merger events will be detected 
in the coming decades. In this era, we should have plenty of information about the statistical 
properties of the merger event distribution in the 2D plane \((m_1, m_2)\). It is the purpose of this 
subsection to review some studies that aim to test the PBH scenario by using the event rate 
distribution that will become available in the future.

Generically speaking, if PBHs exist in the Universe, the merger events of BH binaries we 
observe will be a mixture of PBH binaries and those formed by astrophysical mechanisms. As 
we have seen in the previous sections, because of our ignorance of the PBH abundance and 
mass function and the large uncertainties about the astrophysical processes forming the BH 
binary, there is no definite consensus on which one (primordial or astrophysical) dominates 
the merger events. It is important to keep in mind that the following discussions are based on 
certain simplified assumptions reflecting the aforementioned unknown factors, and their con-
clusions may need to be modified as we become able to reduce the uncertainties in the future.

In [272], considering a possibility that some BH merger events involve PBHs, the case was 
investigated in which the PBH mass function has a sharp spike around a certain mass \(M_0\) with 
\(f_{\text{PBH}} = 1\), and the PBH binaries are formed in the low redshift Universe. The resultant merger 
rates, obtained as a function of the heavier BHs in binaries, was superposed on that predicted 
in the fiducial astrophysical model. It was found that the merger-event histogram exhibits a 
visible bump at around \(M_0\) on top of the smooth astrophysical background if the accumulated 
number of events reaches a few thousand, which is feasible by the advanced LIGO with the 
design sensitivity. According to the analysis in [273], the above PBH scenario with \(f_{\text{PBH}} > 0.5\) 
will be excluded at 99.9 % confidence level by the advanced LIGO. It is also shown that use of 
the data in the two-dimensional (2D) mass plane is more powerful than using only the heavier 
mass in constraining \(f_{\text{PBH}}\).

In [274], the case was investigated where PBHs with extended mass function form binaries 
in the radiation-dominated era, aiming at finding the unique character of the PBH scenario in 
the merger-event distribution. For simplicity, the astrophysical contribution was not included 
in the merger rate. Furthermore, PBHs are assumed to be distributed randomly in space, with 
no correlation between given mass ranges. As explained in section 4.1.1, the major axis and eccentricity of the PBH binary are determined by the gravitational force between the two BHs 
forming the binary, and torque exerted by the outer BHs not involved in the binary, respec-
tively. Denoting by \(m_1\) and \(m_2\) the mass of the individual PBHs that form the binary and by 
\(m_t = m_1 + m_2\) the total mass, \(a\) and \(e\) are given by [274]

\[
a = \frac{1}{1 + z_{\text{eq}}} \frac{\rho_c a \Omega_m}{m_t} x^4, \\
1 - e^2 = \frac{9}{4} \zeta^2, \quad \zeta = |\zeta_i|, \quad \zeta_i = \sum_{i=1}^{N} \frac{x^3 M_i}{m_t} \sin(2\theta_i) \left( \frac{\vec{e}_z \times \vec{e}_i}{|\vec{e}_z \times \vec{e}_i|} \right).
\] (160)

Here, contrary to the analysis in section 4.1.1, not only the closest outer BH but also other 
more distant BHs have been included, to evaluate the eccentricity, for completeness. The upper limit \(N\), which can in practice be taken to be infinity, is the number of PBHs contained
in the Hubble horizon at the binary formation time. Furthermore, $y_i, M_i, \theta_i, \vec{e}_i$ are distance to, mass of, direction to, and angle measured from the major axis of the $i$th outer BH. Then, the merger probability of the PBH binary consisting of BHs with mass $m_1$ and $m_2$ at cosmic time $t$ becomes

$$R_{\text{int}}(m_1, m_2, t) = \int_0^{\infty} \frac{dx}{x} F(x, \zeta(x)) \frac{d\zeta}{dx} \frac{\partial a}{\partial x} \frac{\partial^2 a}{\partial \zeta \partial x},$$  \hspace{1cm} (161)$$

where $\frac{d\zeta}{dx}$ and $\frac{\partial a}{\partial x}$ can be computed from equations (160) and $\frac{\partial a}{\partial \zeta}$ from equation (148). The function $F(x, \zeta)$, which gives the probability density of $x$ and $\zeta$, is formally written as

$$F(x, \zeta) = \Theta(a_{\text{max}} - a) \frac{4\pi x^2(a)}{n_{\text{BH}}} \int \lim_{N \to \infty} \prod_{i=1}^{N} \frac{dV_i}{n_{\text{BH}}} \frac{\sin \theta_i d\theta_i d\phi_i}{4\pi} \Theta(\zeta - \sigma_i)$$

$$\times e^{-\sqrt{2 \pi} m_{\text{BH}} x} \delta(\zeta - g(x, y_i, M_i, \theta_i, \phi_i)).$$  \hspace{1cm} (162)$$

where $\Theta(\cdot)$ is the Heaviside step function, $\delta(\cdot)$ is the Dirac’s delta function, and $g$ is given by

$$g(x, y_i, M_i, \theta_i, \phi_i) \equiv \left| \sum_{i=1}^{N} x^i M_i \frac{\sin(2\theta_i)}{y_i} \frac{e_{\zeta}}{|e_{\zeta} \times e_{\bar{e}}|} \right|.$$ 

Finally, the observable merger rate $\mathcal{R}$ per unit time and unit comoving volume is given by

$$\mathcal{R}(m_1, m_2, t) = \frac{1}{2n_{\text{BH}}} R_{\text{int}}(m_1, m_2, t) f(m_1) f(m_2),$$  \hspace{1cm} (164)$$

where $f(m)$ is the PBH mass function.

In [274], since exact computation of the integral (162) for arbitrary shape of the mass function is impossible, $R_{\text{int}}$ was computed under two different approximations. The first is to include only the closest outer BH ($N = 1$), just as we discussed in section 4.1.1, without specifying the specific shape of $f(m)$. The second is to consider the case with $N \gg 1$. For this case, the analytic form of $F(x, \zeta)$ given by

$$F(x, \zeta) = 6\sqrt{3} \gamma^{1/3} n_{\text{BH}} \sigma^2 \zeta^2 \left( \frac{m_i}{m_{\text{max}}} \right)^2 \left[ \frac{m_i}{m_{\text{max}}} \right]^3 \left( \frac{\zeta^3 + \gamma \hat{\sigma}^6}{\zeta^3 + \gamma \hat{\sigma}^6} \right)^{-1},$$  \hspace{1cm} (165)$$

was adopted. This function was confirmed to correctly reproduce the numerically computed $F(x, \zeta)$ by the Monte Carlo calculations for the flat PBH mass function. Quite interestingly, in both cases, it was found that the merger rate $\mathcal{R}$ depends on the total mass $m_i = m_1 + m_2$ in a specific way almost independent of the mass function. In order to extract this information, a dimensionless quantity $\alpha$ defined by

$$\alpha(m_1, m_2, t) = -m_i^2 \frac{\partial^2}{\partial m_1 \partial m_2} \ln \mathcal{R}(m_1, m_2, t),$$  \hspace{1cm} (166)$$

was introduced. Both cases considered above predict this quantity to be

$$\frac{36}{37} \leq \alpha(m_1, m_2, t) \leq \frac{22}{21},$$  \hspace{1cm} (167)$$

which sharply concentrates around unity. Notice that $\mathcal{R}$ itself strongly depends on the mass function like $\propto f(m_1) f(m_2) m_i^n$, where $f(m)$ is related to $f(m)$ in a non-trivial way, but such part decouples from $\mathcal{R}$ by taking its logarithm and differentiation twice.
The result (167) shows that measuring the quantity $\alpha$ by observation is a powerful method to test the PBH scenario, independently of the PBH mass function. The raw merger rate distribution in the $m_1 - m_2$ plane directly determined by observations will not coincide with $\mathcal{R}$ computed above, since the former is affected by the detector bias. Thus, extracting the correct value of $\alpha$ from observations is a challenging task. When statistically enough merger events have been accumulated in the $m_1 - m_2$ plane in the future, it should in principle be possible to compare the observation with the prediction (167). How to achieve this is an important topic, but we will not discuss it here.

The quantity $\alpha$ can be used not only to test the PBH scenario, but also to discriminate different scenarios of the BH binary formation. For instance, the quantity $\alpha$ for the PBH binaries that formed in the low redshift Universe was also derived and found to take a unique value given by

$$\alpha(m_1, m_2, t) = \frac{10}{7},$$

which is different from (167) predicted by PBH binaries that formed in the radiation-dominated era. The dynamical formation scenario, which is one of the strong astrophysical candidates to form BH binaries, predicts $\alpha \approx 4$ [275]. As these examples vividly demonstrate, different scenarios predict different values of $\alpha$. When the observational data reaches a sensitivity $\Delta \alpha \lesssim 0.1$, $\alpha$ will be used as a powerful discriminator.

4.4.4. Spin distribution. The distribution of BH spin is also useful to constrain the formation scenarios of the BH–BH binaries. Magnitude of the BH spin is commonly expressed in terms of a dimensionless quantity $\chi$ defined by

$$\chi = \frac{|\vec{S}|}{GM_{\text{BH}}},$$

where $\vec{S}$ is the BH spin. The physical requirement that no naked singularity appears restricts the range of $\chi$ as $0 \leq \chi \leq 1$, and BHs with $\chi = 1$ correspond to the maximally rotating BHs.

Spins of the individual BHs in the binary affects the GW waveform primarily through a particular combination given by

$$\chi_{\text{eff}} = \frac{m_1 \chi_1 \cos \theta_1 + m_2 \chi_2 \cos \theta_2}{m_t},$$

where $\theta_i$ ($i = 1, 2$) is the angle between the spin of the $i$th BH and the orbital angular momentum. Thus, the allowed range of $\chi_{\text{eff}}$ is $-1 \leq \chi_{\text{eff}} \leq 1$. At the time of writing this article, five merger events have been robustly detected. The measured values of $\chi_{\text{eff}}$ are

$$\chi_{\text{eff}} = -0.06^{+0.14}_{-0.14}, 0.21^{+0.20}_{-0.10}, -0.12^{+0.21}_{-0.30}, 0.07^{+0.23}_{-0.09}, 0.06^{+0.12}_{-0.12},$$

for GW150914, GW151226, GW170104, GW170608 and GW170814 respectively (with 90% credible intervals). Detection of many more merger events in the future will allow us to construct the distribution of $\chi_{\text{eff}}$.

There are several astrophysical scenarios for the origin of BH–BH binaries. For instance, the isolated field binary scenario, in which stars are formed as a binary and later individual stars collapse to BHs (heavier one first), BH spins are likely to be aligned with the orbital angular momentum, and $\chi_{\text{eff}} > 0$ is a natural consequence in this scenario [24]. Since there has

[^24]: However, we have to keep in mind that the natal kick velocity of a BH due to the supernova explosion can alter this conclusion.
been no robust prediction of the magnitude of $\chi_{\text{eff}}$ yet, low values of $\chi_{\text{eff}}$ of the detected events (171) are used to constrain the properties of the field binaries [276–279]. Tidal lock of the progenitor of the lighter BH from the heavier BH will result in high spin of the lighter BH. Thus the progenitor of the lighter BH must be sufficiently compact, which favors Wolfe–Rayet stars as the progenitors of the lighter BHs [277]. Dynamical formation scenario, in which BH binaries are formed by dynamical interactions among BHs in dense stellar environments such as globular clusters, predicts isotropic distribution of the individual BH spins. Thus, positive and negative $\chi_{\text{eff}}$ are equally probable in this scenario.

Despite the small number of detections of merger events, the comparison between the alignment hypothesis of the BH spins and the isotropic one has already been made against the observed values of $\chi_{\text{eff}}$ in [280]. Considering some simple distribution functions of spins extending to high values, the odds ratio against the alignment hypothesis compared to the isotropic one was found to be 0.015, which corresponds to about $2.4\sigma$. Furthermore, among different models of the isotropic distribution, the most favored one has low spin magnitude distribution, although the statistical significance is not as strong as the above odds ratio. However, see [281], in which slight preference for the alignment hypothesis was suggested. More observational data is definitely needed to establish a robust conclusion about this issue.

As the discussions in the previous sections demonstrate, the PBH binaries, whether they are formed in the radiation-dominated era or in the low redshift Universe, have isotropic spin distribution. What remains an open issue is the distribution of spin magnitude $a$ of individual PBHs. In [282], the PBH spin distribution has been derived under several assumptions in the following way. In the direct collapse scenario of the PBH formation in the radiation-dominated era, in which a nearly spherical overdense region undergoes gravitational collapse upon horizon re-entry, PBHs form when the density contrast exceeds a threshold $\delta_{\text{th}}$. For PBHs that formed out of the density contrast slightly above the threshold, which will be the case in the realistic scenario in which the probability of realizing much higher density contrast is significantly suppressed, the mass and spin may have simple dependence on $\delta_{\text{th}}$ as

$$M_{\text{PBH}} \approx C_M|\delta - \delta_{\text{th}}(q)|^{\gamma M}, \quad S_{\text{PBH}} \approx C_S|\delta - \delta_{\text{th}}(q)|^{\gamma q},$$

$$\delta_{\text{th}}(q) = \delta_{\text{th},0} + Kq^2,$$  \hspace{1cm} (172)

where $q$ is a parameter that characterizes the amount of the PBH spin (see [283]), and $C_M$, $C_S$, $\gamma_M$, $\gamma_q$, $\delta_{\text{th},0}$ and $K$ are all constants. In particular, $K(\approx 5.7 \times 10^{-3})$ is positive. This is natural, since stronger gravity—thus, larger amplitude of the density contrast—is needed to defeat the centrifugal force and to form spinning PBHs.

These scalings were found numerically in asymptotically flat spacetime [283]. In [282], these relations were adopted. Formally, one can write the above relations as $M_{\text{PBH}} = M_{\text{PBH}}(\delta, q)$ and $S_{\text{PBH}} = S_{\text{PBH}}(\delta, q)$. Then, denoting the probability density of $(\delta, q)$ as $P(\delta, q)$, we can convert the probability in the $\delta - q$ plane to that in the $M_{\text{PBH}} - \chi$ plane by

$$dP = P(\delta, q)\delta\delta dq = F(M_{\text{PBH}}, \chi)dM_{\text{PBH}}d\chi,$$  \hspace{1cm} (173)

where $\chi = S_{\text{PBH}}/(GM_{\text{PBH}}^2)$. In [282], it was assumed that there is no correlation between $\delta$ and $q$ and that the probability densities for $\delta$ and $q$ are Gaussian and flat respectively. This latter assumption may be an oversimplification, given the vectorial nature of the angular momentum in three dimensional space. In a more realistic situation, the probability density at small $q$ would be suppressed as $\sim q^2$, and the flatness assumption could be understood as a limiting case in which the initial overdensity can easily have a large angular momentum.

Integration of the above probability over $M_{\text{PBH}}$ was then found to be Gaussian:
Thus, PBHs with higher spins ($\chi \approx 1$) are less likely to be realized than those with lower spins, which can be traced back to the fact that the threshold amplitude of the PBH formation increases as we increase the PBH spin. The value of $\chi_*$ depends on $\delta_{th,0}$ and variance $\sigma^2$ of the density perturbation. In the above equation, $\delta_{th,0} = 1/3$, $\sigma = 0.15 \delta_{th,0}$ were adopted as fiducial values.

Figure 22 shows the PBH spin distribution given by equation (174). As mentioned above, the distribution at low $\chi$ could be significantly modified in a more realistic situation. Nevertheless, it seems robust to conclude that rapidly spinning PBHs ($\chi \approx 1$) are very unlikely. If spin distribution turns out to have a peak at a high value $\chi \approx 1$ in the future, the PBH scenario as the origin of the BH–BH binaries will then be strongly disfavored.

However, we have to keep in mind that even the observational confirmation of the spin distribution (174) does not necessarily prove the PBH scenario. The above result has been derived under a couple of non-trivial assumptions. Neither the Gaussian shape nor the value of $\chi_*$ should be taken seriously as unique predictors of the PBH scenario. For instance, the flatness assumption on $q$ may overestimate the probability for higher values of $q$, and the typical spin magnitude $\chi_*$ could be much smaller than the value in equation (174). Obviously, more studies are needed to clarify the spin distribution of PBHs. At this moment, a general lesson we can learn from this result is that PBHs with lower spin are more favored than those with higher spin, which is reasonable from a physical point of view.

4.4.5. Cross-correlation with galaxies. Cross-correlation between the spatial distribution of BH–BH merger events and that of galaxies offers another method to test the PBH scenario. This method will become powerful when a sufficiently large number of merger events has accumulated to discuss the statistical distribution in space ($\gtrsim 100$).

On large scales, much longer than the typical distance between galaxies ($\sim$Mpc), we can define the number density of galaxies and its density contrast $\delta_g(\vec{x})$. Since galaxies form by the gravitational collapse of baryonic gas where the dark matter density is higher than average,

$$dP \propto \exp \left[ - \left( \frac{\chi}{\chi_*} \right)^2 \right] d\chi, \quad \chi_* \approx 0.46.$$ (174)
there are more galaxies in sites where the dark matter density contrast is higher. It is known that $\delta_g(x)$ does not coincide with the density contrast of matter $\delta_m(x)$, due to the fact that the galaxies are formed at the high density peaks [285]. The mismatch is commonly represented as a bias parameter $b_g$ by

$$\delta_g = b_g \delta_m.$$  \hspace{1cm} (175)

The bias parameter can evolve on cosmological time scales. The observations of the galaxy clustering and weak lensing suggest $b_g = 1.3 - 1.7$ for $z \lesssim 0.5$ and increase of $b_g$ for higher redshift [286].

The fact that the bias parameter $b_g$ is different from unity is good news for the purpose of testing the origin of the BH binaries. The reason for this is simple. If the BH binaries are of astrophysical origin, then $\delta_{BH}$ the number density contrast of the merger events of the BH binaries, should coincide with $\delta_g$. If, on the other hand, BH binaries consist of PBHs, then $\delta_{BH}$ should trace $\delta_m$ unless the BH binary evolution is not significantly affected by the interaction with baryons. When taken together, PBHs and astrophysical BHs predict different magnitudes of cross-correlation between $\delta_{BH}$ and $\delta_g$.

The idea of taking the cross-correlation between $\delta_{BH}$ and $\delta_g$ (as well as weak lensing) to test the spatial clustering of BH binaries was proposed and investigated in [287]. Similar methodology was used for the purpose of testing the PBH scenario in [288], in which the close encounter mechanism for binary formation (see section 4.1.2) was assumed. According to the notation in [288], the cross-correlation between $\delta_{BH}$ and $\delta_g$ in multi-pole decomposition is given by

$$C_{\ell}^{BH,g} = r \int \frac{4\pi dk}{k} \Delta^2(k) W_{\ell}^{BH}(k) W_{\ell}^{g}(k),$$  \hspace{1cm} (176)

where $r$ is the cross-correlation coefficient, which represents how well galaxy distribution actually traces the matter distribution, $\Delta^2$ is the dimensionless power spectrum, and $W_{\ell}^{BH}$ and $W_{\ell}^{g}$ are defined by

$$W_{\ell}^{X} = \int \frac{dN_X}{dz} b_X(z) j_\ell(k\chi(z))dz,$$  \hspace{1cm} (177)

where $X = \{PBH, g\}$, $dN_X/dz$ is the source redshift distribution, $b_X$ is the bias parameter, and $\chi(z)$ is the comoving distance. In [288], constant redshift distribution of galaxies and constant galaxy bias $b_g = 1.4$ were assumed, and the redshift distribution of the PBH binaries given in [12] and the constant BH bias $b_{PBH} = 0.5$ were adopted. The reason why $b_{PBH}$ is smaller than unity is that BH merger events in the scenario in [12] predominantly occur inside the small halos whose spatial distribution are more spread than the heavier halos that host galaxies.

The theoretical prediction of the cross-correlation (176) is the one that can be compared with observational data in the future. The right-hand side of equation (176) is proportional to $b_g$ if the BH binaries are formed by the astrophysical processes, and to $b_{PBH}$ if the BH binaries are PBHs. Thus, the target sensitivity of the future observations to discriminate the PBH scenario from the astrophysical ones is $\delta b \lesssim b_{PBH} - b_{BH} = 0.9$. What then matters is whether future observations are powerful enough to achieve this sensitivity. According to the analysis in [288], such achievement is possible by the Einstein Telescope, and could even be possible by the long-term observations by LIGO, depending on the BH merger rate (see also [289]).

4.4.6. Eccentricity. The eccentricity of the binary orbit also has the potential to discriminate the PBH scenario from the astrophysical scenarios.
First of all, circular and eccentric binaries emit GWs whose waveforms look different. Figure 23 shows two GW waveforms from BH binaries consisting of the equal mass \( m_{\text{BH}} = 30 M_\odot \) with different eccentricities \( e = 0 \) and \( e = 0.5 \) \[290\]. Amplitude of the waveforms are normalized. We find that the waveform for the eccentric orbit is modulated, and shows two phases: gentle hill and tall spike. The spike originates at the periastron, where the acceleration of BHs as well as velocities are higher. This figure demonstrates that it is in principle possible to know the eccentricity of a binary from measurement of the time dependence of the waveform.

While the individual binary formation scenarios predict (in principle) the distribution of the eccentricity at the binary formation time, GW observations do not measure the initial eccentricities, but those in the inspiral phase corresponding to the frequency band to which the detectors are sensitive. This fact is important, since the eccentricity changes as the binary shrinks with the GW emission. According to Peters \[252\], the orbital eccentricity changes due to the gravitational radiation reaction as

\[
\frac{de}{dt} = - \frac{304}{15} \frac{G^3 m_1 m_2 (m_1 + m_2)}{a^4 (1 - e^2)^{5/2}} \left( \frac{121}{304} e^2 \right).
\]

Clearly, \( \frac{de}{dt} \leq 0 \) for any \( e \). Thus, the eccentric orbit is always circularized by the GW emission.

Generally speaking, binaries born with larger separation have smaller eccentricities in the final inspiral phase than those with initially shorter separation, since the former have more time to circularize the orbit. For stellar-mass BH binaries, the frequency band around 100 Hz to which ground-based detectors such as LIGO are the most sensitive corresponds to the last several revolutions before the merger. BH binaries from isolated field binaries, and those dynamically formed in the globular clusters, are expected to have negligible eccentricities in the LIGO band \[291, 292\]. On the other hand, BH binaries formed in the vicinity of the supermassive BHs in galactic nuclei are mostly highly eccentric \( (e \approx 1) \) even at the LIGO band \[293\].

In \[253\], the eccentricity distribution of PBH binaries formed in the low redshift Universe in the LIGO frequency band was investigated. A point of note is that since those binaries are formed with high eccentricity and have much shorter lifetime than the Hubble time, a fraction of them retain some eccentricity even when they enter the LIGO frequency band. According
to the analysis in [253], $O(1)$ merger events with non-zero eccentricity are expected to be
detected over a few years’ observation by LIGO and $O(10)$ events over ten years’ observation
by the Einstein Telescope. On the other hand, the eccentricity distribution for the LIGO band
for the PBH binaries formed in the radiation-dominated era is expected to have a strong spike
at $e = 0$, since their lifetime is about the Hubble time.

Space-based laser interferometers such as eLISA and DECIGO are sensitive to lower fre-
cquency bands than ground-based detectors. They can observe stellar-mass BH binaries in
the inspiral phase long before they merge. The orbits in such low-frequency phases are more
eccentric than in the nearly merging phase, if binaries initially have significant eccentricities.
The analysis in [294, 295] shows that several years’ observations by eLISA have the potential
to distinguish between the field and cluster formation scenarios. However, there is no similar
study for the distinguishability of the PBH scenario.

5. Summary

“I think over the coming decades we will see enormous numbers of things. Just as
electromagnetic astronomy was begun in essence, at least modern astronomy, by Galileo
pointing his telescope in the sky and discovering Jupiter’s moons. This is the same thing
but for gravitational waves …”  
Kip S. Thorne [296]

LIGO’s first observation of GWs finally opened an era of GW astronomy. Direct observations
of the BHs vividly demonstrated that GWs bring us information about the Universe which
can never be obtained from observations of electromagnetic signals. PBHs are no exception
regarding this point. Since the original proposal around 1970, electromagnetic searches for
the PBHs have been performed over decades. Until now, none of these searches have detected
solid evidence for the existence of PBHs, and tight constraints on PBHs with various masses
have been obtained. The situation has been drastically changed by the LIGO event, which
suddenly raised an interesting possibility; LIGO might have detected PBHs for the first time!
The purpose of this article is to deliver recent proposals of the PBHs as the source of the LIGO
events, and give a review of various ideas to test for PBHs using future GW observations,
simultaneously covering the basics of PBH formation, as well as the existing constraints on
non-evaporating PBHs from electromagnetic observations.

After having introduced the basics of PBH formation and the relevant inflation models
in section 2, we reviewed various existing constraints on the PBH abundance of the non-
evaporating PBHs in section 3, as well as the constraints that will be imposed by future
observations. For the sake of completeness, we addressed not only stellar-size PBHs but
also PBHs in wider mass range from $\sim 10^{-16} M_\odot$ to $\sim 10^{10} M_\odot$. In section 4, we reviewed
the PBH scenario as an explanation of the LIGO events and various proposals to test it with
future GW observations.

GW astronomy has just begun. As it progresses, it will continually bring us new findings,
and also stimulate related theoretical studies. Definitely, we will gain more knowledge about
PBHs and the early Universe. PBHs are dark, but the future of PBH research is bright.

Acknowledgments

It is a pleasure to thank Yacine Ali-Haimoud, Christian Byrnes, Bernard Carr, Anne Green,
Tomohiro Harada, Kazunori Kohri, and Masamune Oguri for a careful reading of the manuscript
and their enlightening comments. This work was supported in part by JSPS KEKENHI Grant

Class. Quantum Grav. 35 (2018) 063001
Numbers JP17H06358 (TT), JP17H06359 (TS), JP15H05888 (MS, TS and SY), JP15H02087 (TT), JP16H01103 (SY) JSPS Grant-in-Aid for Young Scientists (B) JP15K17632 (TS) and JP15K17659 (SY), and the Grant-in-Aid for Scientific Research JP26287044 (TT).

Permission for reuse

The following are permissions for reusing the figures/data in the literature to produce some figures in this article. We thank all the persons concerned for having kindly allowed us to reuse their materials.

1. Figure 10

[120]: Tisserand et al, reproduced with permission © ESO.
[123]: Reproduced with permission from Wyrzykowski et al The OGLE view of microlensing towards the Magellanic Clouds IV. OGLE-III SMC data and final conclusions on MACHOs. Published by Oxford University Press on behalf of The Royal Astronomical Society. All rights reserved. The figure is not covered by the Open-Access licence of this publication. For permissions contact Journals.Permissions@OUP.com
[126]: Griest et al, reproduced with permission from the author and © AAS.
[157]: Reproduced with permission from Quinn et al. On the reported death of the MACHO era. Published by Oxford University Press on behalf of The Royal Astronomical Society. All rights reserved. The figure is not covered by the Open-Access licence of this publication. For permissions contact Journals.Permissions@OUP.com
[159]: Brandt, reproduced with permission from the author and © AAS.
[127, 129]: Reproduced with permission from the authors.
[136, 142, 145, 146, 172, 177, 179]: Reproduced with permission from the authors and © APS.
[186]: Reproduced with permission © IOP.

2. Figure 17

[250]: Reproduced with permission from the author and © APS.

3. Figure 19

[263, 269]: Reproduced with permission from the authors and © APS.
[256]: Reproduced with permission © IOP.

4. Figure 20

[269]: Reproduced with permission from the author and © APS.

5. Figure 21

[271]: Reproduced with permission from the author and © APS.

ORCID iDs

Misao Sasaki https://orcid.org/0000-0001-5924-0664

References

[1] Zel’’dovich Y B and Novikov I D 1967 Sov. Astron. 10 602
[2] Hawking S 1971 Mon. Not. R. Astron. Soc. 152 75
[3] Rhoades C E Jr and Ruffini R 1974 Phys. Rev. Lett. 32 324
[4] Carr B J 2005 Primordial black holes: Do they exist and are they useful?, 59th Yamada Conf. on Inflating Horizon of Particle Astrophysics and Cosmology (Tokyo Japan, 20–24 June 2005)
[5] Page D N 1976 Phys. Rev. D 13 198
[6] Miyama S and Sato K 1978 Prog. Theor. Phys. 59 1012
[7] Carr B J, Kohri K, Sendouda Y and Yokoyama J 2010 Phys. Rev. D 81 104019
[8] Carr B, Kühnel F and Sandstad M 2016 Phys. Rev. D 94 083504
[9] Carr B, Raidal M, Tenkanen T, Vaskonen V and Veermäe H 2017 Phys. Rev. D 96 023514
[10] Abbott B P et al (LIGO Scientific and Virgo Collaborations) 2016 Phys. Rev. Lett. 116 061102
[11] Abbott B P et al (LIGO Scientific and Virgo Collaborations) 2016 Astrophys. J. 818 L22
[12] Bird S et al 2016 Phys. Rev. Lett. 116 201301
[13] Clesse S and Garca-Bellido J 2017 Phys. Dark Univ. 15 142
[14] Sasaki M, Suyama T, Tanaka T and Yokoyama S 2016 Phys. Rev. Lett. 117 061101
[15] Nakamura T, Sasaki M, Tanaka T and Thorne K S 1997 Astrophys. J. 487 L139
[16] Abbott B P et al (LIGO Scientific and Virgo Collaborations) 2016 Phys. Rev. Lett. 116 241103
[17] Abbott B P et al (VirGO, LIGO Scientific) 2017 Phys. Rev. Lett. 118 221101
[18] Abbott B P et al (LIGO Scientific and Virgo Collaborations) 2017 Phys. Rev. Lett. 119 141101
[19] Abbott B P et al (VirGO, LIGO Scientific) 2017 Astrophys. J. 851 L35
[20] Punturo M et al 2010 Class. Quantum Grav. 27 194002
[21] Abbott B P et al (LIGO Scientific) 2017 Class. Quantum Grav. 34 044001
[22] Seoane P A et al (eLISA) 2013 arXiv:1305.5720
[23] Seto N, Kawamura S and Nakamura T 2001 Phys. Rev. Lett. 87 221103
[24] Garriga J, Vilenkin A and Zhang J 2016 J. Cosmol. Astropart. Phys. JCAP02(2016)064
[25] Deng H, Garriga J and Vilenkin A 2017 J. Cosmol. Astropart. Phys. JCAP04(2017)050
[26] Deng H and Vilenkin A 2017 J. Cosmol. Astropart. Phys. JCAP12(2017)044
[27] Hawking S W 1989 Phys. Lett. B 231 237
[28] Polnarev A and Zembowicz R 1991 Phys. Rev. D 43 1106
[29] Garriga J and Vilenkin A 1993 Phys. Rev. D 47 3265
[30] Niemeyer J C and Jedamzik K 1998 Phys. Rev. Lett. 80 5481
[31] Shibata M and Sasaki M 1999 Phys. Rev. D 60 084002
[32] Musco I, Miller J C and Rezzolla L 2005 Class. Quantum Grav. 22 1405
[33] Polnarev A G and Musco I 2007 Class. Quantum Grav. 24 1405
[34] Musco I, Miller J C and Polnarev A G 2009 Class. Quantum Grav. 26 235001
[35] Nakama T, Harada T, Polnarev A G and Yokoyama J 2014 J. Cosmol. Astropart. Phys. JCAP01(2014)037
[36] Harada T, Yoo C M and Kohri K 2013 Phys. Rev. D 88 084051
[37] Harada T, Yoo C M and Kohri K 2014 Phys. Rev. D 89 029903 (erratum)
[38] Carr B J 1975 Astrophys. J. 201 1
[39] Yokoyama J 1997 Astron. Astrophys. 318 673
[40] Young S, Byrnes C T and Sasaki M 2014 J. Cosmol. Astropart. Phys. JCAP07(2014)045
[41] Lyth D H, Malik K A and Sasaki M 2005 J. Cosmol. Astropart. Phys. JCAP05(2005)004
[42] Kopp M, Hofmann S and Weller J 2011 Phys. Rev. D 83 124025
[43] Harada T, Yoo C M, Nakama T and Koga Y 2015 Phys. Rev. D 91 084057
[44] Choptuik M W 1993 Phys. Rev. Lett. 70 9
[45] Evans C R and Coleman J S 1994 Phys. Rev. Lett. 72 1782
[46] Koike T, Hara T and Adachi S 1996 Phys. Rev. Lett. 77 5170
[47] Yokoyama J 1998 Phys. Rev. D 58 107502
[48] Kühnel F, Rampf C and Sandstad M 2016 Eur. Phys. J. C76 93
[49] Kühnel F and Sandstad M 2016 Phys. Rev. D 94 063514
[50] Doroshkevich A G 1970 Astrophys. J. 6 581
[51] Bardeen J M, Bond J R, Kaiser N and Szalay A S 1986 Astrophys. J. 304 15
[52] Khlopov M Yu and Polnarev A G 1980 Phys. Lett. 97B 383
[53] Harada T and Jhingan S 2016 PTEP 2016 093E04
[54] Harada T, Yoo C M, Kohri K, Nakao K I and Jhingan S 2016 Astrophys. J. 833 61
[55] Lyth D, Malik K A, Sasaki M and Zaballa I 2006 J. Cosmol. Astropart. Phys. JCAP01(2006)011
[56] Alabidi L, Kohri K, Sasaki M and Sendouda Y 2013 J. Cosmol. Astropart. Phys. JCAP05(2013)033
[57] Torres-Lomas E, Hidalgo J C, Malik K A and Urea-Lpez L A 2014 Phys. Rev. D 89 083008

71
[57] Tada Y and Yokoyama S 2015 Phys. Rev. D 91 123534
[58] Chisholm J R 2006 Phys. Rev. D 73 083504
[59] Clesse S and Garcia-Bellido J 2015 Phys. Rev. D 92 023524
[60] Komatsu E and Spergel D N 2001 Phys. Rev. D 63 063002
[61] Young S and Byrnes C T 2015 Phys. Rev. D 91 083521
[62] Akrami Y, Kuhnel F and Sandstad M 2016 Uncertainties in primordial black-hole constraints on the primordial power spectrum (arXiv:1611.10069)
[63] Bullock J S and Primack J R 1997 Phys. Rev. D 55 7423
[64] Ivanov P 1998 Phys. Rev. D 57 7145
[65] Kawasaki M 1998 Phys. Rev. D 58 083510
[66] Saito R, Kawasaki M and Nagata R 2008 J. Cosmol. Astropart. Phys. JCAP06(2008)024
[67] Young S and Byrnes C T 2013 J. Cosmol. Astropart. Phys. JCAP08(2013)052
[68] Young S, Regan D and Byrnes C T 2016 J. Cosmol. Astropart. Phys. JCAP02(2016)029
[69] Kawasaki M and Tada Y 2016 J. Cosmol. Astropart. Phys. JCAP08(2016)041
[70] Pattison C, Vennin V, Assadullahi H and Wands D 2017 J. Cosmol. Astropart. Phys. JCAP10(2017)046
[71] Kawasaki M, Kusenko A, Tada Y and Yanagida T T 2016 Phys. Rev. D 94 083523
[72] Kodama H, Sasaki M and Sato K 1982 Prog. Theor. Phys. 68 1979
[73] Dolgov A and Silk J 1993 Phys. Rev. D 47 4244
[74] Jedamzik K 1997 Phys. Rev. D 55 5871
[75] Jedamzik K and Niemeyer J C 1999 Phys. Rev. D 59 124014
[76] Khlopov M Yu, Konoplich R V, Rubin S G and Sakharov A S 1999 Grav. Cosmol. 2 S1
[77] Ade P A R et al (Planck) 2016 Astron. Astrophys. 594 A13
[78] Alabidi L, Kohri K, Sasaki M and Sendouda Y 2012 J. Cosmol. Astropart. Phys. JCAP09(2012)017
[79] Drees M and Erñani E 2011 J. Cosmol. Astropart. Phys. JCAP04(2011)005
[80] Drees M and Erñani E 2012 J. Cosmol. Astropart. Phys. JCAP01(2012)035
[81] Leach S M, Grivell I J and Liddle A R 2000 Phys. Rev. D 62 043516
[82] Kohri K, Lyth D H and Melchiorri A 2008 J. Cosmol. Astropart. Phys. JCAP04(2008)038
[83] Alabidi L and Kohri K 2009 Phys. Rev. D 80 063511
[84] Bugaev E and Klimai P 2009 Phys. Rev. D 79 103511
[85] Stewart E D 1997 Phys. Lett. B 391 34
[86] Stewart E D 1997 Phys. Rev. D 56 2019
[87] Ade P A R et al (Planck) 2016 Astron. Astrophys. 594 A20
[88] Garcia-Bellido J and Ruiz Morales E 2017 Phys. Dark Univ. 18 47
[89] Ivanov P, Naselsky P and Novikov I 1994 Phys. Rev. D 50 7173
[90] Ezquiaga J M, Garcia-Bellido J and Ruiz Morales E 2018 Phys. Lett. B 776 345
[91] Kannike K, Marzola L, Raidal M and Veermäe H 2017 J. Cosmol. Astropart. Phys. JCAP09(2017)020
[92] Germani C and Prokopec T 2017 Phys. Dark Univ. 18 6
[93] Motohashi H and Hu W 2017 Phys. Rev. D 96 063503
[94] Gong Y 2017 (arXiv:1707.09578)
[95] Ballesteros G and Taoso M 2018 Phys. Rev. D 97 023501
[96] Hertzberg M P and Yamada M 2017 (arXiv:1712.09750)
[97] Silk J and Turner M S 1987 Phys. Rev. D 35 419
[98] Randall L, Soljacic M and Guth A H 1996 Nucl. Phys. B 472 377
[99] Garcia-Bellido J, Linde A D and Wands D 1996 Phys. Rev. D 54 6040
[100] Tanaka T, Suyama T and Yokoyama S 2010 Class. Quantum Grav. 27 124003
[101] Kawasaki M, Sugiyama N and Yanagida T 1998 Phys. Rev. D 57 6050
[102] Kawasaki M, Takayama T, Yamaguchi M and Yokoyama J 2006 Phys. Rev. D 74 043525
[103] Inomata K, Kawasaki M, Mukaida K and Yanagida T T 2017 arXiv:1711.06129
[104] Inomata K, Kawasaki M, Mukaida K, Tada Y and Yanagida T T 2017 Phys. Rev. D 95 123510
[105] Kawasaki M, Kitajima N and Yanagida T T 2013 Phys. Rev. D 87 063519
[106] Ando K, Inomata K, Kawasaki M, Mukaida K and Yanagida T T 2017 arXiv:1711.08956
[107] Kohri K, Lin C M and Matsuda T 2013 Phys. Rev. D 87 103527
[108] Suyama T and Yokoyama J 2011 Phys. Rev. D 84 083511
[109] Suyama T, Wu Y P and Yokoyama J 2014 Phys. Rev. D 90 043514
[110] Garcia-Bellido J, Peloso M and Unal C 2016 J. Cosmol. Astropart. Phys. JCAP12(2016)031
[111] Josan A S, Green A M and Malik K A 2009 Phys. Rev. D 79 103520
[164] Xu G H and Ostriker J P 1994 Astrophys. J. 427 184
[165] Lacey C G and Ostriker J P 1985 Astrophys. J. 299 633
[166] Nordström B, Mayor M, Andersen J, Holmberg J, Pont F, Jørgensen B R, Olsen E H, Udry S and Mowlavi N 2004 Astron. Astrophys. 418 899
[167] Lacey C G 1984 Mon. Not. R. Astron. Soc. 208 687
[168] Binney J, Dehnen W and Bertelli G 2000 Mon. Not. R. Astron. Soc. 318 658
[169] Carr B J 1981 Mon. Not. R. Astron. Soc. 194 639
[170] Ricotti M, Ostriker J P and Mack K J 2008 Astrophys. J. 680 829
[171] Aloni D, Blum K and Flauger R 2017 J. Cosmol. Astropart. Phys. JCAP05(2017)017
[172] Ali-Haimoud Y and Kamionkowski M 2017 Phys. Rev. D 95 043534
[173] Chen X L and Kamionkowski M 2004 Phys. Rev. D 70 043502
[174] Ali-Haimoud Y and Hirata J M 2011 Phys. Rev. D 83 043513
[175] Blas D, Lesgourgues J and Tram T 2011 J. Cosmol. Astropart. Phys. JCAP07(2011)034
[176] Aghanim N et al (Planck) 2016 Astron. Astrophys. 594 A11
[177] Poulin V, Serpico P D, Calore F, Clésse S and Kohri K 2017 Phys. Rev. D 96 083524
[178] Fujita Y, Inoue S, Nakamura T, Mannhoto T and Nakamura K E 1998 Astrophys. J. 495 L85
[179] Gaggero D, Bertone G, Calore F, Connors R M T, Lovell M, Markoff S and Storm E 2017 Phys. Rev. Lett. 118 241101
[180] Fender R, Maccarone T and Heywood I 2013 Mon. Not. R. Astron. Soc. 430 1538
[181] Kording E, Fender R and Migliari S 2006 Mon. Not. R. Astron. Soc. 369 1451
[182] Merloni A, Heinz S and Di Matteo T 2003 Mon. Not. R. Astron. Soc. 345 1057
[183] Navarro J F, Frenk C S and White S D M 1996 Astrophys. J. 462 563
[184] Muno M P et al 2009 Astrophys. J. Suppl. 181 110
[185] Lazio T J W and Cordes J M 2008 Astrophys. J. Suppl. 174 481
[186] Inoue Y and Kusenko A 2017 J. Cosmol. Astropart. Phys. JCAP10(2017)034
[187] Mineo S, Gifford S and Sunyaev R 2012 Mon. Not. R. Astron. Soc. 419 2095
[188] Read J I, Lake G, Agerz O and Debattista V P 2008 Mon. Not. R. Astron. Soc. 389 1041
[189] Read J I, Mayer L, Brooks A M, Governato F and Lake G 2009 Mon. Not. R. Astron. Soc. 397 44
[190] Agol E and Kamionkowski M 2002 Mon. Not. R. Astron. Soc. 334 553
[191] Matsumoto T, Teraki Y and Ioka K 2017 arXiv:1704.05047
[192] Meszaros P 1975 Astron. Astrophys. 38 5
[193] Afshordi N, McDonald P and Spiergell D N 2003 Astrophys. J. 594 L71
[194] McQuinn M 2016 Ann. Rev. Astron. Astrophys. 54 313
[195] Croft R A C, Weinberg D H, Bolte M, Burles S, Hernquist L, Katz N, Kirkman D and Tytler D 2002 Astrophys. J. 581 20
[196] Kashlinsky A 2016 Astrophys. J. 823 L25
[197] Kashlinsky A 2005 Phys. Rept. 409 361
[198] Hauser M G and Dwek E 2001 Ann. Rev. Astron. Astrophys. 39 249
[199] Kashlinsky A, Arendt R G, Mather J C and Moseley S H 2005 Nature 438 45
[200] Lyth D and Liddle A 2009 The Primordial Density Perturbation (Cambridge: Cambridge University Press)
[201] Nakama T and Suyama T 2015 Phys. Rev. D 92 121304
[202] Nakama T and Suyama T 2016 Phys. Rev. D 94 043507
[203] Ananda K N, Clarkson C and Wands D 2007 Phys. Rev. D 75 123518
[204] Baumann D, Steinhardt P J, Takahashi K and Ichiki K 2007 Phys. Rev. D 76 084019
[205] Saito R and Yokoyama J 2009 Phys. Rev. Lett. 102 161101
[206] Saito R and Yokoyama J 2011 Phys. Rev. Lett. 107 069901 (erratum)
[207] Hobbs G and Dai S 2017 (arXiv:1707.01615)
[208] Detweiler S L 1979 Astrophys. J. 234 1100
[209] Hellings R W and Downs G S 1983 Astrophys. J. 265 L39
[210] Manchester R N et al 2013 Publ. Astron. Soc. Austral. 30 17
[211] McLaughlin M A 2013 Class. Quantum Grav. 30 224008
[212] Kramer M and Champion D J 2013 Class. Quantum Grav. 30 224009
[213] Misner C W, Thorne K S and Wheeler J A 1973 Gravitation (San Francisco: W. H. Freeman and Company)
[214] Saito R and Yokoyama J 2010 Prog. Theor. Phys. 123 867
[215] Saito R and Yokoyama J 2011 Prog. Theor. Phys. 126 351 (erratum)
[266] Ajith P et al 2011 Phys. Rev. Lett. 106 241101
[267] Ioka K, Tanaka T and Nakamura T 1999 Phys. Rev. D 60 083512
[268] Wang S, Wang Y F, Huang Q G and Li T G F 2016 arXiv:1610.08725
[269] Mandic V, Bird S and Cholis I 2016 Phys. Rev. Lett. 117 201102
[270] Nakamura T et al 2016 PTEP 2016 093E01
[271] Martynov D V et al (LIGO Scientific) 2016 Phys. Rev. D 93 112004
[272] Kovetz E D, Cholis I, Breysse P C and Kamionkowski M 2017 Phys. Rev. D 95 103010
[273] Kovetz E D 2017 Phys. Rev. Lett. 119 131301
[274] Kocsis B, Suyama T, Tanaka T and Yokoyama S 2017 arXiv:1709.09007
[275] O’Leary R M, Meiron Y and Kocsis B 2016 Astrophys. J. 824 L12
[276] Nakamura T et al 2016 PTEP 2016 093E01
[277] Hotokezaka K and Piran T 2017 Astrophys. J. 842 111
[278] Postnov K A and Kuranov A G 2017 arXiv:1706.00369
[279] Hotokezaka K and Piran T 2017 arXiv:1707.08978
[280] Farr W M, Stevenson S, Coleman Miller M, Mandel I, Farr B and Vecchio A 2017 Nature 548 426
[281] Farr B, Holz D E and Farr W M 2017 arXiv:1709.07896
[282] Chiba T and Yokoyama S 2017 PTEP 2017 083E01
[283] Baumgarte T W and Gundlach C 2016 Phys. Rev. Lett. 116 221103
[284] Harada T, Yoo C M, Kohri K and Nakao K I 2017 Phys. Rev. D 96 083517
[285] Kaiser N 1984 Astrophys. J. 284 L9
[286] Abbott T M C et al (DES Collaboration) 2017 arXiv:1708.01530
[287] Namikawa T, Nishizawa A and Taruya A 2016 Phys. Rev. D 94 024013
[288] Raccanelli A, Kovetz E D, Bird S, Cholis I and Muñoz J B 2016 Phys. Rev. D 94 023516
[289] Raccanelli A 2017 Mon. Not. R. Astron. Soc. 469 656
[290] Martel K and Poisson E 1999 Phys. Rev. D 60 124008
[291] Kowalska I, Bulik T, Belczynski K, Dominik M and Gondek-Rosinska D 2011 Astron. Astrophys. 527 A70
[292] Rodriguez C L, Chatterjee S and Rasio F A 2016 Phys. Rev. D 93 084029
[293] O’Leary R M, Kocsis B and Loeb A 2009 Mon. Not. R. Astron. Soc. 395 2127
[294] Nishizawa A, Berti E, Klein A and Sesana A 2016 Phys. Rev. D 94 064020
[295] Nishizawa A, Sesana A, Berti E and Klein A 2017 Mon. Not. R. Astron. Soc. 465 4375
[296] Thorne K S 2017 Interview, Nobelprize.org, Nobel Media AB 2014, Web, 25 Oct 20 www.nobelprize.org/nobel_prizes/physics/laureates/2017/thorne-interview.html