Signals of merging supermassive black holes in pulsar timing arrays

Paul Frederik Depta,[1,*] Kai Schmidt-Hoberg,[2,†] Pedro Schwaller,[3,‡] and Carlo Tasillo[2,§]
[1]*Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany*
[2]*Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany*
[3]*PRISMA⁺ Cluster of Excellence and Mainz Institute for Theoretical Physics, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany*

(Received 25 July 2023; revised 8 April 2024; accepted 15 January 2025; published 24 February 2025)

In this work we evaluate whether the gravitational wave background recently observed by a number of different pulsar timing arrays could be due to merging *primordial* supermassive black hole binaries. We find that for homogeneously distributed primordial black holes this possibility is inconsistent with strong cosmological and astrophysical constraints on their total abundance. If the distribution exhibits some clustering, however, the merger rate will, in general, be enhanced, opening the window for a consistent interpretation of the pulsar timing array data in terms of merging primordial black holes, if μ-distortion constraints associated with the formation mechanism can be evaded.

DOI: 10.1103/PhysRevResearch.7.013196

I. INTRODUCTION

The discovery of gravitational waves (GWs) with frequencies $\mathcal{O}(100)$ Hz from binary black hole mergers in 2015 [1] has opened new possibilities for the exploration of our Universe. Depending on the astrophysical or cosmological source, GW signals may exist over a wide range of frequencies and complementary experimental approaches aim to explore much of the available parameter space [2,3]. Pulsar timing arrays (PTAs) in particular are sensitive at frequencies in the nHz range. In 2020, the North American Nanohertz Observatory for Gravitational Waves collaboration (NANOGrav) [3] announced evidence for a common red process hinting at a stochastic nHz gravitational wave background (GWB), with similar results subsequently published by other PTAs [4–6]. Excitingly, the very recent findings from NANOGrav [7], European PTA (EPTA) [8], Parkes PTA (PPTA) [9], and Chinese PTA (CPTA) [10] not only strengthen this signal, but also show mounting evidence for the characteristic quadrupolar Hellings-Downs correlation [11]. It is, therefore, fascinating and timely to investigate possible origins of this newly observed signal in terms of a gravitational wave background (GWB). Mergers of supermassive black hole binaries (SMBHB) are expected to contribute to the GWB in this frequency range [3], although the local SMBHB density would need to be $\mathcal{O}(10)$ larger than previously estimated [12–14] to match the observed signal amplitude. Hence, the validity of this explanation is under active debate [15–18] and it is interesting to consider alternative sources. A nHz GWB could also originate from truly cosmological sources [19,20], such as inflation [21], topological defects [22–27], first-order phase transitions [28–32], or the *production* of subsolar mass primordial black holes (PBHs) [33–37].

In this work we study the possibility that the observed GW signal could be due to inspiraling supermassive PBH binaries. To obtain the correct signal amplitude, a sizable abundance of PBHs is required, which is subject to strong constraints from various observational probes [38]. It is, therefore, a nontrivial question as to whether inspiraling PBHs could constitute a viable explanation. Indeed, we find that it is *not* possible to consistently explain the NANOGrav data with homogeneously distributed PBHs. Specifically, the parameter regions at very large PBH masses which naively allow to fit the PTA data (as done in [39]) no longer result in a stochastic (or even continuous) signal. For clustered PBHs, on the other hand, the formation of binary systems is more efficient and the merger rate is generally enhanced, opening the possibility for an overall consistent description of the observed GWB.

This work is structured as follows. We first review and extend the formalism required to evaluate the GWs from binary mergers for the case of clustered PBHs. After discussing the relevant constraints and how to match the predicted signal to the PTA data we present our results.

II. GRAVITATIONAL WAVE SIGNAL

A stochastic GWB can be specified in terms of the energy density parameter $h^2\Omega_{\rm GW}(f)$ of the gravitational waves (GWs) per logarithmic frequency interval around a frequency f. In our scenario, this GWB is generated from PBH binary mergers occurring at a rate $R(t_{\rm r})$ per comoving volume $V_{\rm c}$ and

*Contact author: frederik.depta@mpi-hd.mpg.de
†Contact author: kai.schmidt-hoberg@desy.de
‡Contact author: pedro.schwaller@uni-mainz.de
§Contact author: carlo.tasillo@desy.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by Max Planck Society.

cosmic time t_r [40]. The frequencies accessible to PTAs can be well below the maximal frequency emitted during the inspiral and might therefore have been radiated long before coalescence, which we need to take into account. In particular, a given binary merging at time $t_{r,\mathrm{merg}}$ emits a frequency f_r in the cosmic rest frame, redshifted to a frequency $f = f_r/(1+z)$ today, at time $t_r = t_{r,\mathrm{merg}} - \tau_{f_r}$, where

$$\tau_{f_r} = \frac{5 \times 2^{1/3}}{256\,\pi^{8/3}}\, f_r^{-8/3}\, (G\, m_{\mathrm{PBH}})^{-5/3} \tag{1}$$

is the time until coalescence for f_r [41]. As $t_{r,\mathrm{merg}}$ is the argument for the *merger* rate, the number of events per comoving volume and cosmic time emitting a frequency f_r in the cosmic rest frame at cosmic time t_r is given by $R(t_r + \tau_{f_r})$. To the best of our knowledge this frequency-dependent rate has not been discussed yet in the literature concerning GWs generated by PBHs. The GW energy density parameter is given by (cf. [42])

$$\Omega_{\mathrm{GW}}(f) = \frac{f}{\rho_{\mathrm{crit}}} \int_0^{t_0} \mathrm{d}t_r \left(R\big(t_r + \tau_{f_r}\big) \frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}f_r} \right)_{f_r=(1+z)f}, \tag{2}$$

where t_0 is the current cosmic time and $\mathrm{d}E_{\mathrm{GW}}/\mathrm{d}f_r$ is the GW spectrum, which we adopt from [43]. Note that $h^2\Omega_{\mathrm{GW}}(f)$ is, in fact, a cosmological average of many PBH binaries. We discuss the effect of only having access to a local realization of the GWB within PTA measurements below.

For the merger rate we adapt the calculation from [44]. We assume a monochromatic initial PBH mass distribution, expecting that an extended distribution does not qualitatively change our results. A PBH binary forms in the early Universe when the gravitational attraction between two neighboring PBHs overcomes the Hubble flow. A third close-by PBH provides angular momentum such that the two PBHs do not simply collide and typically have a large eccentricity [45–47]. After the first version of this work appeared, it was noted that the high eccentricities characteristic for our model setup actually gave a good fit to the observed GW spectral shape [48] and lowered its degree of anisotropy [49].

Clearly the merger rate depends on the local PBH density at binary formation. It is, therefore, interesting to study the effect of clustering, which increases the local density by δ_{dc} compared to the global comoving number density n_{PBH}. The density contrast δ_{dc} can be considered constant on the scales relevant for binary formation [44]. By considering the number density of three-body configurations relevant for the PBH binary formation and their coalescence times (in terms of the distances in the three-body configurations) [44,50,51] one finds [44]

$$\begin{aligned} R(t_r) = {} & \frac{9\,\tilde{N}_{\mathrm{PBH}}^{53/37}}{296\pi\,\delta_{\mathrm{dc}}\,\tilde{x}^3\,\tilde{\tau}} \left(\frac{t_r}{\tilde{\tau}}\right)^{-34/37} \\ & \times \left(\Gamma\left[\frac{58}{37}, \tilde{N}_{\mathrm{PBH}} \left(\frac{t_r}{\tilde{\tau}}\right)^{3/16}\right] \right. \\ & \left. - \Gamma\left[\frac{58}{37}, \tilde{N}_{\mathrm{PBH}} \left(\frac{t_r}{\tilde{\tau}}\right)^{-1/7}\right] \right), \end{aligned} \tag{3}$$

with $\tilde{x}^3 = 6\,m_{\mathrm{PBH}}/(4\pi\, a_{\mathrm{eq}}^3\, \rho_{\mathrm{eq}})$, the scale factor a_{eq} and energy density ρ_{eq} at matter-radiation equality [52], $\tilde{\tau} = 3\,(a_{\mathrm{eq}}\,\tilde{x})^4/[170\,(G\,m_{\mathrm{PBH}})^3]$, $\tilde{N}_{\mathrm{PBH}} = 4\pi\, n_{\mathrm{PBH}}\, \delta_{\mathrm{dc}}\, \tilde{x}^3/3$, and the incomplete gamma function Γ.

Larger clustering generally leads to a larger number of PBH binaries contributing to the merger rate (note that $\tilde{N}_{\mathrm{PBH}} \propto \delta_{\mathrm{dc}}$) as well as a shorter timescale for the mergers (cf. the arguments of the gamma functions). Furthermore, multiple merger steps can be possible for large clustering due to successive hierarchical merging of PBH binaries with increasing mass. When discussing the case of significant clustering we implement multiple merger steps by adding the rates and contributions to the GW energy density parameter for the corresponding steps as detailed in [53]. Note that this shifts our results for significant clustering to, at most, $\mathcal{O}(10\,\%)$ lower PBH abundances compared to only considering a single step.

By substituting $t_r \to t_r + \tau_{f_r}$ in Eq. (3) we can compute the rate $R(t_r + \tau_{f_r})$ for the emission of a frequency f_r. To illustrate the importance of the emission time of a given frequency, in Fig. 1 (left) we compare the rate for binaries merging at t_r (dotted blue line) with the rate for GW emission with frequencies $f = 1\,\mu\mathrm{Hz}$ and 2.5 nHz today (solid green and orange lines), assuming a PBH mass $m_{\mathrm{PBH}} = 10^5\,M_\odot$, a fraction $f_{\mathrm{PBH}} = 1$ of dark matter (DM) in PBHs, and $\delta_{\mathrm{dc}} = 1$. For instance, at $t_r = 10^8\,\mathrm{yr}$ (i.e., $z \approx 30$) one obtains $\tau_{f_r}[(1+z)1\,\mu\mathrm{Hz}] \approx 130\,\mathrm{yr}$ and $\tau_{f_r}[(1+z)2.5\,\mathrm{nHz}] \approx 1.1 \times 10^9\,\mathrm{yr}$. Hence, the rate for the emission of GWs with the larger frequency (solid green line) is very close to the merger rate (dotted blue line), whereas the rate for the emission of GWs with the smaller frequency (solid orange line) differs significantly, i.e., it takes the value that the dotted blue line attains $1.1 \times 10^9\,\mathrm{yr}$ later.

In the right panel of Fig. 1 we show $h^2\Omega_{\mathrm{GW}}(f)$ according to Eq. (2) (solid lines) as well as just inserting the merger rate $R(t_r)$ instead of $R(t_r + \tau_{f_r})$ in the integral (dotted lines). The difference between those calculations is especially important for the low frequencies observed by NANOGrav if the PBH mass is relatively light.

We close the discussion of the GWB signal by mentioning some assumptions in the calculation. These include a monochromatic mass distribution [44], the modeling of PBH clustering for multiple merger steps [53], a circular orbit of the binaries for the GWB spectrum and the time until coalescence for a given frequency, neglecting the effect of other PBHs on the binaries [54–56], as well as further environmental effects, e.g., due to interactions with surrounding inhomogeneities [54], DM halos surrounding the PBH binaries [57], accretion, and late-time formation of binaries [44]. These effects can have an $\mathcal{O}(1)$ impact on the merger rate, but often require further model-dependent assumptions, especially when clustering is involved, as well as expensive N-body simulations that are beyond the scope of this work. Although this may lead to a quantitative change of our results, slightly shifting the best-fit regions and constraints, our results are expected to be qualitatively robust and the conclusions to be unaffected.

III. EXPECTED NUMBER OF BINARIES

The PTA signals are reported as stochastic GWBs. While there have been searches for signals of individual binaries in the data of different PTAs [58–60], no compelling evidence for these signals was found. For sufficiently large PBH masses

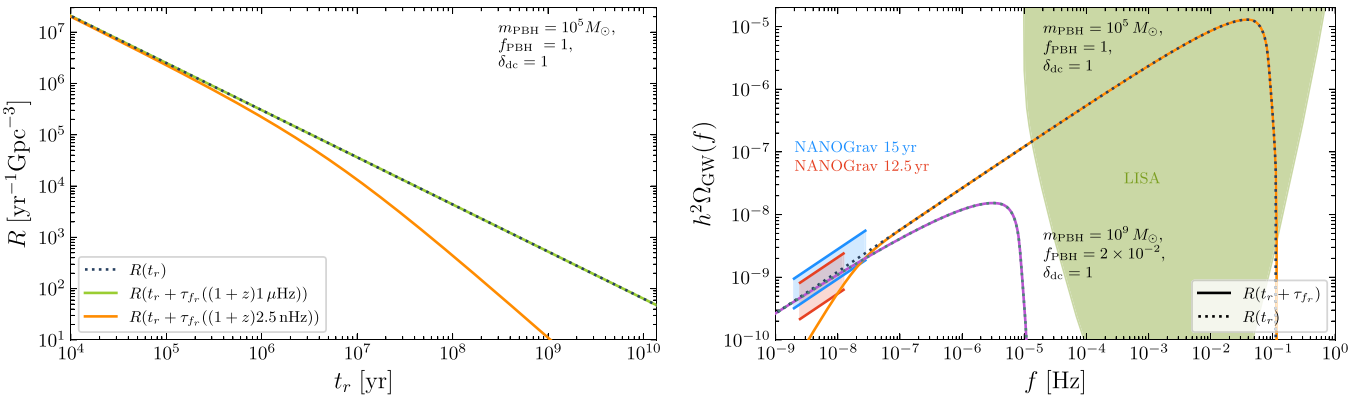


FIG. 1. Left: Rate for binaries merging at $t_{\rm r}$ (dotted line) as well as rate for emitting GWs at $t_{\rm r}$ with different frequencies today (solid green and orange lines). Right: GW energy density $h^2\Omega_{\rm GW}(f)$ obtained using the correct rate $R(t_{\rm r}+\tau_{f_{\rm r}})$ (solid lines) and instead using the merger rate $R(t_{\rm r})$ in Eq. (2) (dotted lines). We also show the regions of the NANOGrav 12.5 yr and 15 yr signals assuming the inspiral of binary BHs. The LISA sensitivity curve is shown in green.

and small abundances the expected signal will, in general, no longer resemble a stochastic background, as only very few binaries will contribute to the signal, causing inconsistency with observations, see [61].

The problem is exacerbated by the question of how well an actual distribution of merging binaries observed by a PTA (corresponding to a local GW energy density $\Omega_{\rm GW,loc}$ on the length scales relevant for PTAs) would reproduce the global average of the GWB $h^2\Omega_{\rm GW}(f) = h^2\langle\Omega_{\rm GW,loc}(f)\rangle$, i.e., how well the global mean is reproduced by the binaries that are in our past light cone. In particular, $h^2\Omega_{\rm GW,loc}(f)$ is *not* deterministic given the model parameters (mass, abundance, and clustering of PBHs), but instead stochastic in itself. The statistics of $h^2\Omega_{\rm GW,loc}(f)$ can be evaluated using Markov chain Monte Carlo methods or moment generating functions [62]. Due to the considerable additional numerical effort we leave this for future studies, use the global average signal prediction $h^2\Omega_{\rm GW}(f)$, and identify regions in parameter space where we expect a significant deviation from the global average.

Most importantly, the distribution of $h^2\Omega_{\rm GW,loc}(f)$ is influenced by the expected number of binaries $\bar{N}(f_-, f_+)$ contributing to the GWB in a frequency range between f_- and f_+. In the Appendix we use [63] to show that

$$\bar{N}(f_-, f_+) = \int_{f_-}^{f_+} \frac{{\rm d}f}{f} \int_0^\infty {\rm d}z\, \frac{32\pi\,\tau_{f_{\rm r}}\,[d_{\rm c}(z)]^2}{3H(z)} R(t_{\rm r} - \tau_{f_{\rm r}})\,, \tag{4}$$

where $d_{\rm c}$ is the comoving distance and H is the Hubble rate. As we fit to the 14 lowest frequencies in the NANOGrav 15 yr data set we use $f_- = 1/T^{15}_{\rm obs} = 1.98\,{\rm nHz}$ and $f_+ = 14/T^{15}_{\rm obs} = 27.7\,{\rm nHz}$.

If $\bar{N} \gg 1$, many PBH binaries contribute to the GW signal and the local GWB is very close to the global value. This holds even though uncertainties due to some close-by binaries can be relevant for $\bar{N} \lesssim 100$ and can potentially cause spikes in the local GWB [62]. If $\bar{N} \sim \mathcal{O}(\text{few})$, i.e., the GW signal is composed of only a few binaries, the uncertainty in the signal prediction is considerable. It would then be more appropriate to search for individual GW events instead of a GWB [58–60]. If $\bar{N} \ll 1$, even having a single PBH binary emitting GWs is unlikely and in most realizations $h^2\Omega_{\rm GW,loc}(f) = 0$. This is a feature of our model setup, which will soon be testable through the anisotropy of a GW background from a low number of sources [61].

IV. PBH PRODUCTION AND CONSTRAINTS

There are many different production mechanisms for PBHs, see, e.g., [64–70]. Highly clustered PBH distributions, in particular, are not expected for Gaussian primordial fluctuations [71,72] [being subject to strong observational constraints [73,74] from Cosmic Background Explorer (COBE) Far Infrared Absolute Spectrophotometer (FIRAS)], but could, e.g., arise due to primordial non-Gaussianities [75,76] or from the collapse of domain walls [65]. In this work we remain agnostic about the origin and spatial distribution of PBHs and concentrate on exploring the phenomenological impact of different assumptions. In [77] it was recently shown that a population of supermassive and clustered PBHs can indeed be produced in inflationary models involving an additional curvaton field, while not violating μ-distortion constraints.

Different astrophysical and cosmological observations place constraints on the abundance of heavy PBHs, which we adopt from [38]. These limits assume a monochromatic mass function as well as a roughly homogeneous spatial distribution (no clustering). While we briefly comment on the expected impact of sizable clustering, a full reevaluation of these limits would require going beyond some of the simple approximations made in the original derivations and is beyond the scope of this work. Also note that many of the different constraints come with different uncertainties and sometimes also with additional caveats.

The most relevant limits in the mass range of interest come from the heating of stars in the Galactic disk [78], the tidal disruption of galaxies [78], the dynamical friction effect on halo objects [78], requiring successful formation of the observed large-scale structure (LSS) [79], and observations of x-ray binaries [80]. Many of these limits require at least one PBH per relevant cosmic structure. In the case of PBH clustering, we expect that some structures will contain more than one PBH while others contain no PBH at all. This may shift limits

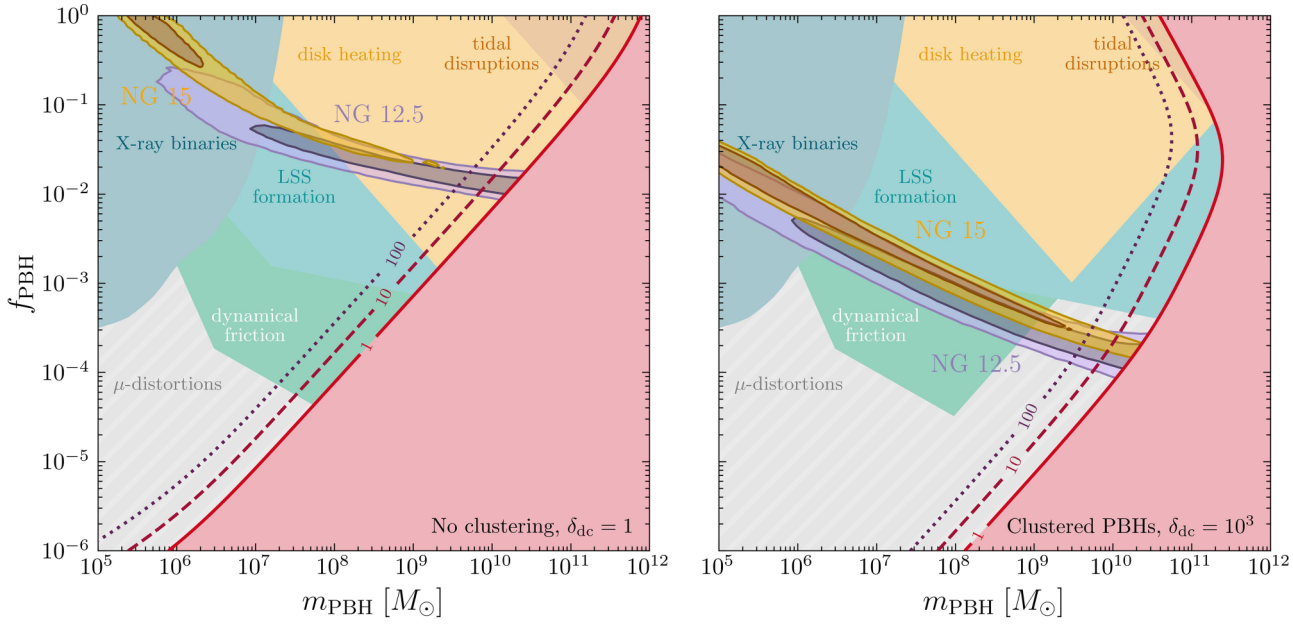


FIG. 2. Best-fit region where the NANOGrav 15 yr (brown) and 12.5 yr (purple) data set can be explained by merging PBH binaries without clustering (left) and with significant clustering (right). We indicate regions, where $\bar{N} < 1$, 10, and 100 PBH binaries are expected to contribute to the signal, and show complementary constraints as discussed in the main text.

to smaller masses roughly by the ratio of individual PBH and cluster masses. A proper evaluation necessitates further assumptions and detailed simulations, which would also induce dependence on the clustering model. We, therefore, only show the constraints cited above from simplified analytical treatments without clustering as estimates.

Depending on the production mechanism of the PBHs, strong limits may also arise from the observation of the cosmic microwave background (CMB). In particular, if PBHs form due to the tail of Gaussian density fluctuations, Silk damping leads to μ distortions and strong constraints over a sizable mass range [81]. However, this limit crucially depends on the production mechanism and may, therefore, even be completely evaded. In fact, as discussed above, large clustering generally requires a different production mechanism, calling into question the relevance of these limits.

Additional constraints from future GW observations could lead to further constraints relevant for the PBH parameter space, see the right panel of Fig. 1. We check explicitly that the Laser Interferometer Space Antenna (LISA) will not be able to probe our best-fit point, but will rather allow independent tests of the parameter space constraints from x-ray binaries [80].

V. PTA DATA ANALYSIS

We fit the GWB spectrum from PBH binary mergers to the NANOGrav 15 yr and 12.5 yr data sets via the interface PTARCADE [82] for CEFFYL [83] using the 14 (five) lowest Fourier modes for the 15 yr (12.5 yr) data set. Given that evidence for a Hellings-Downs correlation is only present in the new data set, we assume a common uncorrelated red noise spectrum for the 12.5 yr data set and only move to Hellings-Downs correlations for the 15 yr data set. We validate our approach by comparison to results obtained using ENTERPRISE and ENTERPRISE-EXTENSIONS [84,85]. We perform calculations with no clustering $\delta_{\rm dc} = 1$, and significant clustering $\delta_{\rm dc} = 10^3$, choosing log priors $f_{\rm PBH} \in [10^{-5}, 1]$ and $m_{\rm PBH} \in [10^5, 10^{12}]\, M_\odot$. When deriving constraints we further add a power-law signal corresponding to SMBHBs. After marginalizing over the amplitude, the region in the remaining $(m_{\rm PBH}, f_{\rm PBH})$ plane excluded with $2\,\sigma$ corresponds to the excluded parameter space in which the expected GWB signal from PBH binaries would exceed observations.

VI. RESULTS

In Fig. 2 we show the regions in PBH mass $m_{\rm PBH}$ and DM fraction $f_{\rm PBH}$, where the NANOGrav signal can be explained assuming merging PBHs with no clustering (left) and significant clustering with $\delta_{\rm dc} = 10^3$ (right). These contours are only expected to be reliable when the expected number of binaries is $\bar{N} \gtrsim 10$ (cf. dashed dark red lines) as otherwise the signal observable in NANOGrav is expected to have significant deviations from the global average GWB signal used in our analysis. Particularly, for $\bar{N} < 1$ (full red line) one would not even expect any GWB signal in most realizations of PBH binary distributions. These effects were neglected in [39], where it was incorrectly concluded that the NANOGrav signal could be explained in a region of parameter space without clustering, where $\bar{N} \lesssim 1$. We find that the case without clustering cannot explain the NANOGrav signal once taking into account cosmological and astrophysical constraints.

Including clustering shifts the signal regions to smaller $f_{\rm PBH}$, enabling a consistent explanation of the PTA data without violating observational constraints, provided that the PBH formation mechanism does not result in significant μ distortions. As demonstrated in [77], certain formation mechanisms for supermassive PBHs can indeed avoid μ-distortion bounds. Note that clustering is also expected to further open up

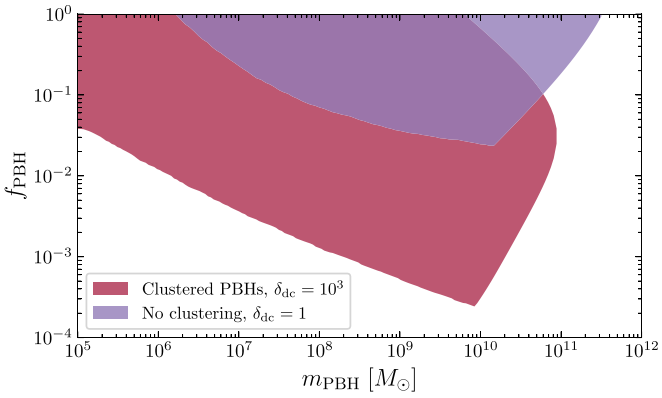


FIG. 3. Constraints on PBHs from PTA observations with no clustering (purple) and significant clustering $\delta_{\rm dc} = 10^3$ (red). We conservatively cut the limits when the expected number of binaries contributing to the signal falls below 10.

parameter space for a consistent explanation as complementary constraints are expected to shift to smaller masses and the LSS constraints quickly lose sensitivity for masses above $\sim 4 \times 10^{10} M_\odot$ [79]. While the exact quantitative impact of clustering on these constraints requires detailed, model-dependent simulations of PBH cluster formation, our analysis indicates that the unconstrained best-fit region identified in our work remains robust.

Comparing the results for the 15 yr and 12.5 yr data sets we note that the signal regions shift to larger $f_{\rm PBH}$ due to the preference for a larger GWB in the new data set [3,7], and there is a slight preference for lower masses, where the earlier emission of GWs leads to an increased slope, cf. Fig. 1, as preferred by the new data [7].

If the GWB signal generated by inspiraling PBH binaries becomes too large, above the regions explaining the PTA data in Fig. 2, the abundance of PBHs can be *constrained* by PTA observations. The constraints are shown in Fig. 3 with and without clustering, conservatively requiring $\bar{N} \geqslant 10$.

VII. DISCUSSION AND CONCLUSIONS

In this work we studied the possibility that the signal observed by various PTAs was due to inspiraling primordial supermassive black hole binaries. If the PBHs were "homogeneously" distributed at their formation, i.e., followed a Poisson distribution, significant cosmological and astrophysical constraints excluded the possibility of explaining the PTA signal with inspiraling PBHs. Instead considering a clustered spatial distribution of PBHs increased the binary merger rate and thus enabled a consistent explanation of the PTA signals with inspiraling PBH binaries. Crucially, we checked that the signal prediction is also reliable in the relevant parameter space by computing the expected number of binaries contributing to the gravitational wave signal. Further, we used PTA data to constrain the PBH parameter space when the GWB generated during the mergers resulted in stronger signal strengths than the one detected.

This work may serve as a motivation for model builders to construct scenarios that can generate clustered supermassive PBHs without running into cosmological and astrophysical constraints, in particular, due to μ distortions of the CMB arising from PBH formation. Moreover, our analysis may also motivate future studies to refine astrophysical constraints on clustered PBHs through detailed simulations, as the most stringent existing bounds typically neglect clustering effects. While the latest PTA data finds no evidence for individual compact binary merger events on top of a stochastic background or anisotropies of the GW spectrum, the situation might change in the future. Such a detection would likely invalidate most other cosmological explanations, but is a prediction of our scenario.

ACKNOWLEDGMENTS

We would like to thank Andrea Mitridate, Xiao Xue, Joachim Kopp, and Felix Kahlhöfer for helpful discussions on related works. This work is funded by the Deutsche Forschungsgemeinschaft (DFG) through Germany's Excellence Strategy EXC 2121 "Quantum Universe" Grants No. 390833306 and No. EXC 2118/1 "Precision Physics, Fundamental Interactions, and Structure of Matter" Grant No. 39083149. We acknowledge the use of PTMCMC [86] for sampling and NUMPY [87], MATPLOTLIB [88], and a customized version of CHAINCONSUMER [89] to analyze and visualize our results.

APPENDIX: NUMBER OF CONTRIBUTING PBH

Here we derive the expression for the expected number of primordial black hole binaries contributing to the gravitational wave background within a given frequency range. To do so, recall that $R(t_{\rm r} + \tau_{f_{\rm r}})$ is the number of binaries per comoving volume $V_{\rm c}$ and per time interval in the cosmic rest frame $t_{\rm r}$ emitting with a frequency $f_{\rm r}$, where $\tau_{f_{\rm r}}$ is the time until coalescence. Hence, the expected number of binaries $\mathrm{d}\bar{N}$ emitting at redshifts between z and $z + \mathrm{d}z$ with a frequency within the logarithmic interval $\mathrm{d}\ln f_{\rm r}$ around $f_{\rm r}$ is [63]

$$R(t_{\rm r} + \tau_{f_{\rm r}}) = \frac{\mathrm{d}z}{\mathrm{d}t_{\rm r}} \frac{\mathrm{d}^2\bar{N}}{\mathrm{d}z\,\mathrm{d}\ln f_{\rm r}} \frac{\mathrm{d}\ln f_{\rm r}}{\mathrm{d}t_{\rm r}} \frac{\mathrm{d}t_{\rm r}}{\mathrm{d}z} \frac{\mathrm{d}z}{\mathrm{d}V_{\rm c}} = \frac{\mathrm{d}^2\bar{N}}{\mathrm{d}z\,\mathrm{d}\ln f_{\rm r}} \left(-\frac{\mathrm{d}\ln f_{\rm r}}{\mathrm{d}\tau_{f_{\rm r}}}\right) \frac{\mathrm{d}z}{\mathrm{d}V_{\rm c}}\,, \tag{A1}$$

where we use that the time in the cosmic rest frame a source is emitting within the frequency interval is given by

$$\mathrm{d}t_{\rm r} \frac{\mathrm{d}\ln f_{\rm r}}{\mathrm{d}t_{\rm r}} = -\mathrm{d}t_{\rm r} \frac{\mathrm{d}\ln f_{\rm r}}{\mathrm{d}\tau_{f_{\rm r}}}\,. \tag{A2}$$

Since the binaries emitting in the logarithmic frequency interval $\mathrm{d}\ln f_{\rm r}$ around $f_{\rm r}$ are detected today in a logarithmic frequency interval $\mathrm{d}\ln f$ around f, where $\mathrm{d}\ln f_{\rm r} = \mathrm{d}f_{\rm r}/f_{\rm r} = \mathrm{d}f/f = \mathrm{d}\ln f$, we have

$$\frac{\mathrm{d}^2\bar{N}}{\mathrm{d}z\,\mathrm{d}\ln f} = \frac{\mathrm{d}^2\bar{N}}{\mathrm{d}z\,\mathrm{d}\ln f_{\rm r}}\,. \tag{A3}$$

With Eq. (1) and the definition of the comoving volume in a spatially flat universe

$$V_c(z) = \frac{4\pi}{3}[d_c(z)]^3 = \frac{4\pi}{3}\left(\int_0^z \frac{dz'}{H(z')}\right)^3 , \tag{A4}$$

where d_c is the comoving distance and H is the Hubble rate, we find

$$\frac{d^2\bar{N}}{dz\, d\ln f} = \frac{8}{3}\,\tau_{f_r}\,\frac{4\pi\,[d_c(z)]^2}{H(z)}\,R(t_r - \tau_{f_r}) . \tag{A5}$$

The average number of binaries $\bar{N}(f_-, f_+)$ contributing to GWs of frequencies between f_- and f_+ is therefore given by

$$\bar{N}(f_-, f_+) = \int_{f_-}^{f_+} \frac{df}{f} \int_0^\infty dz\, \frac{32\pi\,\tau_{f_r}\,[d_c(z)]^2}{3H(z)} R(t_r - \tau_{f_r}) . \tag{A6}$$

Note that the *actual* number of binaries is then Poisson distributed with mean $\bar{N}(f_-, f_+)$.

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Observation of gravitational waves from a binary black hole merger, Phys. Rev. Lett. **116**, 061102 (2016).

[2] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi, J. Camp, C. Caprini, V. Cardoso, M. Colpi, J. Conklin, N. Cornish, C. Cutler, K. Danzmann, R. Dolesi, L. Ferraioli *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[3] Z. Arzoumanian *et al.* (The NANOGrav Collaboration), The NANOGrav 12.5 yr data set: Search for an isotropic stochastic gravitational-wave background, Astrophys. J. Lett. **905**, L34 (2020).

[4] B. Goncharov, R. M. Shannon, D. J. Reardon, G. Hobbs, A. Zic, M. Bailes, M. Curyło, S. Dai, M. Kerr, M. E. Lower, R. N. Manchester, R. Mandow, H. Middleton, M. T. Miles, A. Parthasarathy, E. Thrane, N. Thyagarajan, X. Xue, X.-J. Zhu, A. D. Cameron *et al.*, On the evidence for a common-spectrum process in the search for the nanohertz gravitational-wave background with the Parkes pulsar timing array, Astrophys. J. Lett. **917**, L19 (2021).

[5] S. Chen, R. N. Caballero, Y. J. Guo, A. Chalumeau, K. Liu, G. Shaifullah, K. J. Lee, S. Babak, G. Desvignes, A. Parthasarathy, H. Hu, E. van der Wateren, J. Antoniadis, A.-S. Bak Nielsen, C. G. Bassa, A. Berthereau, M. Burgay, D. J. Champion, I. Cognard, M. Falxa *et al.*, Common-red-signal analysis with 24-yr high-precision timing of the European Pulsar Timing Array: inferences in the stochastic gravitational-wave background search, Mon. Not. Roy. Astron. Soc. **508**, 4970 (2021).

[6] J. Antoniadis, Z. Arzoumanian, S. Babak, M. Bailes, A.-S. Bak Nielsen, P. T. Baker, C. G. Bassa, B. Bécsy, A. Berthereau, M. Bonetti, A. Brazier, P. R. Brook, M. Burgay, S. Burke-Spolaor, R. N. Caballero, J. A. Casey-Clyde, A. Chalumeau, D. J. Champion, M. Charisi, S. Chatterjee *et al.*, The international pulsar timing array second data release: Search for an isotropic gravitational wave background, Mon. Not. Roy. Astron. Soc. **510**, 4873 (2022).

[7] G. Agazie *et al.* (The NANOGrav Collaboration), The NANOGrav 15-year data set: Evidence for a gravitational-wave background, Astrophys. J. Lett. **951**, (2023).

[8] J. Antoniadis *et al.* (EPTA Collaboration and InPTA Collaboration), The second data release from the European Pulsar Timing Array III. Search for gravitational wave signals, Astron. Astrophys. **678**, A50 (2023).

[9] D. J. Reardon, A. Zic, R. M. Shannon, G. B. Hobbs, M. Bailes, V. Di Marco, A. Kapur, A. F. Rogers, E. Thrane, J. Askew, N. D. Ramesh Bhat, A. Cameron, M. Curyło, W. A. Coles, S. Dai, B. Goncharov, M. Kerr, A. Kulkarni, Y. Levin, M. E. Lower *et al.*, Search for an isotropic gravitational-wave background with the Parkes Pulsar Timing Array, Astrophys. J. Lett. **951**, (2023).

[10] H. Xu, S. Chen, Y. Guo, J. Jiang, B. Wang, J. Xu, Z. Xue, R. Nicolas Caballero, J. Yuan, Y. Xu, J. Wang, L. Hao, J. Luo, K. Lee, J. Han, P. Jiang, Z. Shen, M. Wang, N. Wang, R. Xu *et al.*, Searching for the nano-hertz stochastic gravitational wave background with the Chinese Pulsar Timing Array data release I, Res. Astron. Astrophys. **23**, 075024 (2023).

[11] R. W. Hellings and G. W. Downs, Upper limits on the isotropic gravitational radiation background from pulsar timing analysis, Astrophys. J. **265**, L39 (1983).

[12] J. A. Casey-Clyde, C. M. F. Mingarelli, J. E. Greene, K. Pardo, M. Nañez, and A. D. Goulding, A quasar-based supermassive black hole binary population model: Implications for the gravitational wave background, Astrophys. J. **924**, 93 (2022).

[13] L. Z. Kelley, L. Blecha, and L. Hernquist, Massive black hole binary mergers in dynamical galactic environments, Mon. Not. R. Astron. Soc. **464**, 3131 (2017).

[14] L. Z. Kelley, L. Blecha, L. Hernquist, A. Sesana, and S. R. Taylor, The gravitational wave background from massive black hole binaries in illustris: spectral features and time to detection with pulsar timing arrays, Mon. Not. R. Astron. Soc. **471**, 4508 (2017).

[15] H. Middleton, A. Sesana, S. Chen, A. Vecchio, W. Del Pozzo, and P. A. Rosado, Massive black hole binary systems and the NANOGrav 12.5 yr results, Mon. Not. R. Astron. Soc. Lett. **502**, L99 (2021).

[16] D. Izquierdo-Villalba, A. Sesana, S. Bonoli, and M. Colpi, Massive black hole evolution models confronting the n-Hz amplitude of the stochastic gravitational wave background, Mon. Not. R. Astron. Soc. **509**, 3488 (2021).

[17] M. Curyło and T. Bulik, Predictions for LISA and PTA based on SHARK galaxy simulations, Astron. Astrophys. **660**, A68 (2022).

[18] J. J. Somalwar and V. Ravi, The origin of the nano-Hertz stochastic gravitational wave background: the contribution from $z \gtrsim 1$ supermassive black-hole binaries, arxiv:2306.00898v1.

[19] A. Afzal *et al.* (The NANOGrav Collaboration), The NANOGrav 15-year data set: search for signals from new physics, Astrophys. J. Lett. **951**, (2023).

[20] J. Antoniadis *et al.* (EPTA Collaboration and InPTA Collaboration), The second data release from the European Pulsar Timing Array: V. Implications for massive black holes, dark matter and the early universe, A&A **685**, A94 (2024).

[21] S. Vagnozzi, Implications of the NANOGrav results for inflation, Mon. Not. R. Astron. Soc. **502**, L11 (2021).
[22] S. Blasi, V. Brdar, and K. Schmitz, Has NANOGrav found first evidence for cosmic strings? Phys. Rev. Lett. **126**, 041305 (2021).
[23] J. Ellis and M. Lewicki, Cosmic string interpretation of NANOGrav pulsar timing data, Phys. Rev. Lett. **126**, 041304 (2021).
[24] W. Buchmuller, V. Domcke, and K. Schmitz, From NANOGrav to LIGO with metastable cosmic strings, Phys. Lett. B **811**, 135914 (2020).
[25] R. Samanta and S. Datta, Gravitational wave complementarity and impact of NANOGrav data on gravitational leptogenesis, J. High Energ. Phys. 05 (2021) 211.
[26] S.-Y. Guo, M. Khlopov, X. Liu, L. Wu, Y. Wu, and B. Zhu, Footprints of axion-like particle in pulsar timing array data and James Webb Space Telescope observations, Sci. China Phys. Mech. Astron. **67**, 111011 (2024).
[27] S. F. King, D. Marfatia, and M. H. Rahat, Towards distinguishing Dirac from Majorana neutrino mass with gravitational waves, Phys. Rev. D **109**, 035014 (2024).
[28] Y. Nakai, M. Suzuki, F. Takahashi, and M. Yamada, Gravitational waves and dark radiation from dark phase transition: Connecting NANOGrav pulsar timing data and hubble tension, Phys. Lett. B **816**, 136238 (2021).
[29] W. Ratzinger and P. Schwaller, Whispers from the dark side: Confronting light new physics with NANOGrav data, SciPost Phys. **10**, 047 (2021).
[30] Z. Arzoumanian *et al.* (The NANOGrav Collaboration), Searching for gravitational waves from cosmological phase transitions with the NANOGrav 12.5-year dataset, Phys. Rev. Lett. **127**, 251302 (2021).
[31] T. Bringmann, P. F. Depta, T. Konstandin, K. Schmidt-Hoberg, and C. Tasillo, Does NANOGrav observe a dark sector phase transition?, J. Cosmol. Astropart. Phys. 11 (2023) 053.
[32] E. Madge, E. Morgante, C. P. Ibáñez, N. Ramberg, W. Ratzinger, S. Schenk, and P. Schwaller, Primordial gravitational waves in the nano-Hertz regime and PTA data – towards solving the GW inverse problem, J. High Energ. Phys. 10 (2023) 171.
[33] V. De Luca, G. Franciolini, and A. Riotto, NANOGrav data hints at primordial black holes as dark matter, Phys. Rev. Lett. **126**, 041303 (2021).
[34] V. Vaskonen and H. Veermäe, Did NANOGrav see a signal from primordial black hole formation? Phys. Rev. Lett. **126**, 051303 (2021).
[35] K. Kohri and T. Terada, Solar-mass primordial black holes explain NANOGrav hint of gravitational waves, Phys. Lett. B **813**, 136040 (2021).
[36] C. Ünal, E. D. Kovetz, and S. P. Patil, Multimessenger probes of inflationary fluctuations and primordial black holes, Phys. Rev. D **103**, 063519 (2021).
[37] A. Ashoorioon, K. Rezazadeh, and A. Rostami, NANOGrav signal from the end of inflation and the LIGO mass and heavier primordial black holes, Phys. Lett. B **835**, 137542 (2022).
[38] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, Constraints on primordial black holes, Rep. Prog. Phys. **84**, 116902 (2021).
[39] V. Atal, A. Sanglas, and N. Triantafyllou, NANOGrav signal as mergers of Stupendously Large Primordial Black Holes, J. Cosmol. Astropart. Phys. 06 (2021) 022.
[40] Cosmic time refers to the time in the corresponding cosmic rest frame, related to time t measured in the present cosmic rest frame by $\mathrm{d}t = (1+z)\,\mathrm{d}t_\mathrm{r}$.
[41] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments* (Oxford University Press, Oxford, 2007).
[42] E. S. Phinney, A Practical theorem on gravitational wave backgrounds, arXiv:astro-ph/0108028.
[43] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, A. M. Sintes, J. T. Whelan, B. Bruegmann, P. Diener, N. Dorband, J. Gonzalez, M. Hannam, S. Husa, D. Pollney, L. Rezzolla, L. Santamaria, U. Sperhake, and J. Thornburg, Template bank for gravitational waveforms from coalescing binary black holes: Non-spinning binaries, Phys. Rev. D **77**, 104017 (2008) [**79**, 129901(E) (2009)].
[44] M. Raidal, V. Vaskonen, and H. Veermäe, Gravitational waves from primordial black hole mergers, J. Cosmol. Astropart. Phys. 09 (2021) 037.
[45] T. Nakamura, M. Sasaki, T. Tanaka, and K. S. Thorne, Gravitational waves from coalescing black hole MACHO binaries, Astrophys. J. Lett. **487**, L139 (1997).
[46] K. Ioka, T. Chiba, T. Tanaka, and T. Nakamura, Black hole binary formation in the expanding universe: Three body problem approximation, Phys. Rev. D **58**, 063003 (1998).
[47] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, Primordial black hole scenario for the gravitational-wave event GW150914, Phys. Rev. Lett. **117**, 061101 (2016) [**121**, 059901(E) (2018)].
[48] J. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, Eccentricity effects on the supermassive black hole gravitational wave background, Astron. Astrophys. **691**, A212 (2024).
[49] J. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, Statistics of the supermassive black hole gravitational wave background anisotropy, arXiv:2411.19692.
[50] P. C. Peters, Gravitational radiation and the motion of two point masses, Phys. Rev. **136**, B1224 (1964).
[51] Using the coalescence time makes sure that our scenario does not suffer from the so-called "final Parsec problem" [90] as only binaries with sufficiently short coalescence times contribute to the GWB.
[52] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. cosmological parameters, Astron. Astrophys. **641**, A6 (2020).
[53] T. Bringmann, P. F. Depta, V. Domcke, and K. Schmidt-Hoberg, Towards closing the window of primordial black holes as dark matter: The case of large clustering, Phys. Rev. D **99**, 063532 (2019).
[54] Y. Ali-Haïmoud, E. D. Kovetz, and M. Kamionkowski, Merger rate of primordial black-hole binaries, Phys. Rev. D **96**, 123523 (2017).
[55] M. Raidal, C. Spethmann, V. Vaskonen, and H. Veermäe, Formation and evolution of primordial black hole binaries in the early universe, J. Cosmol. Astropart. Phys. 02 (2019) 018.
[56] V. Vaskonen and H. Veermäe, Lower bound on the primordial black hole merger rate, Phys. Rev. D **101**, 043015 (2020).
[57] B. J. Kavanagh, D. Gaggero, and G. Bertone, Merger rate of a subdominant population of primordial black holes, Phys. Rev. D **98**, 023536 (2018).
[58] M. Falxa *et al.* (IPTA Collaboration), Searching for continuous gravitational waves in the second data release of the interna-

tional pulsar timing array, Mon. Not. R. Astron. Soc. **521**, 5077 (2023).

[59] J. Antoniadis *et al.* (EPTA Collaboration and InPTA Collaboration), The second data release from the European pulsar timing array IV. Search for continuous gravitational wave signals, Astron. Astrophys. **690**, A118 (2024).

[60] G. Agazie *et al.* (The NANOGrav Collaboration), The NANOGrav 15 yr data set: Bayesian limits on gravitational waves from individual supermassive black hole binaries, Astrophys. J. Lett. **951**, L50 (2023).

[61] G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, J. G. Baier, P. T. Baker, B. Bécsy, L. Blecha, A. Brazier, P. R. Brook, L. Brown, S. Burke-Spolaor, J. Andrew Casey-Clyde, M. Charisi, S. Chatterjee, T. Cohen, J. M. Cordes, N. J. Cornish, F. Crawford *et al.*, The NANOGrav 15 yr Data Set: Looking for signs of discreteness in the gravitational-wave background, Astrophys. J. **978**, 31 (2025).

[62] J. Ellis, M. Fairbairn, G. Hütsi, M. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, Prospects for future binary black hole GW studies in light of PTA measurements, Astron. Astrophys. **676**, A38 (2023).

[63] A. Sesana, A. Vecchio, and C. N. Colacino, The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with pulsar timing arrays, Mon. Not. R. Astron. Soc. **390**, 192 (2008).

[64] B. J. Carr and S. W. Hawking, Black holes in the early Universe, Mon. Not. R. Astron. Soc. **168**, 399 (1974).

[65] K. M. Belotsky, V. I. Dokuchaev, Y. N. Eroshenko, E. A. Esipova, M. Y. Khlopov, L. A. Khromykh, A. A. Kirillov, V. V. Nikulin, S. G. Rubin, and I. V. Svadkovsky, Clusters of primordial black holes, Eur. Phys. J. C **79**, 246 (2019).

[66] E. Cotner and A. Kusenko, Primordial black holes from scalar field evolution in the early universe, Phys. Rev. D **96**, 103002 (2017).

[67] F. Ferrer, E. Masso, G. Panico, O. Pujolas, and F. Rompineve, Primordial black holes from the QCD axion, Phys. Rev. Lett. **122**, 101301 (2019).

[68] M. Lewicki, P. Toczek, and V. Vaskonen, Primordial black holes from strong first-order phase transitions, J. High Energ. Phys. 09 (2023) 092.

[69] Y. Gouttenoire and T. Volansky, Primordial black holes from supercooled phase transitions, arxiv:2305.04942v1.

[70] M. J. Baker, M. Breitbach, J. Kopp, and L. Mittnacht, Primordial black holes from first-order cosmological phase transitions, arxiv:2105.07481v2.

[71] Y. Ali-Haïmoud, Correlation function of high-threshold regions and application to the initial small-scale clustering of primordial black holes, Phys. Rev. Lett. **121**, 081304 (2018).

[72] V. Desjacques and A. Riotto, Spatial clustering of primordial black holes, Phys. Rev. D **98**, 123533 (2018).

[73] J. C. Mather, E. S. Cheng, D. A. Cottingham, R. E. Eplee, D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jesnsen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, S. H. Moseley, Jr., R. F. Silverberg *et al.*, Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument, Astrophys. J. **420**, 439 (1994).

[74] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, The cosmic microwave background spectrum from the full COBE FIRAS data set, Astrophys. J. **473**, 576 (1996).

[75] S. Young and C. T. Byrnes, Long-short wavelength mode coupling tightens primordial black hole constraints, Phys. Rev. D **91**, 083521 (2015).

[76] T. Matsubara, T. Terada, K. Kohri, and S. Yokoyama, Clustering of primordial black holes formed in a matter-dominated epoch, Phys. Rev. D **100**, 123544 (2019).

[77] D. Hooper, A. Ireland, G. Krnjaic, and A. Stebbins, Supermassive primordial black holes from inflation, J. Cosmol. Astropart. Phys. 04 (2024) 021.

[78] B. J. Carr and M. Sakellariadou, Dynamical constraints on dark compact objects, Astrophys. J. **516**, 195 (1999).

[79] B. Carr and J. Silk, Primordial black holes as generators of cosmic structures, Mon. Not. R. Astron. Soc. **478**, 3756 (2018).

[80] Y. Inoue and A. Kusenko, New X-ray bound on density of primordial black holes, J. Cosmol. Astropart. Phys. 10 (2017) 034.

[81] B. J. Carr and J. E. Lidsey, Primordial black holes and generalized constraints on chaotic inflation, Phys. Rev. D **48**, 543 (1993).

[82] A. Mitridate, D. Wright, R. von Eckardstein, T. Schröder, J. Nay, K. Olum, K. Schmitz, and T. Trickle, PTArcade, arXiv:2306.16377.

[83] W. G. Lamb, S. R. Taylor, and R. van Haasteren, The need for speed: Rapid refitting techniques for Bayesian spectral characterization of the gravitational wave background using pulsar timing arrays, Phys. Rev. D **108**, 103019 (2023).

[84] J. A. Ellis, M. Vallisneri, S. R. Taylor, and P. T. Baker, Enterprise: Enhanced numerical toolbox enabling a robust Pulsar Inference Suite, v3.0.0, Zenodo (2020), https://doi.org/10.5281/zenodo.4059815.

[85] S. R. Taylor, P. T. Baker, J. S. Hazboun, J. Simon, and S. J. Vigeland, ENTERPRISE_EXTENSIONS, v2.3.3 (2021), https://github.com/nanograv/enterprise_extensions.

[86] J. Ellis and R. van Haasteren, PTMCMCSAMPLER: Official release (1.0.0), Zenodo (2017), https://doi.org/10.5281/zenodo.1037579.

[87] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández Del Río, M. Wiebe, P. Peterson, P. Gírard-Marchant *et al.*, Array programming with NUMPY, Nature (London) **585**, 357 (2020).

[88] J. D. Hunter, Matplotlib: A 2D graphics environment, Comput. Sci. Eng. **9**, 90 (2007).

[89] S. R. Hinton, ChainConsumer, J. Open Source Softw. **1**, 45 (2016).

[90] M. Milosavljevic and D. Merritt, The Final parsec problem, AIP Conf. Proc. **686**, 201 (2003).