ANNUAL REVIEWS

Annual Review of Astronomy and Astrophysics

Nanohertz Gravitational Waves

Alberto Sesana[1,2,3] and Daniel G. Figueroa[4]

[1]Dipartimento di Fisica "G. Occhialini," Università degli Studi di Milano-Bicocca, Milano, Italy; email: alberto.sesana@unimib.it

[2]INFN–Sezione di Milano-Bicocca, Milano, Italy

[3]INAF–Osservatorio Astronomico di Brera, Milano, Italy

[4]Instituto de Física Corpuscular, Consejo Superior de Investigaciones Científicas and Universitat de València, Paterna, Valencia, Spain

Annu. Rev. Astron. Astrophys. 2026. 64:537–86

First published as a Review in Advance on June 30, 2026

The *Annual Review of Astronomy and Astrophysics* is online at astro.annualreviews.org

https://doi.org/10.1146/annurev-astro-122325-025132

Keywords

gravitational waves, black hole physics, supermassive black holes, cosmology, early Universe, theory, pulsars, astrometry

Abstract

Evidence of a gravitational wave (GW) signal has emerged in pulsar timing array (PTA) data, opening a new window into the nanohertz GW Universe. We explore the physics of GW signals that may explain the data, with a focus on GW backgrounds (GWBs) considering both astrophysical and cosmological origins. We describe how:

- An astrophysical nanohertz GWB emerges as the superposition of individual signals from inspiraling massive black hole binaries.
- Environment coupling, eccentricity, and sparse sampling cause great uncertainty in the theoretical prediction of the supermassive black hole signal but also offer a way to determine the origin of the signal.
- PTA data offer unprecedented opportunities to constrain high-energy physics beyond the Standard Model by probing early Universe GWBs that originated during or after inflation.
- Different early Universe GWBs, typically created by nonlinear and out-of-equilibrium dynamics, can explain the PTA data (e.g., those from inflation scenarios, first-order phase transitions, or topological defects).
- The PTA detection of GWs opens a new window to explore the Universe, with profound implications for astrophysics and particle physics (probing, e.g., the equation of state of the early Universe, the origin of cosmological perturbations, the nature of dark matter, or whether exotic objects like primordial black holes or cosmic strings exist).

Contents

1. INTRODUCTION

Introduced as "the discovery that shook the world" at the 2017 Nobel Prize ceremony, the direct detection of gravitational waves (GWs) has brought an unprecedented revolution in physics, providing a new cosmic messenger with which to explore the Universe (Abbott et al. 2016). In less than 3 years of effective data taking, up to the O4a catalog, the LIGO and Virgo interferometers observed more than 200 compact binary mergers (Abac et al. 2025), piercing the gravitational Universe for the first time—a realm of dynamical phenomena driven by gravity, often inaccessible via other means, such as electromagnetic (EM) radiation. These observations revealed a restless dark side of the Universe, where black hole binaries whirling at relativistic speeds and crashing into one another offer unique ways to probe stellar evolution (Abbott et al. 2023), the strong-field regime of general relativity, and the underlying theoretical framework of gravity itself (Abac et al. 2025). When GW and EM signals met for the first time in the detection of the merging neutron star binary GW170817 (Abbott et al. 2017), countless new scientific opportunities unfolded, ranging from testing of the γ-ray burst–neutron star binary merger connection and the equation of state of nuclear matter to our understanding of the heavy metal enrichment of the Universe and the measurement of its expansion. Multimessenger astronomy has demonstrated its full potential.

Because of inherent length and seismic noise limitations, ground-based detectors can access only the GW audio band that covers the hertz-to-kilohertz frequency range: a narrow window within an otherwise broad frequency spectrum, which potentially extends over 20 decades, from the inverse age of the Universe up to the megahertz range and possibly higher (Lasky et al. 2016). Similar to how different EM wavelengths reveal different astrophysical phenomena, different GW frequencies probe different mass and energy scales, which could revolutionize our understanding of the Universe. For this reason, a great deal of effort has gone into opening new frequency windows in the so-called GW sky (schematically represented in **Figure 1**). As the sensitivity of ground-based detectors keeps improving, the deployment of LISA (Laser Interferometer Space Antenna) (Amaro-Seoane et al. 2017) is on track to occur in the mid-2030s and is expected to probe the millihertz frequency window. At the same time, observational cosmologists have plans to search (Fuskeland et al. 2023, Wolz et al. 2024) for the imprint of primordial GWs at frequencies below 10^{-15} Hz, in the B-mode polarization of the cosmic microwave background (CMB).

Bridging the gap between the two bands mentioned above, the nanohertz frequency region has been the site of the latest breakthrough in GW astronomy. Several pulsar timing array (PTA) collaborations reported evidence on a GW background (GWB) of unknown origin (Agazie et al. 2023c, EPTA Collaboration et al. 2023b, Reardon et al. 2023, Xu et al. 2023). While a GWB from the incoherent superposition of GWs emitted by inspiraling massive black hole binaries (MBHBs) is naturally expected at PTA frequencies, GWBs can also be produced by various phenomena that may occur in the early Universe, including inflationary scenarios, first-order phase transitions, topological defects, and others (Caprini & Figueroa 2018).

In this review, we consider mostly stochastic GW signals, focusing on GWBs that might explain PTA observations but also delving briefly into deterministic signals. In the next section, we introduce the basic description of stochastic GW signals and their main detection techniques. We focus on PTAs (Section 2.2) but also mention astrometry and the potential of combining the two techniques (Section 2.3). We then summarize the current evidence for a nanohertz signal in PTA data (Section 2.4) and describe the main astrophysical sources of nanohertz GWs, MBHBs (Section 3). We describe in detail the buildup of a GWB from unresolved individual binaries, its main features (Sections 3.1 and 3.2), and the role played by MBHB properties (eccentricity and environmental coupling) in shaping its spectrum and statistical properties (Sections 3.3 and 3.4). We also describe deterministic GWs stemming from MBHBs, including resolvable continuous GWs (CGWs), eccentric bursts, and bursts with memory (Section 3.5). We then comment on the astrophysical interpretation of current PTA observations (Section 3.6). In Section 4 we describe relevant GWBs arising from hypothetical processes in the early Universe. These are referred to as early Universe backgrounds (EUBs), and we classify them into signals coming from inflation (Section 4.2), first-order phase transitions (Section 4.3), cosmic defects (Section 4.4), and others (Section 4.5). We also discuss cosmological deterministic signals (Section 4.6), albeit more briefly. In Section 5 we explore future prospects for disentangling the origin of the signal, before concluding in Section 6.

Pulsar timing array (PTA): systematic monitoring of an ensemble of millisecond pulsars over decades to detect GWs via correlation of their pulses' times of arrival

Massive black hole binaries (MBHBs): with masses $M \approx 10^9$ $M_\odot$, they dominate the nanohertz GW band with wave strains up to $h \sim 10^{-14}$

Resolvable continuous GWs (CGWs): deterministic signals rising on top of the GWB that can dominate the signal at $f \gtrsim 10^{-8}$ Hz

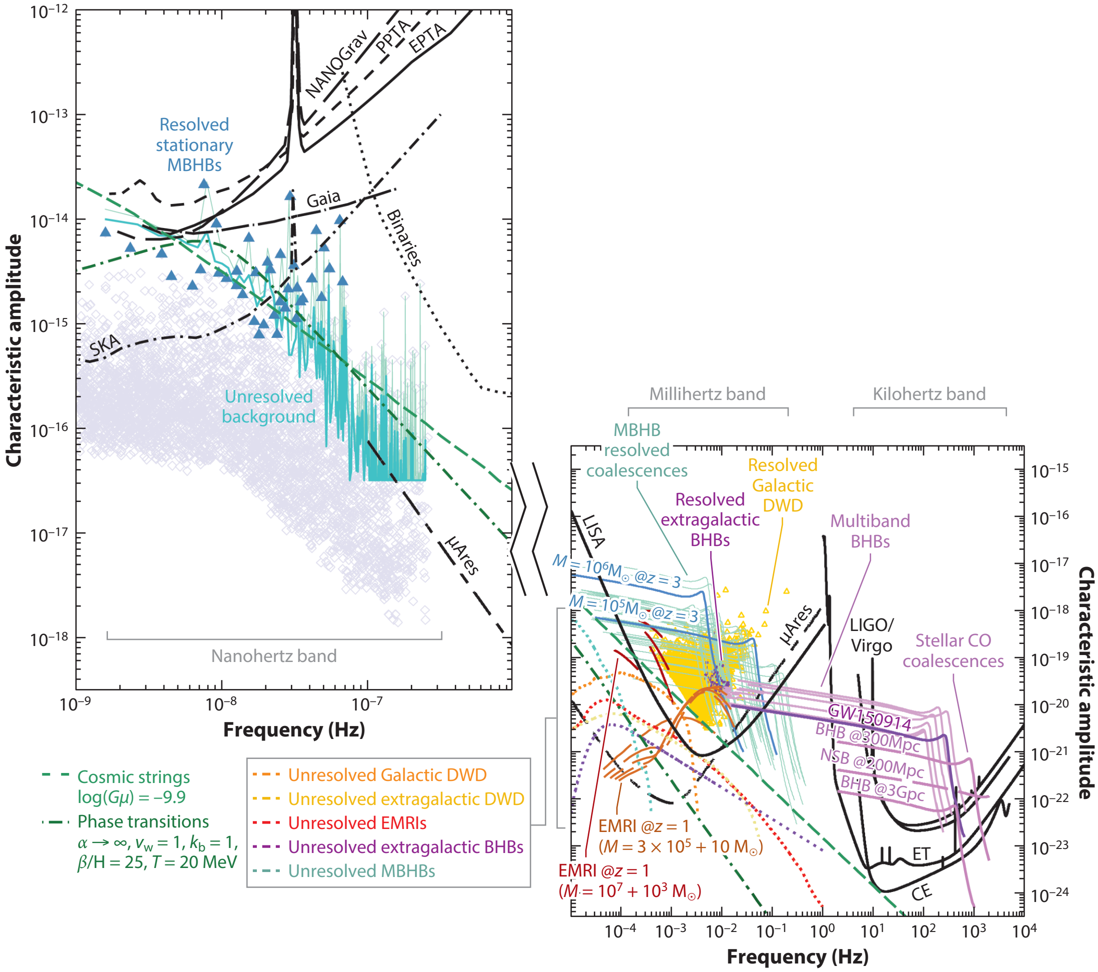


Figure 1

The GW landscape. In the nanohertz-band zoom-in, each lavender diamond represents an individual MBHB. The overall GW signal is marked by the light turquoise jagged line, whereas the residual GWB following subtraction of the loudest source (*dark blue triangles*) at each frequency is marked by the dark turquoise jagged line. Black represents the sensitivity of different nanohertz probes labeled as follows: NANOGrav, PPTA, and EPTA (Agazie et al. 2024); SKA (Truant et al. 2025); Gaia (Vaglio et al. 2025); and binaries (Foster et al. 2025). The figure also shows the millihertz and kilohertz bands for completeness. In each band, GW detectors are depicted in black and various resolved GW signals and unresolved GWBs are shown as colored lines. Abbreviations: BHB, black hole binary; DWD, double white dwarf; EMRI, extreme mass ratio inspiral; GW, gravitational wave; GWB, gravitational wave background; MBHB, massive black hole binary.

Characteristic strain: $h_c(f) = \sqrt{2fS_h(f)}$

Energy density fraction: $\Omega_{GW}(f) \equiv (1/\rho_c)(d\rho_{GW}/d\ln f)$

2. NANOHERTZ GRAVITATIONAL WAVE BASICS: SIGNALS, DETECTORS, AND OBSERVATIONS

We begin with the characterization of a GWB. We then discuss nanohertz GW detection techniques, with a focus on PTAs, including detection principles, noise sources, a definition of the Hellings & Downs (1983) correlation, and a survey of current PTA experiments. We provide a practical dictionary of equations that map between different quantities such as the characteristic strain h_c, energy density fraction Ω_{GW}, power spectral density (PSD) S_h, and root-mean-square

(RMS) residual, which is especially relevant for understanding the PTA results. We also provide simple scalings for estimating the signal-to-noise ratio (S/N) of nanohertz GWs with different detectors. We conclude this section by discussing the current evidence for GWs in PTA data.

Strain PSD: $S_h(f) = (3H_0^2\Omega_{\rm GW})(f)/(4\pi^2 f^3)$

2.1. Basic Concepts Defining a Gravitational Wave Background

In general relativity, a GW propagating in the direction $\hat{\Omega}$ has two polarization modes, h_+ and $h_\times$, which in full generality can be written as a metric perturbation of the form

$$h_{ij}(t,\hat{\Omega}) = e^+_{ij}(\hat{\Omega})h_+(t,\hat{\Omega}) + e^\times_{ij}(\hat{\Omega})h_\times(t,\hat{\Omega}), \tag{1}$$

where the polarization tensors $e^A_{ij}(\hat{\Omega})$ (with $A = +$ or $\times$) are defined as

$$e^+_{ij}(\hat{\Omega}) = \hat{m}_i\hat{m}_j - \hat{n}_i\hat{n}_j, \quad e^\times_{ij}(\hat{\Omega}) = \hat{m}_i\hat{n}_j + \hat{n}_i\hat{m}_j. \tag{2}$$

Here, $\hat{m}$ and $\hat{n}$ are the principal axes that, along with $\hat{\Omega}$, define the GW propagation orthonormal basis. The characterization of h_+ and $h_\times$ depends on the physics of the source of the GW, which is of either astrophysical or cosmological origin.

By construction, a GWB can be expanded—at any point in space—as an incoherent superposition of plane GWs, with different amplitudes, phases, and frequencies. By placing the observer in the origin at $\mathbf{x} = 0$, we can write (Maggiore 2007)

$$h_{ij}(t) = \sum_{A=+,\times}\int_{-\infty}^{\infty}\mathrm{d}f\int_{4\pi}\mathrm{d}\hat{\Omega}\,\tilde{h}(f,\hat{\Omega})e^{-2\pi i f t}e^A_{ij}(\hat{\Omega}). \tag{3}$$

If the GWB is isotropic, stationary, and unpolarized, then

$$\langle\tilde{h}^*_A(f,\hat{\Omega})\tilde{h}'_A(f',\hat{\Omega}')\rangle = \delta(f-f')\frac{\delta^2(\hat{\Omega},\hat{\Omega}')}{4\pi}\delta_{AA'}\frac{1}{2}S_h(f), \tag{4}$$

so that, using Equation 4, we obtain the expectation value[1] (variance) of the GW field as

$$\langle h_{ij}(t)\,h_{ij}(t)\rangle = 4\int_0^{\infty}\mathrm{d}f\,S_h(f) \equiv 2\int_0^{\infty}\mathrm{d}\ln f\,h_c^2(f). \tag{5}$$

Here, we use the fact that $S(-f) = S(f)$ to limit the integral over positive (i.e., physical) frequencies. Equation 5 defines the characteristic strain of the GWB, which is then related to the strain PSD simply as

$$h_c^2(f) = 2f\,S_h(f). \tag{6}$$

A simple characterization of a GWB within a narrow frequency interval is given by a power law (PL) spectrum,

$$h_c(f) = A_{\rm GWB}\left(\frac{f}{f_*}\right)^{\alpha}, \tag{7}$$

where $A_{\rm GWB}$ is the strain amplitude at the reference frequency f_*, which in the PTA band is typically fixed at $f_* = 1$ year^{-1}. As discussed below, the strain of the GWB from a cosmic population of circular, GW-driven binaries can be described as a PL within an extended frequency range (which encompasses the PTA window), with tilt $\alpha = -2/3$.

[1]Throughout this review, we assume a perturbation of the space components of the metric over a flat background. For this reason, we do not differentiate between covariant and contravariant indices, and we always use subscripts.

We can now connect h_c and S_h to the energy density of a GWB, which is given by the 00 component of its stress–energy tensor (Maggiore 2007),

$$\epsilon_{GW} \equiv \rho_{GW} c^2 = T_{00} = \frac{c^2}{32\pi G} \langle \dot{h}_{ij}(t)\, \dot{h}_{ij}(t) \rangle, \qquad 8.$$

where $\dot{h}$ means derivative with respect to the time coordinate. Because in Equation 3 time appears only in the exponent, it is straightforward to demonstrate that

$$\epsilon_{GW} \equiv c^2 \int_0^\infty \mathrm{d}\ln f \left(\frac{\mathrm{d}\rho_{GW}}{\mathrm{d}\ln f} \right) = \frac{\pi c^2}{4G} \int_0^\infty \mathrm{d}\ln f\, f^2 h_c^2(f) = \frac{\pi c^2}{2G} \int_0^\infty \mathrm{d}\ln f\, f^3\, S_h(f). \qquad 9.$$

It follows that

$$\Omega_{GW}(f) \equiv \frac{1}{\rho_c} \frac{\mathrm{d}\rho_{GW}}{\mathrm{d}\ln f} \equiv \frac{8\pi G}{3H_0^2} \frac{\mathrm{d}\rho_{GW}}{\mathrm{d}\ln f} = \frac{2\pi^2}{3H_0^2} f^2 h_c^2(f) = \frac{4\pi^2}{3H_0^2} f^3 S_h(f). \qquad 10.$$

Here, $\Omega_{GW}(f)$ defines the energy content of the GWB per logarithmic frequency, normalized to the critical energy density of the Universe, $\rho_c = 3H_0^2/(8\pi G)$, where $H_0 = 100\ h_0$ km s^{-1} Mpc^{-1} is the Hubble expansion rate and $h_0 \approx 0.67$ (Planck Collaboration et al. 2020a). Therefore, Equation 10 connects the (normalized) spectral energy density spectrum Ω_{GW} to the PSD S_h and the characteristic strain h_c.

The above derivation assumes a GWB that is statistically homogeneous and isotropic, stationary, and unpolarized, as these circumstances apply (at least to a large degree) to most cases (for discussion, see Caprini & Figueroa 2018). However, the possibility of sizable anisotropies might be a key discriminating factor to pin down the origin of the PTA signal as astrophysical, as opposed to cosmological (see Section 5).

2.2. Pulsar Timing Arrays

Millisecond pulsars (MSPs) are magnetized neutron stars that emit beamed radio waves (Lyubarsky 2019). They spin hundreds of times per second along a rotation axis that is typically misaligned with their magnetic axis, making them cosmic lighthouses. If the radio beam intercepts the line of sight to the Earth, we see an extremely precise pulsation: We essentially have a clock in geodesic motion within the gravitational potential of our Galaxy. Any GW passing between the MSP and the Earth perturbs the null geodesic along which the radio waves travel, causing a change in their time of arrival (TOA). MSPs have been monitored for decades with a weekly cadence Δt; therefore, their TOA series are sensitive to GWs with $1/T \approx$ nHz $\lesssim f \lesssim 1/(2\Delta t) \approx$ μHz. As a result, they are ideal nanohertz GW detectors (**Figure 1**) when an ensemble of them are monitored to form a PTA.

2.2.1. Response of a pulsar timing array to a passing gravitational wave.

Radio waves propagate along the GW-perturbed null geodesic, which leads to a characteristic shift[2] $z(t) = [\nu - \nu(t)]/\nu = -\Delta\nu/\nu$. Originally computed for spacecraft tracking (Estabrook & Wahlquist 1975), $z(t)$ takes the form

$$z(t, \hat{\Omega}) = \frac{1}{2} \frac{\hat{p}^i \hat{p}^j}{1 + \hat{p}^i \hat{\Omega}_i} \left\{ e_{ij}^{+}(\hat{\Omega}) \left[h_+(t, 0) - h_+(t - \frac{L}{c}, \mathbf{x}_p) \right] - e_{ij}^{\times}(\hat{\Omega}) \left[h_\times(t, 0) - h_\times(t - \frac{L}{c}, \mathbf{x}_p) \right] \right\}, \qquad 11.$$

[2] We denote the frequency shift by $z(t)$, which is standard in the PTA literature and should not be confused with the cosmological redshift z.

where we place the Earth[3] at $\mathbf{x} = 0$, L is the distance to the pulsar and $\mathbf{x}_\mathrm{p} = L\hat{p}$ is its 3D position vector, $\hat{p}$ is the unit vector pointing from Earth to the pulsar, and $\hat{\Omega}$ is the direction of the incoming wave. Equation 11 shows that the induced shift is due to the difference between the GW strain at Earth at the TOA of the pulse (known as the Earth term) and the strain at the location of the pulsar at the time of emission of the pulse (known as the pulsar term), multiplied by a geometric factor depending on the relative orientation between the pulsar and the Earth. Note also that the pulsar term can equivalently be written by considering the strain at the Earth location at an appropriate time τ, $h_A(t-\tau, 0)$, where $\tau = (L/c)(1 + \hat{p}\cdot\hat{\Omega})$. Equation 11 can be written in compact form as

$$z(t, \hat{\Omega}) = \sum_A F^A(\hat{\Omega})\Delta h_A(t), \qquad 12.$$

where $\Delta h_A(t) = h_A(t, 0) - h_A(t - L/c, L\hat{p})$ and $F^A(\hat{\Omega}) = \hat{p}_i\hat{p}_j e^A_{ij}(\hat{\Omega})/[2(1 + \hat{p}\cdot\hat{\Omega})]$ are the antenna pattern functions representing the geometrical response of the pulsar–Earth system to each GW polarization. This notation separates the physics of the GW from the geometric response due to the projection of the radiation frame defined by the orthonormal unit vectors $(\hat{m}, \hat{n}, \hat{\Omega})$ onto the pulsar–Earth baseline (for the full equations of the coordinate transformations, see Yunes & Siemens 2013). The explicit form of the response functions (or antenna patterns) is universal; that is, it does not depend on the specific GW signal[4] and is given by

$$F^+(\hat{\Omega}) = \frac{1}{2}\frac{(\hat{m}\cdot\hat{p})^2 - (\hat{n}\cdot\hat{p})^2}{1 + \hat{\Omega}\cdot\ \hat{p}}, \qquad F^\times(\hat{\Omega}) = \frac{(\hat{m}\cdot\hat{p})\,(\hat{n}\cdot\hat{p})}{1 + \hat{\Omega}\cdot\hat{p}}. \qquad 13.$$

The timing observable is the residual R, defined as the difference between the expected TOA according to an MSP timing model with and without the presence of the GW. This is simply the integral of $z(t)$ over an observation time T:

$$R(T) = \int_0^T z(t)\,\mathrm{d}t \equiv -\int_0^T \frac{\Delta\nu(t)}{\nu}\mathrm{d}t \sim \int_0^T h(t)\,\mathrm{d}t. \qquad 14.$$

The final tilde ignores geometric factors, highlighting that, for a sinusoidal wave of the form $h \sim h_0\cos(2\pi ft)$, the induced residual is of order $R \approx h_0/(2\pi f)$. Since the best MSPs are monitored with $\sim$100 ns precision, assuming a 10 nHz fiducial frequency, PTAs are sensitive to perturbations $h_0 \sim 10^{-14}$–10^{-15} (**Figure 1**).

2.2.2. Response of a pulsar timing array to a gravitational wave background. If we now consider a GWB expanded as in Equation 3, along with the fact that each component of the expansion produces a frequency shift given by Equation 12, it is straightforward to write the overall shift as (see the definition of τ above Equation 12)

$$z(t) = \sum_{A=+,\times}\int_{-\infty}^{\infty}\mathrm{d}f\int \mathrm{d}^2\hat{\Omega}\, F^A(\hat{\Omega})\tilde{h}_A(f, \hat{\Omega})e^{-2\pi ift}\left[1 - e^{-2\pi if\tau}\right], \qquad 15.$$

[3]In reality, in PTA measurements, the pulse TOAs are computed in the reference frame of the Solar System barycenter by applying an appropriate coordinate transformation that requires knowledge of the Solar System ephemeris. While this process is irrelevant for our discussion above, it is crucial to take it into account in models for timing and GW data analysis.
[4]This is true as long as only general relativity tensor polarizations are considered. In alternative theories of gravity, scalar and vector polarizations might also arise, requiring different response functions (Yunes & Siemens 2013).

where we recognize in the two exponentials the Earth and pulsar terms. Using Equation 4, we find that the expectation value of the frequency shift correlation between MSPs a and b is

Timing residual PSD: $P_h(f) = h_c^2(f)/(12\pi^2 f^3)$

RMS residual spectrum: $\mathrm{RMS}^2(f) = [P_h(f)]/T$

$$\langle z_a(t) z_b(t)\rangle = \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}f S_h(f) \int \frac{\mathrm{d}^2\hat{\Omega}}{4\pi} \sum_{A=+,\times} F_a^A(\hat{\Omega}) F_b^A(\hat{\Omega}), \qquad 16.$$

where we neglect all of the terms involving $e^{-2\pi i f\tau}$ by assuming the short-wavelength limit $\lambda_{\mathrm{GW}} \ll \min(L_a, L_b, \mathbf{x}_a - \mathbf{x}_b)$. This choice is appropriate for PTAs that are sensitive to GW wavelengths of $\sim$1 pc ($\lambda = c/f$, where f is approximately several nanohertz), whereas the closest known MSP is $\sim$150 pc from Earth and typical inter-MSP distances are in the kiloparsec range.[5]

The integral over $\mathrm{d}^2\hat{\Omega}$ in Equation 16, first computed by Hellings & Downs (1983), takes the form

$$C(\zeta_{ab}) \equiv \sum_{A=+,\times} \int \frac{\mathrm{d}^2\hat{\Omega}}{4\pi} F_a^A(\hat{\Omega}) F_b^A(\hat{\Omega}) = \frac{1}{4}\left[1 + \frac{\cos\zeta_{ab}}{3} + 4(1-\cos\zeta_{ab})\ln\left(\sin\frac{\zeta_{ab}}{2}\right)\right], \qquad 17.$$

where ζ_{ab} is the angle between pulsars a and b on the sky. Finally, the observable quantity in the PTAs is the cross-correlation of the timing residuals between two pulsars $r_{ab} = \langle R_a(t) R_b(t)\rangle$, where $R_x(t)$ is defined in Equation 14. Integrating Equation 16 over time, we obtain

$$r_{ab} = \Gamma(\zeta_{ab}) \int_0^{\infty} \mathrm{d}f P_h(f), \qquad 18.$$

where we have renormalized the Hellings & Downs (1983) correlation in Equation 17 as

$$\Gamma(\zeta_{ab}) = \frac{3}{2} C(\zeta_{ab})(1 + \delta_{ab}). \qquad 19.$$

The 3/2 renormalization and the δ_{ab} term ensure that the correlation coefficient is $\Gamma(\zeta_{ab}) = 1$ when $a = b$ (i.e., the GWB has perfect autocorrelation). Note that when $a \neq b$ and $\zeta_{ab} = 0$, $\Gamma_{ab} = 1/2$, because for pulsars at the same location but at different distances only the Earth terms are phase correlated, whereas the pulsar terms act as an additional source of noise. Equation 18 introduces the timing residual PSD as

$$P_h(f) = \frac{h_c^2(f)}{12\pi^2 f^3} = \frac{A_{\mathrm{GWB}}^2}{12\pi^2} \frac{f^{\gamma}}{(1\ \mathrm{year}^{-1})^{2\alpha}}, \qquad \gamma = 3 - 2\alpha, \qquad 20.$$

where in the last equality we have substituted h_c from Equation 7. Note that $P_h(f)$ has dimensions of $[s^3]$, as expected for a spectral density of a time series.

Finally, the total RMS timing residual is obtained by integrating the PSD over frequency: $\mathrm{RMS}^2 = \int \mathrm{d}f P_h(f)$. In PTA analysis, it is standard to divide the frequency domain into bins, $\Delta f_i = f_{i+1} - f_i$ (where $f_i = i/T$, such that $\Delta f_i = 1/T$, $\forall\, i$), and to define the RMS residual spectrum as a function of frequency as

$$\mathrm{RMS}^2(f)_i = \int_{\Delta f_i} \mathrm{d}f P_h(f) \approx P_h(f_i)\Delta f_i = \frac{P_h(f_i)}{T}, \qquad 21.$$

which, as expected, has units of time. The RMS spectrum can then be converted into h_c via Equation 20 and then into S_h and Ω_{GW} via Equation 10.

[5]Note that this assumption might break down if multiple MSPs of a PTA are located within the same globular cluster. In that case, their typical distance is of the order of the cluster core size, which can be less than 1 pc. Here, the correlation of the pulsar terms cannot be neglected, causing an extra correlation in the frequency shifts of the pulsars. Given typical globular cluster distances, the extra correlation can have an impact only for pulsars separated by an angle $\theta_{ab} \ll 1°$.

2.2.3. Pulsar timing array signal-to-noise ratio computations and scalings. The detectability of a signal is generally assessed through the construction of a detection statistic and the definition of an S/N ρ. An in-depth discussion of PTA data analysis techniques is beyond the scope of this review (for an excellent account, see Taylor 2021). Here, we provide simple S/N estimates that stem from contrasting the GWB amplitude within the PTA window with the potential noise affecting the TOAs of the pulses. To do so, we briefly discuss such noise sources.

2.2.3.1. Main noise sources in pulsar timing arrays. The noise PSD in the pulse TOAs is given by

$$P_n(f) = P_{\rm wn} + P_{\rm rn,ac}(f) + P_{\rm rn,c}(f), \tag{22}$$

where:

- $P_{\rm wn} = 2\sigma^2 \Delta t$ is frequency-independent white noise, determined by the cadence of the observations Δt (typically weeks) and by the RMS TOA uncertainty σ. The latter is due mainly to radiometer and jitter noise and has typical values of tens to hundreds of nanoseconds for high-quality MSPs (Cordes 2013).
- $P_{\rm rn,c}(f)$ is chromatic red noise depending on the observed radio-photon frequency ν_γ. It originates from dispersion ($\propto \nu_\gamma^{-2}$) and scattering ($\propto \nu_\gamma^{-4}$) effects in the ionized interstellar medium (Stinebring 2013). Because of its frequency dependence, the chromatic red noise contribution can be separated by use of wideband receivers.
- $P_{\rm rn,ac}(f)$ is achromatic red noise due to intrinsic pulsar spin variations (Shannon & Cordes 2010). These likely arise from complex crust–superfluid interactions and have been observed in several pulsars. Although this noise usually has a steep red spectrum that can mimic a GWB, it is expected to be uncorrelated among different MSPs.

Aside from these noise sources that are uncorrelated among pulsars, additional sources feature specific correlation patterns. Most notably, they include clock offsets that affect all pulsars by inducing a monopole correlated signal, as well as errors in the determination of the Solar System ephemeris, which result in a dipole correlation pattern (Tiburzi et al. 2016). One can show, however, that the GWB-induced Hellings & Downs (1983) correlation has a prominent quadrupolar pattern that is orthogonal to both monopole and dipole (e.g., Taylor 2021).

2.2.3.2. Signal-to-noise ratio for deterministic signals. For deterministic signals, we can apply the match filtering theory and define the S/N as (e.g., Sesana & Vecchio 2010)

$$\rho^2 = \sum_a \rho_a^2, \qquad \rho_a^2 = [R(t)|R(t)] \approx \frac{2}{P_{0,a}} \int_0^{T_a} {\rm d}t\, R_a^2(t), \tag{23}$$

where the sum runs over all pulsars in the PTA, $[\cdot|\cdot]$ is the weighted inner product, and in the last term we assume a monochromatic sine wave. R_a is the residual imprinted in the ath pulsar, and $P_{0,a}$ is the PSD of the noise in the ath pulsar at the frequency of the GW. Equation 23 can also be applied to a generic signal by decomposing it into its harmonics h_n, summing over them, and considering $P(f_n)$ at the observed frequencies f_n of each harmonic (for an application to eccentric MBHBs, see Truant et al. 2025).

The integral in Equation 23 can easily be computed by getting R_a from Equations 11 and 14. If we consider an array of $N_{\rm p}$ equal pulsars dominated by white noise; drop the pulsar term; and average over $F^+, F^\times$, and wave polarizations, then we can solve for the minimum strain h observable at a given S/N threshold:

$$h \approx 6 \times 10^{-15} \frac{\rho}{10} \frac{\sigma}{1\ \mu{\rm s}} \frac{f}{10^{-8}\ {\rm Hz}} \left(\frac{N_{\rm obs}}{500}\right)^{-1/2} \left(\frac{N_{\rm p}}{100}\right)^{-1/2}. \tag{24}$$

Here, $N_{\rm obs} = 500$ is approximately the number of TOAs of each MSP under the assumption of weekly observations for 10 years. Assuming uniform sampling, $N_{\rm obs} \propto T$, we see that the PTA sensitivity to a deterministic GW grows with $\sqrt{N_{\rm p}}$, $\sqrt{T}$, and the RMS precision timing σ. As anticipated, we need strain amplitudes $h > 10^{-15}$ to have a chance of detection. For comparison, typical strains detected in ground-based detectors are $\sim 10^{-22}$–10^{-21}.

2.2.3.3. Signal-to-noise ratio for a gravitational wave background. For stochastic signals, the strategy is to detect the Hellings & Downs (1983) cross-correlated power in the array. The S/N can be written as (Rosado et al. 2015)

$$\rho^2 = 2 \sum_{a=1,M} \sum_{b>a} T_{ab} \int \frac{\Gamma_{ab}^2 P_b^2}{P_{n,\hat{ab}}^2} \mathrm{d}f, \tag{25}$$

where the sums run over all pulsar pairs, T_{ab} is the time span of overlapping observations for pulsars a and b,[6] and Γ_{ab} is defined by Equation 19. The correlated noise term is given by

$$P_{n,\hat{ab}}^2 = P_{n,a}P_{n,b} + P_b[P_{n,a} + P_{n,b}] + P_b^2(1 + \Gamma_{ab})^2. \tag{26}$$

We assume that all pulsars are equal so that $P_{n,x} = P_n$ for each pulsar x. Equation 25 has a different scaling depending on the relative magnitude of P_b and P_n, that is, on whether we are in the weak or strong signal regime. As long as $P_b < P_n$, one can ignore P_b in the denominator of Equation 25. By assuming a red spectrum in the form of Equation 7, and by using an average value $\sqrt{\langle \Gamma^2 \rangle} = 1/(4\sqrt{3})$, we find that the integral is trivial, and we can solve for $A_{\rm GWB}$ to obtain the approximate scaling for the minimum detectable amplitude at $f = 1$ year^{-1}:

$$A_{\rm GWB} \approx 9 \times 10^{-14} \left(\frac{\rho}{10}\right)^{1/2} \frac{\sigma}{1\ \mu\text{s}} \left(\frac{T}{1\ \text{year}}\right)^{\alpha-1} \left(\frac{N_{\rm obs}}{50}\right)^{-1/2} \left(\frac{N_{\rm p}}{100}\right)^{-1/2}. \tag{27}$$

Here we assume weekly observations for 1 year. By observing with the same cadence, we find $A_{\rm GWB} \propto T^{\alpha-3/2}$. A GWB amplitude of $A_{\rm GWB} \approx 10^{-15}$ can be detected on a $T = 10$ year timescale. Once $P_b \gtrsim P_n$, it dominates the noise budget and a crude approximation to the S/N becomes

$$\rho \approx 0.15\, T^{1/2} N_{\rm p} (f_{\rm max} - f_{\rm min})^{1/2} \approx 0.15\, N_{\rm p} \Sigma^{1/2}, \tag{28}$$

where $f_{\rm min} = 1/T$ and $f_{\rm max}$ is the maximum frequency for which the condition $P_b > P_n$ is satisfied. For the last approximation, we divide the frequency domain into resolution bins $\Delta f = 1/T$ and take Σ as the number of frequency bins for which $P_b > P_n$ (usually a few). While the S/N grows linearly with $N_{\rm p}$, it grows only as $\sqrt{T}$ because the uncorrelated part of the GWB acts as an effective common red noise source. Note also that, as long as $P_b > P_n$, the RMS residual σ plays a marginal role by weakly affecting the $f_{\rm max}$ limit (or, alternatively, Σ). The scaling in Equation 28 applies regardless of the GWB spectral shape.

2.2.4. Current pulsar timing array efforts. There exist three main historical PTA collaborations: the European PTA (EPTA; EPTA Collaboration et al. 2023a), the Parkes PTA (PPTA; Zic et al. 2023), and the North American Nanohertz Observatory for Gravitational Waves (NANOGrav; Agazie et al. 2023d). They have been taking data for more than 100 MSPs combined for around 2 decades, and in 2010 they signed an agreement to form the International PTA (IPTA; Perera et al. 2019), a consortium of consortia aiming to enhance the experiments' sensitivity by combining regional PTA data. The IPTA effort was recently joined by the Indian PTA

[6]Note that, in general, MSPs have different observation coverage, depending on when they were discovered, the schedule requirement at observatories, and so forth.

(InPTA; Rana et al. 2025) and by the MeerKAT PTA (MPTA; Miles et al. 2023), the PTA experiment conducted at the MeerKAT radio telescope in South Africa, which has been taking data since 2018. To complete the picture, the Chinese PTA (CPTA; S. Chen et al. 2025) began its observational campaign in 2019 and is now joining the IPTA, and the gamma-ray PTA is exploiting observations with the Fermi telescope (Atwood et al. 2009) to time MSPs in an energy band that is unaffected by the ionized interstellar medium (Valtolina et al. 2026). We discuss the latest PTA results in Section 2.4. **Figure 1** depicts the sensitivity of current and future PTAs.

Astrometry: precise measurement of stellar position and proper motion; can be used to detect GWs arising from the induced perturbation in the star's apparent position

2.3. Other Nanohertz Gravitational Wave Detection Techniques

Although PTAs are currently the most sensitive probes of GWs at nanohertz frequencies, the literature has explored other techniques that could realistically contribute to low-frequency GW science in the future. In this section, we briefly describe these techniques.

2.3.1. Astrometry. Just as GWs alter the path of radio photons emitted by MSPs, they also perturb the path of light coming from stars, causing measurable shifts in their apparent positions in the sky. The derivation of the effect is similar to the frequency shift induced in MSPs' TOAs. If $\hat{n}$ is the incoming photon wave vector, then $\hat{\Omega}$ is the propagation direction of a plane wave GW h_{ij}, and the distance to the star is much greater than the GW wavelength. The astrometric deflection is then (Book & Flanagan 2011)

$$\delta n^i = \left(\frac{n^i + \Omega^i}{2(1+\hat{n}\cdot\hat{\Omega})} n^j n^k - \frac{1}{2}\delta^{ik} n^j \right) h_{jk}, \qquad 29.$$

with h_{jk} evaluated at the Earth location. The *Gaia* mission (Gaia Collaboration et al. 2016) is releasing precise astrometric measurements of billions of stars in the Milky Way over 10 years with an average monthly cadence. As a result, this data set is sensitive to GWs in the 3–100 nHz frequency range, which overlaps with that of PTAs (**Figure 1**).

The δn_i in astrometry are equivalent to $z = -\delta\nu/\nu$ in PTAs. However, there are two degrees of freedom on the celestial sphere, and the expectation value of the correlation between astrometric deflections from two stars a and b is not a scalar function. It takes the form

$$\langle \delta n_a^i \delta n_b^j \rangle \propto \mathcal{T}(\theta) U^{ij}(\hat{n}_a, \hat{n}_b) \times S_h(f), \qquad 30.$$

where $\mathcal{T}(\theta)$ is the astrometric analog of the Hellings & Downs (1983) curve, depending only on the separation angle between the two stars θ, whereas $U^{ij}(\hat{n}_a, \hat{n}_b)$ is a function of the unperturbed star directions that depends on the choice of coordinates (Mihaylov et al. 2018).

A limitation of Equation 30 is its reliance on a fixed reference frame that is extremely hard to determine practically (Pardo et al. 2023, Vaglio et al. 2025). Given a pair a of two stars separated by an angle ψ_{12}^a and a pair b of stars separated by another angle ψ_{12}^b, both ψ_{12}^a and ψ_{12}^b experience a small change due to the GW passage, and one can evaluate the expectation value of the correlation of those changes: $\langle \delta\cos\psi_{12}^a\, \delta\cos\psi_{12}^b \rangle$. Notably, when averaged over the relative orientations of the two star pairs, this correlation is identical to the Hellings & Downs (1983) curve. Note also that both the absolute δn_i and the relative $\delta\cos\psi_{12}^b$ astrometric shifts can be cross-correlated with the frequency shift z to obtain the observables $\langle z\,\delta n_i \rangle$ and $\langle z\,\delta\cos\psi_{12}^b \rangle$, allowing one to combine astrometric and PTA measurements in a coherent framework (Qin et al. 2019, Cruz et al. 2025, Vaglio et al. 2025).

The S/N of a GWB in an astrometric experiment has the same form as Equation 25, except that (*a*) the sums run over all possible pairs of stars (or pairs of stellar pairs, if we consider relative astrometry); (*b*) the Hellings & Downs (1983) curve is replaced by the appropriate astrometric correlation [which is always a $\mathcal{O}(1)$ factor]; (*c*) $P_n = 2\sigma_\star^2 \Delta t_\star$, where $\sigma_\star$ and $\Delta t_\star$ are the astrometric precision and cadence of the measurements, respectively; and (*d*) P_h is replaced by $S_h = h_c^2/(2f)$ because we are dealing with angular rather than time measurement. By solving for $A_{\rm GWB}$ as above, we obtain

$$A_{\rm GWB} \approx 9 \times 10^{-14} \left(\frac{\rho}{10}\right)^{1/2} \frac{\sigma_\star}{0.2\ \rm mas} \left(\frac{T}{1\ \rm year}\right)^{\alpha} \left(\frac{N_{\rm obs}}{20}\right)^{-1/2} \left(\frac{N_\star}{10^9}\right)^{-1/2}, \qquad 31.$$

normalized to the *Gaia* mission's performance (Gaia Collaboration et al. 2016). Strikingly, with the current precision, astrometry needs to monitor $\mathcal{O}(10^9)$ stars to achieve a performance similar to that for PTAs. Note also that, if the same observing cadence is retained, $A_{\rm GWB} \propto T^{\alpha-1/2}$, meaning that astrometry benefits less than PTAs from longer surveys. This is because the latter surveys are sensitive to temporal rather than angular perturbations, resulting in a steeper sensitivity as a function of frequency ($h_c \propto f^{3/2}$ for PTAs versus $h_c \propto f^{1/2}$ for astrometry for $f > 1/T$; **Figure 1**).

When combining PTAs and astrometry, we find that the total S/N is $\rho^2_{\rm tot} = \rho^2_{\rm PTA} + \rho^2_\star + \rho^2_{\star/\rm PTA}$, where the last term is given by the correlation of the two measurements and can also be written by appropriately modifying Equation 25. Beyond the simple scalings derived here, the sensitivity of PTAs, astrometric measurements, and their correlations have been carefully explored by various authors (e.g., Qin et al. 2019, Cruz et al. 2025, Vaglio et al. 2025).

2.3.2. Laser interferometry. Laser interferometry is the most successful technique used to detect kilohertz GWs and is anticipated to open the millihertz frequency band in the mid-2030s with LISA (Amaro-Seoane et al. 2017), TianQin (Luo et al. 2016), and Taiji (Hu & Wu 2017). Going to lower frequencies would require building a Solar System–scale detector, although futuristic mission concepts for satellite constellations forming a triangle the size of the Earth or Mars orbit have been studied (Sesana et al. 2021). Pushing the interferometry detection window deep into the microhertz range, these detectors might achieve a sensitivity of $h_c \approx 10^{-16}$ around 100 nHz, which would complement PTAs at high frequencies (**Figure 1**).

2.3.3. Binary resonances. GWs can resonantly interact with astrophysical binaries, causing small deviations in their orbital period (Hui et al. 2013, Foster et al. 2025). The order-of-magnitude effect that can be detected depends on the precision of the determination of the binary period, its periastron passage, and the number of observed passages (Hui et al. 2013). **Figure 1** shows the projected sensitivity this technique could achieve by the end of the 2030s by monitoring currently known MSPs in binaries (Foster et al. 2025).

2.4. Current Evidence for a Nanohertz Gravitational Wave Signal from Pulsar Timing Array Data

On June 29, 2023, the major PTAs forming the IPTA (EPTA+InPTA,[7] NANOGrav, and PPTA), along with CPTA, coordinated four independent publications reporting evidence for a Hellings & Downs (1983) correlated signal in their respective data sets (Agazie et al. 2023c, EPTA Collaboration et al. 2023b, Reardon et al. 2023, Xu et al. 2023). MPTA followed in early 2025 with a report of evidence for a correlated signal in its data (Miles et al. 2025). The signal is consistent with an isotropic GWB, with evidence ranging from 2σ to 4σ, depending on the data set (**Table 1**). While the EPTA+InPTA, NANOGrav, and PPTA spectra are well-constrained and broadly consistent (**Figure 2**), the CPTA and MPTA data are much less constraining, mostly because of the shorter time span of these data sets (roughly 3 and 4.5 years, respectively, compared with 10 to 20 years for other PTAs). While the FAST (CPTA) and MeerKAT (MPTA) telescopes have a superior sensitivity that allows them to detect a correlated signal with a much

[7]For this analysis, EPTA and InPTA worked together by joining their data into a common EPTA+InPTA data set.

Table 1 A_{GWB} **and** γ **of the HD-correlated common red signal measured by several PTAs**[a]

PTA	$A_{\mathrm{GWB}}/10^{15}$	γ	$A_{\mathrm{GWB}}(\gamma = 13/3)/10^{15}$
EPTA+InPTA	$11.48^{+8.02}_{-7.68}$	$2.71^{+1.18}_{-0.73}$	$2.5^{+0.7}_{-0.7}$ (90%)
NANOGrav[b]	$6.4^{+4.2}_{-2.7}$	$3.2^{+0.6}_{-0.6}$	$2.4^{+0.7}_{-0.6}$ (90%)
PPTA	$3.1^{+1.3}_{-0.9}$	$3.9^{+0.4}_{-0.4}$	$2.0^{+0.3}_{-0.1}$ (68%)
CPTA	$3.98^{+33.82}_{-3.97}$	—	$1.99^{+13.85}_{-1.97}$ (95%)
MPTA	$7.5^{+0.8}_{-0.9}$	$3.52^{+1.12}_{-0.9}$	$4.8^{+0.8}_{-0.9}$ (68%)

[a]The table reports both the best-fit amplitude (A_{GWB}) and slope (γ) (second and third columns) and the A_{GWB} measured by fixing $\gamma = 13/3$ (fourth column) (see Equation 20). The dash indicates that the quantity has not been estimated for that data set.

[b]A re-analysis of the NANOGrav dataset presented by Agarwal et al. (2026b) yielded $A_{\mathrm{GWB}}/10^{15} = 4.68^{+3.45}_{-2.39}$, $\gamma = 3.5^{+0.7}_{-0.6}$, $A_{\mathrm{GWB}}(\gamma = 13/3)/10^{15} = 2.1^{+0.6}_{-0.5}$.

Abbreviations: GWB, gravitational wave background; HD, Hellings & Downs (1983); PTA, pulsar timing array.

shorter observing baseline, more observational time is still required for a properly spectral signal characterization.

Figure 2 depicts the PSD of the residuals (see Equation 21) and the 2D $A_{\mathrm{GWB}} - \gamma$ posterior distributions for EPTA+InPTA, NANOGrav, and PPTA, along with a "face value combination" obtained by stacking them.[8] The emerging picture is a GWB characterized by $A_{\mathrm{GWB}} \in [2 \times 10^{-15}–3 \times 10^{-15}]$ (for a reference $\gamma = 13/3$) and $\gamma \in [2.5–4]$ (1σ confidence). At the current sensitivity level, the signal is broadly consistent with an isotropic, Gaussian GWB; searches for spectral features (Agazie et al. 2025a), anisotropies (Agazie et al. 2023e), and resolved deterministic signals (Agazie et al. 2023a, EPTA Collaboration et al. 2024b) yielded no significant results. Both EPTA and NANOGrav published comprehensive papers dissecting these results and their possible interpretations (Afzal et al. 2023, Agazie et al. 2023b, EPTA Collaboration et al. 2024a), which we discuss in the next two sections.

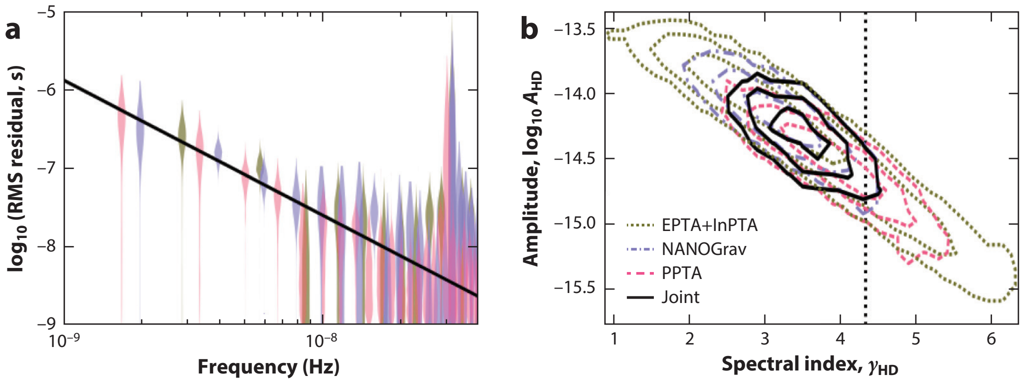


Figure 2

(*a*) Free spectrum showing the power spectral density of the RMS residuals measured by EPTA+InPTA, NANOGrav, and PPTA. (*b*) 2D $A_{\mathrm{GWB}} - \gamma$ posterior distribution of the GWB parameters as determined from the data (68%, 95%, and 99.7% confidence intervals) along with the joint posterior. The vertical dotted line represents $\gamma = 13/3$ for reference. Abbreviations: GWB, gravitational wave background; HD, Hellings & Downs (1983); RMS, root mean square. Figure adapted from Agazie et al. (2024) (CC BY 4.0).

[8]This simple combination is done by taking a flat prior across the three PTAs and assuming the three experiments are independent, which is not the case, as several pulsars are shared among different arrays.

3. ASTROPHYSICAL NANOHERTZ GRAVITATIONAL WAVE SIGNALS

Primary astrophysical GW sources are compact object binaries. GW emission causes energy loss from the pair, and the binary orbit shrinks until the two compact objects eventually coalesce. For a circular binary of black holes with total mass M, the inspiraling phase lasts up to the innermost stable circular orbit (ISCO), which corresponds to an emitted GW frequency

$$f_{\rm ISCO} \approx \left(6^{3/2}\pi \frac{GM}{c^3}\right)^{-1} \approx 4\times 10^{-6}\left(\frac{M}{10^9\,{\rm M}_\odot}\right)^{-1}\ {\rm Hz}. \qquad 32.$$

GWs emitted during the inspiral have $f < f_{\rm ISCO}$. The most massive black holes observed in the Universe have masses of $\sim 10^{10}$ $M_\odot$ (Kormendy & Ho 2013), for which $f_{\rm ISCO} \gtrsim 10^{-7}$ Hz. Since we are interested in nanohertz GWs, and since both PTAs and astrometric techniques are most sensitive at $f < 1$ year$^{-1} \approx 3\times 10^{-8}$ Hz, for the entire astrophysically relevant mass range the signals of interest are adiabatically inspiraling binaries that move at subrelativistic speeds $v \ll c$, thereby allowing the use of the quadrupolar approximation.

3.1. Quadrupolar Gravitational Wave Emission from a Generic Eccentric Binary

The most general system considered in this section is an eccentric nonspinning binary with total mass $M = M_1 + M_2$ and mass ratio $q = M_2/M_1 < 1$, orbiting at a semimajor axis a with orbital frequency $\omega_{\rm K} = 2\pi f_{\rm K} = (GM/a^3)^{1/2}$. The system emits GWs at integer harmonics of the orbital frequency, and $h_{+,\times}$ can be written as (e.g., Taylor 2021)

$$h_+(t) = \sum_{n=1}^{\infty} -\left(1+\cos^2\iota\right)\left[a_n\cos\left(2\gamma(t)\right) - b_n\sin\left(2\gamma(t)\right)\right] + \left(1-\cos^2\iota\right)c_n, \qquad 33.$$

$$h_\times(t) = \sum_{n=1}^{\infty} 2\cos\iota\left[b_n\cos\left(2\gamma(t)\right) + a_n\sin\left(2\gamma(t)\right)\right], \qquad 34.$$

where

$$\begin{aligned} a_n &= -(n/2)A(f_{\rm K})\Big[J_{n-2}(ne) - 2eJ_{n-1}(ne) + (2/n)J_n(ne) \\ &\qquad + 2eJ_{n+1}(ne) - J_{n+2}(ne)\Big]\cos\left[nl(t)\right], \\ b_n &= -(n/2)A(f_{\rm K})\sqrt{1-e^2}\left[J_{n-2}(ne) - 2J_n(ne) + J_{n+2}(ne)\right]\sin\left[nl(t)\right], \\ c_n &= A(f_{\rm K})J_n(ne)\cos[nl(t)]. \end{aligned}$$

Here, n is the harmonic number and $J_n(x)$ are Bessel functions of the first kind (of nth order). The amplitude $A(f_{\rm K})$ is given by

$$\begin{aligned} A(f_{\rm K}) &= 2\frac{\mathcal{M}^{5/3}}{D}\left(2\pi f_{\rm K,r}\right)^{2/3} \equiv 2\frac{\mathcal{M}_z^{5/3}}{D_{\rm l}}\left(2\pi f_{\rm K}\right)^{2/3}, \\ &\approx 3\times 10^{-16}\left(\frac{f_{\rm K,r}}{5\times 10^{-9}\ {\rm Hz}}\right)^{2/3}\left(\frac{\mathcal{M}}{10^9\,{\rm M}_\odot}\right)^{5/3}\left(\frac{D}{1\ {\rm Gpc}}\right)^{-1}. \end{aligned} \qquad 35.$$

Clearly, the GW amplitude can be expressed in the source frame by using ($f_{\rm r}$, $t_{\rm r}$, $\mathcal{M}$, D) or in the observer frame by using (f, t, $\mathcal{M}_z$, $D_{\rm l}$), as in Equation 35. When dealing with GW

NOTE ON REDSHIFT

Due to the Universe expansion, GWs are redshifted just like EM waves. A GW signal can be expressed either in the source rest frame (at the source redshift z) or in the observer frame (at $z = 0$). We adopt the following conventions:

- f is the observed frequency, which is related to the source frame one, f_r, by $f = f_\mathrm{r}/(1+z)$ [and, thus, $\omega = \omega_\mathrm{r}/(1+z)$].
- Similarly, we define the observed t and source frame t_r such that $t = t_\mathrm{r}(1+z)$.
- All plain mass symbols (e.g., $\mathcal{M}$) are defined in the source frame, whereas observed ones carry a z label (e.g., $\mathcal{M}_z$). The relation between the two is $\mathcal{M}_z = \mathcal{M}(1+z)$. Thus, in the observer frame we talk about redshifted masses.
- We adopt the comoving distance D in the source frame, whereas we adopt the luminosity distance $D_\mathrm{l} = D(1+z)$ in the observer frame.

observations, the observer frame is generally preferable; however, rest frame quantities are more intuitive when connecting GW signals with the physical properties of the sources. For example, the mass function of the population is physically more relevant than the redshifted mass function. Below, we generally use rest frame quantities.

In the above set of equations, $\mathcal{M} = (M_1 M_2)^{3/5}/M^{1/5}$ is the chirp mass; ι is the inclination angle, defined as the angle between the GW propagation direction and the binary orbital angular momentum $\hat{L}$; $l(t)$ is the binary mean anomaly, $l(t) \equiv l_0 + 2\pi \int_{t_0}^{t} f_\mathrm{K}(t')\mathrm{d}t'$ [note that $f_\mathrm{K}(t')$ is the observed Keplerian frequency; see the sidebar titled Note on Redshift]; and, finally, $\gamma(t)$ is the angle that measures the direction of the pericenter with respect to the direction $\hat{x}$, defined as $\hat{x} \equiv (\hat{\Omega} + \hat{L}\cos i)/\sqrt{1-\cos^2 i}$, where $\hat{\Omega}$ is the unit vector defining the GW propagation direction (Barack & Cutler 2004). GWs extract energy from the binary at a rate

$$\frac{\mathrm{d}E}{\mathrm{d}t_\mathrm{r}} = L_\mathrm{GW} = \frac{32}{5}\frac{G^{7/3}}{c^5}(2\pi f_\mathrm{K,r}\mathcal{M})^{10/3}F(e), \tag{36}$$

causing the binary orbital elements to evolve at a rate

$$\frac{\mathrm{d}a}{\mathrm{d}t_\mathrm{r}} = -\frac{64}{5}\frac{G^3}{c^5}\frac{\mathcal{M}^{5/3}M^{4/3}}{a^3}\mathcal{F}(e) \quad\rightarrow\quad \frac{\mathrm{d}f_\mathrm{K,r}}{\mathrm{d}t_\mathrm{r}} = \frac{96}{5}(2\pi)^{8/3}\left(\frac{G\mathcal{M}}{c^3}\right)^{5/3} f_\mathrm{K,r}^{11/3}\mathcal{F}(e),$$
$$\frac{\mathrm{d}e}{\mathrm{d}t_\mathrm{r}} = -\frac{(2\pi)^{8/3}}{15}\left(\frac{G\mathcal{M}}{c^3}\right)^{5/3} f_\mathrm{K,r}^{8/3}\mathcal{G}(e). \tag{37}$$

Here, $\mathcal{F}(e)$ and $\mathcal{G}(e)$ take the form

$$\mathcal{F}(e) = \sum_{n=0}^{\infty} g(n,e) = \frac{1 + (73/24)e^2 + (37/96)e^4}{(1-e^2)^{7/2}}, \mathcal{G}(e) = \frac{304e + 121e^3}{(1-e^2)^{5/2}}, \tag{38}$$

and

$$g(n,e) = \frac{n^4}{32}\Bigg[\left(J_{n-2}(ne) - 2eJ_{n-1}(ne) + \frac{2}{n}J_n(ne) + 2eJ_{n+1}(ne) - J_{n+2}(ne)\right)^2$$
$$+ (1-e^2)\left(J_{n-2}(ne) - 2J_n(ne) + J_{n+2}(ne)\right)^2 \frac{4}{3n^2}J_n^2(ne)\Bigg]. \tag{39}$$

Equation 37 indicates that as the binary shrinks, increasing its frequency, it also circularizes. It is therefore expected that comparable mass binaries in the late stage of evolution are rather circular, which is why most early literature on nanohertz GW signals assumed circular binaries (Rajagopal & Romani 1995, Jaffe & Backer 2003, Wyithe & Loeb 2003, Sesana et al. 2004).

3.2. Case Study: The Gravitational Wave Background from a Cosmic Population of Gravitational Wave–Driven Circular Massive Black Hole Binaries

We begin this subsection by discussing circular binaries ($e = 0$). In this case, GWs are emitted only at the second harmonic $n = 2$ and, therefore, at $f = 2f_{\rm K}$. The strain in Equation 33 takes the form

$$h_+(t) = (1 + \cos^2 \iota) A \cos(\omega t + \Phi_0), h_\times(t) = -2 \cos \iota\, A \sin(\omega t + \Phi_0), \tag{40}$$

where $\omega = 2\pi f = 2\omega_{\rm K}$ is the wave angular frequency and Φ_0 is an irrelevant initial phase. Here, $\mathcal{F}(e = 0) = 1$ and $\mathcal{G}(e = 0) = 0$, implying that circular binaries stay circular.

During the adiabatic inspiral, the binary emits GWs with energy spectrum

$$\frac{{\rm d}E}{{\rm d}f_{\rm r}} = \frac{{\rm d}E}{{\rm d}t_{\rm r}}\frac{{\rm d}t_{\rm r}}{{\rm d}f_{\rm r}} = \frac{\pi}{3G}\frac{(G\mathcal{M})^{5/3}}{(\pi f_{\rm r})^{1/3}}. \tag{41}$$

We consider a cosmic population of compact binaries in circular orbits. The population can be generic, but, as discussed below, the nanohertz GW band is dominated by very massive binaries (i.e., MBHBs), so we make all our calculations specific to this case. We define ${\rm d}^2n/({\rm d}z\,{\rm d}\mathcal{M})$ as the merger rate density (i.e., the number of mergers per unit of comoving volume, per unit of redshift, and per unit of chirp mass), and we use ${\rm d}E/{\rm d}\ln f_{\rm r}$ to denote the energy emitted per unit logarithmic rest frame frequency by each source. The GW emission of this population across cosmic time generates a GWB with energy density (Phinney 2001)

$$\epsilon_{\rm GW} = \int {\rm d}\ln f \int {\rm d}z \int {\rm d}\mathcal{M} \frac{1}{1+z}\frac{{\rm d}^2 n}{{\rm d}z\,{\rm d}\mathcal{M}}\frac{{\rm d}E}{{\rm d}\ln f_{\rm r}}, \tag{42}$$

where $1 + z$ is due to the Universe expansion. Combining Equations 9 and 42, we obtain

$$h_{\rm c}^2(f) = \frac{4G}{\pi c^4}\frac{1}{f^2}\int {\rm d}z \int {\rm d}\mathcal{M} \frac{{\rm d}^2 n}{{\rm d}z\,{\rm d}\mathcal{M}}\frac{{\rm d}E}{{\rm d}\ln f_{\rm r}}. \tag{43}$$

If we substitute ${\rm d}E/{\rm d}f$ from Equation 41, assume a monochromatic mass function, and write ${\rm d}^2n/({\rm d}z\,{\rm d}\mathcal{M}) = ({\rm d}n/{\rm d}z)\delta(\mathcal{M} - \mathcal{M}_0)$, then, by defining $n_0 = \int {\rm d}z \int {\rm d}\mathcal{M}\, {\rm d}^2n/({\rm d}z\,{\rm d}\mathcal{M})$, we obtain

$$h_{\rm c}^2(f) = \frac{4(G\mathcal{M}_0)^{5/3}}{3\pi^{1/3}c^2}\frac{1}{f^{4/3}} n_0 \langle (1+z)^{-1/3} \rangle, \tag{44}$$

where n_0 can be interpreted as today's number density of merger relics, and $\langle(1 + z)^{-1/3}\rangle$ is averaged over the redshift distribution of the coalescences producing those relics (Phinney 2001). For example, if we take a coalescence rate that is constant in time, then $\langle(1 + z)^{-1/3}\rangle \approx 0.8$, and if we define $\chi(z) = \langle(1 + z)^{-1/3}\rangle/0.8$, then finally we can write

$$h_{\rm c}(f) \approx 10^{-15} \left(\frac{f}{1\ {\rm year}^{-1}}\right)^{-2/3} \left(\frac{\mathcal{M}}{10^9\ {\rm M}_\odot}\right)^{5/6} \left(\frac{n_0}{10^{-4}\ {\rm Mpc}^{-3}}\right)^{1/2} \chi(z)^{1/2}. \tag{45}$$

As anticipated in Section 2.1, the GWB from a cosmic population of circular, GW-driven, inspiraling compact objects is described by a single PL, $h_{\rm c} \propto f^{-2/3}$. The normalization $n_0 = 10^{-4}$ Mpc^{-3} in Equation 45 is an estimate of the number density of massive black holes of a billion solar masses residing in massive elliptical galaxies today. Equation 45 states that, if each has undergone one major merger since $z = 1$ (which is very reasonable, as discussed below), then the expected GWB strain amplitude at PTA frequencies is $A_{\rm GWB} \sim 10^{-15}$.

3.2.1. Relation between massive black hole binary mass function and characteristic strain. If the late growth of massive black holes in the Universe is merger dominated, then the relic population must correspond to today's massive black hole mass function, $\phi(M) = \mathrm{d}n/\mathrm{dlog}M$ (Sato-Polito et al. 2024). Assuming that today's massive black holes are the result of a single equal-mass merger, Equation 43 can be written as

$$h^2_{\mathrm{c},1}(f) = \int \mathrm{dlog}M \frac{(GM)^{5/3}}{3\pi^{1/3}c^2} \frac{1}{f^{4/3}}\phi(M), \tag{46}$$

where we assume that mergers occur at $z \ll 1$, and the subscript 1 indicates that each massive black hole today is considered the result of a single merger event. By assuming that massive black holes today are the product of a series of subsequent equal-mass mergers, we can reconstruct them backward by doubling the number of involved massive black holes while halving their masses at each round. In the most favorable toy scenario, the series converges to $h_c^{\max} \approx 2h_{\mathrm{c},1}$, which is the upper limit of a GWB produced by a given $\phi(M)$ (I. Ferranti, A. Sesana & D. Izquierdo-Villalba, manuscript in preparation).

Since $\mathrm{d}E/\mathrm{d}\ln f_\mathrm{r} \propto \mathcal{M}^{5/3}$, and since $\phi(M)$ is generally a concave function, Equation 46 implies that the main contribution to h_c^2 (and thus to h_c) comes from masses around the point where $\log\phi(M)$ is tangent to a line with slope $-5/3$. This generally occurs at $M \gtrsim 10^9\ \mathrm{M}_\odot$, where $\phi(M) \approx 10^{-4}\ \mathrm{Mpc}^{-3}$ (see Sato-Polito et al. 2024, their figure 1). By reverse engineering the argument leading to Equation 46, we can consider as a first-order approximation that the GWB from MBHBs is dominated by massive black hole mergers with $\mathcal{M} \sim 10^9\ \mathrm{M}_\odot$, justifying the normalizations used in Equation 45.

3.2.2. The nature of the nanohertz gravitational wave signal from massive black hole binaries. Although a useful approximation, Equation 43 is an accurate description of the GWB only when (*a*) the merger rate density n is large and (*b*) $\mathrm{d}E/\mathrm{dln}f_\mathrm{r}$ is emitted on a timescale much shorter than the duration of observations. Neither condition holds for MBHBs, which feature $n \ll 1$ and evolve at nanohertz frequencies on timescales of thousands of years, while observations (PTAs in particular) have been running for only a couple of decades.

It is more useful to think of nanohertz GW detection as capturing a snapshot of the sky where a finite, discrete number of MBHBs are caught at different stages of their evolution. Therefore, instead of thinking in terms of cosmic merger rate density, it is more useful to think in terms of the number of systems with specific properties emitting GWs at a given frequency. In fact, early estimates of the nanohertz GW signal were obtained in this way (Rajagopal & Romani 1995, Jaffe & Backer 2003). Sesana et al. (2008) formalized the relation between this method and Equation 43 by rewriting as follows:

$$\frac{\mathrm{d}^2 n}{\mathrm{d}z\,\mathrm{d}\mathcal{M}} = \frac{\mathrm{d}^3 N}{\mathrm{d}z\,\mathrm{d}\mathcal{M}\,\mathrm{d}\ln f_\mathrm{r}} \frac{\mathrm{d}\ln f_\mathrm{r}}{\mathrm{d}t_\mathrm{r}} \frac{\mathrm{d}t_\mathrm{r}}{\mathrm{d}z} \frac{\mathrm{d}z}{\mathrm{d}V_\mathrm{c}}. \tag{47}$$

Substituting Equation 47 into Equation 43, deriving $\mathrm{d}\ln f_\mathrm{r}/\mathrm{d}t_\mathrm{r}$ from Equation 37, and using standard cosmological relations for $\mathrm{d}V_\mathrm{c}/\mathrm{d}z$ and $\mathrm{d}z/\mathrm{d}t_\mathrm{r}$, one arrives at the alternative formulation

$$h^2_\mathrm{c}(f) = \int_0^\infty \mathrm{d}z \int_0^\infty \mathrm{d}\mathcal{M}\, \frac{\mathrm{d}^3 N}{\mathrm{d}z\,\mathrm{d}\mathcal{M}\,\mathrm{d}\ln f_\mathrm{r}}\, h^2(f_\mathrm{r}) \equiv \int_0^\infty \mathrm{d}z \int_0^\infty \mathrm{d}\mathcal{M}\, \frac{\mathrm{d}^3 N}{\mathrm{d}z\,\mathrm{d}\mathcal{M}\,\mathrm{d}t_\mathrm{r}} \frac{\mathrm{d}t_\mathrm{r}}{\mathrm{d}\ln f_\mathrm{r}}\, h^2(f_\mathrm{r}). \tag{48}$$

Here, h is the inclination-polarization-averaged strain amplitude, given by

$$h(f_\mathrm{r}) = \sqrt{\frac{32}{5}} \frac{(G\mathcal{M})^{5/3}}{c^4 D} (\pi f_\mathrm{r})^{2/3}. \tag{49}$$

Both ways of writing h_c^2 in Equation 48 are useful. The left side highlights that h_c^2 is simply the integral over mass and redshift of the individual h^2s of the sources emitting per unit logarithmic frequency. The right side separates the role of the MBHB merger rate, $\mathrm{d}^3N/(\mathrm{d}z\,\mathrm{d}\mathcal{M}\,\mathrm{d}t_\mathrm{r})$, from its dynamical evolution, captured by the residence time, $t_\mathrm{res} = \mathrm{d}t_\mathrm{r}/\mathrm{d}\ln f_\mathrm{r}$. This separation is particularly useful because cosmological models for MBHB evolution usually output merger rates of the form $\mathrm{d}^3N/(\mathrm{d}z\,\mathrm{d}\mathcal{M}\,\mathrm{d}t_\mathrm{r})$, from which the GWB can be reconstructed via a specific model for the MBHB evolution, as discussed in Section 3.3.

3.2.3. Discretization. In reality, Equation 48 is the continuum limit of a discrete distribution of sources. Moreover, if we collect data for a time span T, the nominal frequency resolution bin is $\Delta f = 1/T$, which means that the GWB spectrum is evaluated in discrete frequency bins, $\Delta f_j = [j/T, (j+1)/T]$, centered at $f_j = (2j+1)/(2T)$ for $j = 1, \ldots, N_\mathrm{bin}$, where N_bin is the number of frequency bins considered. Equation 48 can then be discretized as

$$h_\mathrm{c}^2(f_j) = \sum_z \sum_\mathcal{M} \frac{\Delta^3 N}{\Delta z \Delta \mathcal{M} \Delta f_j} f_j h^2(f_\mathrm{r}) \Theta\left(\frac{f_\mathrm{r}}{1+z}, \Delta f_j\right) \Delta z \Delta \mathcal{M}, \qquad 50.$$

where $\Theta = 1$ if $f_\mathrm{r}/(1+z) \in \Delta f_j$ and $\Theta = 0$ otherwise, and we use the fact that $\mathrm{d}\ln f_\mathrm{r} = \mathrm{d}\ln f$.[9] If we now use index i to label each discrete MBHB contributing to the cosmic population, we can simplify Equation 50 to

$$h_\mathrm{c}^2(f_j) = \sum_{f\in\Delta f_j} \frac{f_j h_i^2(f_\mathrm{r})}{\Delta f_j} = \sum_{f\in\Delta f_j} h_i^2(f_\mathrm{r}) f_j T = \sum_{f\in\Delta f_j} h_i^2(f_\mathrm{r}) N_\mathrm{cyc} = \sum_{f\in\Delta f_j} h_{\mathrm{c},i}^2(f_\mathrm{r}), \qquad 51.$$

where we define the ith binary characteristic strain $h_{\mathrm{c},i}$ as the inclination-polarization-averaged strain[10] multiplied by the square root of the number of observed GW cycles within the observing time T. Equation 51 simplifies the interpretation of the GWB strain produced by a cosmic population of MBHBs to the bone, highlighting several key facts:

- The GWB strain is the sum in quadrature of the characteristic strains of all binaries that emit at a rest frame frequency $f_\mathrm{r} = f(1+z)$, such that $f \in \Delta f_j$.
- Assuming multiple sources per frequency bin, $h_\mathrm{c}(f_j)$ is, on average, independent of T. While the number of binaries in each frequency bin scales as $N_\mathrm{B}(\Delta f) \propto \Delta f = 1/T$, the contribution of each individual binary to h_c^2 scales as $N_\mathrm{cyc} \propto T$.
- For circular, GW-driven binaries, the number of sources contributing to a frequency bin scales as $\mathrm{d}N/\mathrm{d}f = (\mathrm{d}N/\mathrm{d}t)(\mathrm{d}t/\mathrm{d}f) \propto f^{-11/3}$ and the contribution of each source to h_c^2 scales as $\propto f^{7/3}$. Combining these two scalings yields $h_\mathrm{c}^2 \propto f^{-4/3}$. At low frequencies, the signal is the superposition of a large number of dim sources, so a stationary unpolarized and isotropic GWB with $h_\mathrm{c} \propto f^{-2/3}$ emerges. Conversely, at high frequencies, none of the assumptions made to describe a stationary unpolarized and isotropic GWB apply: Sparse loud binaries dominate and can be individually resolved, the spectrum becomes spiky, and the GWB level steepens (Sesana et al. 2008).
- The properties of the detector play an active role in the definition of the nature of the signal. In fact, if T increases and Δf decreases, the GWB is better resolved in its individual

[9]While we have written the right-hand sides of Equations 48 and 50 in terms of intrinsic, rest frame quantities, they can also be written by adopting redshifted, observed quantities.

[10]When the signal is discretized, one can go beyond the inclination-polarization-averaged approximation by substituting, for each individual binary, the factor $\sqrt{32/5}$ in Equation 49 with $\sqrt{2(a^2+b^2)}$, where $a = 1 + \cos^2\iota$ and $b = -2\cos\iota$. This better captures the variance of the signal arising from the specific inclinations of the loudest sources (Rosado et al. 2015).

components, there is more variance in the strain between adjacent frequency bins, and eventually more sources become resolvable.[11]

As a useful order-of-magnitude exercise, let us consider the signal produced by a cosmic population of MBHBs with $\mathcal{M} = 10^9\ \mathrm{M}_\odot$, residing in large ellipticals with number density $n_0 = 10^{-4}\ \mathrm{Mpc}^{-3}$ (see Equation 45). We arbitrarily limit our calculation to $z = 1$.[12] Within $z = 1$ there are $\sim 10^7$ massive ellipticals. If each massive black hole residing in their centers has experienced one major merger since $z = 1$, approximating the lookback time to $z = 1$ at ~ 10 Gyr, then the MBHB merger rate would be $\mathrm{d}N_\mathrm{B}/\mathrm{d}t = 10^7/(10^{10}\ \text{years}) = 10^{-3}\ \text{year}^{-1}$. The number of MBHBs emitting per frequency bin is given by $N_\mathrm{B}(\Delta f) = (\mathrm{d}N_\mathrm{b}/\mathrm{d}t)(\mathrm{d}t/\mathrm{d}f)\Delta f$, where $\Delta f = 1/T$. If we now substitute $\mathrm{d}f/\mathrm{d}t$ from Equation 37 assuming $\mathcal{M} = 10^9\ \mathrm{M}_\odot$, use $T = 30$ years, and solve for $N_\mathrm{B}(\Delta f) = 1$, we obtain $f_\mathrm{c} \approx 1.8 \times 10^{-8}$ Hz. At $f > f_\mathrm{c}$, there is less than one source per frequency bin and the loudest signals can be resolved as deterministic CGWs. Conversely, at $f < f_\mathrm{c}$, multiple sources contribute to the signal, which can be considered stochastic. Note that because $\mathrm{d}t/\mathrm{d}f \propto f^{-11/3}$, changing the number of sources by an order of magnitude causes f_c to change by less then a factor of two. Therefore, a fairly robust estimate is that most resolvable sources should arise at $f \gtrsim 10^{-8}$ Hz, consistent with **Figure 3**.

3.2.4. On the dominant role of massive black hole binaries at nanohertz gravitational wave frequencies. Using Equation 51, one can quickly demonstrate that no other astrophysical population of binaries can compete with MBHBs in the nanohertz GW band. Let us assume that two binary populations, A and B, have monochromatic mass functions and share the same distance distribution. At any given frequency, the ratio of the characteristic strain produced by them is approximately

$$\frac{h^2_{\mathrm{c},A}}{h^2_{\mathrm{c},B}} = \frac{N_A}{N_B}\left(\frac{\mathcal{M}_A}{\mathcal{M}_B}\right)^{10/3}. \qquad 52.$$

Let A be the cosmic MBHB population. At $f = 10^{-8}$ Hz, $N_A \approx 1$, while $\mathcal{M}_A = 10^9\ \mathrm{M}_\odot$. Let B be the population of all stars in the observable Universe. If we assume that all these stars were in binaries emitting at $f = 10^{-8}$ Hz, then $N_B \approx 10^{23}$ and $\mathcal{M}_B = 1\ \mathrm{M}_\odot$. Substituting in Equation 52 leads to $h^2_{\mathrm{c},A}/h^2_{\mathrm{c},B} \approx 10^9$, demonstrating that the MBHB GWB is dominant.[13]

Other astrophysical GW sources could be relevant for particular PTA configurations. For example, a PTA featuring many precise MSPs in the Mily Way center might detect early extreme mass ratio inspirals around SgrA* (Kocsis et al. 2012); likewise, a PTA featuring many MSPs residing in a globular cluster might be used to detect an intermediate-mass black hole binary lurking within that same cluster (X. Chen et al. 2025).

3.3. Eccentricity and Environmental Effects: Dynamical Evolution of Massive Black Hole Binaries

When evaluating Equation 48, we considered circular binaries driven by GWs. This simplifying assumption is likely inaccurate for astrophysical MBHBs evolving in dense galactic nuclei for two reasons: (*a*) MBHBs might have significant eccentricity, if anything because they are the product of galaxy mergers that usually occur on almost radial orbits (Fastidio et al. 2024), and (*b*) binaries

[11] For PTAs, source resolvability also depends on the number of pulsars (Boyle & Pen 2012).

[12] For most astrophysical models, $\gtrsim 90\%$ of the GW signal originates at $z < 1$ (Sesana et al. 2008).

[13] One might argue that the distance distributions of stars and MBHBs are not comparable, primarily because there are 10^{11} stars within the Milky Way (i.e., at kiloparsec distances) whereas the closest MBHB is likely several megaparsecs away. A similar calculation assuming that all stars in our Galaxy are in binaries emitting GWs at $f = 10^{-8}$ Hz still yields a strain much smaller than that by MBHBs.

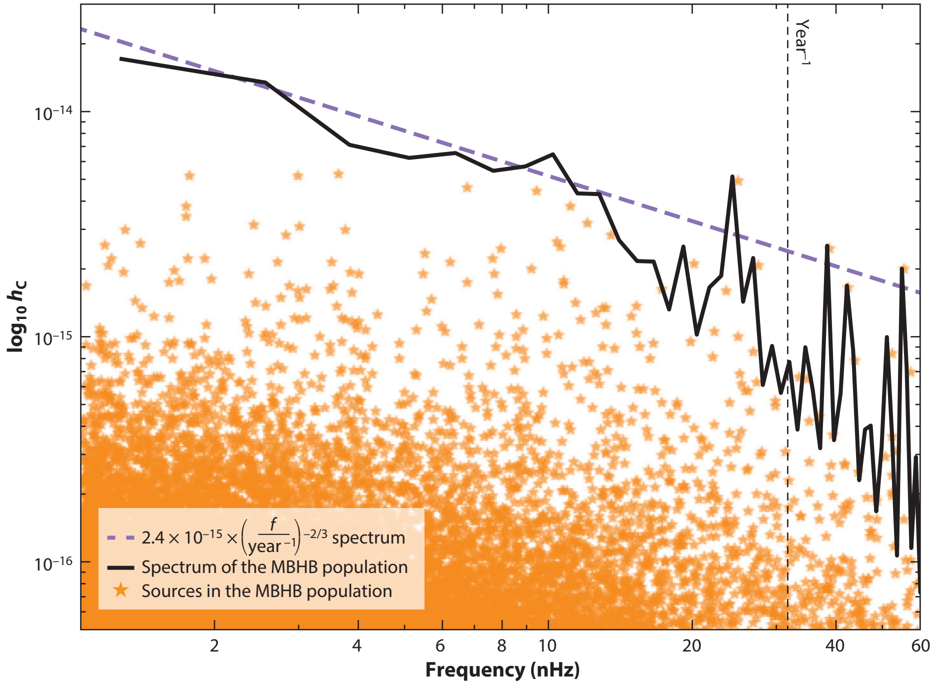


Figure 3

Example of a GWB generated from a population of circular, GW-driven MBHBs, assuming $T \approx 25$ years. The overall GWB (*black*) is built from the sum of each individual source (*orange stars*). The spectrum becomes appreciably spiky and departs from the $f^{-2/3}$ power law (*dashed purple line*) above 10 nHz. Abbreviations: GW, gravitational wave; GWB, gravitational wave background; MBHB, massive black hole binary. Figure adapted from Ferranti et al. (2025) (CC BY 4.0).

exchange energy and angular momentum with the dense environment, which might significantly contribute to their dynamical evolution. In fact, the interaction with the environment must be dominant at a sufficiently low frequency, where the expected GW coalescence timescale becomes longer than the Hubble time t_H. By integrating Equation 37 backward from coalescence to a time t_0, we find

$$f_{K,0} \approx 5 \times 10^{-11} \left(\frac{t_0}{10^{10}\ \text{years}}\right)^{-3/8} \left(\frac{\mathcal{M}}{10^9\ \mathrm{M}_\odot}\right)^{-5/8}\ \text{Hz}, \tag{53.}$$

which, for a reference circular equal-mass binary of $M_1 = M_2 = 10^9\ \mathrm{M}_\odot$, corresponds to a semi-major axis of $a_0 \approx 0.5$ pc. If MBHBs did not lose energy to their environment at parsec scales, they would stall and would never coalesce because of GW emission. To fully model the nanohertz GWB, therefore, we need to understand MBHB dynamical evolution.

MBHBs form in the aftermath of galaxy mergers. The journey of the two massive black holes, from kiloparsec separations down to final coalescence, was first described by Begelman et al. (1980), who identified three evolutionary stages (discussed in Sections 3.3.1–3.3.3, below), with a focus on massive systems relevant to PTAs. For the sake of argument, we model the galaxy merger remnant as a singular isothermal sphere (SIS), characterized by a density profile $\rho(r) = \sigma^2/(2\pi G r^2)$, where σ is the stellar velocity dispersion (Binney 2008).[14] We define the sphere of

[14]Not to be confused with the RMS residual.

influence of an MBHB of total mass M located at the center of the galaxy as

$$r_{\mathrm{inf}} = \frac{GM}{\sigma^2} \tag{54}$$

and connect the massive black hole mass to the stellar bulge velocity dispersion through the M–σ relation of the form $(M/10^6\ \mathrm{M}_\odot) = (\sigma/70\ \mathrm{km\ s^{-1}})^4$ (Kormendy & Ho 2013).[15]

The typical PTA system we consider comprises two massive black holes of $M_1 = M_2 = 10^9\ \mathrm{M}_\odot$, evolving in a merger remnant with $\sigma = 300\ \mathrm{km\ s^{-1}}$, so that $r_{\mathrm{inf}} = 50$ pc and $\rho_{\mathrm{inf}} = \rho(r_{\mathrm{inf}}) = 10^3\ \mathrm{M}_\odot\ \mathrm{pc}^{-3}$. Although the SIS is a rough approximation for the density profile of massive galaxies, it suffices to provide an order-of-magnitude estimate of the relevant physical processes. We emphasize, however, that central density profiles of massive galaxies are extremely diverse, ranging from flat cores to steep cusps. As a result, r_{inf} and ρ_{inf} can vary significantly, with 10 pc $< r_{\mathrm{inf}} <$ 500 pc, and $10\ \mathrm{M}_\odot\ \mathrm{pc}^{-3} < \rho_{\mathrm{inf}} < 10^5\ \mathrm{M}_\odot\ \mathrm{pc}^{-3}$. In general, galaxies with larger cores have higher r_{inf} and lower ρ_{inf}.

3.3.1. Phase 1. Dynamical friction: 10 kpc → 10 pc.

Right after the merger of two massive galaxies, massive black holes residing at the center of the parent galaxies are found in the outskirts of the galaxy merger remnant and independently sink to the center because of dynamical friction (DF) on a timescale (Binney 2008)

$$t_{\mathrm{DF},i} \approx 1.1 \left(\frac{\ln\Lambda}{10}\right)^{-1} \left(\frac{r_{0,i}}{10\ \mathrm{kpc}}\right)^{2} \left(\frac{\sigma}{300\ \mathrm{km\ s^{-1}}}\right) \left(\frac{M_i}{10^9\ \mathrm{M}_\odot}\right)^{-1}\ \mathrm{Gyr}, \tag{55}$$

where $i = 1$ or 2, r_0 is the initial distance from the center, and $\ln\Lambda \approx \mathcal{O}(10)$. Equation 55 is oversimplified in many ways, as (*a*) it includes only stars moving slower than the massive black hole and faster stars also contribute to DF (Dosopoulou & Antonini 2017); (*b*) the initial massive black hole orbit can be very eccentric (Fastidio et al. 2024), implying a shorter t_{DF}; and (*c*) the sinking massive black holes are initially surrounded by the respective progenitor galaxy stellar nuclei, making the effective sinking masses M_i bigger and t_{DF} shorter.

Note that t_{DF} can be greater than t_{H} for light black holes or large initial r_0. In fact, light massive black holes generally fail to reach the remnant center and wander in the galaxy outskirts (e.g., McWilliams et al. 2014). Thus, DF preferentially selects the formation of massive, equal-mass binaries. When two massive black holes come within r_{inf} of each other, they bind into an MBHB. As the binary shrinks, its circular velocity V_{c} increases to $V_{\mathrm{c}} \gg \sigma$ (i.e., the binary becomes hard), making DF ineffective (Binney 2008). If there were no further physical mechanisms to extract energy and angular momentum from the MBHB, it would stall at a separation of several parsecs, much larger than the value at which GWs can make it coalesce within t_{H} (see Equation 53).

3.3.2. Phase 2. Binary hardening: 10 pc → 0.1 pc.

There are at least three ways to drive an MBHB to a GW-dominated inspiral. We discuss them separately in the following subsections.

3.3.2.1. Stellar hardening.

A hard binary loses energy and angular momentum to surrounding stars via three-body scatterings, evolving according to (Sesana 2010)

$$\frac{\mathrm{d}a}{\mathrm{d}t_{\mathrm{r}}} = -\frac{a^2 G\rho_{\mathrm{inf}}}{\sigma} H \quad \rightarrow \quad \frac{\mathrm{d}f_{\mathrm{K,r}}}{\mathrm{d}t} = \frac{3G^{4/3}M^{1/3}H\rho}{2(2\pi)^{2/3}\sigma} f_{\mathrm{K,r}}^{1/3},$$
$$\frac{\mathrm{d}e}{\mathrm{d}t_{\mathrm{r}}} = \frac{aG\rho_{\mathrm{inf}}}{\sigma} HK, \tag{56}$$

[15] In general, these relations hold for single massive black holes, while we assume here that galaxy merger remnants correlate with MBHBs. For the sake of this discussion, the difference is negligible.

where H and K are dimensionless rates derived from numerical three-body scattering experiments. In particular, $H \approx 15$–20 depending on q and e, while $K \to 0$ for $e \to 0 \wedge e \to 1$ is generally positive and has a maximum $K \approx 0.1$–0.4 (depending on q) for $e \approx 0.6$. Equation 56 shows that stars shrink the MBHB while increasing its eccentricity. Stars interacting with the MBHB are ejected from the galactic nucleus, and the evolution of the system depends on how efficiently new stars can be supplied to the binary, that is, at the pace at which the so-called binary loss cone can be repopulated (Milosavljevic & Merritt 2003). The literature on this subject is vast, but Sesana (2010) showed that using $\rho_{\rm inf}$ in Equation 56 corresponds to assuming a full loss cone at the influence radius, which is typically the case for triaxial merger remnants (Khan et al. 2011, Preto et al. 2011).

3.3.2.2. Gas hardening. Galaxy mergers trigger cold gas infall toward the center of the merger remnant (Mihos & Hernquist 1996). Part of this gas reaches the nucleus, where it can form a circumbinary disk (CBD) at subparsec scales. The MBHB–CBD energy and angular momentum exchange is an extremely complex topic that has recently been addressed by several authors (for the status of the field, see Duffell et al. 2024). The outcome of the interaction depends on the detailed dynamics in the CBD cavity; on the role of self-gravity (Franchini et al. 2021); on the binary parameters, disk aspect ratio, and equation of state (e.g., Duffell et al. 2020); and on whether the disk rotation is prograde or retrograde with respect to the binary orbital angular momentum (Nixon et al. 2011). Interaction with a prograde CBD drives the MBHB toward $q = 1$ (due to preferential accretion onto the secondary massive black hole), while e tends to saturate to values $0.4 < e < 0.7$ (Siwek et al. 2023).

As an example, we consider the case of a prograde CBD, with a cavity maintained by the torque exerted by the binary onto the disk and with no mass allowed to enter the cavity. Ivanov et al. (1999) showed that the binary evolution can be approximated as (we neglect $\mathrm{d}e/\mathrm{d}t_{\rm r}$ here)

$$\frac{\mathrm{d}a}{\mathrm{d}t_{\rm r}} = 2(aa_0)^{1/2}\frac{(1+q)^2}{q}\frac{\eta_{\rm E}}{t_{\rm S}} \quad \rightarrow \quad \frac{\mathrm{d}f_{\rm K,r}}{\mathrm{d}t_{\rm r}} = 3\pi^{1/3}a_0^{1/2}\frac{(1+q)^2}{q}\frac{\eta_{\rm E}}{t_{\rm S}}(GM)^{-1/6}f_{\rm K,r}^{4/3}, \qquad 57.$$

where a_0 is the semimajor axis at which the mass of the unperturbed disk equals the mass of the secondary black hole:

$$a_0 = 0.05\left(\frac{\alpha}{0.1}\right)^{4/7}\left(\frac{M_1}{10^9\ \mathrm{M}_\odot}\right)^{1/7}\eta_{\rm E}^{-3/7}q^{5/7}\ \mathrm{pc}. \qquad 58.$$

Here, $\eta_{\rm E} = \dot{M}/\dot{M}_{\rm E}$ is the accretion rate normalized to the Eddington rate, $t_{\rm S} = M/\dot{M} = 0.44$ Gyr is the Salpeter time, and α is the disk viscosity parameter. Note that when $a = a_0$ the residence time $\mathrm{d}t_{\rm r}/\mathrm{d}\ln a \sim qt_{\rm S}/\eta_{\rm E}$, which means that equal-mass binaries accreting at the Eddington rate evolve on a Salpeter timescale.

3.3.2.3. Massive black hole triple interactions. In massive, gas-poor galaxies with shallow cores, the hardening timescale can exceed $t_{\rm H}$, causing binary stalling. A subsequent merger may bring a third massive black hole to the center, where triple interactions can trigger a merger—via secular or chaotic processes—or eject one massive black hole, leaving a wide binary (Blaes et al. 2002). Mergers often follow the brief formation of tight, highly eccentric binaries with short GW coalescence times, making high eccentricity a signature of triple interactions. Cosmological models have shown that up to 20% of MBHBs form triplets (Bonetti et al. 2018b, Sayeb et al. 2023), and Bonetti et al. (2018a) find that ~40% of these triplets merge, implying that up to ~10% of all MBHB mergers might result from triple interactions.

3.3.3. Phase 3. Gravitational wave emission: 0.1 pc → $a_{\rm ISCO}$. Eventually, GW emission becomes efficient at small separations, causing the binary to evolve according to Equation 37 until coalescence. To determine whether nanohertz MBHBs are driven primarily by GWs or by interaction with their environment, we can simply equate the GW orbital frequency evolution rate of Equation 37 to its stellar and gas-driven counterparts in Equations 56 and 57 and solve for the frequency, obtaining

$$f_{\mathrm{K},\star/\mathrm{GW}} \approx 5 \times 10^{-10} \left(\frac{M_1}{10^9\ \mathrm{M}_\odot}\right)^{-5/8} \left(\frac{\rho_{\mathrm{inf}}}{\rho_{\mathrm{inf,SIS}}}\right)^{3/10} q^{-3/10} \mathcal{F}(e)^{-3/10}\ \mathrm{Hz},$$

$$f_{\mathrm{K,gas/GW}} \approx 5 \times 10^{-10} \left(\frac{\alpha}{0.1}\right)^{-6/49} \left(\frac{M_1}{10^9\ \mathrm{M}_\odot}\right)^{-37/49} q^{-69/98} \eta_{\mathrm{E}}^{-33/98} \mathcal{F}(e)^{-6/14}\ \mathrm{Hz}. \qquad 59.$$

From the decoupling point, the lifetime of the MBHB, estimated using $\mathrm{d}t_{\mathrm{r}}/\mathrm{d}\ln f_{\mathrm{K}}$, is

$$t_{\star/\mathrm{GW}} \approx 10^8 \left(\frac{\rho_{\mathrm{inf}}}{\rho_{\mathrm{inf,SIS}}}\right)^{4/5} q^{-1/5} \mathcal{F}(e)^{-1/5}\ \text{years},$$

$$t_{\mathrm{gas/GW}} \approx 10^8 \left(\frac{\alpha}{0.1}\right)^{-16/49} \left(\frac{M_1}{10^9\ \mathrm{M}_\odot}\right)^{-17/49} q^{-43/49} \eta_{\mathrm{E}}^{44/49} \mathcal{F}(e)^{1/7}\ \text{years}, \qquad 60.$$

where for the stellar case we also retain the dependence on ρ_{inf}, keeping in mind that, for our reference binary, $\rho_{\mathrm{inf,\,SIS}} = 10^3\ \mathrm{M}_\odot\ \mathrm{pc}^{-3}$. Let us highlight a few key points (**Figure 4**):

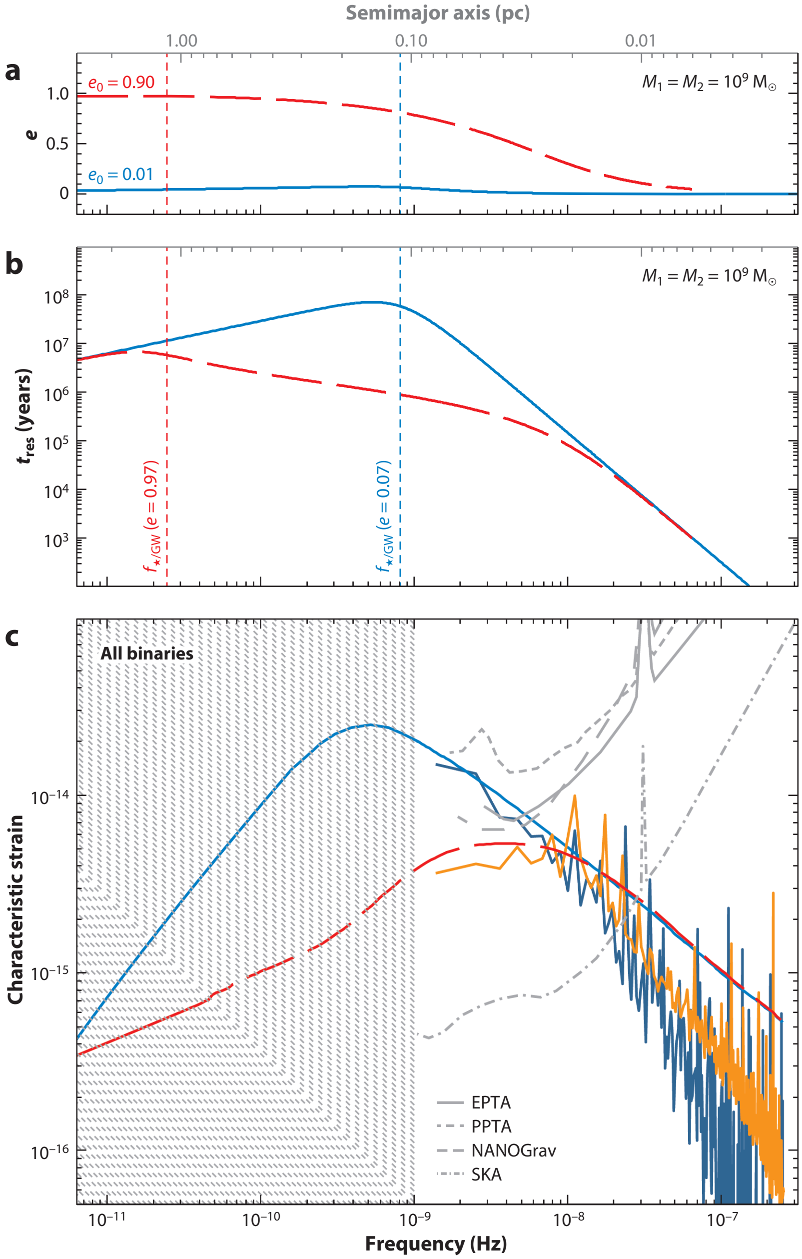
- By substituting $\mathrm{d}f/\mathrm{d}t$ from Equations 56 and 57 into Equation 48, we find $h_{\mathrm{c}} \propto f$ and $h_{\mathrm{c}} \propto f^{1/2}$ for stellar and gas-driven binaries, respectively. Therefore, in both cases there is a low-frequency turnover in the h_{c} spectrum.
- For $e \approx 0$, the turnover occurs around $f \approx 1$ nHz, at the lower end of the PTA band.
- The turnover frequency has a very weak dependence on ρ_{inf} and η_{E}. Pushing it to 10 nHz, where PTAs today are most sensitive, would require $\rho_{\mathrm{inf}} \gtrsim 10^6\ \mathrm{M}_\odot\ \mathrm{pc}^{-3}$ or $\eta_{\mathrm{E}} \gtrsim 10^3$, both of which are extremely unlikely in massive, low-redshift galaxies.
- e decreases both the decoupling frequency and the remaining binary lifetime at such frequency. Thus, when binaries are very eccentric, fewer of them are distributed over a much larger orbital frequency range. Therefore, even without considering the redistribution of GW power at higher harmonics, high eccentricity suppresses the low-frequency GW signal, as clearly shown in **Figure 4**.

3.4. Putting Everything Together: The Full Gravitational Wave Signal from Massive Black Hole Binaries

By adding all the relevant physics, including multiple hardening mechanisms and eccentricity evolution, we can write the characteristic strain $h_{\mathrm{c}}(f)$ of the GWB as

$$h_{\mathrm{c}}^2(f) = \int_0^\infty \mathrm{d}z \int_0^\infty \mathrm{d}M_1 \int_0^1 \mathrm{d}q \int_0^1 \mathrm{d}e \frac{\mathrm{d}^5 N}{\mathrm{d}z\, \mathrm{d}M_1\, \mathrm{d}q\, \mathrm{d}t\, \mathrm{d}e_0} \left[\frac{\mathrm{d}t}{\mathrm{d}\ln f_{\mathrm{K,r}}} \frac{\mathrm{d}e_0}{\mathrm{d}e}\right] \times$$
$$h^2(f_{\mathrm{K,r}}) \sum_{n=1}^{+\infty} \frac{g(n,e)}{(n/2)^2}\Bigg|_{f_{\mathrm{K,r}} = f(1+z)/n}, \qquad 61.$$

where each individual MBHB now contributes a formally infinite series of harmonics of the orbital frequency $f_{\mathrm{K,r}}$, adjusted by $1 + z$ to account for cosmological redshift and thus observed at frequency $f = nf_{K,r}/(1+z)$. The normalization of each harmonic is given by $h_c^2(f_{K,r})$ per Equation 49, weighted by the $g(n, e)$ factor given in Equation 39. As in the circular case, Equation 61 is the continuum limit of an inherently discrete problem. Using the same procedure of Section 3.2.3, we find that the GW signal is discretized as



(Caption for **Figure 4** appears on following page)

Figure 4 (*Figure appears on preceding page*)

GW signal and MBHB dynamics. Solid blue and dashed red lines represent MBHBs with $e_0 = 0.01$ and $e_0 = 0.9$ at formation. (*a*) Evolution of e. (*b*) Evolution of $t_{\rm res} = {\rm d}t/{\rm d}\ln f$ of a fiducial MBHB with $M_1 = M_2 = 10^9\ {\rm M}_\odot$, as a function of f (lower x axes) and a (gray upper x axes). Binaries are initially hardened by three-body scattering and reach up to $e \approx 0.07$ (for $e_0 = 0.01$) and $e \approx 0.97$ (for $e_0 = 0.9$), until GW emission takes over, leading to circularization. $t_{\rm res}$ peaks around $f_{\star/{\rm GW}}$ (see Equation 60). (*c*) GW strain $h_c(f)$ from the overall cosmic population of MBHBs integrated over masses and redshift. Smooth lines represent GWB calculations from Equation 61, whereas spiky lines represent actual realizations from the discrete population according to Equation 62 for $e_0 = 0.01$ (*orange*) and $e_0 = 0.9$ (*steel blue*). Also shown are PTA sensitivities and the area inaccessible with $T < 30$ years (*shaded area*). Abbreviations: GW, gravitational wave; GWB, gravitational wave background; MBHB, massive black hole binary; PTA, pulsar timing array.

(Amaro-Seoane et al. 2010)

$$h_c^2(f_j) = \sum_{i=1}^{N_B}\sum_{n=1}^{+\infty} h_i^2(f_{\rm K,r})\frac{g(n,e_i)}{(n/2)^2}\frac{n f_{{\rm K,r}_i}}{\Delta f(1+z)}\Theta\left(\frac{n f_{{\rm K,r}_i}}{1+z},\Delta f_j\right), \qquad 62.$$

where $\Theta = 1$ if $n f_{{\rm K,r}_i}/(1+z) \in \Delta f_j$ and $\Theta = 0$ otherwise, and the first sum is over the N_B MBHBs of the population, each contributing to the signal with the sum of n harmonics. For circular binaries, $g(n, e = 0) = 1$ if $n = 2$ and 0 otherwise; Equation 62 simplifies to Equation 51, which, in the continuum limit, for GW-driven binaries returns the well-known PL of Equation 45. In Equation 61, we have separated the merger rate from dynamical evolution by writing

$$\frac{{\rm d}^5 N}{{\rm d}z\,{\rm d}M_1\,{\rm d}q\,{\rm d}\ln f_{\rm K,r}\,{\rm d}e} = \frac{{\rm d}^5 N}{{\rm d}z\,{\rm d}M_1\,{\rm d}q\,{\rm d}t\,{\rm d}e_0}\left[\frac{{\rm d}t}{{\rm d}\ln f_{\rm K,r}}\frac{{\rm d}e_0}{{\rm d}e}\right]. \qquad 63.$$

3.4.1. Gravitational wave background normalization: cosmic merger rate. The first term on the right-hand side of Equation 63, ${\rm d}^5N/({\rm d}z\,{\rm d}M_1\,{\rm d}q\,{\rm d}t\,{\rm d}e_0)$, represents the differential cosmic merger rate of MBHBs as a function of the relevant binary parameters and determines mainly the normalization $A_{\rm GWB}$ of the GWB spectrum. Several authors have derived this quantity, either theoretically from semianalytic models (Rajagopal & Romani 1995, Jaffe & Backer 2003, Wyithe & Loeb 2003, Sesana et al. 2004, McWilliams et al. 2014) and cosmological simulations (Kelley et al. 2017, Siwek et al. 2020, Izquierdo-Villalba et al. 2021) or empirically by combining observations of the galaxy mass function, pair fraction, and massive black hole mass–host galaxy relation to obtain an estimated MBHB merger rate (Sesana 2013, Ravi et al. 2015, Simon 2023). **Figure 5** presents a compilation of the estimated $A_{\rm GWB}$ from models predating the first observational evidence of a signal in PTA data (Arzoumanian et al. 2020). The uncertainty of the underlying physics and the diversity of techniques adopted result in a wide range of $10^{-16} \lesssim A_{\rm GWB} \lesssim 3\times 10^{-15}$. Note that the upper limit is essentially dictated by the massive black hole mass function today, as discussed in Section 3.2.1.[16]

3.4.2. Gravitational wave background spectral shape: binary dynamics. The term in square brackets in Equation 63 represents a map $(t, e_0) \to (f_{\rm K,r}, e)$, transforming a merger rate described by an eccentricity distribution e_0 into the number of systems emitting at all frequencies with their

[16]A notable outlier is the prediction by McWilliams et al. (2014). However, their $A_{\rm GWB}$ estimate appears to be higher than the upper limit imposed by the argument of Sato-Polito et al. (2024) applied to the mass function they use.

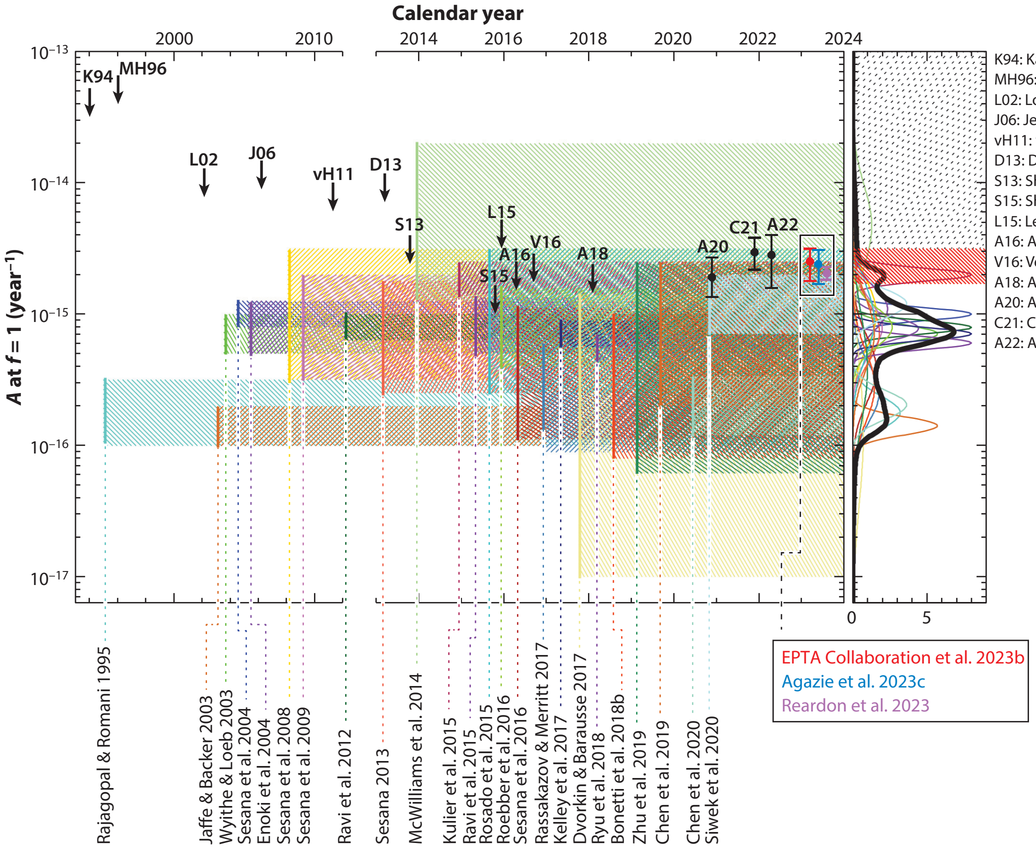


Figure 5

Progression of predictions, limits, and measurements of the GWB produced by MBHBs. The main plot shows the predicted amplitude A_{GWB}, under the assumption $h_c \propto f^{-2/3}$ (Equation 7; *vertical colored lines*) up to 2020; upper limits placed by several PTAs up to 2018 (*downward pointing arrows*); evidence of common red noise since 2020 (*thick black dots with error bars*); and evidence of a correlated GWB from EPTA+InPTA, NANOGrav, and PPTA (*colored marks with error bars*, shifted with respect to one another for clarity). By giving equal credit to each prediction, the right panel shows the PDF of the theoretically expected A_{GWB} (*thick black line*). The red-shaded area marks the A_{GWB} range consistent with PTA measurements, whereas the black-shaded area marks the excluded A_{GWB} region. Abbreviations: GWB, gravitational wave background; MBHB, massive black hole binary; PDF, probability distribution function; PTA, pulsar timing array.

eccentricities. This map must be obtained by numerically solving the equations

$$\frac{\mathrm{d}f_{\mathrm{K,r}}}{\mathrm{d}t_{\mathrm{r}}} = \sum_i \left.\frac{\mathrm{d}f_{\mathrm{K,r}}}{\mathrm{d}t_{\mathrm{r}}}\right|_i, \qquad \frac{\mathrm{d}e}{\mathrm{d}t_{\mathrm{r}}} = \sum_i \left.\frac{\mathrm{d}e}{\mathrm{d}t_{\mathrm{r}}}\right|_i, \tag{64}$$

where the sum is over the relevant physical mechanisms driving the binary: GW emission and interaction with stars and gas, as given by Equations 37, 56, and 57. **Figure 4** depicts examples of the full GW signal for a quasi-circular and very eccentric population of MBHBs. The figure shows the connection between e, t_{res}, and h_c shape. Note that while the function $g(n, e)$ redistributes the GW emission at higher harmonics, the main contribution to the GWB drop comes from the much shorter t_{res}, driven by the $\mathcal{F}(e)$ function of Equation 38. Conversely, the pileup of harmonics for multiple sources at high frequency causes a slight increase of the GWB compared with the circular case (**Figure 4c**) while decreasing its variance. We emphasize that the high-frequency drop is due to sparse sampling of high-frequency sources. The square averaged total GW signal computed over 1,000 realizations of the MBHB population sits around the $h \propto f^{-2/3}$ PL at $f > 10^{-8}$ Hz.

3.5. Resolvable Signals from Massive Black Hole Binaries

The incoherent superposition of the signals produced by the cosmic MBHB population is expected to result in a stochastic GWB. However, a number of possible deterministic signals are emerging in the nanohertz frequency band, as we discuss in this section.

3.5.1. Continuous gravitational waves. As mentioned in Section 3.2.3, particularly loud or nearby MBHBs can be resolved individually as deterministic signals, known as CGWs. CGWs have been studied mainly in the context of PTA detection, where the signal in the data is obtained by inserting Equation 33 into Equation 11 (for full expressions, see Taylor 2021, Truant et al. 2025). Circular CGWs have been studied extensively (Gardiner et al. 2025). The signal in each MSP is composed of an Earth and a pulsar term, the latter evaluated at a time τ in the past (see Equation 11). During the time τ, the MBHB has evolved according to Equation 37, such that the GW frequency of the two terms is different and $f_p < f_E$:

- If $f_E - f_p > \Delta f$, then the pulsar and Earth terms fall at different frequencies. All Earth terms can be added coherently and properly matched with a deterministic filter.
- If $f_E - f_p < \Delta f$, then the pulsar and Earth terms fall in the same frequency bin and their superposition partially destroys the coherency of the signal.

For a typical 10^9 $M_\odot$ MBHB and for typical PTA observing times $\mathcal{O}$(10 years), the separation between the two classes of sources occurs at $\sim 10^{-8}$ Hz (Rosado et al. 2015). Although CGWs have not yet been detected, a fully operational SKA is expected to resolve tens of them (Boyle & Pen 2012, Truant et al. 2025), with a sky localization precision of tens of square degrees (Sesana & Vecchio 2010). Such precision might still be sufficient to allow effective EM counterpart searches (Burke-Spolaor 2013), since CGWs are expected to be sourced by MBHBs hosted in massive and nearby elliptical galaxies (Zhou et al. 2025).

3.5.2. Eccentric bursts. Another class of deterministic nanohertz GW signals consists of bursts from very eccentric binaries, especially during triple interactions. At every periastron passage, the system emits a burst of GW with peak frequency (Wen 2003)

$$f_b = f_K (1 - e^2)^{-3/2} (1 + e)^{1.1954}. \tag{65}$$

Note that $f_b \approx 2 f_K (1 - e)^{-3/2}$; in other words, the frequency of a circular binary of the size of the periastron $r_p = a(1 - e)$. The duration of the burst is $\Delta t_b \approx 1/f_b$. This means that binaries with orbital frequencies much lower than $1/T$ can produce localized bursts well within the PTA band. For example, an MBHB with a period of 100 years and an eccentricity 0.99 at periastron generates a burst with $f_b \approx 3 \times 10^{-7}$ Hz with a duration of $\Delta t_b \approx 3 \times 10^6$ s $\sim$1 month. Amaro-Seoane et al. (2010) found that a future PTA sensitivity of approximately 1 ns is needed in order to resolve GW bursts from triplets.

3.5.3. Bursts with memory. GW bursts can feature nonoscillatory components that leave a permanent deformation of space-time, known as memory (Christodoulou 1991). Such bursts, termed bursts with memory (BWM), occur because GW emission induces lasting second time derivatives in the source's mass–energy quadrupole moments, causing the metric to relax to a new configuration. At nanohertz frequencies, the most promising BWM sources are MBHB mergers. Under the assumption of a circular binary inspiral, the permanent metric shift has a vanishing $h_\times$ component and approximately follows (e.g., van Haasteren & Levin 2010)

$$h_+ \approx \frac{1-\sqrt{8}/3}{24}\frac{\mu}{D_\mathrm{l}}\sin^2\iota(17+\cos^2\iota)\left[1+\mathcal{O}(\mu^2/M^2)\right] \approx 1.5\times10^{-15}\frac{\mu}{10^9\ \mathrm{M}_\odot}\left(\frac{D_\mathrm{l}}{1\ \mathrm{Gpc}}\right)^{-1}, \quad 66.$$

where μ is the redshifted reduced mass and in the last term we have averaged over ι. This shift rapidly grows at merger on a timescale $t_\mathrm{r} \approx 2\pi R_\mathrm{S}/c \approx 1\ \mathrm{day}\,(M/10^9\ \mathrm{M}_\odot)$, where R_S is the Schwarzschild radius. The wave form is therefore described by $h(t) = h_+\Theta(t-t_0)$, where $\Theta(t-t_0)$ is the Heaviside step function; the perturbation is null until time t_0 and quickly jumps to the value given by Equation 66. Since the induced timing residual R is the integral of a Heaviside function, it takes the form of a linear drift, similar to a pulsar glitch, which is partially absorbed by fitting a quadratic function to the pulsar spin and spin derivative, leaving behind a cusplike burst shape similar to the one caused by the Shapiro delay. BWM are expected to be observable if sourced by $\sim10^9\ \mathrm{M}_\odot$ binaries at $z<1$. Their cosmic rate is equal to the merger rate of these MBHBs, which is expected to be $\ll$1 year^{-1}. Searches for BWM have been performed in several PTA data sets, yielding null results (Agazie et al. 2025b).

3.6. Astrophysical Interpretation of Pulsar Timing Array Observations

Although A_GWB lies at the upper end of the range shown in **Figure 5** and $\gamma < 13/3$, a cosmic population of MBHBs remains a natural explanation of the observed PTA signal presented in Section 2.4. **Figure 6** shows highlights from EPTA's interpretation paper (EPTA Collaboration et al. 2024a) as an example. The RMS residual PSD aligns quite well with the expected distribution from empirical models (**Figure 6*b***), as expected. When the model presented by Chen et al. (2019) is fitted to the data, the high value of A_GWB translates into a short average MBHB coalescence time ($\tau_0 < 1$ Gyr with 90% confidence) and a high normalization M_*, implying $M_\mathrm{BH}/M_\mathrm{bulge} = 2.5^{+1.3}_{-1.3}\times10^{-3}$ for $M_\mathrm{bulge} = 10^{11}$. Likewise, the addition of the evolution of the M_BH–M_bulge relation into the picture favors positive redshift evolution, implying more massive black holes for a given galaxy mass in the past (Matt et al. 2026). Although Sato-Polito et al. (2024) argued that such a high amplitude is difficult to reconcile with the observed local massive black hole mass function, this observation has been disputed (Liepold & Ma 2024). Note that the maximum A_GWB is very sensitive to the high-mass end of the local massive black hole mass function (Izquierdo-Villalba et al. 2021), which is determined by a handful of sparse measurements and suffers from considerable uncertainties. The low γ would naturally arise from either high eccentricity or strong environmental coupling (see Section 3.3), the latter expressed by the parameter ζ_0. Although there is a slight preference for the high e–ζ_0 corner, the marginalized posteriors of both parameters show only a tentative preference for such high values, leaving the analysis inconclusive.

Finally, the results discussed above are conditioned on the astrophysical priors included in the Chen et al. (2019) model. Assuming less informative priors might result in differences in the inferred MBHB observables (Agazie et al. 2023b). This is a hint that astrophysical inference is still dominated by the prior choice, meaning that current PTA observations are not yet very informative for constraining underlying signal models (I. Ferranti, A. Sesana & D. Izquierdo-Villalba, manuscript in preparation).

4. COSMOLOGICAL GRAVITATIONAL WAVE SIGNALS

The evidence (Agazie et al. 2023c, EPTA Collaboration et al. 2023b, Reardon et al. 2023, Xu et al. 2023) for a GW signal in PTA data offers unprecedented opportunities for exploring high-energy physics beyond the Standard Model (BSM) embedded in early Universe cosmology

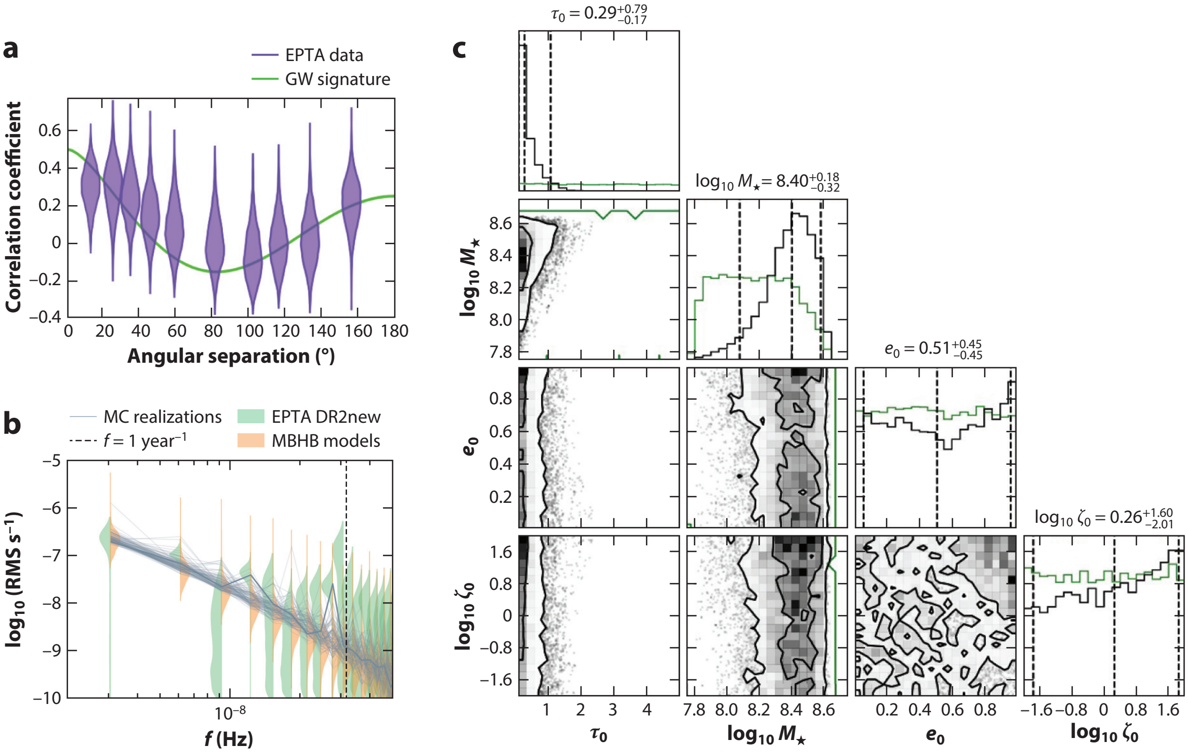


Figure 6

Bird's-eye view of EPTA results. (*a*) Measured correlation (*purple violins*) compared with the expectation value from Hellings & Downs (1983) (*green line*). (*b*) Measured RMS PSD (*orange violins*) compared with the expected distribution from the theoretical MBHB models from Chen et al. (2019) (*green violins*). The jagged lines show 100 realizations of the PSD from the models. (*c*) Triangle plots of the posterior distribution (*black*; priors shown in *green*) of four key parameters of the theoretical models: the average coalescence time τ_0 of an MBHB from 20 kpc separation, the normalization M_* of the $M_{\rm BH}$–$M_{\rm bulge}$ relation at $M_{\rm bulge} = 10^{11}\ {\rm M}_\odot$, the eccentricity e_0 at MBHB formation, and the normalized density of the environment $\zeta_0 = \rho_{\rm inf}/\rho_{\rm inf,\,SIS}$ (see Equation 60). Abbreviations: GW, gravitational wave; MBHB, massive black hole binary; MC, Monte Carlo; PSD, power spectral density; RMS, root mean square. Figure adapted from EPTA Collaboration et al. (2024a).

(Caprini & Figueroa 2018). The key point is that early Universe dynamics, typically operating at very high energies, created GWBs that carry information about their source(s). These are referred to as cosmological backgrounds, in contrast to the astrophysical ones introduced in Section 3. A plethora of cosmological backgrounds might be permeating the Universe (Caprini & Figueroa 2018), including signals that may have originated during or after inflation, typically created by nonlinear and out-of-equilibrium field dynamics, as in first-order phase transitions or topological defect evolution.

While the PTA signal is likely to be due to MBHBs (see Section 3.6), cosmological backgrounds also represent viable candidates that could explain the data. In fact, the parameter space of models producing a GWB that can fit, significantly exceed, or simply fail to explain the PTA signal can be clearly identified. Remarkably, these parameter constraints are independent of the origin of the PTA signal. For example, parameter regions in certain models may lead to signals in conflict with the PTA data, resulting in excluded regions and upper limits on the individual parameters. In other scenarios, PTA constraints can complement existing bounds, probing previously unexplored parameter regions.

The constraining power of PTA data in the search for new physics marks the dawn of early Universe GW cosmology as a field. With the advent of PTA astronomy, we can start addressing fundamental questions, such as:

- What is the parameter space region (or regions) of BSM models that can explain the observed data?
- Is there a preference for either an astrophysical or a cosmological explanation for the signal?
- Could the observed signal arise simultaneously from a combination of MBHBs and a cosmological source?

In this section, we review the state of the art of the answer(s) to the first question, as current PTA data already allow for a quantitative and rich analysis in this respect. We focus mainly on cosmological GWB signals, although for completeness we also discuss deterministic signals from dark matter models, albeit more briefly. We postpone discussion of the second and third questions, and others, to Section 5, as the current data are not yet precise enough to address these aspects.

4.1. Early Universe Backgrounds

The present Universe may harbor a wide variety of cosmological backgrounds, including signals that may have originated during inflation (e.g., from vacuum fluctuations, from strong particle production, or induced by large scalar fluctuations), signals that could have been generated after inflation but are still related to the inflationary setup (e.g., from preheating dynamics or enhanced by kination domination), and signals potentially generated after inflation (typically unrelated to the inflationary setup; e.g., from thermal plasma motions, oscillon dynamics, first-order phase transitions, or cosmic defect dynamics). We refer to all these signals as EUBs (for a comprehensive review of EUBs, see Caprini & Figueroa 2018 and references therein). Below, we discuss how leading EUB models may succeed or fail in explaining current PTA data and review the current constraints on their parameter space, mostly following Afzal et al. (2023); EPTA Collaboration et al. (2024a); Ellis et al. (2024), and Figueroa et al. (2024).

4.2. Inflation

Today's spectrum of a GWB from inflation can be written as (Caprini & Figueroa 2018)

$$h^2\Omega_{\rm GW}^{(0)}(k) = \mathcal{T}_{\rm rh}(k)\mathcal{G}_{\rm RD}\frac{h^2\Omega_{\rm rad}^{(0)}}{24}\mathcal{P}_{\rm t}(k) \simeq 1.19\cdot 10^{-16}\left(\frac{r}{0.032}\right)\left(\frac{f}{f_{\rm pivot}}\right)^{n_{\rm t}}. \qquad 67.$$

Here, the first equality represents an exact expression, where $\mathcal{P}_{\rm t}(k)$ is the inflationary tensor power spectrum, $\mathcal{T}_{\rm rh}(k)$ is the reheating transfer function connecting the end of inflation with the onset of radiation domination (RD), and $\mathcal{G}_{\rm RD} \equiv (g_{\rm RD}/g_0)(g_{s,0}/g_{s,{\rm RD}})^{4/3}$ is a factor indicating the changes in the relativistic degrees of freedom from RD until today. The second equality is valid only for "vanilla" inflationary models (i.e., canonically normalized, single-field, slow-roll models, assuming general relativity), where a primordial tensor spectrum is expected in the form of a PL. The latter is parameterized around CMB scales as $\mathcal{P}_{\rm t}(k) = rA_{\rm s}(k/k_{\rm pivot})^{n_{\rm t}}$, where $k_{\rm pivot} = 0.05$ Mpc^{-1}, $n_{\rm t}$ is the spectral tilt, r is the tensor-to-scalar ratio bounded as $r \leq 0.032$ (Tristram et al. 2022), and $A_{\rm s}$ is the primordial scalar perturbation measured as $A_{\rm s} \simeq 2.1\cdot 10^{-9}$ (Planck Collaboration et al. 2020c). For convenience, we assume instant reheating ($\mathcal{T}_{\rm rh} = 1$) and only Standard Model degrees of freedom during RD ($\mathcal{G}_{\rm RD} \simeq 0.39$), and we convert $k_{\rm pivot}$ into today's frequency, $f_{\rm pivot} \simeq 8\cdot 10^{-17}$ Hz.

As the inflationary signal can be at most $h^2\Omega_{\rm GW} \simeq 10^{-16}$ at CMB scales (see Equation 67), explaining a signal with higher amplitude at PTA frequencies clearly requires a blue tilt, $n_{\rm t} > 0$ (the smaller r is, the larger the required tilt will be). Vanilla scenarios, however, must verify the so-called consistency condition, $n_t = -r/8$; therefore, current CMB constraints allow these scenarios to admit at most a tiny red tilt, $-n_t \lesssim 0.0035$. Therefore, more complex scenarios are needed in order to explain the PTA data, particularly models incorporating a mechanism that leads to a large positive tilt at small scales. This is the case, for example, for models with exponential particle production during inflation (see, e.g., Guzzetti et al. 2016 and references therein). In such scenarios, a sizable running of n_t is also expected. Thus, Equation 67 represents only a simplifying parameterization, which helps us understand the averaged positive tilt required to explain the PTA data by an inflationary blue-tilted signal.

Fitting $\Omega_{\rm GW}^{(0)} = \mathcal{A}_{\rm year}(f/f_{\rm year})^{n_t}$ to PTA data, with free amplitude $\mathcal{A}_{\rm year}$ and tilt n_t, leads straightforwardly to a direct constraint on r (at the CMB scales) via Equation 67. For example, EPTA Collaboration et al. (2024a) obtained $\log_{10} r = -12.18^{+8.81}_{-7.00}$ and $n_t = 2.29^{+0.87}_{-1.11}$ at 90% CL by using EPTA data, whereas Figueroa et al. (2024) found $r_{0.05} \in [4.37 \cdot 10^{-14}, 3.47 \cdot 10^{-9}]$ and $n_t = 2.08^{+0.32}_{-0.30}$ at 68% CL by combining NG15 and EPTA data. Using NG15 data, Afzal et al. (2023) and Ellis et al. (2024) obtained similar contour plots in the plane $(n_t, \log_{10} r)$ with a reheating transfer function $\mathcal{T}_{\rm rh}(k) \neq 1$ from Kuroyanagi et al. (2015, 2021), finding a strong correlation as $n_t \simeq -0.14 \log_{10} r + 0.58$. Ellis et al. (2024) highlighted $n_t \simeq 2.6$ and $\log_{10} r \simeq -14$ as their best fit.

The highly blue tilt required can be in tension with extraradiation bounds, typically parameterized by the number of extra neutrino species ΔN_ν (Allen & Romano 1999). This bound translates into a constraint on the reheating temperature (Figueroa et al. 2024),

$$T_{\rm rh} \lesssim 0.68\ \text{MeV} \times \left[1 + (4.3 \cdot 10^9)\, n_t \left(\frac{f_{\rm pivot}}{f_{\rm BBN}}\right)^{n_t} \left(\frac{0.032}{r}\right)\right]^{\frac{1}{n_t}}, \qquad 68.$$

where we assume only Standard Model fields during RD, fix $\Delta N_\nu \simeq 0.2$ and $A_s = 2.1 \cdot 10^{-9}$ (Planck Collaboration et al. 2020a), and introduce $f_{\rm BBN} \simeq 1.84 \cdot 10^{-11}$ Hz as the frequency today, corresponding to the Hubble scale at the onset of Big Bang nucleosynthesis (BBN). Using Equation 68, Figueroa et al. (2024) find that a low reheating temperature, $2.65\ \text{GeV} \lesssim T_{\rm rh} \lesssim 23.0\ \text{GeV}$ at 95% CL, must be imposed to respect $\Delta N_{\rm eff}$ constraints. This fact is visualized in **Figure 7**, which shows the 2D posterior of $\{n_t, h^2\Omega_{\rm GW}(f_{\rm year})\}$ obtained from the combination of EPTA and NG15 data.

Overall, all analyses point to blue spectral tilts $n_t \gtrsim 1$ and tiny tensor-to-scalar ratios, orders of magnitude below the current CMB constraint of $r < 0.032$ (Tristram et al. 2022). An inflationary signal that strongly violates the consistency relation $n_t = -r/8$, with a blue tilt as large as required by PTA data, $n_t \simeq 2$, is not impossible to conceive. In fact, Vagnozzi (2023) discusses several inflationary (and noninflationary) scenarios that lead to blue tensor spectra. However, such scenarios either do not predict a sufficiently large amplitude at PTA scales (or a sufficiently blue tilt) or are simply ruled out as a result of an excessive amplitude or non-Gaussianity level of scalar perturbations. The added requirement for a very low reheating temperature makes it even more challenging to construct a viable model. Thus, an inflationary explanation of the PTA signal seems rather implausible (however, for an explicit model claiming otherwise, see Unal et al. 2024).

4.2.1. Kination domination after inflation. Equations 67 and 68, above, were obtained under the assumption of instant reheating. If, however, there is an extended reheating period between the end of inflation and the onset of RD, then the GWB from inflation today is affected through the corresponding transfer function of this period. If the latter is modeled by a dominant energy density characterized by a constant equation-of-state ω, then the spectrum receives a tilt (Figueroa & Tanin 2019),

$$\Omega_{\rm GW}^{(0)}(f) \propto f^{n_t + \Delta n_t(\omega)}, \quad \text{with} \quad \Delta n_t \equiv -2\frac{(1-3\omega)}{(1+3\omega)}, \qquad 69.$$

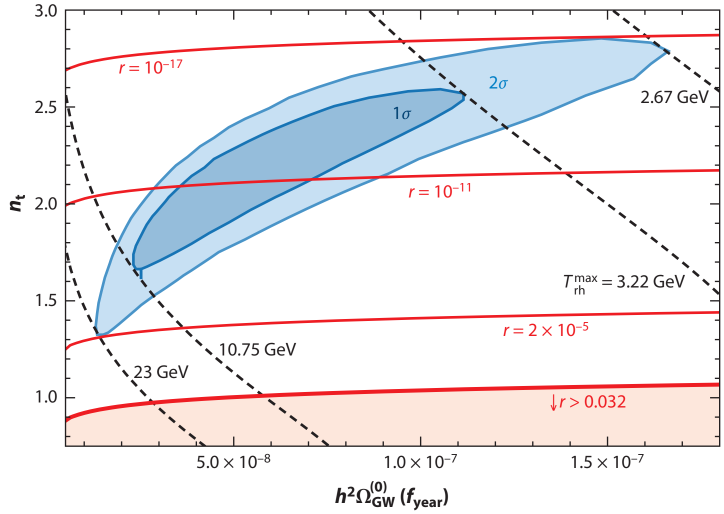


Figure 7

Parameter space n_t versus $h^2\Omega_{GW}(f_{year})$, which is required to explain the EPTA+NG15 data together, under the assumption of an inflationary signal as in Equation 67. The CMB bound on the tensor-to-scalar ratio (*red*) excludes the region with $r > 0.032$. The ΔN_{eff} bound requires a low-scale reheating, $T_{rh} < 23$ GeV, at 95% CL. Abbreviation: CMB, cosmic microwave background. Figure adapted from Figueroa et al. (2024).

thus becoming either more blue for a period of kination[17] domination (KD) or more red for a stage of matter domination (MD). If we assume that between the end of inflation and RD there is an MD phase with $w \simeq 0$ followed by a KD period with $w \simeq 1$, then we can explain the PTA data with the resulting broken PL GWB signal, which grows as $\Omega_{GW}^{(0)}(f) \propto f^{n_t+1}$ in the PTA frequency range but decays as $\propto f^{n_t-2}$ at higher frequencies, so that the BBN constraint is fully respected. Li et al. (2025) analyzed this model, finding a spectral tilt as $n_t = 1.29^{+0.13}_{-0.42}$ at 68% CL and a correlation as $n_t \simeq -0.13\log_{10} r + 0.7$, while allowing for higher reheating temperatures than in the case of instant reheating.

4.2.2. Second-order-induced gravitational waves due to large scalar fluctuations. Primordial scalar curvature fluctuations ζ can generate a GWB at second order in perturbation theory (see Domènech 2021 and references therein). Detectable scalar-induced GWs (SIGWs) at PTA scales from this mechanism require enhanced curvature perturbations that enter the horizon near the quantum chromodynamics (QCD) phase transition. If ζ is Gaussian, then the spectrum today reads

$$\Omega_{GW}^{(0)}(f) = \frac{\Omega_{rad}^{(0)}\mathcal{G}(\eta_c)}{24}\left(\frac{2\pi f}{a(\eta_c)H(\eta_c)}\right)^2 \overline{\mathcal{P}_h^{ind}(\eta_c, 2\pi f)}, \qquad 70.$$

$$\overline{\mathcal{P}_h^{ind}(\eta, k)} = 2\int_0^\infty dt \int_{-1}^1 ds \left[\frac{t(2+t)(s^2-1)}{(1-s+t)(1+s+t)}\right]^2 \overline{I^2(u, v, k, \eta)}\,\mathcal{P}_\zeta(ku)\,\mathcal{P}_\zeta(kv), \qquad 71.$$

[17] Kination refers to the fact that the energy budget of the Universe is dominated by the kinetic energy of a scalar field, which is much larger than its potential energy, leading to the equation of state $1/3 < \omega \leq 1$ (e.g., Figueroa & Tanin 2019).

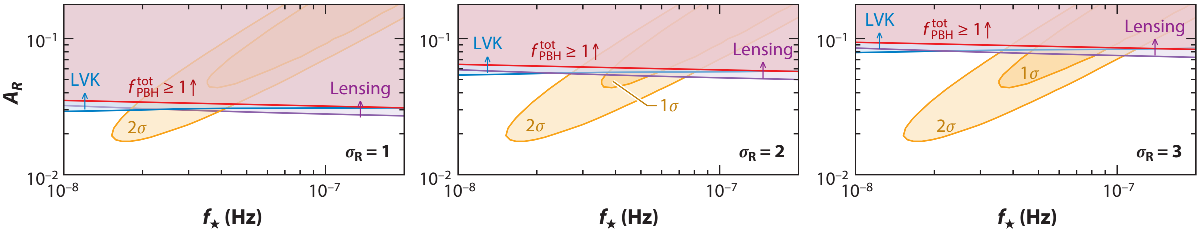


Figure 8

Constraints on the Gaussian scalar perturbation bumps in Equation 72 from PBH abundance (overabundance, lensing, LVK mergers) and other constraints (μ distortion). The parameters are $\{\sigma, f_*, A_R\}$. The second-order scalar-induced GW interpretation of the PTA signal (*yellow*) survives but cannot account for the dark matter totality. Abbreviations: GW, gravitational wave; PBH, primordial black hole; PTA, pulsar timing array. Figure adapted from Figueroa et al. (2024).

where c indicates horizon crossing, η is the conformal time, $\mathcal{P}_\zeta(k)$ is the power spectrum of the curvature perturbation, $u = (1 + s + t)/2$, and $v = (1 - s + t)/2$. Here, $\overline{I^2}$ is given by Kohri & Terada (2018, equation 27).

A broad class of primordial curvature spectra can be characterized by a single peak (bump) (e.g., Kohri & Terada 2021). This is the case in ultra-slow-roll (USR) scenarios (see, e.g., Figueroa et al. 2021, Franciolini et al. 2024, and references therein), which can also lead to the formation of primordial black holes (PBHs) (Hawking 1971) that could constitute all (or a fraction) of the dark matter in the Universe or provide the seed for massive black holes (Carr et al. 2021). As bumplike spectra are naturally generated by USR models, we naturally include SIGW signals in Section 4.2 on inflation. Note, however, that SIGW signals simply require a large scalar power spectrum at scales smaller than those of the CMB, independently of whether the origin of the large scalar power spectrum is inflationary or not. USR inflation is simply a natural scenario in which the necessary ingredients occur.

A bumplike power spectrum around a characteristic scale k_* can be parameterized as

$$\mathcal{P}_\zeta(k) = \frac{A_R}{\sqrt{2\pi}\sigma} e^{-\frac{(\ln k/k_*)^2}{2\sigma^2}}, \qquad 72.$$

where A_R controls the peak amplitude and σ is the width of the bump. Using NG15 data, Ellis et al. (2024) find a best fit of $\log_{10}(\text{Mpc}\, k_*) = 7.7$, $A_R = 0.63$, and $\sigma = 0.21$, whereas Afzal et al. (2023) report $A_R \gtrsim 0.037$. Using NG15+EPTA data, Figueroa et al. (2024) report a peak location $k_* \in [4.57, 88] \cdot 10^7\ \text{Mpc}^{-1}$, $A_R \in [0.10, 2.82]$, and $\sigma \in [0.11, 0.82]$, all at 68% CL. The 2D posteriors of the parameter space $\{A_R, \sigma, k_*\}$ are presented by Ellis et al. (2024, figure 10), Afzal et al. (2023, figure 7), and Figueroa et al. (2024, figure 1).

4.2.2.1. Primordial black hole constraints. Sufficiently large scalar perturbations can form PBHs, with different populations expected to be formed when a broad scalar power spectrum as in Equation 72 is considered. For a typical scale $k_* \sim 10^8\ \text{Mpc}^{-1}$ of SIGWs in the PTA band ($f_* \sim$ 1–10 nHz), PBHs are expected in the mass range $M_{\rm H}/M_\odot \sim$ 0.4–40. While the most conservative constraint is the PBH overabundance $f_{\rm PBH}^{\rm tot} > 1$, many observations also exclude the PBH abundance for $f_{\rm PBH}^{\rm tot} < 1$ over a wide range of masses (Green & Kavanagh 2021). Furthermore, there are two relevant constraints for the PBH mass range that are compatible with PTA data: LVK mergers (both individual events and stochastic background) and gravitational lensing. **Figure 8** shows these three constraints on the log-normal scalar perturbation (Figueroa et al. 2024); the bounds on the monochromatic PBH mass (Green & Kavanagh 2021) have been translated for a PBH mass distribution, using the method outlined by Carr et al. (2017).

4.2.2.2. *Cosmic microwave background and μ distortions.*

While the scalar perturbation spectrum considered above has a peak at PTA scales, its low-momentum tail at scales larger than PTA ($k \lesssim 10^6$ Mpc^{-1}) can also be constrained either via μ distortions in the CMB spectrum (Chluba et al. 2012) or directly via CMB temperature anisotropies when the spectrum is too broad (Hunt & Sarkar 2015). The latter constraint sets an upper bound on the width $\sigma \lesssim 3$ in Equation 72 (**Figure 8**).

Fitting NG15+EPTA data with SIGWs and choosing $\sigma < 3$ (Figueroa et al. 2024) require a peak frequency $f_* \in [0.07, 1.44]$ μHz at 68% CL, which leads to a PBH distribution peaked around masses $M_{\rm H}/M_\odot \in [1.5 \cdot 10^{-5}, 5.6 \cdot 10^{-3}]$. **Figure 8** shows that the region of the PTA best fit is cut by the PBH overabundance criterion, as well as by gravitational lensing and LKV bounds (Green & Kavanagh 2021). In light of **Figure 8**, the PTA signal cannot correspond to an interpretation of PBHs as the dark matter totality for $\sigma \lesssim 1$. Note, however, that to obtain the total fraction of PBHs today, $f_{\rm PBH}^{\rm tot}$, one needs to account for different ingredients, such as collapse threshold considerations, nonlinearities between ζ and the density contrast δ_m, and other effects (for a discussion, see Ellis et al. 2024 and references therein). Studies analyzing PTA data in light of SIGW signals with a bump spectrum have incorporated different ingredients, so the (in)compatibility of PBH bounds with a bump signal fitting PTA data varies from paper to paper. For example, **Figure 8** is based on the considerations of Figueroa et al. (2024), who used the δ_m–ζ nonlinear relation, critical collapse, and peak theory (Young et al. 2019), with a real-space top-hat window function and fixed threshold. Ellis et al. (2024, figure 10), however, incorporated the effects of the QCD phase transition and state-of-the-art threshold criteria, reaching the conclusion that an SIGW signal from a bump spectrum is ruled out at (almost) 3σ by PBH overabundance. Future improvements in PBH abundance calculations (incorporating, e.g., effects beyond gradient expansion at superhorizon scales and nonlinear transfer functions), and certainly more data from LVK, lensing, and PTA observations, will help clarify whether an SIGW signal explaining the origin of the PTA signal can also explain the abundance of dark matter in the Universe, or at least a fraction of it.

4.2.2.3. *Primordial non-Gaussianity.*

For broader generality, some studies have considered a curvature perturbation with local primordial non-Gaussianity (pNG) as $\zeta(\mathbf{x}) = \zeta_{\rm G}(\mathbf{x}) + f_{\rm nl}[\zeta_{\rm G}^2(\mathbf{x}) - \langle\zeta_{\rm G}^2(\mathbf{x})\rangle]$, with $\zeta_{\rm G}$ Gaussian. In this case, the resulting GWB spectrum is the sum of a series of terms (Cai et al. 2019, Unal 2019), a Gaussian contribution as in Equation 72, and non-Gaussian contributions $\propto f_{\rm nl}^2$ and $\propto f_{\rm nl}^4$ (which can be found explicitly in equations 2.33–2.39 of Adshead et al. 2021). Note that the calculation of the PBH abundance in the presence of pNG also needs to be modified (Young 2022). By fitting NG15 data with an SIGW signal with local pNG and a Gaussian component with a bump spectrum, Ellis et al. (2024) conclude that current data cannot constrain $f_{\rm nl}$ at PTA scales. In contrast, Wang et al. (2024) find—using the same data—a stringent constraint as $|f_{\rm nl}| \lesssim 4.1$ at 68% CL. Using the same modeling but fitting to NG15+EPTA+PPTA data, Liu et al. (2024) find that $|f_{\rm nl}| \lesssim 13.9$. Figueroa et al. (2024) claim, however, that pNG can be constrained at PTA scales only if the perturbativity condition $A_R f_{\rm nl}^2 < 1$ is imposed in the analysis, in which case NG15+EPTA data provide a stringent constraint on the pNG parameter as $|f_{\rm nl}| \lesssim 2.34$ at 95% CL (and otherwise remain unconstrained). While more research is needed to reach a clear conclusion, it is remarkable that pNG could be constrained at PTA scales that are approximately eight orders of magnitude smaller than CMB scales (Planck Collaboration et al. 2020b).

4.2.3. Postinflationary preheating.

In many models of particle physics, the reheating of the Universe can initially be driven by preheating effects, typically characterized by exponential particle production (e.g., Kofman et al. 1997). Exponential particle production leads to very efficient production of GWs, although the details depend strongly on the scenario considered (see, e.g., Figueroa & Torrenti 2017 and references therein). In the case of parametric resonance, the inflaton is assumed to have a potential $V(\phi) \propto |\phi|^p$ after inflation and to be coupled via $g^2\phi^2\chi^2$ (where g^2 is the coupling constant) with a secondary daughter (massless scalar) field χ. If the resonance parameter $q_\star \equiv g^2\phi_\star^2/\omega_\star^2$ is $\gg 1$, where $\phi_\star$ is the amplitude at the onset of oscillations and $\omega_\star \equiv V''(\phi_*)$, then fluctuations of the daughter field will grow exponentially via broad resonances for modes $k \lesssim k_\star \sim q_*^{1/4}\omega_\star$. The spectrum of the GWB produced during this process has a peak at approximately the frequency (Figueroa & Torrenti 2017)

$$f_{\rm peak} \simeq 8 \cdot 10^{-9}\ {\rm Hz} \times \left(\frac{\omega_\star}{\rho_\star^{1/4}}\right) \epsilon_\star^{\frac{1}{4}} q_\star^{\frac{1}{4}+\eta}, \tag{73}$$

where $\rho_\star$ is the total initial energy density in the Universe, $\eta \sim 0.3$–0.4 is a parameter accounting for nonlinear effects, and $\epsilon_\star$ parameterizes the period between the end of inflation and the onset of RD. Figueroa & Torrenti (2017) find, for $V(\phi) \propto \phi^2$ and $q_\star \in (10^4, 10^6)$ (assuming $\epsilon_\star = 1$), a GWB peak amplitude today as large as $\Omega_{\rm GW,\,0} \simeq 10^{-12}$–$10^{-11}$, at a frequency $f_{\rm peak} \simeq 10^8$–10^9 Hz. Similarly, for $V(\phi) \propto \phi^4$ with $q_\star \in (1, 10^4)$, Figueroa & Torrenti (2017) find a multipeak GWB spectrum with a leading peak amplitude $\Omega_{\rm GW,\,0} \simeq 10^{-13}$–$10^{-11}$ and frequency today $f_{\rm peak} \simeq 10^7$–10^8 Hz. Clearly, given the high frequencies involved, the PTA data cannot be explained by these phenomena.

GWs can also be efficiently produced when preheating leads to oscillons, which are long-lived compact scalar field configurations (e.g., Amin & Shirokoff 2010). Their dynamics can efficiently source GWs arising from field interactions and collisions among one another (Helfer et al. 2019), but the GWB peak today is typically well above LVK frequencies. GWs can also be efficiently produced in preheating by fields with spin $\neq 0$. For example, GWs can be generated during the out-of-equilibrium production of fermions after inflation for both spin-1/2 and spin-3/2 fields. Similarly, GWs can be generated when the particles thus produced are Abelian or non-Abelian gauge fields, when coupled to charged scalar fields (Dufaux et al. 2010), or when coupled to a pseudoscalar field via axial coupling. Preheating can be remarkably efficient in the second case, resulting in a GW energy density as large as $\Omega_{\rm GW} \lesssim \mathcal{O}(10^{-6})$ for certain couplings (Adshead et al. 2020), but still peaked at very high frequencies. Neither case can produce a GWB signal within PTA scales, even though they constitute very efficient early Universe GW generation mechanisms.

4.3. Phase Transitions

First-order phase transitions in the early Universe occur whenever a potential barrier emerges that separates higher (true vacuum) and lower (false vacuum) minima in a scalar field potential. A phase transition is triggered when quantum tunneling or thermal fluctuations take the scalar field amplitude over the potential barrier in some localized region. This leads to the nucleation of spherical bubble configurations, within which the scalar field reaches the true vacuum whereas outside the scalar field still resides in the false vacuum. If the nucleated bubbles have a radius above a critical scale, they expand and eventually collide with one another. The expansion and collision of the bubbles, the sound waves generated in the plasma as a result of its coupling to the scalar field, and possible turbulent motions of the plasma can all source GWs very efficiently, leading to a large and potentially observable GWB (for a review, see Hindmarsh et al. 2021 and references therein).

The resulting GWB produced during a phase transition can therefore be a superposition of the bubble collision, sound-wave, and turbulence contributions (Hindmarsh et al. 2014). In the case of magnetohydrodynamic (MHD) turbulence at the QCD scale, GWs are generated exactly around

Table 2 Parameters and functions characterizing the bubble, sound-wave, and turbulence contributions to the gravitational wave background spectrum from a first-order phase transition[a]

	Bubbles	Sound waves	Turbulence
$\mathcal{N}_i$	1	0.16	20
$\Delta_i(v_{\mathrm{w}})$	$\frac{0.11\, v_{\mathrm{w}}^3}{0.42+v_{\mathrm{w}}^2}$	v_{w}	v_{w}
p_i	2	2	3/2
q_i	2	1	1
$s_i(f, f_*)$	$\frac{3.8(f/f_*)^{2.8}}{1+2.8(f/f_*)^{3.8}}$	$\left(\frac{f}{f_*}\right)^3 \left[\frac{7}{4+3(f/f_*)^2}\right]^{\frac{7}{2}}$	$\frac{(f/f_*)^3}{(1+f/f_*)^{11/3}(1+8\pi f/\tilde{H}_*)}$
f_*	$11\ \mathrm{nHz}\left(\frac{0.62}{1.8-0.1 v_{\mathrm{w}}+v_{\mathrm{w}}^2}\right)\mathcal{F}$	$13\ \mathrm{nHz}\ v_{\mathrm{w}}^{-1}\mathcal{F}$	$19\ \mathrm{nHz}\ v_{\mathrm{w}}^{-1}\mathcal{F}$

[a] $$\left\{\begin{array}{l} \mathcal{F} \equiv \left(\frac{\beta}{H_*}\right)\left(\frac{T_*}{0.1\ \mathrm{GeV}}\right)\left(\frac{g_*(T_*)}{10.75}\right)^{\frac{1}{2}}\left(\frac{g_{*,\mathrm{s}}(T_*)}{10.75}\right)^{-\frac{1}{3}}, \\ \tilde{H}_* \equiv H_*\left(\frac{a_*}{a_0}\right) \simeq 11\ \mathrm{nHz}\left(\frac{T_*}{0.1\ \mathrm{GeV}}\right)\left(\frac{g_*(T_*)}{10.75}\right)^{\frac{1}{2}}\left(\frac{g_{*,\mathrm{s}}(T_*)}{10.75}\right)^{-\frac{1}{3}} \end{array}\right\}.$$
Bubble, sound-wave, and turbulence contributions are from Huber & Konstandin (2008), Hindmarsh et al. (2015), and Caprini et al. (2009b), respectively.

the PTA (approximately nanohertz) frequencies (Neronov et al. 2021), whereas in the case of a dark sector, the GWB could peak across a wide frequency range (Schwaller 2015). The spectrum today of the GWB from a first-order phase transition can be written as

$$\Omega_{\mathrm{GW}}^{(0)} = \Omega_{\mathrm{rad}}^{(0)}\mathcal{G}(T_*)\sum_i \Omega_{\mathrm{GW},i}(f), \quad \Omega_{\mathrm{GW},i}(f) \equiv \mathcal{N}_i\, \Delta_i(v_{\mathrm{w}})\left(\frac{\kappa_i\, \alpha}{1+\alpha}\right)^{p_i}\left(\frac{H}{\beta}\right)^{q_i} s_i(f), \qquad 74.$$

where $\Omega_{\mathrm{rad}}^{(0)} \simeq 9.2 \cdot 10^{-5}$; $\mathcal{G}(T) \equiv [g(T)/g_0][g_{\mathrm{s},0}/g_{\mathrm{s}}(T)]^{4/3}$ accounts for the number of relativistic degrees of freedom at temperature T; the index i stands for b (bubbles), sw (sound waves), or tb (turbulence); and α, β, and T_* denote the strength, (inverse) duration, and temperature of the phase transition, respectively. **Table 2** summarizes the normalization factors $\mathcal{N}_i$, exponents p_i and q_i, efficiency factor κ_i, and function $\Delta_i(v_{\mathrm{w}})$, depending on the wall velocity v_{w}. The spectral shapes $s_i(f)$, also presented in **Table 2**, are essentially a broken PL ($i =$ b, sw) or a broken double PL ($i =$ tb). The total spectrum can contain various peaks arising from the various contributions. However, we focus on scenarios dominated by a single contribution (the peak frequency today, f_*, shown in **Table 2**). Using both the peak position and amplitude of each contribution, we can convert parameters extracted from a broken (double) PL fit to PTA data into phase transition parameters from different scenarios, considering either strong or weak to mild first-order phase transitions. For an example of a single-peaked signal corresponding to a strong first-order phase transition, see **Figure 1** (see also Section 4.3.1).

4.3.1. Strong first-order phase transition. If there is a supercooling phase with $\alpha \gg 1$, a strong first-order phase transition takes place in vacuum, and the bubble collisions are the dominant contribution to the GWB. According to the envelope approximation (Huber & Konstandin 2008), the shape of the IR tail of the spectrum scales as $\Omega_{\mathrm{GW}} \propto f^{2.8}$ and the UV tail as $\Omega_{\mathrm{GW}} \propto f^{-1}$ (**Table 2**). Although other studies have found different slope values (in the range $f^{[-3,-0.9]}$ for the IR tail and $f^{[-3,-1]}$ for the UV slopes), we stick to the envelope approximation simply to make a concrete choice. Using **Table 2** for the peak frequency f_* and Equation 74 for the GWB amplitude $h^2\Omega_{\mathrm{GW}}^{(0)}(f_*)$, we can translate a broken PL fit into the duration and temperature of a strong first-order phase transition as follows (Figueroa et al. 2024):

$$\left(\frac{\beta}{H_*}\right)_{\rm b} \simeq 206\left[\Delta_{\rm b}(v_{\rm w})\kappa_{\rm b}^2\mathcal{G}(T_*)\left(\frac{10^{-9}}{h^2\Omega_{\rm GW}^{(0)}(f_*)}\right)\right]^{1/2}, \tag{75}$$

$$(T_*)_{\rm b} \simeq 0.48\ {\rm MeV}\left(\frac{f_*}{10\ {\rm nHz}}\right)\left(\frac{h^2\Omega_{\rm GW}^{(0)}(f_*)}{10^{-9}}\right)^{\frac{1}{2}}\left(\frac{g_{*,s}(T_*)}{g_*(T_*)}\right)\left(\frac{1.8-0.1v_{\rm w}+v_{\rm w}^2}{0.62\kappa_{\rm b}\Delta_{\rm b}^{1/2}}\right). \tag{76}$$

Here, the vacuum energy is efficiently converted into bubble energy (i.e., $\kappa_{\rm b}=1$), the wall velocity reaches the speed of light ($v_{\rm w}\to 1$), and $\Delta_{\rm b}\approx 0.08$. For relativistic bubble velocities, the average bubble separation in units of the Hubble radius at the time of percolation is related to the bubble nucleation rate β through the relation $H_*R_* = (8\pi)^{1/3}H_*/\beta$, so alternatively constraints on β/H_* or H_*R_* can be provided.

Fitting NG15+EPTA for bubble collisions with $\alpha \gg 1$, Figueroa et al. (2024) find $T_* \in [6.2\cdot 10^{-3}, 9.0]$ GeV and $\beta/H_* \in [2.2, 500]$, at 68% CL, with a best-fit temperature $T_* \simeq 20$ MeV. Fitting only NG15 data for the same case, Ellis et al. (2024) find a best fit $T_* = 0.34$ GeV and $\beta/H = 6.0$. Fitting NG15 but allowing α to vary, Afzal et al. (2023) find $\alpha > 1.1$ (0.29), $H_*R_* > 0.28$ (0.14), and $T_* \in [0.047, 0.41]$ ([0.023,1.75]) GeV at 68% (95%) CL. However, a strong phase transition with low temperature, $T_* \sim$ 10–100 MeV, as required by PTA data, could affect BBN (Bringmann et al. 2023) and would be in tension with lattice QCD, which predicts a crossover (Aoki et al. 2006), not a first-order phase transition.

4.3.2. Weak to moderate first-order phase transition. When a phase transition is not too strong, if sound waves are created, they typically dominate GW production over other contributions. This occurs when bubbles expand in a thermal medium, reaching a terminal velocity thanks to their interaction with the plasma (i.e., a nonrunaway scenario). Hindmarsh et al. (2015) find the shape of the GWB as a PL with $\propto f^3$ and $\propto f^{-4}$ IR and UV tails. Although recent simulations (Roper Pol et al. 2024) have found an intermediate slope associated with the width of the sound shell, behaving as $\propto f$ close to the peak, we stick to the above broken PL for simplicity. The peak amplitude and position in Equation 74 (**Table 2**) can be translated into phase transition parameters via

$$\left(\frac{\beta}{H_*}\right)_{\rm sw} \simeq 68\left(\frac{10^{-9}}{h^2\Omega_{\rm GW}^{(0)}(f_*)}\right)\mathcal{G}(T_*)v_{\rm w}\left[\frac{(1+\alpha)^2\alpha^{-2}\kappa_{\rm sw}^{-2}(\alpha)}{100}\right]^{-1}, \tag{77}$$

$$(T_*)_{\rm sw} \simeq 1.5\ {\rm MeV}\left(\frac{f_*}{10\ {\rm nHZ}}\right)\left(\frac{h^2\Omega_{\rm GW}^{(0)}(f_*)}{10^{-9}}\right)\left[\frac{10.75}{g_*(T_*)}\right]^{\frac{3}{2}}\left[\frac{g_{*,s}(T_*)}{10.75}\right]^{\frac{5}{3}}\left[\frac{(1+\alpha)}{10\alpha\kappa_{\rm sw}(\alpha)}\right]^2, \tag{78}$$

where the efficiency factor is $\kappa_{\rm sw}(\alpha) \simeq \alpha(0.73+0.083\sqrt{\alpha}+\alpha)^{-1}$ (Espinosa et al. 2010). While the degeneracy among α, β/H_*, and T_* can be broken by considering UV completions of a low-T first-order phase transition, PTA data by themselves cannot break such degeneracy.

Turbulent motion can be a relevant GW contribution during a weak to mild first-order phase transition. Therefore, it is common to expect both sound waves and turbulence in this case, typically with $\kappa_{\rm turb} < \kappa_{\rm sw}$. In the case of MHD turbulence, so-called MHD eddies form during the phase transition, but MHD turbulence can also arise as a result of other, unrelated mechanisms (Kamada & Shin 2020). Either way, GWs in this case are sourced by the magnetic field present, characterized by an energy density fraction at the time of production t_*, scaling as $\Omega^*_{\rm GW} \sim 3(\Omega^*_B)^2(H_*\mathcal{T})^2$ (Neronov et al. 2021), where $\Omega^*_B = B_*^2/(2\rho_{{\rm c},*})$ is the magnetic energy density fraction and $\mathcal{T} \sim \lambda_*/v_{\rm A} \lesssim H_*^{-1}$ is the turbulence-sourcing timescale (with $v_{\rm A} \simeq \sqrt{2\Omega^*_{\rm B}}$

the Alfvén speed and λ_* the characteristic scale of turbulence). The GWB amplitude scales as $\Omega^*_{\rm GW} \sim (\Omega^*_B)^n (H_*\lambda_*)^2$, where $n \in [1, 2]$, depending on the details. The GWB spectrum therefore has a peak amplitude and frequency today of $\Omega^{(0)}_{\rm GW} \sim \Omega^{(0)}_{\rm rad}\mathcal{G}(T_*)(\Omega^*_B)^n(H_*\lambda_*)^2$ and $f_* = 2/l_* \cdot a_*/a_0$ (Neronov et al. 2021). Recent numerical simulations of GW production from MHD turbulence (Roper Pol et al. 2022) have shown that the envelope of the GWB thus produced can be estimated analytically under the assumption that the total anisotropic stress varies more slowly than the production of GWs (for validation of this assumption in simulations of purely kinetic turbulence, see also Auclair et al. 2022). Caprini et al. (2009b) determined the GWB spectrum in the case of turbulence dynamics as a double-broken PL, with an IR causality tail $\Omega_{\rm GW} \propto f^3$, an intermediate slope $\Omega_{\rm GW} \propto f^2$ within scales $\tilde{H}_* < f < f_*$, and a UV tail $\Omega_{\rm GW} \propto f^{-5/3}$. Roper Pol et al. (2022) estimated that the UV tail scales instead as $\Omega_{\rm GW} \propto f^{-8/3}$.

For the largest processed eddies (i.e., those with maximum length scale $l_* = v_{\rm A}/H_*$), the peak amplitude and frequency of the GWB spectrum are related to the magnetic energy fraction and temperature at the time of GW production (Neronov et al. 2021, Figueroa et al. 2024):

$$\Omega^*_B \simeq \left[1.4 \cdot 10^{-5} \left(\frac{h^2 \Omega^{(0)}_{\rm GW}(f_*)}{10^{-9}}\right)\left(\frac{10.75}{g_*(T_*)}\right)\left(\frac{g_{*,s}(T_*)}{10.75}\right)^{\frac{4}{3}}\right]^{1/(n+1)}, \quad 79.$$

$$f_* \simeq 16.1\ \mathrm{nHz} \left(\frac{T_*}{\mathrm{MeV}}\right)\left(\frac{10^{-2}}{\Omega^*_B}\right)^{\frac{1}{2}}. \quad 80.$$

Alternatively, today's magnetic field strength B_0 and correlation length scale $l_B \equiv l_*(a_0/a_*)$ are given by (Neronov et al. 2021, Figueroa et al. 2024)

$$B_0 = 0.4\ \mu\mathrm{G} \left(\frac{\Omega^*_B}{10^{-2}}\right)^{\frac{1}{2}} \left(\frac{g_*(T_*)}{10.75}\right)^{\frac{1}{2}} \left(\frac{10.75}{g_{*,s}(T_*)}\right)^{\frac{2}{3}}, \quad l_B \simeq 1.95\ \mathrm{pc}\left(\frac{10\ \mathrm{nHz}}{f_*}\right). \quad 81.$$

Fitting a signal from sound waves and turbulence to NG15+EPTA data, Figueroa et al. (2024) find $T_* \in [5.5, 360]$ MeV, $\beta/H_* \in [1.0, 15]$, and $\alpha \in [0.017, 0.060]$ at 68% CL, with a best-fit temperature $T_* \simeq 94$ MeV. Fitting to NG15+EPTA data a contribution from MHD turbulence alone, they find $T_* \in [10^{-3}, 3.2]$ GeV at 68% CL, which, by translating the best-fit peak position into a primordial magnetic field (Neronov et al. 2021), leads to a strength $B_0 \simeq 23.3$ μG and correlation length $l_B \simeq 0.25$ pc—close to current constraints on primordial magnetic fields (Acciari et al. 2023, Neronov et al. 2023). Fitting NG15 data alone to a signal from only sound waves, Afzal et al. (2023) find $\alpha_* > 0.42$ (0.37), $H_*R_* \in [0.053, 0.27]$ ([0.046,0.89]), and $T_* \in [4.7, 33]$ ([2.7,93]) MeV at 68% (95%) CL. Fitting EPTA data alone to a signal from MHD turbulence only, EPTA Collaboration et al. (2024a, figure 18) show 2D posteriors for λ_*H_*, Ω^*_B, and T_*. They comment that MHD turbulence provides a good fit to the data only in the limit of large Ω_B (close to the upper bound of their prior $\Omega_B \leq 1$) and observe a clear degeneracy between λ_*H_* and Ω_B, as expected from a signal scaling as $\Omega^*_{\rm GW} \sim (\Omega^*_B)^n(H_*\lambda_*)^2$.

4.4. Topological Defects

Topological defects are stable configurations that may form upon spontaneous symmetry breaking during a phase transition (Kibble 1976). Depending on the topology of the associated vacuum manifold (characterized by the homotopy groups Π_n; $n = 1, 2, 3, \ldots$), different types of defects (or no defects at all) may appear. Independently of the type, defects can be global or local, depending on whether the symmetry broken is global or gauge. Either way, in all cosmologically viable cases (i.e., those that do not overclose the Universe), defects appear everywhere in the Universe upon completion of the phase transition, forming a network. The defect network soon reaches a scaling regime, where the mean separation between defects tracks the horizon during cosmic evolution. Cosmic strings (CSs), corresponding to 1D configurations, are topological solutions that appear in a phase transition when $\Pi_1 \neq \mathcal{I}$. Domain walls (DWs), in contrast, emerge in phase transitions driven by the breaking of a discrete symmetry and correspond to 2D topological solutions in the form of sheets. Other types of defects can emerge, depending on the topology of the vacuum manifold.

Independently of the type of defect, as the network's energy–momentum tensor adapts itself to maintaining scaling, the network emits GWs (Figueroa et al. 2013). This leads to a scale-invariant GWB with amplitude $\Omega_{\mathrm{GW}} \propto (v/m_{\mathrm{p}})^4$ for scaling networks (except in the case of DWs), where v is the vacuum expectation of the scalar field within $\mathcal{M}$. Current CMB constraints set $v \lesssim 10^{-3}$ (Lizarraga et al. 2016, Lopez-Eiguren et al. 2017), making this GWB too small to explain the PTA signal. In the case of CSs, however, in addition to the GWs emitted by the network (infinite) strings during scaling, there is expected to be a much stronger emission from the loops (Vilenkin 1981), which might explain PTA data (for a review on GWB signals from CS networks, see Gouttenoire et al. 2020, and for a classification of different string network scenarios, see Dimitriou et al. 2025). DWs, in contrast, are not viable defects unless a mechanism is in place such that they annihilate at a given energy scale, before their energy density overtakes the energy budget of the Universe. In viable cases, DW networks emit GWs until their disappearance at an annihilation temperature $T = T_{\mathrm{ann}}$ (Gelmini et al. 1989). The resulting GWB emitted before annihilation might also explain PTA data.

4.4.1. Local strings and superstrings. Local strings may emerge after symmetry breaking in gauge theories. They have a finite core's width, corresponding to the inverse mass scale of the particles involved (both scalar and gauge fields). As this width is much smaller than the typical length scale of the strings, the motion of the strings is expected to be well-approximated by the Nambu–Goto (NG) dynamics of infinitely thin strings (Vilenkin & Shellard 2000). Under this assumption, the spectrum of the GWB emitted by a network of local CSs, in particular by the loop distribution along cosmic history, can be written as (Blanco-Pillado & Olum 2017)

$$\Omega^{(0)}_{\mathrm{GW}}(f) \equiv \frac{16\pi\,\Gamma(G\mu)^2}{3p\zeta(q)H_0^2}\frac{1}{f}\sum_{j=1}^{\infty}\frac{\mathcal{C}_j(f)}{j^{q-1}}, \quad \mathcal{C}_j(f) = \int \frac{dz'}{H(z')(1+z')^6}\, n\!\left(\frac{2j}{(1+z')f}, t_{\mathrm{e}}(z')\right), \qquad 82.$$

where G is Newton's constant, $G\mu$ is the string tension in Planck units, $\zeta(q)$ is the Riemann zeta function, p is the intercommutation probability, $H(z)$ and H_0 are the Hubble rate at redshift z and today, and $n[l_{\mathrm{e}}(z)$ and $t_{\mathrm{e}}(z)]$ is the number density of loops of size $l_{\mathrm{e}}(z)$ at emission time $t_{\mathrm{e}}(z)$. Because Equation 82 is valid for NG strings, the formulation also applies to superstrings. Therefore, one considers $p = 1$ for field theory local strings and $10^{-3} \leq p < 1$ for superstrings.

To evaluate the signal, Figueroa et al. (2024) consider a velocity one-scale modeling, with loop birth length $l = 0.1/H$ (Blanco-Pillado et al. 2014); $\Gamma = 50$ (Blanco-Pillado & Olum 2017); and $q = 4/3$, corresponding to cusps. Fitting NG15+EPTA data to local strings represented by NG strings with $p = 1$ and the above parameters leads to $\log_{10}(G\mu) = -9.90^{+0.11}_{-0.19}$ at 68% CL (this signal is plotted in **Figure 1**). This value translates into a very tight constraint on the associated energy scale as $\sqrt{\mu} \simeq 1.32^{+0.20}_{-0.24} \cdot 10^{14}$ GeV, which can be accommodated in a variety of particle physics models (Vilenkin & Shellard 2000) despite being $\mathcal{O}(10^{-2})$ smaller than the standard Grand Unified Theory scale $\sim 10^{16}$ GeV. This GWB spectrum, if viable, would be accessible to higher-frequency experiments and thus would allow us to probe the cosmological history above the BBN $\sim$MeV scale (Cui et al. 2018, Gouttenoire et al. 2021, Auclair et al. 2023). However, this signal does not fit the PTA data well. Fitting the NG15+EPTA data instead to a superstring signal (Damour & Vilenkin 2005), represented by NG strings with $p < 1$, Figueroa et al. (2024) find $\log_{10}(G\mu) = -11.83^{+0.27}_{-0.15}$ and $\log_{10} p = -2.63^{+0.49}_{-0.31}$ at 68% CL. This result fits current PTA data very well because it accommodates a smaller energy scale, $\sqrt{\mu} \simeq 1.56^{+0.48}_{-0.39} \cdot 10^{13}$ GeV, while suppressing the intercommutation probability. Afzal et al. (2023) consider more CS modelings, including NG strings with only cusps ($q = 4/3$) or only kinks ($q = 5/3$) in loops, NG strings emitting GWs only at the fundamental frequency $f = 2/l_e$ of loops, and strings and GW emission as calibrated from NG simulations (Blanco-Pillado et al. 2011, 2015). Fitting all these modelings against NG15 data, they obtain peaked distributions centered around values $\log_{10} G\mu \sim -(10.5\ldots10.0)$. Extending the analysis to superstrings, they find a posterior density maximized at small intercommutation probabilities $\log_{10} p \sim -3$ (at the edge of their prior range) and CS tensions $\log_{10} G\mu \sim -12$. However, this parameter region is in tension with the LVK bounds (Abbott et al. 2021). Afzal et al. (2023) also analyze metastable strings, which are string networks that are unstable against the nucleation of Grand Unified Theory monopoles (Buchmuller et al. 2021). Fitting NG15 data, they find a decay parameter value of around $\sqrt{\kappa} \sim 8$, along with a large tension of $\log_{10} G\mu \sim -(7\ldots4)$.

4.4.2. Global strings. Global strings form after the breaking of a global symmetry with a nontrivial fundamental homotopy group [e.g., breaking of a global $U(1)$ symmetry]. Unlike local strings, they have a spatially extended core profile, with a tension scaling radially as $\propto \log\,[r/\eta^{-1}]$. As a consequence, they naturally sustain long-range interactions, and Goldstone boson emission is subjected to a strong $\Delta N_{\rm eff}$ bound.

Global strings are naturally generated in BSM models featuring an additional global $U(1)$ symmetry, which is spontaneously broken by the vacuum expectation value of a complex scalar field, thereby providing a Nambu–Goldstone boson. The Peccei–Quinn $U(1)$ symmetry that is proposed to solve the strong-*CP* problem, and its associated axion particle, is the paradigmatic example. Because the $U(1)$ symmetry is also broken explicitly at later times, the axion acquires a mass, and at that moment DWs are also formed (Sikivie 1982). If $U(1)$ is broken after inflation, then CSs generate axion particles throughout the cosmic history (Davis 1986). In this case, the resulting population of loops generate a GWB, while the DWs provide additional GW production (discussed in Section 4.4.3, below).

Servant & Simakachorn (2023) fitted NG15 data to the GWB predictions from a postinflationary axion model that naturally leads to global strings. They found bounds for both low and high axion mass ranges, which are associated with signals from axionic global strings ($N_{\rm DW} = 1$) and DWs ($N_{\rm DW} > 1$), respectively, where $N_{\rm DW}$ is the number of DWs. In the low axion mass region, they found the strongest constraint on the axion decay constant as $f_a < 2.8 \times 10^{15}$ GeV for an axion mass $m_a \ll 10^{-17}$ eV. This value is competitive with the $\Delta N_{\rm eff}$ bound. At high masses, $0.1\ {\rm GeV} \lesssim m_a \lesssim 10^3$ TeV, a substantial region corresponding to $m_a(f_a/N_{\rm DW})^2 \gtrsim 2 \times 10^{11}\ {\rm GeV}^3$ can be excluded when DWs decay in the temperature range $T_* \propto \sqrt{V_{\rm bias}} \sim 1$–300 MeV. Below, we comment more on direct fits of the data to DWs.

4.4.3. Domain walls. DWs (Vilenkin 1985) may emerge in the breaking of a discrete symmetry during a phase transition. While they reach scaling, like any other cosmic defect, their energy density ratio to the critical energy evolves as $\rho_{\rm DW}/\rho_{\rm c} = \mathcal{O}(1)\sigma H/m_{\rm p}^2 H^2 \propto t$ (Kibble 1976), where σ is the wall tension (i.e., surface mass density). If they remain in scaling for too long, they would overclose the Universe. Thus, a bias term in the Lagrangian is needed so that they annihilate at a given temperature, $T = T_{\rm ann}$ (Gelmini et al. 1989), before they dominate the energy budget of the Universe. DWs emit GWs until they annihilate, with a peak frequency $f_{\rm p} = f_{\rm H}(T_{\rm ann})$ and a high-frequency tail $\propto f^{-1}$. The GW spectrum at annihilation can be approximated by (Ferreira et al. 2023)

$$\Omega_{\rm DW}(f,T_{\rm ann})=\frac{3\epsilon\alpha_*^2}{8\pi}\left(\frac{1}{4}\left[\frac{\Omega_{\rm CT}(f_{\rm p})}{\Omega_{\rm CT}(f)}\right]^{1/\delta}+\frac{3}{4}\left[\frac{f}{f_{\rm p}}\right]^{1/\delta}\right)^{-\delta}, \quad 83.$$

where $\epsilon = 0.7$, $\delta = 1$, and $\alpha_{\equiv}\rho_{\rm DW}(T_{\rm ann})/\rho_{\rm rad}(T_{\rm ann})$ is the DW–to–radiation energy density ratio.

If DWs annihilate into dark radiation, one obtains $\Delta N_{\rm eff} = 13.6$ and $g_*(T_{\rm ann})^{-1/3}\alpha_*$ (Ferreira et al. 2023), which must respect BBN and CMB constraints. If DWs decay into Standard Model particles, then BBN and CMB impose $T_{\rm ann} \gtrsim 5$ MeV (de Salas et al. 2015). Ellis et al. (2024) fitted the DW model (see Equation 83) to NG15 data, obtaining as best fits $T_{\rm ann} = 0.85$ GeV and $\alpha_* = 0.11$. While this result is consistent with annihilation into Standard Model particles, it is in tension with $\Delta N_{\rm eff}$ bounds if DWs decay into dark radiation (although smaller α_* values within 1σ remain viable). Like first-order phase transitions, DW annihilation may produce PBHs (Ferrer et al. 2019, Gelmini et al. 2023). If efficient, PBH formation could exclude regions favored by NG15 data, though further analysis is needed.

4.5. Other Signals

In addition to GWBs from inflation, phase transitions, and cosmic defects, other backgrounds may have been generated in the early Universe. For instance, the thermal plasma formed after reheating is guaranteed to produce a GWB. The spectrum of this background is determined by the particle content and the maximum temperature $T_{\rm max}$ reached by the plasma. The GW energy density spectrum scales as $\Omega_{\rm GW} \propto f^3$ in the IR regime and reaches a maximum at a peak frequency (Ghiglieri & Laine 2015) $f^{\Omega}_{\rm peak} \sim 80$ GHz. While the amplitude of this GWB can be very large at such a high-frequency peak, it is extremely suppressed at smaller frequencies as a result of the $\propto f^3$ scaling. Clearly, therefore, under no circumstances can this background explain a signal at PTA frequencies.

Another possibility is the production of a dark $U(1)$ gauge field coupled to an axion-like particle via $\alpha/(4f_a)\phi F_{\mu\nu}\tilde{F}^{\mu\nu}$ (Machado et al. 2019). This mechanism is similar to particle production during preheating, but in a different context. Specifically, long after the end of inflation and reheating, the oscillations of this axion-like particle lead to chiral instability of the gauge field, which sources a large GWB. This scenario is referred to as audible axions. The peak frequency and amplitude of the resulting GWB depend on the axion mass m_a and the decay constant f_a (Machado et al. 2019). The spectrum today can be fitted as

$$\Omega^{(0)}_{\rm GW}(f)=\frac{6.3\,\mathcal{A}_*}{\left[\frac{f}{2f_*}\right]^{-3/2}+\exp\left[12.9\left(\frac{f}{2f_*}-1\right)\right]}, \quad 84.$$

where $f_* \simeq 3.8 \cdot 10^{-9}$ Hz$(\alpha\theta/10)^{2/3}[m_a/(10^{-12}$ meV$)]^{1/2}$ and $\mathcal{A}_* \simeq 3.41 \cdot 10^{-6}\left(f_a/m_{\rm p}\right)^4\theta^{8/3}(10/\alpha)^{4/3}$, with $\alpha \gtrsim \mathcal{O}(10)$ and $\theta \sim \mathcal{O}(1)$ the initial misalignment angle (Machado et al. 2019). The spectrum thus has an IR tail $\propto f^{3/2}$, which extends to high frequencies until it dies off exponentially at $f > f_*$.

Using Equation 84, Figueroa et al. (2024) fitted $\{f_*, \mathcal{A}_*\}$ to NG15+EPTA data, obtaining the following lower bounds on the axion mass and decay constant at 68% CL: $m_a \gtrsim 8.0 \cdot 10^{-11}$ meV and $f_a \gtrsim 1.3 \cdot 10^{18}$ GeV. Imposing a sub-Planckian $f_a \le m_{\rm p}$ motivated by string theory turns this fit into a very tight constraint on $m_a = 0.32^{+1.48}_{-0.24}$ PeV. However, this preferred parameter region is in tension with axion overproduction and $\Delta N_{\rm eff}$ constraints (Geller et al. 2024). Ellis et al. (2024) multiplied the profile in Equation 84 by a spectral function that accounts for the causality tail of the spectrum at superhorizon frequencies (Caprini et al. 2009a). By fitting to the NG15 data, they found best-fit values of $m_a = 3.1 \times 10^{-11}$eV and $f_a = 0.87\,M_{\rm P}$. They found a larger range for the axion mass below super-Planckian f_a values, although part of this region is ruled out by superradiance constraints (Zhang et al. 2021). The discrepancy between the results obtained by Ellis et al. (2024) and Figueroa et al. (2024) is likely due to the use of NG15 versus NG15+EPTA data as well as the use of the causality tail.

4.6. Deterministic Signals

In addition to the GWB signals discussed in Sections 4.2–4.5, BSM scenarios can lead to deterministic signals in PTAs. A paradigmatic example is ultralight dark matter (ULDM)—also known as fuzzy dark matter—which must be a bosonic species with a tiny mass, bounded from below by the CMB as $m_\phi > 10^{-24}$ eV (Hlozek et al. 2018). ULDM models generically suppress structures on small scales, potentially solving the so-called small-scale structure problem of the standard ΛCDM paradigm (Bullock & Boylan-Kolchin 2017). PTAs can probe their tiny mass because the frequency of the PTA signal they cause is proportional to the ULDM mass, $f_* \sim m_\phi/2\pi$, so the sensitivity window of PTA observations falls within 10^{-23} eV $\lesssim m_\phi \lesssim 10^{-20}$ eV. PTA searches for ULDM in such a mass range (e.g., Porayko et al. 2018) can be viewed as complementary to other astrophysical searches, such as measurements of the Lyman-α forest, galactic subhalo mass functions, stellar kinematics, or the lack of observation of superradiance at supermassive black holes.

ULDM models can create various effects in PTA data. In a seminal paper, Khmelnitsky & Rubakov (2014) showed that the travel time of a pulsar radio beam is affected by the metric fluctuations induced by ULDM particles, making PTAs excellent probes for such species. This effect is due to the oscillations of the ULDM field, which induce fluctuations in the local stress–energy tensor $T_{\mu\nu}$ and, hence, generate fluctuations in the metric perturbations through Einstein's equations. Such metric perturbations, which are not GWs but rather time-dependent gravitational potentials, nonetheless affect the photon geodesic on its path from the pulsar, generating a timing residual. Other effects from ULDM models might cause timing residuals, such as Doppler $U(1)$ forces for vector ULDM and pulsar spin fluctuations or reference clock shifts, which are specific to scalar ULDM (for discussion, see Afzal et al. 2023 and references therein).

Furthermore, various theories of dark matter leave unique fingerprints on primordial perturbations on smaller scales, giving rise to different predictions for the amount of subgalactic dark matter substructures. Consequently, measuring local dark matter substructures can be help determine the correct dark matter scenario. For example, a very simple example of such small-scale dark matter substructures is that of PBHs, which can form in inflationary setups with large density fluctuations on small scales (see Section 4.2.2). Galactic PBH populations can be investigated through analysis of the Doppler and Shapiro signals they can leave in PTA data. In particular, the Doppler signal is due to the apparent shift in the pulsar spin frequency, generated by the acceleration induced by the gravitational pull of a passing PBH. The Shapiro signal, in contrast, refers to shifts in the TOAs caused by metric perturbations along the photon geodesic, produced by PBHs passing along the observer's line of sight.

Afzal et al. (2023) searched NG15 data for deterministic signals generated by models of ULDM and dark matter substructures. They found no significant evidence for either signal, although they set constraints on the parameter space of both models. For a wide range of ULDM models, the constraints compete with (or outperform) laboratory constraints in the 10^{-23} eV $\lesssim m_\phi \lesssim 10^{-20}$ eV mass range. However, the signal from dark matter substructures is harder to detect, so Afzal et al. (2023) could set only very weak constraints on the local abundance of such structures.

EPTA Collaboration et al. (2024a) also searched for ULDM signatures in EPTA data. They obtained a prominent peak in the posterior of the ULDM particle mass around

$\log_{10}(m_\phi/\mathrm{eV}) \approx -23$, which corresponds to a field oscillation frequency of $\log_{10}(f_*/\mathrm{Hz}) \approx -8.3$, consistent with the CGW candidate they also searched for. However, by performing a joint ULDM+GWB search, EPTA Collaboration et al. (2024a) found that the peak in the ULDM mass posterior distribution vanishes, as the EPTA data strongly prefer the presence of a Hellings & Downs (1983) correlation in the common power. Therefore, these authors concluded that an ULDM origin of the PTA signal is disfavored, placing a direct constraint on the abundance of ULDM. In particular, they found that, within our Galaxy, only ~80% of the dark matter density in the Solar Neighborhood can be attributed to ULDM, with $-24 < \log_{10}(m_\phi/\mathrm{eV}) < -23.7$ (for more stringent constraints on ULDM using EPTA data, see Smarra et al. 2023).

5. DISCRIMINATING THE ORIGIN OF THE SIGNAL

Current PTA measurements are broadly consistent with a cosmic population of MBHBs as well as with a variety of cosmological signals, as described in Sections 3 and 4. Countless papers have been devoted to the interpretation of those measurements, sometimes disfavoring the MBHB interpretation because of the $\gamma = 13/3$ prediction. Although the caveats in these analyses are generally well spelled out, we nonetheless warn against overinterpretation of the current PTA results by noting the following points:

- Evidence in the PTA data is only at the 2–4σ level, and uncertainty in the spectral parameters is considerable. For example, Valtolina et al. (2024) showed that the injection of a smooth $\gamma = 13/3$ PL into simulated data might return offset posteriors simply because of the particular realization of MSPs' stochastic noise.
- In general, the signal produced by MBHBs is not smooth. Even in the absence of eccentricity or strong environmental coupling, bright yet unresolved high-frequency CGWs might flatten the measured spectrum to $\gamma < 13/3$ (Valtolina et al. 2024).
- The algorithms employed in the analysis might be affected by the specific prior used (Goncharov et al. 2025) or by leakage (Crisostomi et al. 2025). Estimates of γ might suffer from bias, and its actual value might be closer to 13/3.

Although an astrophysical origin of the observed signal is a natural consequence of hierarchical galaxy formation, currently there is no observational evidence to substantiate this hypothesis. In the future, aside from solidifying the evidence of the detection, the primary goal of PTA experiments will be to pin down the origin of the observed signal. Currently, both MBHBs and various cosmological signals could explain the data.

As the theoretical prediction of the MBHB signal is quite uncertain (given its dependency on modeling assumptions such as eccentricity, environmental feedback, and mass distribution), simultaneously fitting MBHB and cosmological signals would make the inference on the cosmological background parameters strongly dependent on the astrophysical modeling. For this reason, in this review we have considered the analyses on cosmological signals in Section 4 separately from the MBHB analysis in Section 3. We have also discussed the constraints on either astrophysical or cosmological signals separately, always assuming that only one of them is present in the data.

While cosmological GWBs are anticipated to be isotropic, Gaussian, and stationary, characterized by a smooth spectrum, these properties are expected to break down for an astrophysical signal (see Section 3.2.3). Below, we enumerate methods to differentiate between cosmological and astrophysical GWBs and describe the current status of their testing.

5.1. Spectral Shape

As discussed above, the spectral shape alone is likely insufficient to discriminate among astrophysical and cosmological processes because of MBHB eccentricity, environmental coupling, and variance due to sparse sources. What could help instead is the smoothness of the spectrum. While cosmic GWBs are expected to have smooth, continuous spectra, MBHBs produce spiky features, especially at $f > 10^{-8}$ Hz, as a result of loud sparse sources. Searches for deviations from a smooth spectrum in the NANOGrav 15 year data set (Agazie et al. 2025a) have yielded inconclusive excursions at the $\sim 2\sigma$ level.

5.2. Anisotropies and Hot Spots

The supermassive black hole signal is expected to be dominated by a few loud sources (**Figure 3**), implying a Poisson rather than Gaussian power distribution across the sky. Such a distribution is characterized by a flat, nonnull power at all multipoles of the spherical harmonics decomposition (Mingarelli et al. 2013). Moreover, hot spots corresponding to the location of the brightest sources might arise (Taylor et al. 2020). When normalized to the monopole, the expected power in the higher multiples varies by orders of magnitude in the range $10^{-4} \lesssim C_l/C_0 \lesssim 0.3$ for most astrophysical models (Sah et al. 2024, Sato-Polito & Kamionkowski 2024), increasing for higher frequencies. Searches for anisotropies in the NANOGrav 15 year data sets constrained $C_l/C_0 \lesssim 0.2$ in the five lowest-frequency bins (Agazie et al. 2023e), a result that is consistent with most models.

5.3. Cross-Correlation with Large-Scale Structure

MBHBs form in the centers of massive galaxies, found mostly in galaxy clusters. As such, the GWB power in the sky must correlate to some extent with large-scale structure. Although the GWB anisotropy is expected to show a Poissonian noise-like structure because the signal is likely dominated by few massive sources (Allen et al. 2024), correlation with large-scale structure might be observable at high multipoles by an SKA-like array (Semenzato et al. 2024).

5.4. Non-Gaussianity

Another consequence of the discreteness of the MBHB population is the non-Gaussian distribution of the GWB power. In their study of the detectability of non-Gaussianity in PTA data, Falxa & Sesana (2026) found that a Gaussian mixture model is effective at capturing non-Gaussianity, both when a loud source is present in the data and when the GWB is produced by numerous unresolved MBHBs. Searches for non-Gaussianity have not yet been performed on real data.

5.5. Nonstationarity

Another feature that likely arises from a cosmic population of sparse sources is nonstationarity. This is especially true if the population is dominated by a few eccentric binaries, emitting loud GW bursts at periastron over timescales of months to years (see Section 3.3). Falxa et al. (2025) showed that dim eccentric MBHBs with S/N ≈ 1 might produce detectable nonstationarity in simulated PTA data. Moreover, eccentric MBHBs emit GWs across a broad frequency spectrum. If the signal is dominated by a few MBHBs, then the reconstructed GW power emitted across the sky at different frequencies is expected to be correlated (Moreschi et al. 2025, Sah et al. 2025). Moreschi et al. (2025) have shown that an SKA-like array might detect such a GW power cross-correlation at S/N >3.

5.6. Resolvable Continuous Gravitational Waves

Finally, as more MSPs are added to the PTAs, their RMS residual is reduced, and the observation time increases, individual CGWs will eventually emerge from the data (Rosado et al. 2015). Detection of deterministic GWBs that are compatible with an MBHB origin will prove beyond a doubt that there is an astrophysical component of the nanohertz detected GW signal, while identification of EM counterparts will enable multimessenger studies of MBHBs and their galactic hosts, shedding light on the dynamics of the most massive dense nuclei and on the hierarchical buildup of the most massive galaxies in the Universe. Although searches for CGWs in current PTA data have so far proved inconclusive, a recent analysis targeting EM-selected MBHB candidates yielded marginal positive evidence for two systems (Agarwal et al. 2026a).

6. CONCLUSIONS AND OUTLOOK

The nanohertz GW window offers a new, invaluable opportunity to explore the Universe, from its first instants during the inflationary phase to the most massive black holes populating our local cosmological neighborhood. On one hand, the only astrophysical plausible source of the observed PTA signal is a cosmic population of MBHBs, naturally emerging in the framework of hierarchical clustering of cosmic structures. In this case, the signal amplitude, shape, and statistical properties offer important information about the formation of the most massive galaxies and the merger-driven assembly of the most massive black holes in the Universe, adding an important missing piece to the puzzle of cosmic structure formation. On the other hand, the possibility of a signal originating in the early Universe opens a new door to address fundamental physics questions, such as the state of the early Universe, the origin of cosmological perturbations, the nature of dark matter, and whether exotic objects like PBHs or cosmic defects exist at all. While the current signal reconstruction does not carry enough information to characterize GWB properties beyond an approximate estimate of the amplitude and slope of its spectrum, the parameter space of cosmological GWBs that are (in)compatible with PTA data can be clearly identified. This identification, which, remarkably, is independent of the origin of the signal, marks the dawn of early Universe GW cosmology as a research field.

PTA collaborations are relentlessly observing, night after night, amassing data of ever-increasing quality. Eventually, the IPTA, bringing together the EPTA, InPTA, NANOGrav, PPTA, MPTA, and CPTA, and taking advantage of the superior capabilities of the new generation of radio telescopes, such as MeerKAT, FAST, and eventually SKA, will collect enough high-quality data to reveal the still-hidden properties of the signal, including anisotropy, nonstationarity, non-Gaussianity, hot spots, and CGWs. Such data will allow us to pin down the origin of the currently observed nanohertz GW signal, heralding a new era in the exploration of the gravitational Universe.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships or funding that might affect the objectivity of this review.

AUTHOR CONTRIBUTIONS

This review covers nanohertz GW signal generation mechanisms across cosmic time, from the early Universe to today. As such, it requires a range of expertise, to which both authors contributed. The authors' ordering follows the astrophysics community standard, which is not necessarily alphabetical order. A.S. covered the description of astrophysical GW sources (MBHBs; Section 3), and D.G.F. covered the mechanisms generating GWBs in the early Universe (EUBs; Section 4).

ACKNOWLEDGMENTS

A.S. thanks G. Shaifullah, I. Ferranti, and A. Franchini for Figures 1, 3, and 6. Research by A.S. is supported by a European Union Horizon 2020 ERC Advanced Grant, "PINGU" (grant agreement 101142079). D.G.F. thanks Peera Simakachorn for generating customized versions of Figures 7 and 8. Research by D.G.F. was supported by grants EUR2022-134028 and PID2023-148162NB-C22 (from the Spanish Ministry of Science and Education), PROMETEO/2021/083 and ASFAE/2022/020 (from Generalitat Valenciana), and PRTR-C17.I01 (from NextGenerationEU).

LITERATURE CITED

Abac AG, Abouelfettouh I, Acernese F et al. 2025. *Phys. Rev. Lett.* 135:111403
Abbott BP, Abbott R, Abbott TD, et al. 2016. *Phys. Rev. Lett.* 116:061102
Abbott BP, Abbott R, Abbott TD, et al. 2017. *Phys. Rev. Lett.* 119:161101
Abbott R, Abbott TD, Abraham S, et al. 2021. *Phys. Rev. D* 104:022004
Abbott R, Abbott TD, Ancernese F, et al. 2023. *Phys. Rev. X* 13:011048
Acciari VA, Agudo I, Aniello T, et al. 2023. *Astron. Astrophys.* 670:A145
Adshead P, Giblin JT, Pieroni M, Weiner ZJ. 2020. *Phys. Rev. D* 101:083534
Adshead P, Lozanov KD, Weiner ZJ. 2021. *J. Cosmol. Astropart. Phys.* 10:080
Afzal A, Agazie G, Anumarlapudi A, et al. 2023. *Ap. J. Lett.* 951:L11. Erratum. 2024. *Ap. J. Lett.* 971:L27
Agarwal N, Agazie G, Anumarlapudi A, et al. 2026a. *Ap. J. Lett.* 998:L11
Agarwal N, Gabriella A, Amosso A, et al. 2026b. Preprint, arXiv:2606.28554 [astro-ph.CO]
Agazie G, Anumarlapudi A, Archibald AM, et al. 2023a. *Ap. J. Lett.* 951(2):L50
Agazie G, Anumarlapudi A, Archibald AM, et al. 2023b. *Ap. J. Lett.* 952(2):L37
Agazie G, Anumarlapudi A, Archibald AM, et al. 2023c. *Ap. J. Lett.* 951:L8
Agazie G, Alam MF, Anumarlapudi A, et al. 2023d. *Ap. J. Lett.* 951:L9
Agazie G, Anumarlapudi A, Archibald AM, et al. 2023e. *Ap. J. Lett.* 956:L3
Agazie G, Antoniadis J, Anumarlapudi A, et al. 2024. *Ap. J.* 966:105
Agazie G, Anumarlapudi A, Archibald AM, et al. 2025a. *Ap. J.* 978:31
Agazie G, Anumarlapudi A, Archibald AM, et al. 2025b. *Ap. J.* 987:5
Allen B, Agarwal D, Romano JD, Valtolina S. 2024. *Phys. Rev. D* 110:123507
Allen B, Romano JD. 1999. *Phys. Rev. D* 59:102001
Amaro-Seoane P, Audley H, Babak S, et al. 2017. Preprint, arXiv:1702.00786 [astro-ph.IM]
Amaro-Seoane P, Sesana A, Hoffman L, et al. 2010. *MNRAS* 402:2308–20
Amin MA, Shirokoff D. 2010. *Phys. Rev. D* 81:085045
Antoniadis J, Arzoumanian Z, Babak S, et al. 2022. *MNRAS* 510(4):4873–87
Aoki Y, Endrodi G, Fodor Z, et al. 2006. *Nature* 443:675–78
Arzoumanian Z, Baker PT, Blumer H, et al. 2020. *Ap. J. Lett.* 905(2):L34
Arzoumanian Z, Baker PT, Brazier A, et al. 2018. *Ap. J.* 859:47
Arzoumanian Z, Brazier A, Burke-Spolaor S, et al. 2016. *Ap. J.* 821:13
Atwood WB, Abdo AA, Ackermann M, et al. 2009. *Ap. J.* 697(2):1071–102
Auclair P, Caprini C, Cutting D, et al. 2022. *J. Cosmol. Astropart. Phys.* 09:029
Auclair P, Leyde K, Steer DA. 2023. *J. Cosmol. Astropart. Phys.* 04:005
Barack L, Cutler C. 2004. *Phys. Rev. D* 69:082005
Begelman MC, Blandford RD, Rees MJ. 1980. *Nature* 287:307–9
Binney JT. 2008. *Galactic Dynamics*. Princeton University Press. 2nd ed.
Blaes O, Lee MH, Socrates A. 2002. *Ap. J.* 578(2):775–86
Blanco-Pillado JJ, Olum KD. 2017. *Phys. Rev. D* 96:104046
Blanco-Pillado JJ, Olum KD, Shlaer B. 2011. *Phys. Rev. D* 83:083514
Blanco-Pillado JJ, Olum KD, Shlaer B. 2014. *Phys. Rev. D* 89:023512
Blanco-Pillado JJ, Olum KD, Shlaer B. 2015. *Phys. Rev. D* 92:063528

Bonetti M, Haardt F, Sesana A, Barausse E. 2018a. *MNRAS* 477(3):3910–26
Bonetti M, Sesana A, Barausse E, Haardt F. 2018b. *MNRAS* 477(2):2599–612
Book LG, Flanagan EE. 2011. *Phys. Rev. D* 83:024024
Boyle L, Pen UL. 2012. *Phys. Rev. D* 86:124028
Bringmann T, Depta PF, Konstandin T, et al. 2023. *J. Cosmol. Astropart. Phys.* 11:053
Buchmuller W, Domcke V, Schmitz K. 2021. *J. Cosmol. Astropart. Phys.* 12:006
Bullock JS, Boylan-Kolchin M. 2017. *Annu. Rev. Astron. Astrophys.* 55:343–87
Burke-Spolaor S. 2013. *Class. Quant. Gravity* 30:224013
Cai R-G, Pi S, Sasaki M. 2019. *Phys. Rev. Lett.* 122:201101
Caprini C, Durrer R, Konstandin T, Servant G. 2009a. *Phys. Rev. D* 79:083519
Caprini C, Durrer R, Servant G. 2009b. *J. Cosmol. Astropart. Phys.* 12:024
Caprini C, Figueroa DG. 2018. *Class. Quant. Gravity* 35:163001
Carr B, Kohri K, Sendouda Y, Yokoyama J. 2021. *Rep. Prog. Phys.* 84:116902
Carr B, Raidal M, Tenkanen T, et al. 2017. *Phys. Rev. D* 96:023514
Chen S, Caballero RN, Guo YJ, et al. 2021. *MNRAS* 508(4):4970–93
Chen S, Sesana A, Conselice CJ. 2019. *MNRAS* 488:401–18
Chen S, Xu H, Guo Y, et al. 2025. *Astron. Astrophys.* 699:A165
Chen X, Vázquez-Aceves V, Chen S, et al. 2025. *Phys. Rev. Res.* 7:033300
Chen Y, Yu Q, Lu Y. 2020. *Ap. J.* 897:86
Chluba J, Erickcek AL, Ben-Dayan I. 2012. *Ap. J.* 758:76
Christodoulou D. 1991. *Phys. Rev. Lett.* 67:1486–89
Cordes JM. 2013. *Class. Quant. Gravity* 30:224002
Crisostomi M, van Haasteren R, Meyers PM, Vallisneri M. 2025. Preprint, arXiv:2506.13866 [astro-ph.IM]
Cruz NMJ, Malhotra A, Tasinato G, Zavala I. 2025. *Phys. Rev. D* 112:083558
Cui Y, Lewicki M, Morrissey DE, Wells JD. 2018. *Phys. Rev. D* 97:123505
Damour T, Vilenkin A. 2005. *Phys. Rev. D* 71:063510
Davis RL. 1986. *Phys. Lett. B* 180:225–30
de Salas PF, Lattanzi M, Mangano G, et al. 2015. *Phys. Rev. D* 92:123534
Demorest PB, Ferdman RD, Gonzalez ME, et al. 2013. *Ap. J.* 762:94
Dimitriou A, Figueroa DG, Simakachorn P, Zaldivar B. 2025. Preprint, arXiv:2508.05395 [astro-ph.CO]
Domènech G. 2021. *Universe* 7(11):398
Dosopoulou F, Antonini F. 2017. *Ap. J.* 840:31
Dufaux JF, Figueroa DG, Garcia-Bellido J. 2010. *Phys. Rev. D* 82:083518
Duffell PC, Dittmann AJ, D'Orazio DJ, et al. 2024. *Ap. J.* 970(2):156
Duffell PC, D'Orazio D, Derdzinski A, et al. 2020. *Ap. J.* 901:25
Dvorkin I, Barausse E. 2017. *MNRAS* 470(4):4547–56
Ellis J, Fairbairn M, Franciolini G, et al. 2024. *Phys. Rev. D* 109:023522
Enoki M, Inoue KT, Nagashima M, Sugiyama N. 2004. *Ap. J.* 615:19–28
EPTA Collaboration, Antoniadis J, Babak S, et al. 2023a. *Astron. Astrophys.* 678:A48
EPTA Collaboration, InPTA Collaboration, Antoniadis J, et al. 2023b. *Astron. Astrophys.* 678:A50
EPTA Collaboration, InPTA Collaboration, Antoniadis J, et al. 2024a. *Astron. Astrophys.* 685:A94
EPTA Collaboration, InPTA Collaboration, Antoniadis J, et al. 2024b. *Astron. Astrophys.* 690:A118
Espinosa JR, Konstandin T, No JM, Servant G. 2010. *J. Cosmol. Astropart. Phys.* 06:028
Estabrook FB, Wahlquist HD. 1975. *Gen. Relativ. Gravit.* 6(5):439–47
Falxa M, Quelquejay Leclere H, Sesana A. 2025. *Phys. Rev. D* 111:023047
Falxa M, Sesana A. 2026. *Phys. Rev. D* 113:043047
Fastidio F, Gualandris A, Sesana A, et al. 2024. *MNRAS* 532:295–304
Ferranti I, Shaifullah G, Chalumeau A, Sesana A. 2025. *Astron. Astrophys.* 694:A194
Ferreira RZ, Notari A, Pujolas O, Rompineve F. 2023. *J. Cosmol. Astropart. Phys.* 02:001
Ferrer F, Masso E, Panico G, et al. 2019. *Phys. Rev. Lett.* 122:101301
Figueroa DG, Hindmarsh M, Urrestilla J. 2013. *Phys. Rev. Lett.* 110:101302
Figueroa DG, Pieroni M, Ricciardone A, Simakachorn P. 2024. *Phys. Rev. Lett.* 132:171002

Figueroa DG, Raatikainen S, Rasanen S, Tomberg E. 2021. *Phys. Rev. Lett.* 127:101302
Figueroa DG, Tanin EH. 2019. *J. Cosmol. Astropart. Phys.* 08:011
Figueroa DG, Torrenti F. 2017. *J. Cosmol. Astropart. Phys.* 10:057
Foster JW, Blas D, Bourgoin A, et al. 2025. Preprint, arXiv:2504.15334 [astro-ph.CO]
Franchini A, Sesana A, Dotti M. 2021. *MNRAS* 507:1458–67
Franciolini G, Iovino A Jr., Taoso M, Urbano A. 2024. *Phys. Rev. D* 109:123550
Fuskeland U, Aumont J, Aurlien R, et al. 2023. *Astron. Astrophys.* 676:A42
Gaia Collaboration, Prusti T, de Bruijne JHJ, et al. 2016. *Astron. Astrophys.* 595(Gaia Data Release 1):A1
Gardiner EC, Bécsy B, Kelley LZ, Cornish NJ. 2025. *Ap. J.* 988(2):222
Geller M, Ghosh S, Lu S, Tsai Y. 2024. *Phys. Rev. D* 109:063537
Gelmini GB, Gleiser M, Kolb EW. 1989. *Phys. Rev. D* 39:1558–66
Gelmini GB, Hyman J, Simpson A, Vitagliano E. 2023. *J. Cosmol. Astropart. Phys.* 06:055
Ghiglieri J, Laine M. 2015. *J. Cosmol. Astropart. Phys.* 07:022
Goncharov B, Sardana S, Sesana A, et al. 2025. *Nat. Commun.* 16:9692
Gouttenoire Y, Servant G, Simakachorn P. 2020. *J. Cosmol. Astropart. Phys.* 07:032
Gouttenoire Y, Servant G, Simakachorn P. 2021. Preprint, arXiv:2111.01150 [hep-ph]
Green AM, Kavanagh BJ. 2021. *J. Phys. G* 48:043001
Guzzetti MC, Bartolo N, Liguori M, Matarrese S. 2016. *Riv. Nuovo Cim.* 39(9):399–495
Hawking S. 1971. *MNRAS* 152:75–78
Helfer T, Lim EA, Garcia MAG, Amin MA. 2019. *Phys. Rev. D* 99:044046
Hellings RW, Downs GS. 1983. *Ap. J. Lett.* 265:L39–42
Hindmarsh M, Huber SJ, Rummukainen K, Weir DJ. 2014. *Phys. Rev. Lett.* 112:041301
Hindmarsh M, Huber SJ, Rummukainen K, Weir DJ. 2015. *Phys. Rev. D* 92:123009
Hindmarsh MB, Lüben M, Lumma J, Pauly M. 2021. *SciPost Phys. Lect. Notes* 24:1
Hlozek R, Marsh DJE, Grin D. 2018. *MNRAS* 476(3):3063–85
Hu WR, Wu YL. 2017. *Natl. Sci. Rev.* 4(5):685–86
Huber SJ, Konstandin T. 2008. *J. Cosmol. Astropart. Phys.* 09:022
Hui L, McWilliams ST, Yang IS. 2013. *Phys. Rev. D* 87:084009
Hunt P, Sarkar S. 2015. *J. Cosmol. Astropart. Phys.* 12:052
Ivanov PB, Papaloizou JCB, Polnarev AG. 1999. *MNRAS* 307:79–90
Izquierdo-Villalba D, Sesana A, Bonoli S, Colpi M. 2021. *MNRAS* 509(3):3488–503
Jaffe AH, Backer DC. 2003. *Ap. J.* 583:616–31
Jenet FA, Hobbs GB, van Straten W, et al. 2006. *Ap. J.* 653:1571–76
Kamada K, Shin CS. 2020. *J. High Energy Phys.* 04:185
Kaspi VM, Taylor JH, Ryba MF. 1994. *Ap. J.* 428:713–28
Kelley LZ, Blecha L, Hernquist L, et al. 2017. *MNRAS* 471(4):4508–26
Khan FM, Just A, Merritt D. 2011. *Ap. J.* 732:89
Khmelnitsky A, Rubakov V. 2014. *J. Cosmol. Astropart. Phys.* 02:019
Kibble TWB. 1976. *J. Phys. A* 9:1387–98
Kocsis B, Ray A, Portegies Zwart S. 2012. *Ap. J.* 752:67
Kofman L, Linde AD, Starobinsky AA. 1997. *Phys. Rev. D* 56:3258–95
Kohri K, Terada T. 2018. *Phys. Rev. D* 97:123532
Kohri K, Terada T. 2021. *Phys. Lett. B* 813:136040
Kormendy J, Ho LC. 2013. *Ann. Rev. Astron. Astrophys.* 51:511–653
Kulier A, Ostriker JP, Natarajan P, et al. 2015. *Ap. J.* 799(2):178
Kuroyanagi S, Takahashi T, Yokoyama S. 2015. *J. Cosmol. Astropart. Phys.* 02:003
Kuroyanagi S, Takahashi T, Yokoyama S. 2021. *J. Cosmol. Astropart. Phys.* 01:071
Lasky PD, Mingarelli CMF, Smith TL, et al. 2016. *Phys. Rev. X* 6:011035
Lentati L, Taylor SR, Mingarelli CMF, et al. 2015. *MNRAS* 453(3):2576–98
Li B, Meyers J, Shapiro PR. 2025. *Ap. J.* 985:117
Liepold ER, Ma CP. 2024. *Ap. J. Lett.* 971(2):L29
Liu L, Chen ZC, Huang QG. 2024. *Phys. Rev. D* 109:061301

Lizarraga J, Urrestilla J, Daverio D, et al. 2016. *J. Cosmol. Astropart. Phys.* 10:042
Lommen AN. 2002. New limits on gravitational radiation using pulsars. In *270th WE-Heraeus Seminar on Neutron Stars, Pulsars and Supernova Remnants*, ed. W Becker, H Lesch, J Trümper. Max-Planck-Institut für extraterrestrische Physik
Lopez-Eiguren A, Lizarraga J, Hindmarsh M, Urrestilla J. 2017. *J. Cosmol. Astropart. Phys.* 07:026
Luo J, Chen L-S, Duan H-Z, et al. 2016. *Class. Quant. Gravity* 33:035010
Lyubarsky Y. 2019. *MNRAS* 483(2):1731–36
Machado CS, Ratzinger W, Schwaller P, Stefanek BA. 2019. *J. High Energy Phys.* 01:053
Maggiore M. 2007. *Gravitational Waves*, Vol. 1: *Theory and Experiments*. Oxford University Press
Matt C, Gültekin K, Kelley LZ, et al. 2026. *Ap. J.* 997(2):188
McHugh MP, Zalamansky G, Vernotte F, Lantz E. 1996. *Phys. Rev. D* 54:5993–6000
McWilliams ST, Ostriker JP, Pretorius F. 2014. *Ap. J.* 789:156
Mihaylov DP, Moore CJ, Gair JR, et al. 2018. *Phys. Rev. D* 97:124058
Mihos JC, Hernquist L. 1996. *Ap. J.* 464:641–63
Miles MT, Shannon RM, Bailes M, et al. 2023. *MNRAS* 519(3):3976–91
Miles MT, Shannon RM, Reardon DJ, et al. 2025. *MNRAS* 536(2):1489–500
Milosavljevic M, Merritt D. 2003. *Ap. J.* 596:860–78
Mingarelli CMF, Sidery T, Mandel I, Vecchio A. 2013. *Phys. Rev. D* 88:062005
Moreschi BE, Valtolina S, Sesana A, et al. 2025. Preprint, arXiv:2506.14882 [astro-ph.GA]
Neronov A, Roper Pol A, Caprini C, Semikoz D. 2021. *Phys. Rev. D* 103:041302
Neronov A, Semikoz D, Kalashev O. 2023. *Phys. Rev. D* 108:103008
Nixon CJ, Cossins PJ, King AR, Pringle JE. 2011. *MNRAS* 412:1591–98
Pardo K, Chang TC, Doré O, Wang Y. 2023. Preprint, arXiv:2306.14968 [astro-ph.GA]
Perera BBP, DeCesar ME, Demorest PB, et al. 2019. *MNRAS* 490(4):4666–87
Phinney ES. 2001. Preprint, arXiv:astro-ph/0108028
Planck Collaboration, Aghanim N, Akrami Y, et al. 2020a. *Astron. Astrophys.* 641:A6. Erratum. 2021. *Astron. Astrophys.* 652:C4
Planck Collaboration, Akrami Y, Arroja F, et al. 2020b. *Astron. Astrophys.* 641:A9
Planck Collaboration, Akrami Y, Arroja F, et al. 2020c. *Astron. Astrophys.* 641:A10
Porayko NK, Zhu X, Levin Y, et al. 2018. *Phys. Rev. D* 98:102002
Preto M, Berentzen I, Berczik P, Spurzem R. 2011. *Ap. J. Lett.* 732:L26
Qin W, Boddy KK, Kamionkowski M, Dai L. 2019. *Phys. Rev. D* 99:063002
Rajagopal M, Romani RW. 1995. *Ap. J.* 446:543–49
Rana P, Tarafdar P, Nobleson K, et al. 2025. *Publ. Astron. Soc. Aust.* 42:e108
Rasskazov A, Merritt D. 2017. *Phys. Rev. D* 95:084032
Ravi V, Wyithe JSB, Hobbs G, et al. 2012. *Ap. J.* 761:84
Ravi V, Wyithe JSB, Shannon RM, Hobbs G. 2015. *MNRAS* 447(3):2772–83
Reardon DJ, Zic A, Shannon RM, et al. 2023. *Ap. J. Lett.* 951:L6
Roebber E, Holder G, Holz DE, Warren M. 2016. *Ap. J.* 819:163
Roper Pol A, Caprini C, Neronov A, Semikoz D. 2022. *Phys. Rev. D* 105:123502
Roper Pol A, Procacci S, Caprini C. 2024. *Phys. Rev. D* 109:063531
Rosado PA, Sesana A, Gair J. 2015. *MNRAS* 451(3):2417–33
Ryu T, Perna R, Haiman Z, et al. 2018. *MNRAS* 473(3):3410–33
Sah MR, Maurya A, Mukherjee S, et al. 2025. *Ap. J.* 993:118
Sah MR, Mukherjee S, Saeedzadeh V, et al. 2024. *MNRAS* 533(2):1568–82
Sato-Polito G, Kamionkowski M. 2024. *Phys. Rev. D* 109:123544
Sato-Polito G, Zaldarriaga M, Quataert E. 2024. *Phys. Rev. D* 110:063020
Sayeb M, Blecha L, Kelley LZ. 2023. *MNRAS* 527(3):7424–37
Schwaller P. 2015. *Phys. Rev. Lett.* 115:181101
Semenzato F, Casey-Clyde JA, Mingarelli CMF, et al. 2024. Preprint, arXiv:2411.00532 [astro-ph.CO]
Servant G, Simakachorn P. 2023. *Phys. Rev. D* 108:123516
Sesana A. 2010. *Ap. J.* 719:851–64

Sesana A. 2013. *MNRAS* 433:L1–5
Sesana A, Haardt F, Madau P, Volonteri M. 2004. *Ap. J.* 611:623–32
Sesana A, Korsakova N, Arca Sedda M, et al. 2021. *Exp. Astron.* 51(3):1333–83
Sesana A, Shankar F, Bernardi M, Sheth RK. 2016. *MNRAS* 463:L6–11
Sesana A, Vecchio A. 2010. *Class. Quant. Gravity* 27:084016
Sesana A, Vecchio A, Colacino CN. 2008. *MNRAS* 390:192–209
Sesana A, Vecchio A, Volonteri M. 2009. *MNRAS* 394(4):2255–65
Shannon RM, Cordes JM. 2010. *Ap. J.* 725:1607–19
Shannon RM, Ravi V, Coles WA, et al. 2013. *Science* 342(6156):334–37
Shannon RM, Ravi V, Lentati LT, et al. 2015. *Science* 349(6255):1522–25
Sikivie P. 1982. *Phys. Rev. Lett.* 48:1156–59
Simon J. 2023. *Ap. J. Lett.* 949(2):L24
Siwek M, Weinberger R, Hernquist L. 2023. *MNRAS* 522(2):2707–17
Siwek MS, Kelley LZ, Hernquist L. 2020. *MNRAS* 498:537–47
Smarra C, Goncharov B, Barausse E, et al. 2023. *Phys. Rev. Lett.* 131:171001
Stinebring D. 2013. *Class. Quant. Gravity* 30:224006
Taylor SR. 2021. Preprint, arXiv:2105.13270 [astro-ph.HE]
Taylor SR, van Haasteren R, Sesana A. 2020. *Phys. Rev. D* 102:084039
Tiburzi C, Hobbs G, Kerr M, et al. 2016. *MNRAS* 455(4):4339–50
Tristram M, Banday AJ, Górski KM, et al. 2022. *Phys. Rev. D* 105:083524
Truant RJ, Izquierdo-Villalba D, Sesana A, et al. 2025. *Astron. Astrophys.* 694:A282
Unal C. 2019. *Phys. Rev. D* 99:041301
Unal C, Papageorgiou A, Obata I. 2024. *Phys. Lett. B* 856:138873
Vaglio M, Falxa M, Mentasti G, et al. 2025. Preprint, arXiv:2507.18593 [gr-qc]
Vagnozzi S. 2023. *J. High Energy Astrophys.* 39:81–98
Valtolina S, Clark CJ, van Haasteren R, et al. 2026. *Phys. Rev. D* 113(6):063061
Valtolina S, Shaifullah G, Samajdar A, Sesana A. 2024. *Astron. Astrophys.* 683:A201
van Haasteren R, Levin Y. 2010. *MNRAS* 401(4):2372–78
van Haasteren R, Levin Y, Janssen GH, et al. 2011. *MNRAS* 414(4):3117–28. Erratum. 2012. *MNRAS* 425:1597
Verbiest JPW, Lentati L, Hobbs G, et al. 2016. *MNRAS* 458(2):1267–88
Vilenkin A. 1981. *Phys. Lett. B* 107:47–50
Vilenkin A. 1985. *Phys. Rep.* 121:263–315
Vilenkin A, Shellard EPS. 2000. *Cosmic Strings and Other Topological Defects*. Cambridge University Press
Wang S, Zhao ZC, Li JP, Zhu QH. 2024. *Phys. Rev. Res.* 6:L012060
Wen L. 2003. *Ap. J.* 598:419–30
Wolz K, Azzoni S, Hervías-Caimapo C, et al. 2024. *Astron. Astrophys.* 686:A16
Wyithe JSB, Loeb A. 2003. *Ap. J.* 590:691–706
Xu H, Chen S, Guo Y, et al. 2023. *Res. Astron. Astrophys.* 23:075024
Young S. 2022. *J. Cosmol. Astropart. Phys.* 05:037
Young S, Musco I, Byrnes CT. 2019. *J. Cosmol. Astropart. Phys.* 11:012
Yunes N, Siemens X. 2013. *Living Rev. Relativ.* 16:9
Zhang J, Lyu Z, Huang J, et al. 2021. *Phys. Rev. Lett.* 127:161101
Zhou Y, Di Matteo T, Chen N, et al. 2025. *Ap. J. Lett.* 988(2):L74
Zhu X-J, Cui W, Thrane E. 2019. *MNRAS* 482(2):2588–96
Zic A, Reardon DJ, Kapur A, et al. 2023. *Publ. Astron. Soc. Aust.* 40:e049