



The NANOGrav 15 yr Data Set: Looking for Signs of Discreteness in the Gravitational-wave Background

Gabriella Agazie¹ , Akash Anumalapudi¹ , Anne M. Archibald² , Zaven Arzoumanian³ , Jeremy George Baier⁴ , Paul T. Baker⁵ , Bence Bécsey⁴ , Laura Blecha⁶ , Adam Brazier^{7,8} , Paul R. Brook⁹ , Lucas Brown¹⁰ , Sarah Burke-Spolaor^{11,12,69} , J. Andrew Casey-Clyde¹³ , Maria Charisi¹⁴ , Shami Chatterjee⁷ , Tyler Cohen¹⁵ , James M. Cordes⁷ , Neil J. Cornish¹⁶ , Fronefield Crawford¹⁷ , H. Thankful Cromartie¹⁸ , Kathryn Crowter¹⁹ , Megan E. DeCesar²⁰ , Paul B. Demorest²¹ , Heling Deng⁴ , Timothy Dolch^{22,23} , Elizabeth C. Ferrara^{24,25,26} , William Fiore^{11,12} , Emmanuel Fonseca^{11,12} , Gabriel E. Freedman¹ , Nate Garver-Daniels^{11,12} , Peter A. Gentile^{11,12} , Joseph Glaser^{11,12} , Deborah C. Good²⁷ , Kayhan Gültekin²⁸ , Jeffrey S. Hazboun⁴ , Ross J. Jennings^{11,12,70} , Aaron D. Johnson^{1,29} , Megan L. Jones¹ , Andrew R. Kaiser^{11,12} , David L. Kaplan¹ , Luke Zoltan Kelley³⁰ , Matthew Kerr³¹ , Joey S. Key³² , Nima Laal⁴ , Michael T. Lam^{33,34,35} , William G. Lamb¹⁴ , Bjorn Larsen³⁶ , T. Joseph W. Lazio³⁷ , Natalia Lewandowska³⁸ , Tingting Liu^{11,12} , Duncan R. Lorimer^{11,12} , Jing Luo^{39,71} , Ryan S. Lynch⁴⁰ , Chung-Pei Ma^{30,41} , Dustin R. Madison⁴² , Alexander McEwen¹ , James W. McKee^{43,44} , Maura A. McLaughlin^{11,12} , Natasha McMann¹⁴ , Bradley W. Meyers^{19,45} , Patrick M. Meyers²⁹ , Chiara M. F. Mingarelli³⁶ , Andrea Mitridate⁴⁶ , Priyamvada Natarajan^{47,48} , Cherry Ng⁴⁹ , David J. Nice⁵⁰ , Stella Koch Ocker^{29,51} , Ken D. Olum¹⁰ , Timothy T. Pennucci⁵² , Benetge B. P. Perera⁵³ , Nihan S. Pol⁵⁴ , Henri A. Radovan⁵⁵ , Scott M. Ransom⁵⁶ , Paul S. Ray³¹ , Joseph D. Romano⁵⁴ , Jessie C. Runnøe¹⁴ , Shashwat C. Sardesai¹ , Ann Schmiedekamp⁵⁷ , Carl Schmiedekamp⁵⁷ , Kai Schmitz⁵⁸ , Brent J. Shapiro-Albert^{11,12,59} , Xavier Siemens^{1,4} , Joseph Simon^{60,72} , Magdalena S. Siwek⁶¹ , Sophia V. Sosa Fiscella^{34,35} , Ingrid H. Stairs¹⁹ , Daniel R. Stinebring⁶² , Kevin Stovall²¹ , Abhimanyu Susobhanan⁶³ , Joseph K. Swiggum^{50,70} , Stephen R. Taylor¹⁴ , Jacob E. Turner⁴⁰ , Caner Unal^{64,65,66} , Michele Vallisneri^{29,37} , Sarah J. Vigeland¹ , Haley M. Wahl^{11,12} , London Willson¹³ , Caitlin A. Witt^{67,68} , David Wright⁴ , and Olivia Young^{34,35}

¹ Center for Gravitation, Cosmology and Astrophysics, Department of Physics, University of Wisconsin-milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA

² Newcastle University, NE1 7RU, UK

³ X-Ray Astrophysics Laboratory, NASA Goddard Space Flight Center, Code 662, Greenbelt, MD 20771, USA

⁴ Department of Physics, Oregon State University, Corvallis, OR 97331, USA

⁵ Department of Physics and Astronomy, Widener University, One University Place, Chester, PA 19013, USA

⁶ Physics Department, University of Florida, Gainesville, FL 32611, USA

⁷ Cornell Center for Astrophysics and Planetary Science and Department of Astronomy, Cornell University, Ithaca, NY 14853, USA

⁸ Cornell Center for Advanced Computing, Cornell University, Ithaca, NY 14853, USA

⁹ Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

¹⁰ Institute of Cosmology, Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA

¹¹ Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA

¹² Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, USA

¹³ Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA; andrew.casey-clyde@uconn.edu

¹⁴ Department of Physics and Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA

¹⁵ Department of Physics, New Mexico Institute of Mining and Technology, 801 Leroy Place, Socorro, NM 87801, USA

¹⁶ Department of Physics, Montana State University, Bozeman, MT 59717, USA

¹⁷ Department of Physics and Astronomy, Franklin & Marshall College, P.O. Box 3003, Lancaster, PA 17604, USA

¹⁸ National Research Council Research Associate, National Academy of Sciences, Washington, DC 20001, USA resident at Naval Research Laboratory, Washington, DC 20375, USA

¹⁹ Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada

²⁰ George Mason University, Fairfax, VA 22030, resident at the U.S. Naval Research Laboratory, Washington, DC 20375, USA

²¹ National Radio Astronomy Observatory, 1003 Lopezville Road, Socorro, NM 87801, USA

²² Department of Physics, Hillsdale College, 33 E. College Street, Hillsdale, MI 49242, USA

²³ Eureka Scientific, 2452 Delmer Street, Suite 100, Oakland, CA 94602-3017, USA

²⁴ Department of Astronomy, University of Maryland, College Park, MD 20742, USA

²⁵ Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, MD 20771, USA

²⁶ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

²⁷ Department of Physics and Astronomy, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA

²⁸ Department of Astronomy and Astrophysics, University of Michigan, Ann Arbor, MI 48109, USA

²⁹ Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

³⁰ Department of Astronomy, University of California, Berkeley, 501 Campbell Hall #3411, Berkeley, CA 94720, USA

³¹ Space Science Division, Naval Research Laboratory, Washington, DC 20375-5352, USA

³² University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011, USA

³³ SETI Institute, 339 N Bernardo Avenue Suite 200, Mountain View, CA 94043, USA

³⁴ School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA

³⁵ Laboratory for Multiwavelength Astrophysics, Rochester Institute of Technology, Rochester, NY 14623, USA

³⁶ Department of Physics, Yale University, New Haven, CT 06520, USA

³⁷ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

³⁸ Department of Physics and Astronomy, State University of New York at Oswego, Oswego, NY 13126, USA

³⁹ Department of Astronomy & Astrophysics, University of Toronto, 50 Saint George Street, Toronto, ON M5S 3H4, Canada

⁴⁰ Green Bank Observatory, P.O. Box 2, Green Bank, WV 24944, USA

⁴¹ Department of Physics, University of California, Berkeley, CA 94720, USA

⁴² Department of Physics, University of the Pacific, 3601 Pacific Avenue, Stockton, CA 95211, USA

- ⁴³ E.A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Kingston-upon-Hull, HU6 7RX, UK
- ⁴⁴ Centre of Excellence for Data Science, Artificial Intelligence and Modeling (DAIM), University of Hull, Cottingham Road, Kingston-upon-Hull, HU6 7RX, UK
- ⁴⁵ International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia
- ⁴⁶ Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany
- ⁴⁷ Department of Astronomy, Yale University, 52 Hillhouse Avenue, New Haven, CT 06511, USA
- ⁴⁸ Black Hole Initiative, Harvard University, 20 Garden Street, Cambridge, MA 02138, USA
- ⁴⁹ Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
- ⁵⁰ Department of Physics, Lafayette College, Easton, PA 18042, USA
- ⁵¹ The Observatories of the Carnegie Institution for Science, Pasadena, CA 91101, USA
- ⁵² Institute of Physics and Astronomy, Eötvös Loránd University, Pázmány P. s. 1/A, 1117 Budapest, Hungary
- ⁵³ Arecibo Observatory, HC3 Box 53995, Arecibo, PR 00612, USA
- ⁵⁴ Department of Physics, Texas Tech University, Box 41051, Lubbock, TX 79409, USA
- ⁵⁵ Department of Physics, University of Puerto Rico, Mayagüez, PR 00681, USA
- ⁵⁶ National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
- ⁵⁷ Department of Physics, Penn State Abington, Abington, PA 19001, USA
- ⁵⁸ Institute for Theoretical Physics, University of Münster, 48149 Münster, Germany
- ⁵⁹ Giant Army, 915A 17th Ave, Seattle WA 98122, USA
- ⁶⁰ Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA
- ⁶¹ Center for Astrophysics, Harvard University, 60 Garden Street, Cambridge, MA 02138, USA
- ⁶² Department of Physics and Astronomy, Oberlin College, Oberlin, OH 44074, USA
- ⁶³ Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Callinstrasse 38, D-30167, Hannover, Germany
- ⁶⁴ Department of Physics, Middle East Technical University, 06531 Ankara, Turkey
- ⁶⁵ Department of Physics, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel
- ⁶⁶ Feza Gursey Institute, Bogazici University, Kandilli, 34684, Istanbul, Turkey
- ⁶⁷ Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, Evanston, IL 60208, USA
- ⁶⁸ Adler Planetarium, 1300 S. DuSable Lake Shore Drive, Chicago, IL 60605, USA

Received 2024 April 10; revised 2024 September 27; accepted 2024 November 4; published 2024 December 23

Abstract

The cosmic merger history of supermassive black hole binaries (SMBHBs) is expected to produce a low-frequency gravitational wave background (GWB). Here we investigate how signs of the discrete nature of this GWB can manifest in pulsar timing arrays (PTAs) through excursions from, and breaks in, the expected $f_{\text{GW}}^{-2/3}$ power law of the GWB strain spectrum. To do this, we create a semianalytic SMBHB population model, fit to North American Nanohertz Observatory for Gravitational Waves (NANOGrav's) 15 yr GWB amplitude, and with 1000 realizations, we study the populations' characteristic strain and residual spectra. Comparing our models to the NANOGrav 15 yr spectrum, we find two interesting excursions from the power law. The first, at 2 nHz, is below our GWB realizations with a p -value significance $p = 0.05\text{--}0.06$ ($\approx 1.8\sigma\text{--}1.9\sigma$). The second, at 16 nHz, is above our GWB realizations with $p = 0.04\text{--}0.15$ ($\approx 1.4\sigma\text{--}2.1\sigma$). We explore the properties of a loud SMBHB that could cause such an excursion. Our simulations also show that the expected number of SMBHBs decreases by 3 orders of magnitude, from $\sim 10^6$ to $\sim 10^3$, between 2 and 20 nHz. This causes a break in the strain spectrum as the stochasticity of the background breaks down at 26^{+28}_{-19} nHz, consistent with predictions pre-dating GWB measurements. The diminished GWB signal from SMBHBs at frequencies above the 26 nHz break opens a window for PTAs to detect continuous GWs from individual SMBHBs or GWs from the early Universe.

Unified Astronomy Thesaurus concepts: Gravitational wave astronomy (675); Gravitational waves (678); Quasars (1319); Supermassive black holes (1663)

1. Introduction

Massive galaxies should host supermassive black holes (SMBHs) at their centers (J. Kormendy & D. Richstone 1995). As galaxies merge, their central SMBHs are expected to eventually gravitationally bind, forming SMBH binaries (SMBHBs; M. C. Begelman et al. 1980). As these binaries inspiral, they radiate away energy via low-frequency gravitational waves (GWs). The incoherent superposition of GWs from SMBHBs forms a GW background (GWB; e.g., M. Rajagopal & R. W. Romani 1995).

GWs from these sources, including the GWB itself, are the primary detection target of pulsar timing array (PTA) experiments (R. S. Foster & D. C. Backer 1990). Current PTA experiments, including the North American Nanohertz Observatory for Gravitational Waves (NANOGrav; G. Agazie et al. 2023c), the European PTA (EPTA) in collaboration with the Indian PTA (InPTA; EPTA Collaboration et al. 2023), the Parkes PTA (PPTA; D. J. Reardon et al. 2023), and the Chinese PTA (H. Xu et al. 2023), have now reported evidence for a GWB in their pulsar data, opening a new window to study the SMBHB population. New data from MeerKAT PTA (MPTA, M. T. Miles et al. 2023) are expected to be important for GWB characterization and continuous wave (CW) searches.

Due to the slowly evolving nature of SMBHBs, we expect hundreds of thousands to millions of them to be emitting in the nanoHertz regime (B. Bécsy et al. 2022). The majority of these are expected to be unresolvable; however, we can probe their ensemble properties, such as mass and redshift distribution, via the GWB's amplitude (J. A. Casey-Clyde et al. 2022; G. Agazie et al. 2023b; EPTA Collaboration et al. 2024).

⁶⁹ Sloan Fellow.

⁷⁰ NANOGrav Physics Frontiers Center Postdoctoral Fellow.

⁷¹ Deceased.

⁷² NSF Astronomy and Astrophysics Postdoctoral Fellow.



Moreover, there are two ways in which discrete aspects of the GWB may manifest: first, a few nearby and/or very massive SMBHBs could manifest as excursions in the GWB strain spectrum before being individually resolvable with CW searches. Second, due to the discrete nature of the GWB (if indeed sourced by SMBHBs), there should be a frequency in its strain spectrum where the stochasticity of the GWB breaks down (A. Sesana et al. 2008; V. Ravi et al. 2012; C. M. F. Mingarelli et al. 2013; E. Roebber et al. 2016; L. Z. Kelley et al. 2017). Here we model and search for both excursions in the strain and residual spectra, and a break in the GWB’s power-law behavior at high frequencies.

We use a state-of-the-art, semianalytic SMBHB model to calculate the expected properties of the GWB to search for these signs of discreteness. Searching for a knee, we fit the expected GWB strain spectrum with a double power-law model. We then compare the expected GWB signal to the actual timing residual spectrum observed by NANOGrav (G. Agazie et al. 2023c), searching for excursions from the power-law behavior in the measured GWB spectra.

This paper is organized as follows: In Section 2 we briefly review both analytic and discrete methods for modeling the GWB. In Section 3 we present the analytic models that we fit to our discretely sampled GWB spectra. In Section 4 we present the results of these fits, including the frequency where the spectrum diverges from the expected power-law slope. We also compare our discretely sampled GWB spectra to the spectrum observed by G. Agazie et al. (2023c) to search for evidence of discreteness in the data. In Section 5 we summarize our results and discuss their implications for both future GWB and CW searches.

Throughout this work, we use units where $G = c = 1$.

2. GWB Population Model

We adopt an SMBHB population model derived from the major merger rate of galaxies (S. Chen et al. 2019). Briefly, in this model the galaxy major merger rate assumes a galaxy stellar mass function, an observed galaxy pairing fraction, and a theoretical galaxy merger timescale. We then compute the SMBHB merger rate from the galaxy merger rate by assuming empirical scaling relations between galaxy stellar mass, bulge mass, and SMBH mass (S. Chen et al. 2019). The SMBHB merger rate, $\dot{\phi}_{\text{BHB}} = d^3\Phi_{\text{BHB}}/(dM dq dt_r)$, describes the differential comoving number density of SMBHB mergers, Φ_{BHB} , per unit proper time, t_r , binary total mass, $M = M_1 + M_2$, and mass ratio, $q = M_2/M_1 \leq 1$, where M_1 and M_2 are the masses of the primary and secondary SMBHBs in the binary, respectively.

The SMBHB merger rate is related to the characteristic strain, $h_c(f_{\text{GW}})$, of the GWB by E. S. Phinney (2001)

$$h_c^2(f_{\text{GW}}) = \frac{4}{3\pi} \frac{1}{f_{\text{GW}}^{4/3}} \iiint \dot{\phi}_{\text{BHB}} \frac{dt_r}{dz} \frac{\mathcal{M}^{5/3}}{(1+z)^{1/3}} dM dq dz, \quad (1)$$

where $\mathcal{M}^{5/3} = M^{5/3}q/(1+q)^2$ is the chirp mass, f_{GW} is the observed GW frequency, z is the redshift, and dt_r/dz is given by cosmology (D. W. Hogg 1999). Equation (1) is frequently written as a power law, $h_c(f_{\text{GW}}) = A_{\text{yr}}(f_{\text{GW}}/f_{\text{yr}})^\alpha$, where A_{yr} is the GWB amplitude at a reference frequency $f_{\text{yr}} = 1 \text{ yr}^{-1}$. For circular SMBHBs and GW-driven evolution, we expect $\alpha = -2/3$ (E. S. Phinney 2001). While Equation (1) can be used to model the GWB analytically, it does not account for the fact that the GWB arising from SMBHBs is composed of

discrete sources. This is important at high frequencies, where we expect the stochastic, power-law behavior of the real GWB to break down as finite number effects cause the GWB spectrum to decrease more steeply than Equation (1) predicts.

To model these effects, we generate 1000 realizations of discretized SMBHB populations (A. Sesana et al. 2008). We can then compute realistic GWB spectra built up from the individual SMBHB strains in each population realization. Specifically, each population realization is generated by sampling SMBHBs from the differential number of SMBHBs per unit M , q , z , and f_{GW} (S. Chen et al. 2019),

$$\mathcal{N}_{\text{BHB}} \equiv \frac{d^4 N_{\text{BHB}}}{dM dq dz df_{\text{GW}}} = \dot{\phi}_{\text{BHB}} \frac{dV_c}{dz} \frac{dt_r}{df_r} \frac{df_r}{df_{\text{GW}}}, \quad (2)$$

where dV_c/dz is the comoving volume per unit redshift (D. W. Hogg 1999), $f_r = f_{\text{GW}}(1+z)$ is the rest-frame GW frequency, and $dt_r/df_r = (5/96)\pi^{-8/3}\mathcal{M}^{-5/3}f_r^{-11/3}$ is the differential residence timescale per rest-frame frequency (P. C. Peters & J. Mathews 1963). We use a model of $\dot{\phi}_{\text{BHB}}$ derived from galaxy major merger rates, which has been fit via Equation (1) to the GWB amplitude measured in NANOGrav’s 15 yr data set, $A_{\text{yr}} = 2.4_{-0.6}^{+0.7} \times 10^{-15}$, using a Markov Chain Monte Carlo (MCMC; G. Agazie et al. 2023c; J. A. Casey-Clyde et al. 2024). The resulting GWB realizations also reflect uncertainty in A_{yr} . The SMBH mass function implicit in our model is consistent with the local SMBH mass function (e.g., A. Marconi et al. 2004; F. Shankar et al. 2009; M. Vika et al. 2009; see G. Sato-Polito et al. 2023). In the right panel of Figure 1, we show the median number of SMBHBs generated at each frequency over all 1000 realizations.

We compute the sky- and polarization-averaged strain, h , of each sampled SMBHB as (K. S. Thorne 1987; A. Sesana et al. 2008; P. A. Rosado et al. 2015)

$$h = \frac{8}{\sqrt{10}} \frac{\mathcal{M}_0^{5/3}(\pi f_{\text{GW}})^{2/3}}{D_L(z)}, \quad (3)$$

where $D_L(z)$ is the luminosity distance to a binary at z , given by standard cosmology, and where $\mathcal{M}_0 = \mathcal{M}(1+z)$ is the observer frame chirp mass. In each realization, we then compute $h_c(f_{\text{GW}})$ at each frequency via $h_c^2(f_{\text{GW}}) = \sum_k h_k^2 f_k / \Delta f_{\text{GW}}$, where $\Delta f_{\text{GW}} = 1/T_{\text{obs}}$ is the frequency sampling interval, set by the total PTA observation time, T_{obs} .

The spectral energy density (SED) of the GWB is given by L. Z. Kelley et al. (2017):

$$S_h(f_{\text{GW}}) = \frac{h_c^2(f_{\text{GW}})}{12\pi^2 f_{\text{GW}}^3}. \quad (4)$$

We can easily move between the characteristic strain and the residuals induced by the GWB via the amplitude spectral density (ASD), $\sqrt{S_h(f)}/T_{\text{obs}}$. The frequency dependence of the ASD is therefore $\text{ASD} \propto f_{\text{GW}}^{-13/6}$, compared to the scaling of $h_c \propto f_{\text{GW}}^{-2/3}$. Excursions from an $f_{\text{GW}}^{-2/3}$ power law in characteristic strain space remain excursions from an $f_{\text{GW}}^{-13/6}$ in ASD space.

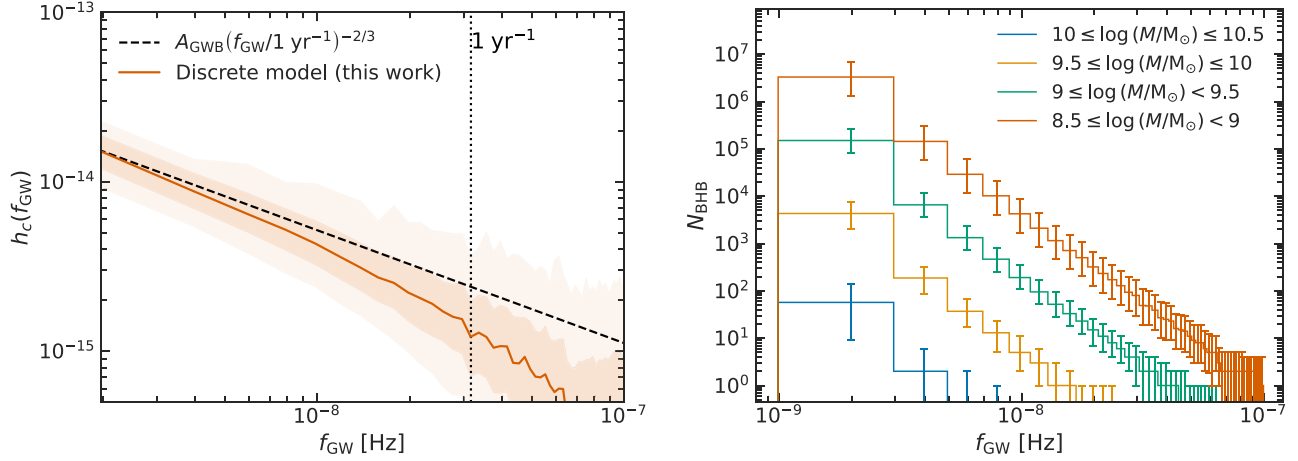


Figure 1. Characteristic strain spectra of 1000 SMBHB populations, each generating a GWB. Left panel: the expected $h_c(f_{\text{GW}})$ induced by a GWB with $A_{\text{yr}} = 2.4^{+0.7}_{-0.6} \times 10^{-15}$ at $f = 1 \text{ yr}^{-1}$ (red). The solid line shows the median GWB, while the inner and outer shaded regions denote the 68% and 95% confidence intervals, respectively. These confidence intervals reflect uncertainty from NANOGrav’s GWB measurement and from differences between individual GWB realizations. At high frequencies, the discretely generated GWB diverges from the power-law behavior of the analytic spectrum (black dashed line), which does not account for the discrete nature of SMBHBs. Right panel: the number of SMBHBs generated at each frequency. Histograms show the median number of binaries generated in each frequency interval, with colors denoting mass range. Error bars denote 68% confidence intervals. Colored lines show the number of binaries at each frequency broken down by mass. It is clear that the vast majority of binaries at any f_{GW} have low masses.

3. The High-frequency Knee

A. Sesana et al. (2008) pointed out a high-frequency GWB “knee” feature, where the paucity of GW sources reduces the amplitude of the strain spectrum. Using a range of SMBHB merger rates derived from different dark matter halo merger tree prescriptions, they found this knee to occur at $37^{+15}_{-13} \text{ nHz}$. Here, we instead assume an SMBHB merger rate that has been fit to the GWB amplitude measured in NANOGrav’s 15 yr data set to compute the location of the knee. We interpret this knee frequency, f_{knee} , as the frequency at which the stochasticity of the GWB breaks down on average, leading to deviations from an $f_{\text{GW}}^{-2/3}$ power law.

We consider a double power-law model of the average characteristic strain spectrum, which we implement here for the first time. We parameterize the characteristic strain as,

$$h_c^{\text{dpl}}(f_{\text{GW}}) = \frac{2A_{\text{knee}}}{\left(\frac{f_{\text{GW}}}{f_{\text{knee}}}\right)^{-\alpha_1} + \left(\frac{f_{\text{GW}}}{f_{\text{knee}}}\right)^{-\alpha_2}}, \quad (5)$$

where $\alpha_1 = -2/3$ is the low-frequency slope of a GWB arising from circular SMBHBs undergoing GW-driven evolution, α_2 is the high-frequency slope, and A_{knee} is the amplitude of the background at f_{knee} . We take $A_{\text{knee}} = \lim_{f_{\text{GW}} \ll f_{\text{knee}}} (A_{\text{yr}}/2)(f_{\text{knee}}/f_{\text{yr}})^{\alpha_1} [1 + (f_{\text{GW}}/f_{\text{knee}})^{\alpha_1 - \alpha_2}]$ such that Equation (5) reduces to the standard power-law model at $f_{\text{GW}} \ll f_{\text{knee}}$.

We additionally consider a physically motivated model of $h_c(f_{\text{GW}})$, which accounts for the fact that contributions to the GWB at each f_{GW} must come from an integer number of SMBHBs. Specifically, A. Sesana et al. (2008) posited that in a given frequency interval, there is a characteristic mass, $\tilde{M}(f_{\text{GW}})$, such that less than one SMBHB with $M > \tilde{M}$ contributes to the GWB on average. For example, in the right panel of Figure 1, the median number of SMBHBs with $M > 10^{10} M_{\odot}$ in each frequency interval above $\sim 6 \text{ nHz}$ is < 1 . In a typical realization of the Universe, we therefore do not expect to have SMBHBs more massive than $10^{10} M_{\odot}$ emitting GWs at $f_{\text{GW}} \gtrsim 6 \text{ nHz}$. By assuming that contributions to

$h_c(f_{\text{GW}})$ come from binaries with $M < \tilde{M}(f_{\text{GW}})$ and that the mass dependence of $\dot{\phi}_{\text{BHB}}$ can be approximated as a power law, the characteristic strain of the GWB can be modeled as

$$h_c^{\text{phys}}(f_{\text{GW}}) = A_{\text{yr}} \left(\frac{f_{\text{GW}}}{f_{\text{yr}}} \right)^{\alpha_1} \times \left[1 + \left(\frac{f_{\text{GW}}}{f_{\text{knee}}} \right)^{11/3} \right]^{3(\alpha_2 - \alpha_1)/11}, \quad (6)$$

where $\alpha_1 = -2/3$ is the slope at $f_{\text{GW}} \ll f_{\text{knee}}$, and where α_2 is the slope of the average strain spectrum at $f_{\text{GW}} \gg f_{\text{knee}}$ (see Appendix B and A. Sesana et al. 2008 for details).

We constrain the double power-law and physical models via MCMC sampling. We use Gaussian kernel density estimators (KDEs) to estimate the probability density function (PDF) of the $h_c(f_k)$ realizations at $k = 1, \dots, 30$ frequencies, $f_k \approx k/15 \text{ yr}^{-1}$. These frequencies were chosen to match the GWB spectrum frequencies from NANOGrav’s 15 yr data (G. Agazie et al. 2023c). The KDE at f_k estimates the PDF as a sum of Gaussian kernels centered on each of the $h_{c,i}(f_k)$ realizations (M. Rosenblatt 1956; E. Parzen 1962):

$$\text{PDF}[h_c(f_k)] = \frac{1}{n\sigma_k\sqrt{2\pi}} \times \sum_{i=1}^n \exp \left\{ -\frac{1}{2} \frac{[h_c(f_k) - h_{c,i}(f_k)]^2}{\sigma_k^2} \right\}, \quad (7)$$

where $n = 1000$ is the total number of GWB realizations, and σ_k is the kernel bandwidth. We choose σ_k at each frequency using cross-validated grid searches, which maximize the likelihood each KDE could have generated the $h_{c,i}(f_k)$ realizations (F. Pedregosa et al. 2011; A. Géron 2017). The KDEs are then used to calculate the MCMC posterior log-likelihood during fitting.

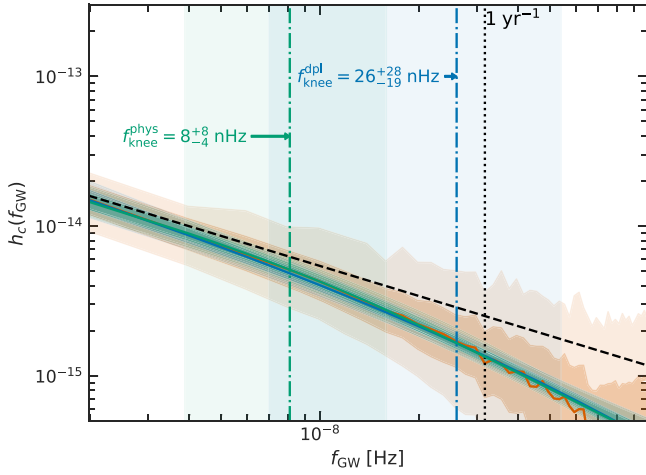


Figure 2. Discreteness creates a knee in the GWB strain spectrum. The $h_c(f_{\text{GW}})$ spectrum from a population of discrete SMBHBs is shown in red, while the double power-law and physical model fits are shown in blue and green, respectively. The solid line shows the median value of the expected $h_c(f_{\text{GW}})$, while inner and outer shaded regions show the 1σ and 2σ confidence intervals, as in Figure 1. The dotted vertical lines show the location of the “knee” frequency for each fitted model, while the vertical shaded regions show the 68% confidence intervals.

4. Results

Here we present the results of our GWB spectrum analyses. We start by estimating the knee frequency of the characteristic strain spectrum in our simulated GWB spectra. We then assess how significant single-frequency excursions from a power law in NANOGrav’s 15 yr residual spectrum are compared to the spectra we compute from discrete SMBHB populations. We next compare NANOGrav’s 15 yr residual spectrum to NANOGrav’s 12.5 yr residual spectrum to assess whether or not excursions in the 15 yr spectrum were present in previous data sets. Finally, we consider two hypotheses to explain an observed excursion in the 15 yr spectrum around 16 nHz: we first determine whether this excursion could be due to excess white noise in the 15 yr residual spectrum. We then consider a loud binary hypothesis, wherein the 16 nHz excursion is due to a loud SMBHB that sits below the CW detection threshold.

4.1. The Knee

With 1000 realizations of the cosmic population of SMBHBs, we find that the average strain spectrum clearly follows an $f_{\text{GW}}^{-2/3}$ power law at low frequencies (Figure 1), as expected. However, as we move to higher GW frequencies, we clearly see the stochasticity of the GWB breaking down, as discrete sources start to become important. Looking at the right panel of Figure 1, we can see that this breakdown of stochasticity is due to a dramatic decrease in the number of SMBHBs at higher frequencies. For example, at 2 nHz, there are $\mathcal{O}(10^6)$ SMBHBs contributing to the background, while at 20 nHz, there are fewer than $\mathcal{O}(10^3)$ SMBHBs. The decreasing number of SMBHBs at higher f_{GW} results in there being too few SMBHBs to maintain an average $f_{\text{GW}}^{-2/3}$ power-law scaling at these higher frequencies. This can be seen in the decreased median value of our discrete model of the GWB, as it falls below the $f_{\text{GW}}^{-2/3}$ power law.

In Figure 2 we present the results of fitting the double power-law model to the generated characteristic strain spectra. We can see that there is a bend in the strain spectrum at

$f_{\text{knee}} = 26^{+28}_{-19}$ nHz. At this point, the GWB deviates from the $f_{\text{GW}}^{-2/3}$ power law to $f_{\text{GW}}^{-1.3^{+0.2}_{-0.3}}$: approximately twice as steep as the $f_{\text{GW}}^{-2/3}$ power-law behavior expected at lower frequencies.

To determine if a double power law is preferred by the data over a single power law, we calculate a Savage–Dickey Bayes factor (J. M. Dickey 1971). This can be done because the double power-law model reduces to a single power law when $\alpha_1 = \alpha_2$ (Equation (5)). The Bayes factor is thus computed as the ratio of the prior and posterior probability densities of α_2 at $\alpha_2 = -2/3$. We find that a double power-law model is preferred to a single power-law with a Bayes factor of 270.

We similarly fit a broken power-law model with $h_c(f_{\text{GW}}) = A_{\text{yr}}(f_{\text{GW}}/f_{\text{yr}})^{\alpha_1}(1 + f_{\text{GW}}/f_{\text{knee}})^{\alpha_2 - \alpha_1}$, as in A. Sesana et al. (2008), to our 1000 generated SMBHB spectra. The resulting fit is nearly identical to the double power-law model, with a bend at $f_{\text{knee}} = 25^{+42}_{-19}$ nHz and a high-frequency power-law slope of $f_{\text{GW}}^{-1.4^{+0.3}_{-0.5}}$. The Bayes factor for a broken power law over a single power law is 171, indicating a preference for the broken power-law model. See Appendix A for details on the broken power-law model and its comparison to the double power-law model.

We also fit the physical model in Equation (6) to our generated spectra, finding $f_{\text{knee}} = 8^{+8}_{-4}$ nHz and $h_c \propto f_{\text{GW}}^{-1.1^{+0.1}_{-0.2}}$ at $f_{\text{GW}} \gg f_{\text{knee}}$. This is a lower knee than found using the double or broken power-law models, but is still within the 1σ confidence intervals of those models. The high-frequency slope of the physical model is also shallower than the high-frequency slopes of the double and broken power-law models, but is ultimately consistent with their 1σ confidence intervals. The Bayes factor for the physical model over a single power-law model is 324. We do not attempt to characterize f_{knee} for NANOGrav’s 15 yr GWB spectrum, which appears to be dominated by white noise above ~ 20 nHz.

Finally, we compare the double power-law, broken power-law, and physical models using the Bayesian information criterion (BIC). The BIC is defined as $\text{BIC}_X = 2 \ln(\hat{L}_X) - k_X \ln n$, where $\hat{L}$ is the maximum likelihood estimate for model X , k_X is the number of parameters in model X , and n is the number of data points used for fitting (G. Schwarz 1978). Differences in BIC values, $\Delta(\text{BIC})_{12} = \text{BIC}_1 - \text{BIC}_2$, can be used to compare models, with positive $\Delta(\text{BIC})_{12}$ indicating model 1 is favored over model 2 (R. E. Kass & A. E. Raftery 1995). R. E. Kass & A. E. Raftery (1995) suggested $\Delta(\text{BIC})_{12} \gtrsim 10$ as a threshold for selecting model 1 over model 2.

We find that the double power-law, broken power-law, and physical model all have $\text{BIC} = 27$. Thus, $\Delta(\text{BIC}) = 0$ between any of these models, indicating none of these models are favored over the others. By contrast, a single power law has $\text{BIC} = 8$. The double power-law, broken power-law, and physical models are thus all preferred over a single power law with $\Delta(\text{BIC}) = 19$.

4.2. Power-law Excursions

We next determine if excursions from $f_{\text{GW}}^{-2/3}$ in the Hellings–Downs-correlated GWB spectrum are more signs of discreteness. To do so, we compare NANOGrav’s observed 15 yr GWB spectrum to the GWB spectra we compute from discrete SMBHBs (Section 2). We consider the difference between the observed S_h , $S_{h,k}^{\text{obs}}$, and the S_h of our generated spectra, $S_{h,k}^{\text{gen}}$, at each f_k . If the observed characteristic strain spectrum is similar to our spectral realizations, then the difference between the

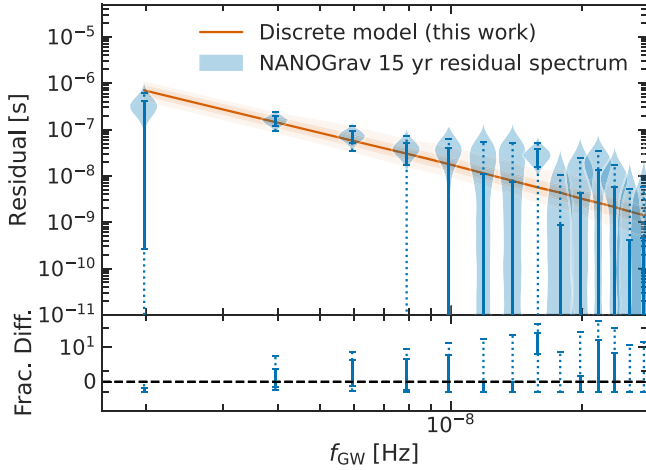


Figure 3. The 15 yr GWB residual spectrum compared to 1000 spectral realizations. The top panel shows the observed 15 yr residual spectrum (blue) compared to 1000 realizations of the GWB (red). We show the 68% and 95% confidence intervals at each frequency as solid and dotted error bars, respectively. We can see that these confidence intervals are heavily skewed toward small residual values. The bottom panel shows the fractional difference, $(S_h^{\text{obs}} - S_h^{\text{gen}})/S_h^{\text{gen}}$, between the observed SED, S_h^{obs} , and that of the generated spectra, S_h^{gen} .

observed and the realizations’ SEDs, $\Delta S_{h,k} = S_{h,k}^{\text{obs}} - S_{h,k}^{\text{gen}}$, should be zero.

The $S_{h,k}^{\text{obs}}$ and $S_{h,k}^{\text{gen}}$ are probability distributions; thus, the $\Delta S_{h,k}$ are also probability distributions, $P_k(\Delta S_h)$. We therefore calculate the two-sided p -value for the null hypothesis that ΔS_h is consistent with zero at each frequency, $p = 2 \min\{P_k(\Delta S_h \geq 0), P_k(\Delta S_h \leq 0)\}$. We focus the presentation of our results on the GWB spectrum values at 2, 12, 14, and 16 nHz, which were highlighted in G. Agazie et al. (2023c), though all 30 GWB spectrum frequencies were analyzed. We find that the observed residual spectrum value at 2 nHz is below our GWB realizations with $p = 0.05$, corresponding to a 1.9σ excursion. The values at 12 nHz and 14 nHz are below our GWB with $p = 0.31$ (1.0σ) for each. Finally, the spectrum value at 16 nHz is above our GWB with $p = 0.15$ (1.4σ , Figure 3).

We notice, however, that the 95% confidence intervals for the residual spectrum—shown by the dotted error bars in Figure 3—skew heavily toward small residuals. We investigate this further in Figure 4, which shows the log-PDFs of the residual spectrum posteriors. We find that the residual spectrum posteriors appear to be log-uniform at $\lesssim 0.1$ –1 ns, reflecting the choice of prior. This indicates that very weak GWB signals are not heavily disfavored at most frequencies the same way that very loud signals are disfavored. Consequentially, the significance of excursions from an expected $f_{\text{GW}}^{-2/3}$ power law can sometimes depend on where we cut these PDFs.

To assess how the significance of excursions at 2 nHz, 12 nHz, 14 nHz, and 16 nHz change with the prior range for the residuals, we artificially cut the residual spectrum PDFs at 0.1 ns and 1 ns. The lower value of 0.1 ns is chosen to reflect the residual values where the PDFs appear to transition to prior dominated, log-uniform distributions. The upper value of 1 ns corresponds to the prior bound used to analyze NANOGrav’s 12.5 yr data set (Z. Arzoumanian et al. 2020a). To make these cuts, we consider only the portion of the residual spectrum

PDFs above 0.1 ns and 1 ns. We then renormalize each PDF so they integrate to unity above these limits.

We find the residual spectrum at 2 nHz remains below our GWB spectral realizations with $p = 0.05$ (1.9σ) when using a residual cut at 0.1 ns, and with $p = 0.06$ (1.8σ) when using a 1 ns residual cut. With a 0.1 ns residual cut, the spectrum values at 12 nHz and 14 nHz are below our spectral realizations with $p = 0.78$ (0.3σ) and $p = 0.92$ (0.2σ), respectively. With a 1 ns residual cut, they are both above our spectral realizations with $p = 0.82$ (0.2σ). The spectrum at 16 nHz lies above our GWB realizations with $p = 0.06$ (1.9σ) using a 0.1 ns residual cut, and with $p = 0.04$ (2.1σ) using a 1 ns residual cut. All other frequencies are $\lesssim 1\sigma$ from the median of the GWB realizations, regardless of prior choice. Results for the 2 nHz, 12 nHz, 14 nHz, and 16 nHz frequencies are summarized in Table 1.

It is worth emphasizing that analytic models of the GWB strain spectrum, including an $f_{\text{GW}}^{-2/3}$ power law and more complicated power laws, are only statistical descriptions of the strain spectrum arising from the expected statistical distribution of SMBHBs over f_{GW} (E. S. Phinney 2001; A. H. Jaffe & D. C. Backer 2003; A. Sesana et al. 2008; W. G. Lamb & S. R. Taylor 2024). The spectrum of a single realization of the GWB depends only on the underlying SMBHB population. For example, in Figure 5 we show a single characteristic strain spectrum realization out of the 1000 we generated. This realization includes two excursions from the expected $f_{\text{GW}}^{-2/3}$ power-law behavior near 20 nHz. Each excursion appears to be associated with having $\mathcal{O}(1)$ a more massive or nearby SMBHB than expected from naïvely extrapolating the $f_{\text{GW}}^{-2/3}$ power law to higher frequencies. This demonstrates that excursions from the expected $f_{\text{GW}}^{-2/3}$ power-law behavior of the GWB are unsurprising. Even a single massive or nearby binary in excess of the power-law expectation can lead to excursions from an $f_{\text{GW}}^{-2/3}$ power law.

4.3. Comparison to 12.5 yr Spectrum

It is also interesting to compare the 15 yr residual spectrum to the residual spectrum from NANOGrav’s 12.5 yr data set (Z. Arzoumanian et al. 2020a). Assuming the GWB is a stationary signal (i.e., it does not change appreciably from the 12.5 yr data set to the 15 yr data set), any differences between the two sets of spectra should be due to differences between the 12.5 yr and 15 yr data sets, including increased pulsar timing baselines, the inclusion of more pulsars, and improved pulsar noise modeling in the 15 yr data set compared to the 12.5 yr data set (G. Agazie et al. 2023a, 2023d). By comparing the residual spectra, we can thus determine if excursions from an expected $f_{\text{GW}}^{-2/3}$ power law in the 15 yr data set are consistent with previous NANOGrav data sets or if they have only emerged recently as NANOGrav data sets have improved.

The frequency intervals in both the 12.5 yr and 15 yr residual spectra are $\Delta f_{\text{GW}} = 1/T_{\text{obs}}$, corresponding to the minimum frequency resolution in each data set. Since the 15 yr data set has a longer observation time than the 12.5 yr data set, the interval, $\Delta f_{15} = 2.0$ nHz, in the 15 yr data set is smaller than the interval in the 12.5 yr data set, $\Delta f_{12.5} = 2.5$ nHz. To consistently compare the residual spectra measured in each data set, we re-compute the 15 yr residual spectrum using the coarser $\Delta f_{12.5}$ intervals (see Figure 6). This is equivalent to decomposing the GWB signal in the 15 yr data set on the 12.5 yr Fourier basis frequencies, $f_k \approx k/12.5 \text{ yr}^{-1}$ for $k = 1, \dots, 30$.

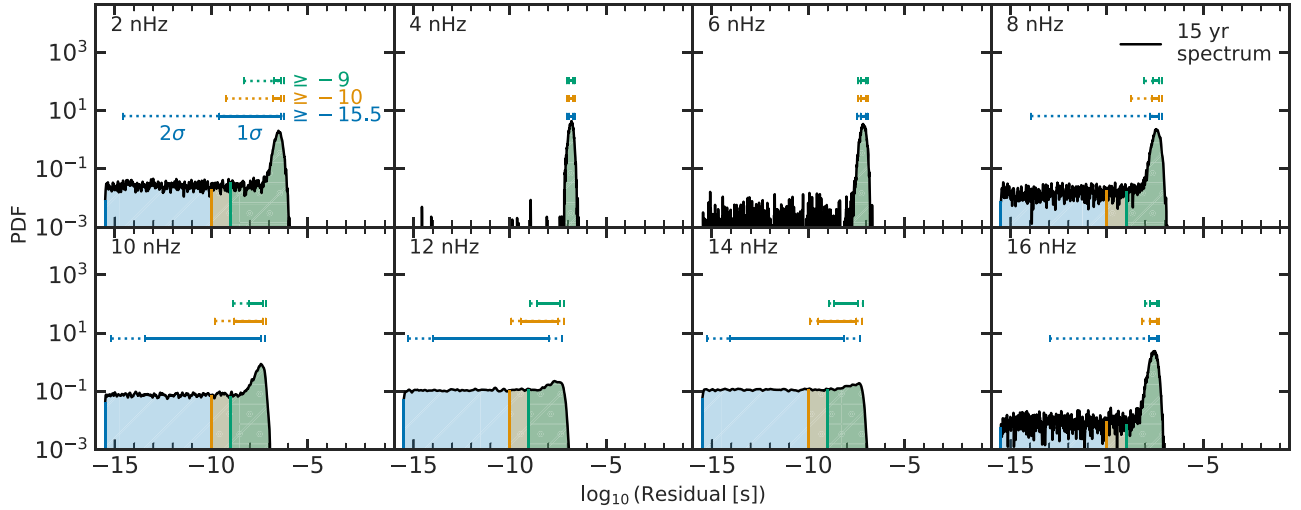


Figure 4. The broad, log-uniform support in the residual posterior distributions skews the value of the 15 yr GWB residual spectrum at each frequency. This can in turn affect our interpretation of the spectrum. Here we show a detailed view of the first eight GWB residual spectrum frequencies (black), which broadly tend toward log-uniform distributions at small residual values. The best constrained frequencies are 4 and 6 nHz, which exhibit much lower support at small residuals than other frequencies. Small “spikes” in the PDFs of these frequencies are due to the small number of MCMC samples contributing to the KDE distributions at these low residual values. Error bars above each distribution show the 68% and 95% confidence intervals (solid and dotted) under different posterior cuts, which are color coordinated with the shaded regions under each probability distribution. The original, $\log_{10}(\text{Residual [s]}) \geq -15.5$ cut is shown in blue, while a cut at 0.1 ns ($\log_{10}(\text{Residual [s]}) \geq -10$) is shown in yellow, and a cut at 1 ns ($\log_{10}(\text{Residual [s]}) \geq -9$) is shown in green. It is clear that the choice of priors appreciably affect the long tails we see in, e.g., Figure 3.

Table 1

Significance of Excursions in the Observed GWB Residual Spectrum from the Range of Spectra Computed from Discrete SMBHB Populations for the Four Frequencies Highlighted in G. Agazie et al. (2023c)

f_{GW} (nHz)	$p_{15.5} (\sigma_{15.5})$	$p_{10} (\sigma_{10})$	$p_9 (\sigma_9)$
2	0.05 (−1.9 σ)	0.05 (−1.9 σ)	0.06 (−1.8 σ)
12	0.31 (−1.0 σ)	0.78 (−0.3 σ)	0.92 (0.1 σ)
14	0.31 (−1.0 σ)	0.82 (−0.2 σ)	0.82 (0.2 σ)
16	0.15 (1.4 σ)	0.05 (1.9 σ)	0.04 (2.1 σ)

Note. p -values reflect the consistency of ΔS_h with zero, while corresponding σ significances are given in parentheses. Subscripts X on p_X and σ_X in the last three columns denote the lower log-residual cut, i.e., $\log_{10}(\text{Residual}) \geq -X$. The 15 yr GWB residual spectrum at 2 nHz is $\sim 2\sigma$ below the median value of our realizations regardless of the prior cut used, while at 16 nHz, the spectrum is up to 2.1 σ above the median value, depending on the prior cut (see Figure 4). At 12 nHz and 14 nHz, the spectrum is consistent with our GWB realizations.

We see that the 16 nHz excursion remains consistently $\sim 2\sigma$ above the median value of our GWB realizations. At frequencies higher than 16 nHz, the 12.5 yr and 15 yr spectra appear to have similar white noise. Below 16 nHz, the spectra differ appreciably. These differences may be due to differences between the data sets themselves, such as 3 yr longer timing baselines in the 15 yr data set compared to the 12.5 yr data set, and the addition of 21 new pulsars with ≥ 2.5 yr timing baselines in the 15 yr data set.

4.4. Excess White Noise Hypothesis

We next assess the significance of excursions at 2, 12, 14, and 16 nHz under the assumption that NANOGrav’s 15 yr residual spectrum includes excess white noise. We compute a phenomenological residual spectrum model by adding constant white noise, σ_{WN} , which is uncorrelated over f_{GW} to the 1000 GWB spectra computed from discrete SMBHBs, consistent with the flat residual spectrum above ~ 20 nHz. Specifically, we

model the characteristic strain spectrum of each of the 1000 generated spectra and at each f_k as a chi-distributed random variable with two degrees of freedom, centered on $h_c(f_k)$ with scale parameter σ_{WN} . We then fit this model to NANOGrav’s 15 yr residual spectrum, finding $\sigma_{\text{WN}} \sim 7$ ns.

We compare our generated residual spectra plus additional white model to NANOGrav’s 15 yr residual spectrum as in Section 4.2. We find that the excursion at 2 nHz is $\sim 2\sigma$ below the discrete spectrum with additional noise ($p = 0.04$ – 0.05 , depending on prior cut), the excursions at 12 and 14 nHz are 1.3σ – 0.3σ below the discrete spectrum with additional noise ($p = 0.2$ – 0.8), and the excursion at 16 nHz is $\sim 1\sigma$ above the discrete spectrum with additional noise ($p = 0.3$ – 0.4). Full results for these frequencies are summarized in Table 2.

We note that NANOGrav’s analysis pipeline already includes three white noise parameters encompassing different white noise sources. The inclusion of these parameters yields a reduced χ^2 near unity for the fit of the pulsar timing model to timing residuals without necessitating additional sources of noise (G. Agazie et al. 2023d). Thus, 7 ns of excess white noise correlated between pulsars in the residual spectrum is difficult to justify. The white noise component assumed here therefore lacks a clear physical motivation.

It is possible this excess results from the fact that NANOGrav uses the maximum a posteriori white noise parameter values for each pulsar in their GWB search and characterization, as marginalizing over white noise parameters is computationally infeasible (G. Agazie et al. 2023d). Alternatively, this excess noise may be due to using a one-size-fits-all approach for pulsar noise modeling, which does not consider additional sources of noise in individual pulsars (L. Lentati et al. 2016; M. Falxa et al. 2023; B. Larsen et al. 2024). A deeper investigation of noise in individual pulsars in NANOGrav’s 12.5 yr data set is currently underway (J. Simon et al. 2024, in preparation). The custom noise model techniques developed in J. Simon et al. (2024, in preparation) will then be

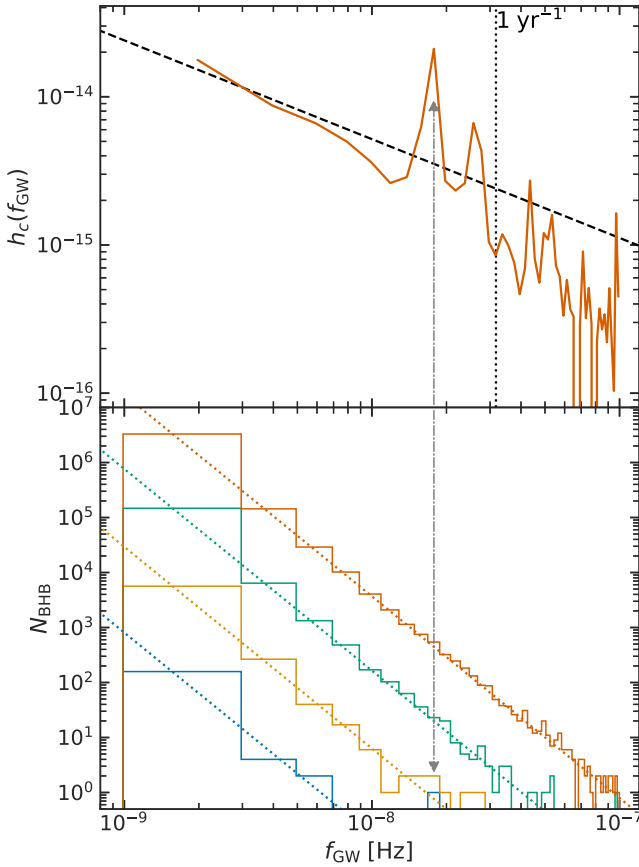


Figure 5. Example of expected excursion(s) in the GWB strain spectrum. Top panel: a single GWB realization. The dashed black line shows the $f_{\text{GW}}^{-2/3}$ power-law expectation, while the dotted gray arrow highlights one of two power-law excursions present in this realization. Bottom panel: corresponding histogram showing the number of SMBHBs in each frequency interval, as in the right side of Figure 1 with dotted lines showing the expected analytic distribution. Colors correspond to mass bins, specifically $8.5 \leq \log_{10} M < 9.0$ (red), $9.0 \leq \log_{10} M < 9.5$ (green), $9.5 \leq \log_{10} M < 10.0$ (yellow), and $10.0 \leq \log_{10} M < 10.5$ (blue). In this realization, the excursion below ~ 20 nHz (gray arrow) appears to come from a combination of two SMBHBs: one with $M \gtrsim 10^{9.5} M_{\odot}$ and one with $M \gtrsim 10^{10} M_{\odot}$. The excursion above ~ 20 nHz may result from one extra SMBHB with $M \gtrsim 10^{9.5} M_{\odot}$.

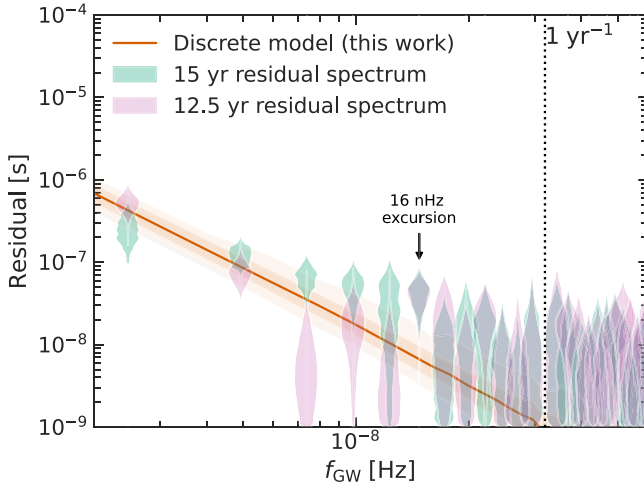


Figure 6. Comparison of 12.5 yr and 15 yr residual spectra, both using 12.5 yr frequency intervals. Specifically, we have re-computed the 15 yr residual spectrum using the 15 yr pulsar timing data with $\Delta f = 2.5$ nHz. We can see that the excursion at 16 nHz is remarkably consistent from the 12.5 yr data set to the 15 yr data set.

Table 2

Significance of Excursions in the Observed GWB Residual Spectrum as in Table 1, but with an Additional 7 ns of Excess White Noise

f_{GW} (nHz)	$p_{15.5} (\sigma_{15.5})$	$p_{10} (\sigma_{10})$	$p_9 (\sigma_9)$
2	0.04 (-2.0σ)	0.05 (-2.0σ)	0.05 (-2.0σ)
12	0.21 (-1.2σ)	0.52 (-0.6σ)	0.73 (-0.3σ)
14	0.21 (-1.3σ)	0.55 (-0.6σ)	0.79 (-0.3σ)
16	0.39 (0.9σ)	0.31 (1.0σ)	0.29 (1.1σ)

applied to the full 15 yr data set (G. Agazie et al. 2024, in preparation).

4.5. Loud Binary Hypothesis

A possible explanation for the excess at 16 nHz is the presence of a loud SMBHB at that frequency, such as the example in Figure 5. We test this hypothesis by constraining the additional strain amplitude a loud binary would contribute to the GWB to be consistent with the 16 nHz excursion. We then compare this hypothetical strain to constraints from NANOGrav’s 15 yr CW search (G. Agazie et al. 2023e), which found no CWs. Other possible sources for this excursion, such as mis-modeled pulsar noise or multiple SMBHBs emitting at 16 nHz, will require further investigation. See Section 5 for further discussion.

If we assume the 16 nHz excursion is due to a loud binary in addition to an underlying population of fainter binaries, we find it would need to have a sky- and polarization-averaged strain of $h = (3.8^{+2.3}_{-3.2}) \times 10^{-15}$, where the central given value is the median, and upper and lower uncertainties denote the 68% confidence interval. For comparison, the hypothetical binary CW 95% upper limit is 7.8×10^{-15} at this frequency. In Figure 7, we show the range of $\mathcal{M}$ and luminosity distances, D_L , which could produce this strain.

Importantly, NANOGrav did not find significant evidence of CWs in the 12.5 yr and 15 yr data sets (G. Agazie et al. 2023e; Z. Arzoumanian et al. 2023). Thus, we compare the hypothetical binary strain to constraints on CW strain near 16 nHz from NANOGrav’s 15 yr CW search (see Figure 7). We consider two sets of constraints from NANOGrav’s 15 yr CW search: one using uniform strain amplitude priors and one with log-uniform priors. Interestingly, both sets of constraints have peaks near the hypothetical binary strain peak, as shown in the right side of Figure 7. For each set of constraints, we calculate the p -value for the null hypothesis that the 15 yr CW constraints are consistent with the hypothetical binary strain. We find $p = 0.69$ for the uniform prior constraints and $p = 0.65$ for the log-uniform prior constraints. Thus, if the 16 nHz excursion is due to a binary, it would not have been detected in NANOGrav’s 15 yr CW search.

There are several extant SMBHB candidates that have $\mathcal{M}$ and D_L that are consistent with the 16 nHz excursion. Here we show three such candidate SMBHBs, HS 1630+2355, HS 0926+3608, and SDSS J114857.33+1600. These were identified by C. Xin et al. (2021) as interesting candidates, due to their optical periodic quasar light curves in the Catalina Real-time Transient Survey (M. J. Graham et al. 2015), and large masses. Hydrodynamical simulations show that SMBHBs can induce periodicity in quasar light curves, though the specific relationship between binary periodicity and light-curve periodicity is uncertain (B. D. Farris et al. 2014; J. R. Westerman-Schneider et al. 2022; F. Cocchiararo et al. 2024). We use FWHM binary

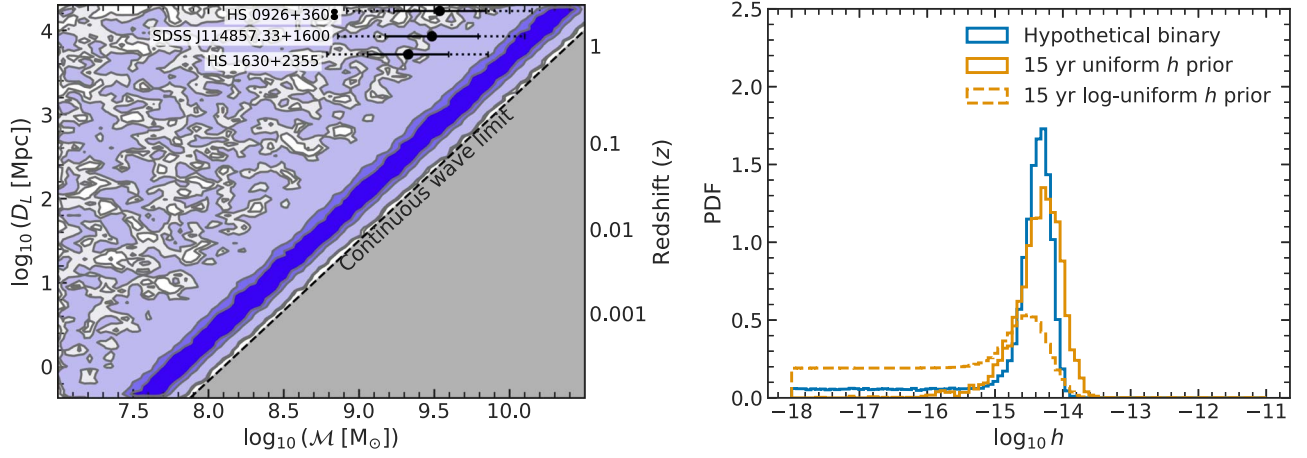


Figure 7. The 16 nHz excursion can be explained by the presence of an additional SMBHB at that frequency. Left panel: allowed luminosity distance, D_L , and chirp mass, $\mathcal{M}$, distribution of single binaries that could reproduce the excess at 16 nHz. Contours and blue shading show the regions of parameter space containing (from darkest to lightest) 50%, 68%, 95%, and 99% of MCMC samples. The dashed black line and gray region show the parameter space that has already been ruled out by the nondetection of continuous waves in NANOGrav’s 15 yr data set (G. Agazie et al. 2023e). Black dots show the median $\mathcal{M}$ estimates and D_L of four SMBHB candidates. Solid error bars show the 68% $\mathcal{M}$ confidence intervals for each candidate, while dotted error bars show 95% $\mathcal{M}$ confidence intervals. Right panel: the strain of a hypothetical binary at 16 nHz (blue) is comparable to CW constraints from NANOGrav’s 15 yr data set (yellow). NANOGrav’s 15 yr CW constraints were set using two sets of priors on strain amplitude: uniform (solid yellow) and log-uniform priors (dashed yellow). Interestingly, both sets of 15 yr CW constraints have peaks near the hypothetical binary strain peak. However the differences between the hypothetical binary strain and the CW constraints are only marginal. Therefore, a binary comparable to the hypothetical binary we consider here may not have been detectable in the 15 yr data set.

mass estimates, M , from C. Xin et al. (2021) to calculate $\mathcal{M}$ with Monte Carlo uncertainty estimates, drawing the binary mass ratio q from a uniform distribution between 0.1 and 1. Uncertainties on $\mathcal{M}$ are dominated by the ~ 0.3 dex Gaussian uncertainties on M .

An SMBHB’s GW emission frequency evolves as $\dot{f}_{\text{GW}} = (96/5) \pi^{8/3} \mathcal{M}^{5/3} f_r^{11/3}$, such that a binary emitting at 16 nHz evolves at a rate of $\sim 5 \times 10^{-4} (\mathcal{M}/10^9 M_\odot)^{5/3} (f_{\text{GW}}/16 \text{ nHz})^{11/3} \text{ nHz yr}^{-1}$. At the same GW frequency, a rarer $\mathcal{M} = 10^{10} M_\odot$ SMBHB (Figure 1), will evolve at a rate of $\dot{f}_{\text{GW}} \approx 0.02 \text{ nHz yr}^{-1}$. If a single loud binary is the source of the 16 nHz excursion, we thus expect it to persist in future NANOGrav data sets.

5. Discussion

We carried out a suite of Monte Carlo realizations to determine how the discreteness of the GWB manifests in its strain and residual spectra. We predicted a knee in the GWB strain spectrum at $26_{-19}^{+28} \text{ nHz}$ using a double power-law model, and at 8_{-4}^{+8} nHz using a physically motivated model of the characteristic strain spectrum. A similar knee was first expected at $37_{-13}^{+15} \text{ nHz}$ by A. Sesana et al. (2008), and is also subtly visible, though not highlighted, in simulated strain spectra from V. Ravi et al. (2012), E. Roebber et al. (2016), L. Z. Kelley et al. (2017), and S. R. Taylor et al. (2017). The fact that the double power-law model and the physical model predict different values for f_{knee} suggest this is somewhat model dependent. However, the uncertainties in f_{knee} for both models are broad, and the median value of f_{knee} predicted with the physical model lies within the 68% confidence interval of the f_{knee} predicted with the double power-law model. It would be interesting to explore systematic differences between these models in future work to aid interpretation of f_{knee} . The 8_{-4}^{+8} nHz knee predicted by the physically motivated model suggests that a knee in the GWB characteristic strain spectrum could be by characterized by future PTA data sets.

We next assessed the level of significance of excursions from an average $f_{\text{GW}}^{-2/3}$ power law in NANOGrav’s 15 yr GWB residual spectrum (G. Agazie et al. 2023c). To determine the significance of these excursions, we compared them to the distribution of 1000 realizations of GWB residual spectra. We paid particular attention to the frequencies 2 nHz, 12 nHz, 14 nHz, and 16 nHz, which were previously noted as potentially interesting in G. Agazie et al. (2023c). We found that the excursion from an average expected $f_{\text{GW}}^{-2/3}$ power law seen at 2 nHz is below the median value of our GWB realizations with $p = 0.05$, the excursions at 12 nHz and 14 nHz are below the median with $p = 0.31$, and the power at 16 nHz is above the median with $p = 0.15$ (Table 1).

We also find, however, that constraints on the residual spectrum do not strongly rule out weak GW signals at most frequencies. Consequentially, the significance of excursions from an $f_{\text{GW}}^{-2/3}$ power law depends on the priors. By cutting the residual spectrum PDFs at higher residuals, we find the excursion at 2 nHz below the median value of our GWB realizations with $p = 0.05$ – 0.06 , the excursions at 12 nHz and 14 nHz are consistent with the median value of our GWB realizations with $p = 0.78$ – 0.92 , and the excursion at 16 nHz is above the median of our GWB realizations with $p = 0.04$ – 0.05 .

We additionally compared the spectra measured by NANOGrav in the 12.5 yr and 15 yr data sets, re-computing the 15 yr spectrum on the coarser 12.5 yr frequency intervals for a more consistent comparison. Both data sets show excursions from an average expected $f_{\text{GW}}^{-2/3}$ power law at $\sim 16 \text{ nHz}$. The constraints on the GWB spectrum at this frequency are very consistent from the 12.5 yr data set to the 15 yr data set and lie $\sim 2\sigma$ above the median value of our GWB realizations. Below $\sim 16 \text{ nHz}$, there are noticeable differences between both sets of spectra, with the 15 yr spectrum appearing to be better constrained than the 12.5 yr spectrum at these frequencies. Above 16 nHz, both spectra appear fully consistent with white noise.

The excursion in the 15 yr GWB spectrum from an $f_{\text{GW}}^{-2/3}$ power law at 2 nHz is below our GWB realizations with

$p = 0.05\text{--}0.06$ ($\approx 1.8\sigma\text{--}1.9\sigma$). This may indicate a turnover in the GWB spectrum due to environmental coupling between SMBHBs and their host environments at low frequencies (see, e.g., L. Sampson et al. 2015). This scenario has previously been shown to be consistent with the measured GWB spectrum (G. Agazie et al. 2023c, 2023b). Specifically, G. Agazie et al. (2023b) reported that the 2 nHz excursion is consistent with a turnover in the characteristic strain spectrum due to interactions between SMBHBs and their host galaxy environments. We thus focus our discussion on the 16 nHz excursion.

One plausible explanation for the excursion from an expected average $f_{\text{GW}}^{-2/3}$ power law at 16 nHz is the existence of a single sufficiently massive and/or nearby SMBHB emitting at ~ 16 nHz. This interpretation is consistent with a GWB sourced by a discrete population of SMBHBs, as an $f_{\text{GW}}^{-2/3}$ power law only holds on average over many realizations of the GWB. Indeed, we find that any random realization of the GWB can include large excursions from an $f_{\text{GW}}^{-2/3}$ power law if the underlying SMBHB population includes even one more SMBHB at a higher mass than the analytic average expectation (Figure 5).

Assuming the excursion at 16 nHz is due to a single loud binary, we decomposed the excess strain (compared to the distribution of generated GWB spectra) at this frequency into the corresponding $\mathcal{M}$ and D_L (Figure 7). This is consistent with a GWB sourced by a population of discrete SMBHBs, as individual realizations of the GWB do not need to adhere to an average $f_{\text{GW}}^{-2/3}$ power law. Interestingly, if the excursion from an $f_{\text{GW}}^{-2/3}$ power law at 16 nHz is a single binary, it would not be detected in NANOGrav’s 15 yr CW search. The 16 nHz excursion could alternatively be sourced by multiple binaries, which would also be undetectable as CWs in NANOGrav’s 15 yr data. This scenario may be less likely than a single loud SMBHB, however, as it would require multiple very massive SMBHBs—which are expected to be rare (Figure 1)—to be coincidentally emitting at similar frequencies. For example, if we assume that HS 1630+2355, HS 0926+3608, and SDSS J114857.33+1600 are all SMBHBs emitting near 16 nHz, the characteristic strain of their combined GW emission would still only be $(6.6^{+8.3}_{-3.6}) \times 10^{-16}$ —which could not source the 16 nHz excursion.

One promising avenue for follow-up is targeted CW searches, which can be up to an order-of-magnitude more sensitive than all-sky CW searches (Z. Arzoumanian et al. 2020b). Targeted searches for CWs in the 15 yr data set are currently underway. These searches target SMBHB candidates that have been identified electromagnetically via, e.g., apparent quasar light-curve periodicity (D. J. D’Orazio et al. 2013; B. D. Farris et al. 2014; R. Miranda et al. 2017; D. J. Muñoz et al. 2020). Existing SMBHB candidate catalogs may contain SMBHBs with orbital periods $P_{\text{orb}} \lesssim 6$ yr (M. J. Graham et al. 2015), corresponding to $f_{\text{GW}} \gtrsim 10$ nHz (since $f_{\text{GW}} = 2/P_{\text{orb}}$). Thus, an SMBHB with $f_{\text{GW}} \sim 16$ nHz may already be present in the extant SMBHB candidate catalogs.

Additionally, the signal-to-noise ratio of the GWB improves proportionally to the number of pulsars in the array (X. Siemens et al. 2013). Constraints on the GWB spectrum at 16 nHz—which corresponds to GW periods of ~ 2 yr—can thus be improved by adding pulsars with $\gtrsim 2$ yr of timing data to future NANOGrav data sets. The International PTA (IPTA), which combines data from PTA experiments around the globe (J. P. W. Verbiest et al. 2016; B. B. P. Perera et al. 2019), could accomplish this in its

anticipated third data release by adding data from the MeerKAT PTA, which included 78 pulsars with ~ 2.5 yr of pulsar timing data in its first data release (M. T. Miles et al. 2023).

Finally, we must also consider the possibility that the 16 nHz excursion does not have a GW origin at all. Indeed, neither the PPTA nor the EPTA+InPTA spectra have clear excursions at 16 nHz compared to NANOGrav’s 15 yr GWB spectrum (EPTA Collaboration et al. 2023; D. J. Reardon et al. 2023; G. Agazie et al. 2024a). We thus also considered a phenomenological model of the characteristic strain spectrum that includes ~ 7 ns of excess white noise in the residual spectrum of the GWB. We find that the excursion at 16 nHz is above our generated GWB realizations with a reduced significance of $p = 0.3\text{--}0.4$, corresponding to a $\sim 1\sigma$ excursion. However, we emphasize that current NANOGrav analyses already consider white noise when fitting the pulsar timing model to timing residual data, with a reduced χ^2 near unity.

While searching for CWs in the IPTA’s second data release, M. Falxa et al. (2023) found that a one-size-fits-all approach to pulsar noise modeling can increase the probability of a false alarm. They further found that custom pulsar noise models reduce false alarms and improve our ability to constrain CWs from individual SMBHBs. PTA experiments could test whether or not the 16 nHz excursion is due to mis-modeled pulsar noise by incorporating custom pulsar noise models in future data sets (L. Lentati et al. 2016; A. Chalumeau et al. 2022; B. Goncharov et al. 2024; B. Larsen et al. 2024). It will be interesting to see how custom noise models for individual pulsars in the 15 yr data set will affect apparent white noise levels (G. Agazie et al. 2024, in preparation), and thus potentially change our results. Determining the specific source of the 16 nHz power-law excursion—including whether or not it is due to a GW source at all—will require further investigation.

Our study is also important to help assess if SMBHBs may indeed be likely sources of the GWB, alongside searches for CWs (G. Agazie et al. 2023e; Z. Arzoumanian et al. 2023) and anisotropy (C. M. F. Mingarelli et al. 2013; S. R. Taylor & J. R. Gair 2013; N. J. Cornish & R. van Haasteren 2014; J. Gair et al. 2014; S. R. Taylor et al. 2015, 2020; N. Pol et al. 2022; G. Agazie et al. 2023f; E. C. Gardiner et al. 2024; G. Sato-Polito & M. Kamionkowski 2024). Without a careful exploration of the strain spectrum, it will not be clear if excursions from the expected average power-law behavior are consistent with what we expect from a population of discrete SMBHBs, or are sourced by some other physics, e.g., A. Afzal et al. (2023). Indeed, primordial GWB are expected to be isotropic and also follow a power law (L. P. Grishchuk 2005; P. D. Lasky et al. 2016; A. Afzal et al. 2023; G. Agazie et al. 2024b). Signs of discreteness in the GWB residual spectrum are therefore an important signature of SMBHBs, and may be observed before GWB anisotropy or CWs.

We show conclusively that excursions in NANOGrav’s 15 yr GWB spectrum are well within the range of spectra generated by a population of discrete SMBHBs, needing no other physics to explain it. Furthermore, we predict that the GWB breaks down at 26^{+28}_{-19} nHz. If white noise can be reduced, the rest of the available high-frequency parameter space will be ideal for constraining CWs. Finally, our study also highlights the importance of spectral analyses, such as those carried out by NANOGrav, which reveal interesting features in the GWB. Recent advancements in spectral analysis techniques have reduced the computational cost associated with constraining the

GWB spectrum, making spectral analyses more accessible for PTAs to carry out, e.g., W. G. Lamb et al. (2023).

Acknowledgments

L.B.I. acknowledges support from the National Science Foundation under award AST-1909933 and from the Research Corporation for Science Advancement under Cottrell Scholar Award No. 27553. P.R.B. is supported by the Science and Technology Facilities Council, grant No. ST/W000946/1. S.B. gratefully acknowledges the support of a Sloan Fellowship, and the support of NSF under award No. 1815664. M.C. and S.R.T. acknowledge support from NSF AST-2007993. M.C. was supported by the Vanderbilt Initiative in Data Intensive Astrophysics (VIDA) Fellowship. Support for this work was provided by the NSF through the Grote Reber Fellowship Program administered by Associated Universities, Inc./National Radio Astronomy Observatory. M.E.D. acknowledges support from the Naval Research Laboratory by NASA under contract S-15633Y. T.D. and M.T.L. are supported by an NSF Astronomy and Astrophysics grant (AAG) award No. 2009468. E.C.F. is supported by NASA under award No. 80GSFC21M0002. G.E. F., S.C.S., and S.J.V. are supported by NSF award PHY-2011772. K.A.G. and S.R.T. acknowledge support from an NSF CAREER award No. 2146016. A.D.J. and M.V. acknowledge support from the Caltech and Jet Propulsion Laboratory President's and Director's Research and Development Fund. A.D.J. acknowledges support from the Sloan Foundation. The work of N.La., X.S., and D.W. is partly supported by the George and Hannah Bolinger Memorial Fund in the College of Science at Oregon State University. N.La. acknowledges the support from Larry W. Martin and Joyce B. O'Neill Endowed Fellowship in the College of Science at Oregon State University. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). M.A.M. is supported by NSF No. 1458952. M.A.M. is supported by NSF No. 2009425. C.M.F.M. was supported in part by the National Science Foundation under grant Nos. NSF PHY-1748958 and AST-2106552. A.Mi. is supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy—EXC 2121 Quantum Universe—390833306. K.D.O. was supported in part by NSF grant No. 2207267. T.T.P. acknowledges support from the Extragalactic Astrophysics Research Group at Eötvös Loránd University, funded by the Eötvös Loránd Research Network (ELKH), which was used during the development of this research. H.A.R. is supported by NSF Partnerships for Research and Education in Physics (PREP) award No. 2216793. S.M.R. and I.H. S. are CIFAR Fellows. Portions of this work performed at NRL were supported by ONR 6.1 basic research funding. J.D.R. also acknowledges support from start-up funds from Texas Tech University. J.S. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-2202388, and acknowledges previous support by the NSF under award 1847938. Pulsar research at UBC is supported by an NSERC Discovery Grant and by CIFAR. C.U. acknowledges support from BGU (Kreitman fellowship), and the Council for Higher Education and Israel Academy of Sciences and Humanities (Excellence fellowship). C.A.W. acknowledges support from CIERA, the Adler Planetarium, and the Brinson Foundation through a CIERA-Adler postdoctoral fellowship. O.Y. is supported by the National Science Foundation Graduate Research Fellowship under grant No. DGE-2139292.

Software: arviz (R. Kumar et al. 2019), astropy (Astropy Collaboration et al. 2013, 2018, 2022), jupyter (T. Kluyver et al. 2016), matplotlib (J. D. Hunter 2007), nHzGWs (C. Mingarelli 2017), numpy (C. R. Harris et al. 2020), pandas (W. McKinney 2010; The pandas development team 2022), pymc (T. Wiecki et al. 2023), scikit-learn (F. Pedregosa et al. 2011), scipy (P. Virtanen et al. 2020), seaborn (M. Waskom 2021)

Author Contributions

This paper is the result of the work of many people and uses data from over a decade of pulsar timing observations. J.A.C.C. wrote and developed new python codes to perform the analysis, created all of the Figures and Tables, and wrote a majority of the text. C.M.F.M. conceived of the project, supervised the analysis, helped write and develop the manuscript, and advised on the analysis, the Figures, and interpretation of the results. W.G.L. resampled the 15 yr data in Figure 6 (left panel). L.Z.K, K.D.O, L.Br., P.N., B.B., S.V., and D.K. provided insights into interpreting the results, and comments on the manuscript.

G.A., A.A., A.M.A., Z.A., P.T.B., P.R.B., H.T.C., K.C., M.E.D, P.B.D., T.D., E.C.F, W.F., E.F., G.E.F., N.G.D., D.C.G., P.A.G., J.G., R.J.J., M.L.J., D.L.K., M.K., M.T.L., D.R.L., J.L., R.S.L., A.M., M.A.M., N.M., B.W.M., C.N., D.J.N., T.T.N., B.B.P.P., N.S.P., H.A.R., S.M.R., P.S.R., A.S., C.S., B.J.S.A., I.H.S., K.S., A.S., J.K.S., and H.M.W. all ran observations and developed timing models for the NANOGrav 15 yr data set.

Appendix A Broken Power Law

Here we fit a broken power-law model to our simulated GWB spectra. This model was first employed by A. Sesana et al. (2008) to fit the knee of GWB realizations derived from halo merger trees. The broken power-law characteristic strain spectrum model is:

$$h_c(f_{\text{GW}}) = A_{\text{yr}} \left(\frac{f_{\text{GW}}}{f_{\text{yr}}} \right)^{\alpha_1} \left(1 + \frac{f_{\text{GW}}}{f_{\text{knee}}} \right)^{\alpha_2 - \alpha_1}. \quad (\text{A1})$$

Here, α_1 and α_2 are the low- and high-frequency characteristic strain spectrum slopes, respectively. $\alpha_1 = -2/3$ for circular SMBHBs. Fitting this to our GWB realizations, we find a bend

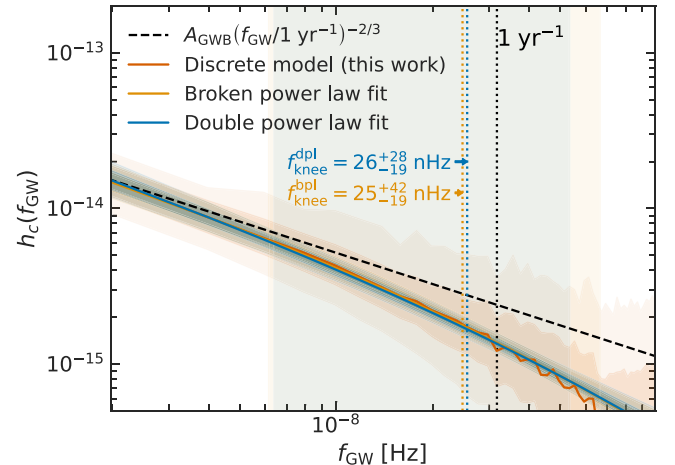


Figure 8. The broken power-law model (yellow) is nearly identical to the double power-law model (blue).

at $f_{\text{knee}} = 25_{-19}^{+42}$ nHz with a high-frequency power-law slope $f_{\text{GW}}^{-1.4 \pm 0.3}$. The results of this fit are shown in yellow in Figure 8, and are nearly identical to the double power-law fit.

Appendix B Physical Spectrum

Here we consider a physically motivated model of $h_c(f_{\text{GW}})$ that accounts for the fact that the total number of binaries in any frequency interval must be an integer; i.e., one cannot have fractions of an SMBHB contributing to the GWB (A. Sesana et al. 2008). Specifically, the total number of SMBHBs in frequency interval Δf_{GW} centered on frequency f_0 is

$$N_0 = \int_0^\infty \int_{f_0 - \Delta f_{\text{GW}}/2}^{f_0 + \Delta f_{\text{GW}}/2} \frac{d^2 N_{\text{BHB}}}{dM df_{\text{GW}}} dM df_{\text{GW}}, \quad (\text{B1})$$

where $d^2 N_{\text{BHB}}/(dM df_{\text{GW}}) = \int_0^\infty \int_0^1 d^2 N_{\text{BHB}}/(dM df_{\text{GW}}) dq dz$. In general, N_0 is not an integer. This is because N_0 is really the expected number of binaries in Δf_{GW} over many realizations of the Universe.

A. Sesana et al. (2008) therefore posited a characteristic mass, $\tilde{M}(f_0)$, such that over infinite realizations of the Universe, there is less than one SMBHB with $M > \tilde{M}(f_0)$ on average, i.e.,

$$\begin{aligned} \int_{\tilde{M}(f_0)}^\infty \int_{f_0 - \Delta f_{\text{GW}}/2}^{f_0 + \Delta f_{\text{GW}}/2} \frac{d^2 N_{\text{BHB}}}{dM df_{\text{GW}}} dM df_{\text{GW}} < 1 \\ \Rightarrow \int_{\tilde{M}(f_0)}^\infty p(M) dM \int_{f_0 - \Delta f_{\text{GW}}/2}^{f_0 + \Delta f_{\text{GW}}/2} \frac{dN_{\text{BHB}}}{df_{\text{GW}}} df_{\text{GW}} < 1, \end{aligned} \quad (\text{B2})$$

where in the second line, $dN_{\text{BHB}}/df_{\text{GW}}$ is the differential number of SMBHBs per f_{GW} and $p(M)$ is the PDF of SMBHBs with mass M , such that $d^2 N_{\text{BHB}}/(dM df_{\text{GW}}) = p(M) \times dN_{\text{BHB}}/df_{\text{GW}}$. SMBHBs with $M > \tilde{M}(f_0)$ thus do not contribute to the GWB in a typical realization of the Universe. To account for this discreteness effect, we model the effective characteristic strain spectrum as $h_{c,\text{eff}}(f_{\text{GW}}) = h_c(f_{\text{GW}}) \sqrt{1 - Z(f_{\text{GW}})}$, where

$$Z(f_{\text{GW}}) = \frac{\int_{\tilde{M}(f_{\text{GW}})}^\infty \int_0^1 \int_0^\infty \dot{\phi}_{\text{BHB}} \frac{d\tau}{dz} \frac{\mathcal{M}^{5/3}}{(1+z)^{1/3}} dM dq dz}{\int_0^\infty \int_0^1 \int_0^\infty \dot{\phi}_{\text{BHB}} \frac{d\tau}{dz} \frac{\mathcal{M}^{5/3}}{(1+z)^{1/3}} dM dq dz} \quad (\text{B3})$$

is the fraction of contributions to $h_c^2(f_{\text{GW}})$ coming from less than one SMBHB (A. Sesana et al. 2008).

We can derive an analytic expression for $Z(f_{\text{GW}})$ —and thus for $h_{c,\text{eff}}(f_{\text{GW}})$ —by assuming that $\dot{\phi}_{\text{BHB}}$ is power law distributed over mass (A. Sesana et al. 2008). Indeed, halo merger tree models of SMBHBs (e.g., M. Volonteri et al. 2003; A. Sesana et al. 2008) find $d^2 N_{\text{BHB}}/(dM d\tau) \propto M^{-\beta}$, with $1.5 \lesssim \beta \lesssim 2$ (A. Sesana et al. 2008). We also find a similar distribution in our generated SMBHB populations. Since $\dot{\phi}_{\text{BHB}} \propto d^2 N_{\text{BHB}}/(dM d\tau)$ (Equation (2)), this suggests that $\dot{\phi}_{\text{BHB}}$ can also be approximated as a power law, i.e.,

$$\dot{\phi} \approx \begin{cases} A(q, z) M^{-\beta} & M_{\min} \leq M \leq M_{\max}, \\ 0 & \text{otherwise} \end{cases} \quad (\text{B4})$$

where $A(q, z)$ is a model-dependent normalization factor that does not depend on M (and therefore cancels out in Equation (B3)), and where $M_{\min} = 10^{8.5} M_\odot$ and $M_{\max} = 10^{10.5} M_\odot$ are the bounds of integration. Plugging Equations (B4) into (B3), we find

(see A. Sesana et al. 2008)

$$\begin{aligned} Z(f_{\text{GW}}) &= \frac{\int_{\tilde{M}(f_{\text{GW}})}^{M_{\max}} M^{5/3-\beta} dM}{\int_{M_{\min}}^{M_{\max}} M^{5/3-\beta} dM} \\ &\simeq 1 - \left(\frac{\tilde{M}(f_{\text{GW}})}{M_{\max}} \right)^{8/3-\beta}, \end{aligned} \quad (\text{B5})$$

where, in the second line of Equation (B5), we have assumed $\beta < 8/3$ and $\tilde{M}(f_{\text{GW}}) \gg M_{\min}$. This is true for our generated populations at most frequencies and is in general true for a real population of SMBHBs, which can include small contributions to the characteristic strain from lower-mass SMBHBs than those considered in this work (J. A. Casey-Clyde et al. 2022).

We next solve for $\tilde{M}(f_{\text{GW}})$ before finally arriving at an analytic expression for $Z(f_{\text{GW}})$. Since $d\tau/df_{\text{GW}} \propto M^{-5/3} f_{\text{GW}}^{-11/3}$ (P. C. Peters & J. Mathews 1963), it follows from our power-law formulation of $\dot{\phi}_{\text{BHB}}$ that $d^2 N_{\text{BHB}}/(dM df_{\text{GW}}) \propto M^{-\beta-5/3} f_{\text{GW}}^{-11/3}$ (A. Sesana et al. 2008). The total number of SMBHBs per frequency interval Δf_{GW} centered on frequency f_{GW} is thus $N(f_{\text{GW}}) = N_0 (f_{\text{GW}}/f_0)^{-11/3}$ (A. Sesana et al. 2008), where f_0 is an arbitrary reference frequency, and N_0 is set by Equation (B1). Combining this expression for $N(f_{\text{GW}})$ with Equations (B1) and (B2) and solving for $\tilde{M}$, we find (see A. Sesana et al. 2008)

$$\begin{aligned} \tilde{M}(f_{\text{GW}}) &= M_{\max} \left[1 + \frac{\beta + 2/3}{BN_0 M_{\max}^{-\beta-2/3}} \right. \\ &\quad \times \left. \left(\frac{f_{\text{GW}}}{f_0} \right)^{11/3} \right]^{1/(-\beta-2/3)}, \end{aligned} \quad (\text{B6})$$

where B is a normalization constant such that $\int_0^\infty p(M) dM = \int_0^\infty BM^{-\beta-5/3} dM = 1$. Combining Equations (B5) and (B6), and defining $f_{\text{knee}} \equiv f_0 [(\beta + 2/3)/(BN_0 M_{\max}^{-\beta-2/3})]^{-3/11}$, the characteristic strain spectrum can be parameterized as

$$\begin{aligned} h_{c,\text{eff}} &= A_{\text{yr}} \left(\frac{f_{\text{GW}}}{f_{\text{yr}}} \right)^{\alpha_1} \\ &\quad \times \left[1 + \left(\frac{f_{\text{GW}}}{f_{\text{knee}}} \right)^{11/3} \right]^{3(\alpha_2 - \alpha_1)/11}, \end{aligned} \quad (\text{B7})$$

where $\alpha_1 = -2/3$ is the slope at $f_{\text{GW}} \ll f_{\text{knee}}$, and where $\alpha_2 = \alpha_1 + (11/6)(3\beta - 8)/(3\beta + 2)$ is the slope of the average strain spectrum at $f_{\text{GW}} \gg f_{\text{knee}}$. At $f_{\text{GW}} \gg f_{\text{knee}}$, we therefore have $h_{c,\text{eff}} \propto f_{\text{GW}}^{-2/3 + (11/6)(3\beta - 8)/(3\beta + 2)}$, which for $1.5 \lesssim \beta \lesssim 2$ is $h_{c,\text{eff}} \propto f_{\text{GW}}^{-1.6} - f_{\text{GW}}^{-1.1}$, in good agreement with both the double and broken power-law models. In fact, by fitting Equation (B7) to the 1000 spectra we generate from discrete SMBHB populations, we find $h_c \propto f_{\text{GW}}^{-1.1 \pm 0.1}$. This corresponds to $\beta = 2.1_{-0.2}^{+0.1}$, which is consistent with expectations from halo merger tree models (A. Sesana et al. 2008) and the SMBHB populations used in this work.

ORCID iDs

Gabriella Agazie  <https://orcid.org/0000-0001-5134-3925>
 Akash Anumalapudi  <https://orcid.org/0000-0002-8935-9882>
 Anne M. Archibald  <https://orcid.org/0000-0003-0638-3340>
 Jeremy George Baier  <https://orcid.org/0000-0002-4972-1525>
 Paul T. Baker  <https://orcid.org/0000-0003-2745-753X>
 Bence Bécsy  <https://orcid.org/0000-0003-0909-5563>
 Laura Blecha  <https://orcid.org/0000-0002-2183-1087>
 Adam Brazier  <https://orcid.org/0000-0001-6341-7178>
 Paul R. Brook  <https://orcid.org/0000-0003-3053-6538>
 Lucas Brown  <https://orcid.org/0009-0008-8199-580X>
 Sarah Burke-Spolaor  <https://orcid.org/0000-0003-4052-7838>
 J. Andrew Casey-Clyde  <https://orcid.org/0000-0002-5557-4007>
 Maria Charisi  <https://orcid.org/0000-0003-3579-2522>
 Shami Chatterjee  <https://orcid.org/0000-0002-2878-1502>
 Tyler Cohen  <https://orcid.org/0000-0001-7587-5483>
 James M. Cordes  <https://orcid.org/0000-0002-4049-1882>
 Neil J. Cornish  <https://orcid.org/0000-0002-7435-0869>
 Fronefield Crawford  <https://orcid.org/0000-0002-2578-0360>
 H. Thankful Cromartie  <https://orcid.org/0000-0002-6039-692X>
 Kathryn Crowter  <https://orcid.org/0000-0002-1529-5169>
 Megan E. DeCesar  <https://orcid.org/0000-0002-2185-1790>
 Paul B. Demorest  <https://orcid.org/0000-0002-6664-965X>
 Timothy Dolch  <https://orcid.org/0000-0001-8885-6388>
 Elizabeth C. Ferrara  <https://orcid.org/0000-0001-7828-7708>
 William Fiore  <https://orcid.org/0000-0001-5645-5336>
 Emmanuel Fonseca  <https://orcid.org/0000-0001-8384-5049>
 Gabriel E. Freedman  <https://orcid.org/0000-0001-7624-4616>
 Nate Garver-Daniels  <https://orcid.org/0000-0001-6166-9646>
 Peter A. Gentile  <https://orcid.org/0000-0001-8158-683X>
 Joseph Glaser  <https://orcid.org/0000-0003-4090-9780>
 Deborah C. Good  <https://orcid.org/0000-0003-1884-348X>
 Kayhan Gültekin  <https://orcid.org/0000-0002-1146-0198>
 Jeffrey S. Hazboun  <https://orcid.org/0000-0003-2742-3321>
 Ross J. Jennings  <https://orcid.org/0000-0003-1082-2342>
 Aaron D. Johnson  <https://orcid.org/0000-0002-7445-8423>
 Megan L. Jones  <https://orcid.org/0000-0001-6607-3710>
 Andrew R. Kaiser  <https://orcid.org/0000-0002-3654-980X>
 David L. Kaplan  <https://orcid.org/0000-0001-6295-2881>
 Luke Zoltan Kelley  <https://orcid.org/0000-0002-6625-6450>
 Matthew Kerr  <https://orcid.org/0000-0002-0893-4073>
 Joey S. Key  <https://orcid.org/0000-0003-0123-7600>
 Nima Laal  <https://orcid.org/0000-0002-9197-7604>
 Michael T. Lam  <https://orcid.org/0000-0003-0721-651X>
 William G. Lamb  <https://orcid.org/0000-0003-1096-4156>
 Bjorn Larsen  <https://orcid.org/0000-0001-6436-8216>
 Natalia Lewandowska  <https://orcid.org/0000-0003-0771-6581>
 Tingting Liu  <https://orcid.org/0000-0001-5766-4287>
 Duncan R. Lorimer  <https://orcid.org/0000-0003-1301-966X>
 Jing Luo  <https://orcid.org/0000-0001-5373-5914>
 Ryan S. Lynch  <https://orcid.org/0000-0001-5229-7430>
 Chung-Pei Ma  <https://orcid.org/0000-0002-4430-102X>
 Dustin R. Madison  <https://orcid.org/0000-0003-2285-0404>
 Alexander McEwen  <https://orcid.org/0000-0001-5481-7559>
 James W. McKee  <https://orcid.org/0000-0002-2885-8485>
 Maura A. McLaughlin  <https://orcid.org/0000-0001-7697-7422>
 Natasha McMann  <https://orcid.org/0000-0002-4642-1260>
 Bradley W. Meyers  <https://orcid.org/0000-0001-8845-1225>
 Patrick M. Meyers  <https://orcid.org/0000-0002-2689-0190>
 Chiara M. F. Mingarelli  <https://orcid.org/0000-0002-4307-1322>

Andrea Mitridate  <https://orcid.org/0000-0003-2898-5844>
 Priyamvada Natarajan  <https://orcid.org/0000-0002-5554-8896>
 Cherry Ng  <https://orcid.org/0000-0002-3616-5160>
 David J. Nice  <https://orcid.org/0000-0002-6709-2566>
 Stella Koch Ocker  <https://orcid.org/0000-0002-4941-5333>
 Ken D. Olum  <https://orcid.org/0000-0002-2027-3714>
 Timothy T. Pennucci  <https://orcid.org/0000-0001-5465-2889>
 Benetge B. P. Perera  <https://orcid.org/0000-0002-8509-5947>
 Nihan S. Pol  <https://orcid.org/0000-0002-8826-1285>
 Henri A. Radovan  <https://orcid.org/0000-0002-2074-4360>
 Scott M. Ransom  <https://orcid.org/0000-0001-5799-9714>
 Paul S. Ray  <https://orcid.org/0000-0002-5297-5278>
 Joseph D. Romano  <https://orcid.org/0000-0003-4915-3246>
 Jessie C. Runnoe  <https://orcid.org/0000-0001-8557-2822>
 Shashwat C. Sardesai  <https://orcid.org/0009-0006-5476-3603>
 Ann Schmiedekamp  <https://orcid.org/0000-0003-4391-936X>
 Carl Schmiedekamp  <https://orcid.org/0000-0002-1283-2184>
 Kai Schmitz  <https://orcid.org/0000-0003-2807-6472>
 Brent J. Shapiro-Albert  <https://orcid.org/0000-0002-7283-1124>
 Xavier Siemens  <https://orcid.org/0000-0002-7778-2990>
 Joseph Simon  <https://orcid.org/0000-0003-1407-6607>
 Magdalena S. Siwek  <https://orcid.org/0000-0002-1530-9778>
 Sophia V. Sosa Fiscella  <https://orcid.org/0000-0002-5176-2924>
 Ingrid H. Stairs  <https://orcid.org/0000-0001-9784-8670>
 Daniel R. Stinebring  <https://orcid.org/0000-0002-1797-3277>
 Kevin Stovall  <https://orcid.org/0000-0002-7261-594X>
 Abhimanyu Susobhanan  <https://orcid.org/0000-0002-2820-0931>
 Joseph K. Swiggum  <https://orcid.org/0000-0002-1075-3837>
 Stephen R. Taylor  <https://orcid.org/0000-0003-0264-1453>
 Jacob E. Turner  <https://orcid.org/0000-0002-2451-7288>
 Caner Unal  <https://orcid.org/0000-0001-8800-0192>
 Michele Vallisneri  <https://orcid.org/0000-0002-4162-0033>
 Sarah J. Vigeland  <https://orcid.org/0000-0003-4700-9072>
 Haley M. Wahl  <https://orcid.org/0000-0001-9678-0299>
 London Willson  <https://orcid.org/0009-0001-4123-3957>
 Caitlin A. Witt  <https://orcid.org/0000-0002-6020-9274>
 David Wright  <https://orcid.org/0000-0003-1562-4679>
 Olivia Young  <https://orcid.org/0000-0002-0883-0688>

References

- Afzal, A., Agazie, G., Anumalapudi, A., et al. 2023, *ApJL*, 951, L11
 Agazie, G., Alam, M. F., Anumalapudi, A., et al. 2023a, *ApJL*, 951, L9
 Agazie, G., Antoniadis, J., Anumalapudi, A., et al. 2024a, *ApJ*, 966, 105
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023b, *ApJL*, 952, L37
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023c, *ApJL*, 951, L8
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023d, *ApJL*, 951, L10
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023e, *ApJL*, 951, L50
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023f, *ApJL*, 956, L3
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2024b, arXiv:2408.10166
 Arzoumanian, Z., Baker, P. T., Blecha, L., et al. 2023, *ApJL*, 951, L28
 Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2020a, *ApJL*, 905, L34
 Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2020b, *ApJ*, 900, 102
 Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
 Bécsy, B., Cornish, N. J., & Kelley, L. Z. 2022, *ApJ*, 941, 119
 Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, 287, 307
 Casey-Clyde, J. A., Mingarelli, C. M. F., Greene, J. E., et al. 2022, *ApJ*, 924, 93
 Casey-Clyde, J. A., Mingarelli, C. M. F., Greene, J. E., et al. 2024, arXiv:2405.19406

- Chalumeau, A., Babak, S., Petiteau, A., et al. 2022, *MNRAS*, **509**, 5538
- Chen, S., Sesana, A., & Conselice, C. J. 2019, *MNRAS*, **488**, 401
- Cocchiararo, F., Franchini, A., Lupi, A., & Sesana, A. 2024, *A&A*, **691**, A250
- Cornish, N. J., & van Haasteren, R. 2014, arXiv:1406.4511
- Dickey, J. M. 1971, *Ann. Math. Statist.*, **42**, 204
- D’Orazio, D. J., Haiman, Z., & MacFadyen, A. 2013, *MNRAS*, **436**, 2997
- EPTA CollaborationInPTA Collaboration, Antoniadis, J., et al. 2023, *A&A*, **678**, A50
- EPTA CollaborationInPTA Collaboration, Antoniadis, J., et al. 2024, *A&A*, **685**, A94
- Falxa, M., Babak, S., Baker, P. T., et al. 2023, *MNRAS*, **521**, 5077
- Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2014, *ApJ*, **783**, 134
- Foster, R. S., & Backer, D. C. 1990, *ApJ*, **361**, 300
- Gair, J., Romano, J. D., Taylor, S., & Mingarelli, C. M. F. 2014, *PhRvD*, **90**, 082001
- Gardiner, E. C., Kelley, L. Z., Lemke, A.-M., & Mitridate, A. 2024, *ApJ*, **965**, 164
- Géron, A. 2017, *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (1st ed.; Sebastopol, CA: O’Reilly & Associates)
- Goncharov, B., Sardana, S., Sesana, A., et al. 2024, arXiv:2409.03627
- Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, *MNRAS*, **453**, 1562
- Grishchuk, L. P. 2005, *PhyU*, **48**, 1235
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- Hogg, D. W. 1999, arXiv:astro-ph/9905116
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jaffe, A. H., & Backer, D. C. 2003, *ApJ*, **583**, 616
- Kass, R. E., & Raftery, A. E. 1995, *JASA*, **90**, 773
- Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2017, *MNRAS*, **471**, 4508
- Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. 2016, *Jupyter Notebooks—A Publishing Format for Reproducible Computational Workflows* (Amsterdam: IOS Press), **87**
- Kormendy, J., & Richstone, D. 1995, *ARA&A*, **33**, 581
- Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, *JOSS*, **4**, 1143
- Lamb, W. G., & Taylor, S. R. 2024, *ApJL*, **971**, L10
- Lamb, W. G., Taylor, S. R., & van Haasteren, R. 2023, *PhRvD*, **108**, 103019
- Larsen, B., Mingarelli, C. M. F., Hazboun, J. S., et al. 2024, *ApJ*, **972**, 49
- Lasky, P. D., Mingarelli, C. M. F., Smith, T. L., et al. 2016, *PhRvX*, **6**, 011035
- Lentati, L., Shannon, R. M., Coles, W. A., et al. 2016, *MNRAS*, **458**, 2161
- Marconi, A., Risaliti, G., Gilli, R., et al. 2004, *MNRAS*, **351**, 169
- McKinney, W. 2010, *Proc. 9th Python in Science Conf. (SciPy 2010)* (Austin, TX) ed. S. van der Walt & J. Millman, 56
- Miles, M. T., Shannon, R. M., Bailes, M., et al. 2023, *MNRAS*, **519**, 3976
- Mingarelli, C. 2017, *Chiaramingarelli/Nanohertz_Gws: First Release! v1.0*, Zenodo, doi:10.5281/zenodo.838712
- Mingarelli, C. M. F., Sidery, T., Mandel, I., & Vecchio, A. 2013, *PhRvD*, **88**, 062005
- Miranda, R., Muñoz, D. J., & Lai, D. 2017, *MNRAS*, **466**, 1170
- Muñoz, D. J., Lai, D., Kratter, K., & Miranda, R. 2020, *ApJ*, **889**, 114
- Parzen, E. 1962, *Ann. Math. Statist.*, **33**, 1065
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *Journal of Machine Learning Research*, **12**, 2825
- Perera, B. B. P., DeCesar, M. E., Demorest, P. B., et al. 2019, *MNRAS*, **490**, 4666
- Peters, P. C., & Mathews, J. 1963, *PhRv*, **131**, 435
- Phinney, E. S. 2001, arXiv:astro-ph/0108028
- Pol, N., Taylor, S. R., & Romano, J. D. 2022, *ApJ*, **940**, 173
- Rajagopal, M., & Romani, R. W. 1995, *ApJ*, **446**, 543
- Ravi, V., Wyithe, J. S. B., Hobbs, G., et al. 2012, *ApJ*, **761**, 84
- Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, *ApJL*, **951**, L6
- Roebber, E., Holder, G., Holz, D. E., & Warren, M. 2016, *ApJ*, **819**, 163
- Rosado, P. A., Sesana, A., & Gair, J. 2015, *MNRAS*, **451**, 2417
- Rosenblatt, M. 1956, *Ann. Math. Statist.*, **27**, 832
- Sampson, L., Cornish, N. J., & McWilliams, S. T. 2015, *PhRvD*, **91**, 084055
- Sato-Polito, G., & Kamionkowski, M. 2024, *PhRvD*, **109**, 123544
- Sato-Polito, G., Zaldarriaga, M., & Quataert, E. 2023, arXiv:2312.06756
- Schwarz, G. 1978, *AnSta*, **6**, 461
- Sesana, A., Vecchio, A., & Colacino, C. N. 2008, *MNRAS*, **390**, 192
- Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2009, *ApJ*, **690**, 20
- Siemens, X., Ellis, J., Jenet, F., & Romano, J. D. 2013, *CQGra*, **30**, 224015
- Taylor, S. R., & Gair, J. R. 2013, *PhRvD*, **88**, 084001
- Taylor, S. R., Mingarelli, C. M. F., Gair, J. R., et al. 2015, *PhRvL*, **115**, 041101
- Taylor, S. R., Simon, J., & Sampson, L. 2017, *PhRvL*, **118**, 181102
- Taylor, S. R., van Haasteren, R., & Sesana, A. 2020, *PhRvD*, **102**, 084039
- The pandas development team 2022, *Pandas-Dev/Pandas: Pandas v1.5.0*, Zenodo, doi:10.5281/zenodo.7093122
- Thorne, K. S. 1987, in *Three Hundred Years of Gravitation*, ed. S. W. Hawking & W. Israel (Cambridge: Cambridge Univ. Press), 330
- Verbiest, J. P. W., Lentati, L., Hobbs, G., et al. 2016, *MNRAS*, **458**, 1267
- Vika, M., Driver, S. P., Graham, A. W., & Liske, J. 2009, *MNRAS*, **400**, 1451
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Volonteri, M., Haardt, F., & Madau, P. 2003, *ApJ*, **582**, 559
- Waskom, M. 2021, *JOSS*, **6**, 3021
- Westernacher-Schneider, J. R., Zrake, J., MacFadyen, A., & Haiman, Z. 2022, *PhRvD*, **106**, 103010
- Wiecki, T., Salvatier, J., Vieira, R., et al. 2023, *Pymc-Devs/Pymc: v5.5.0*, Zenodo, doi:10.5281/zenodo.8021093
- Xin, C., Mingarelli, C. M. F., & Hazboun, J. S. 2021, *ApJ*, **915**, 97
- Xu, H., Chen, S., Guo, Y., et al. 2023, *RAA*, **23**, 075024