Finite Populations & Finite Time: The Non-Gaussianity of a Gravitational Wave Background

William G. Lamb,[1] Jeremy M. Wachter,[2] Andrea Mitridate,[3] Shashwat C. Sardesai,[4]
Bence Bécsy,[5] Emily L. Hagen,[2] Stephen R. Taylor,[1] and Luke Zoltan Kelley[6]

[1] *Department of Physics and Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA*
[2] *School of Sciences & Humanities, Wentworth Institute of Technology, Boston, MA 02115*
[3] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*
[4] *University of Wisconsin Milwaukee, Milwaukee, WI 53211*
[5] *Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
[6] *Astrophysics Working Group, NANOGrav Collaboration, Berkeley, CA, USA*

(Dated: November 14, 2025)

Strong evidence for an isotropic, Gaussian gravitational wave background (GWB) has been found by multiple pulsar timing arrays (PTAs). The GWB is expected to be sourced by a finite population of supermassive black hole binaries (SMBHBs) emitting in the PTA sensitivity band, and astrophysical inference of PTA data sets suggests a GWB signal that is at the higher end of GWB spectral amplitude estimates. However, current inference analyses make simplifying assumptions, such as modeling the GWB as Gaussian, assuming that all SMBHBs only emit at frequencies that are integer multiples of the total observing time, and ignoring the interference between the signals of different SMBHBs. In this paper, we build analytical and numerical models of an astrophysical GWB without the above approximations, and compare the statistical properties of its induced PTA signal to those of a signal produced by a Gaussian GWB. We show that interference effects introduce non-Gaussianities in the PTA signal, which are currently unmodeled in PTA analyses.

I. INTRODUCTION

Multiple pulsar timing arrays (PTAs) have found strong evidence for a nanohertz-frequency gravitational-wave background (GWB) [1–5], with significant effort being coordinated by individual PTAs and collaboratively as the International Pulsar Timing Array [6] to determine its source(s). However, analyses to determine its (dominant) source have so far been inconclusive [7–15].

The two families of sources that could produce a GWB at nanohertz frequencies are cosmological and astrophysical. The cosmological description results from early universe phenomena (see, for example, Refs. [16, 17]), and a detection would result in the first direct probe of the pre-recombination universe. Alternatively, the astrophysical description identifies the source as a finite population of supermassive black hole binaries (SMBHBs) inspiralling from gravitational wave (GW) radiation [18–22]. The GWs emitted by unresolved binaries contribute to the GWB, while a number of resolved sources will result in continuous GW signals that are louder than the nanohertz GWB and may be resolvable [22–26].

We can characterize the GWB by expanding the metric perturbation it generates in plane waves:

$$h_{ij}(t,\vec{x}) = \sum_A \int df \int d\hat{\Omega}\, \tilde{h}_A(f,\hat{\Omega}) e^{i2\pi f(t-\hat{\Omega}\cdot\vec{x})} e^{A}_{ij}(\hat{\Omega})\,, \tag{1}$$

where f is the GW frequency, $\hat{\Omega}$ the direction of propagation of the plane waves, $A = +, \times$ labels the two GW polarizations, e^A_{ij} are the GW polarization tensors, and $\tilde{h}_A(f,\hat{\Omega})$ are two complex functions (one for each GW polarization) satisfying $\tilde{h}^*_A(f,\hat{\Omega}) = \tilde{h}_A(-f,\hat{\Omega})$. A GWB is characterized by the fact that the functions $\tilde{h}_A(f,\hat{\Omega})$ can be treated as random variables, whose distribution is set by the properties of the GWB source. Typically, several assumptions are made when describing the statistics of the $\tilde{h}^*_A(f,\hat{\Omega})$ functions:

(i) Gaussianity—since the GWB is expected to arise as a central-limit-theorem process, it is common to model it as a Gaussian process. Specifically, the $h^*_A(f,\hat{\Omega})$ functions are assumed to be zero-mean Gaussian random variables.

(ii) Isotropy—the statistical properties of $\tilde{h}^*_A(f,\hat{\Omega})$ are assumed to be the same in all sky directions.

(iii) Unpolarized—the statistical properties of both GW polarizations are assumed to be the same.

(iv) Stationary—the properties of the GWB are assumed to be time-independent.

With all these assumptions, the two-point functions for the $\tilde{h}^*_A(f,\hat{\Omega})$ functions reads:

$$\left\langle \tilde{h}^*_A(f,\hat{\Omega})\tilde{h}_A(f',\hat{\Omega}') \right\rangle = \delta_{AA'}\delta(f-f')\delta^2(\Omega,\Omega')\frac{S_h(f)}{16\pi}, \tag{2}$$

where $S_h(f)$ is the GW strain power spectral density (PSD).

These assumptions are valid for both cosmologically sourced GWBs [16, 27] and for an astrophysical GWB if there is a large number of unresolvable sources [28, 29]. However, for a low number of sources, GW shot noise will drive non-Gaussianity [28–32]. GW shot noise will also result in anisotropies in addition to clustering of sources by large-scale structure [31, 33], while inclined binaries will generate polarization in the GWB [31].

Comparison of current GWB spectrum measurements with predictions from SMBHB population models suggests a possible underestimate of the number density or the masses of sources [7, 11, 34–36], or mismodeled noise in PTA analyses [37–39]. Model selection studies prefer a Gaussian cosmological source for the GWB over the fiducial power-law model in the NANOGrav 15yr data set [8]; however, as we will show in this paper, the power-law does not properly model the astrophysical GWB, which hinders any conclusions regarding the source of the GWB.

Deviations from the power-law spectrum are driven by Poisson variance of the finite population [26, 30, 40], and mismodeling the GWB can result in biased parameter estimation. For example, if there are resolvable continuous waves above a Gaussian power-law GWB, failure to model the continuous wave will result in over-estimation of the power-law amplitude and under-estimation of its spectral index [41].

Instead, we can model a superposition of continuous sources above the power-law [41, 42]. Unfortunately, this is difficult to analyze with real data because of its slow likelihood evaluation time, unknown number of resolvable continuous wave sources, and very large parameter space. Instead, simpler spectral models that are computationally tractable are often used. For example, Poisson variance can be modeled as 'excursions' from a power-law, either by using data-driven weights to model deviations [1, 43] or by making no assumptions about the shape of the spectrum [14, 44]. These spectra can be useful to point towards any continuous wave sources; however, the intrinsic Gaussianity of these spectra will result in mismodeling and biased parameter estimation [45].

More involved astrophysical inference techniques compare ensemble distributions of GWB spectra, obtained through density estimation of many ensemble realizations of semi-analytic SMBHB population models, to posteriors from Bayesian analyses such as Markov Chain Monte Carlo (MCMC) [7, 11, 46–49]. Such analyses introduce approximations, such as binary orbit circularity, and GW frequencies of sources observed at harmonics of $1/T$, where T is the observation time of the pulsars, which are not accurate. The same assumptions are also intrinsic to the weighted-Poisson GWB spectral model introduced by Sato-Polito and Zaldarriaga [34] (see also [30]). Additionally, a finite observation time introduces frequency covariances that are currently not included in astrophysical inference [37].

Xue *et al.* [45] presents the distribution of the squared magnitude of the pulsar redshift Fourier coefficient in terms of a semi-analytic model (SAM) of circular SMBHBs. This work shows how a GWB spectrum distribution can be numerically computed by Fourier transforming the model's characteristic function. However, they approximate the distribution of observed GW frequencies as a step-function, which does not reflect the predicted power-law distribution for a GW-hardened binary [19, 50]. More recently, Falxa and Sesana [51] introduces an agnostic Gaussian mixture model to represent the distribution of Fourier coefficients from a PTA time-series to efficiently model non-Gaussianities. However, it is unclear how this model can be used for downstream astrophysical parameter estimation.

Additionally, these models do not model the correlated response of two pulsars to a discrete GW source and its associated uncertainty over realizations of sources and pulsars. For an isotropic, unpolarized, and Gaussian GWB, the mean correlated response between two pulsars is described by the Hellings & Downs (HD) curve [52]. However, the correlated response between two pulsars to a GW point source will differ from the mean HD curve in individual realizations of the astrophysical GWB [28] (see also [29, 53, 54]). Additionally, the finite observation time of an experiment will result in confusion between discrete sources, which will vary over realizations of the GWB. This is an example of *cosmic variance* [28].

Previous analysis on the astrophysical GWB focuses on the timing-residual power spectral density $S_{\delta t}(f)$ (or quantities related to it, such as the squared characteristic strain $h_c^2(f)$). However, these quantities implicitly require a Gaussian prior on the Fourier coefficient of the GWB component of the timing residual, with non-Gaussianities introduced as "corrections" to this model (see e.g. [45]). In this paper, we investigate the accuracy of the Gaussian prior by extending the simple model of Allen [28] to Fourier coefficients of a time-series measured by a pulsar in response to a GWB from a finite population of SMBHBs. We show that non-Gaussianity in the Fourier coefficient is only expressed in the fourth moment, from which we develop a kurtosis metric to measure said non-Gaussianity and show how Poisson variance, HD variance, and finite observation times affect the moments found in Lamb and Taylor [30]. We also analyze the distribution of the argument of the complex Fourier coefficient and show how it can be used to investigate non-Gaussianity. Finally, we conclude by discussing how our work can be utilized for faster and more accurate simulation and analysis of GWBs.

Throughout this paper, we will present our model in the context of PTAs and their sources. However, our framework is applicable across GW detectors, with differences arising from different antenna responses and sources, resulting in different prefactors to our analytical results.

For the busy reader, here is an overview of our results:

- We derive the Fourier coefficients of the timing residuals from a pulsar p responding to a GWB from a finite population of circular SMBHBs that are hardened by gravitational wave emission only.

- We derive the first four moments of the Fourier coefficient for the discrete-source GWB, and discuss different origins of those non-Gaussianities.

- We investigate the argument of the (complex) Fourier coefficients. The tangent of the argument for a Gaussian, isotropic, and unpolarized GWB has a standard Cauchy distribution, and so looking at the tangent of the argument of the Fourier coefficients provides a complementary probe of non-Gaussianity.

The paper is laid out as follows. In section II, we introduce the GWB timing-residual model from which we derive Fourier coefficients. We then find the moments of the coefficients to open section III before looking at the limiting cases of one GW source (section III A) and many GW sources (section III B), and close this section by considering the distribution of the argument of the Fourier coefficients in section III C. We compare our analytical results to numerical simulations in section IV for both a toy model (section IV A) and an astrophysical model (section IV B). We conclude in section V.

Further details on ensemble averaging, complex random variables, window functions, and semi-analytical models can be found in the Appendices. Throughout, we will use natural units, $c = G = 1$.

II. OBSERVING THE ASTROPHYSICAL GRAVITATIONAL WAVE BACKGROUND

We will model the astrophysical GWB as being formed by a superposition of GWs from a finite number N of sources with unique frequencies f_j, propagation directions $\hat{\Omega}_j$, and phases ϕ_j, with all astrophysical parameters in a vector $\vec{\theta}_j$, such as the redshift of the source z_j and its chirp mass $\mathcal{M}_j$. We will also consider how the GWB interacts with a detector – for concrete purposes, a PTA – over a finite observation time T. This setup lets us identify two major sources of non-Gaussianity in measurements of the GWB: *Poisson variance* [30], due to the fact that N is a random (Poisson) variable, and *shot noise* due, to the fact that the GWB is produced by a finite number of interfering GW sources.

Assuming that all GW sources are far enough from the Earth-pulsar system, we will describe their metric perturbations as plane waves. These plane waves are uniquely characterized by two functions of a single variable, one for each GW polarization, which we call h^+ and $h^\times$:

$$h_j^+(t) = \mathcal{A}_j \cos\left(2\pi f_j t + \phi_j\right) \,, \tag{3a}$$

$$h_j^\times(t) = \mathcal{A}_j \sin\left(2\pi f_j t + \phi_j\right) \,, \tag{3b}$$

where the amplitude of the metric perturbation, $\mathcal{A}(f_j, \vec{\theta}_j) \equiv \mathcal{A}_j$, depends on the astrophysical parameters and frequency of its source, and we have assumed that the SMBHB orbit is circular and face-on to the Earth line of sight (from which follows that the relative phase between the two polarizations is $\pi/2$). The rest-frame frequency f_j is related to the observed GW frequency $f_{j,\mathrm{obs}}$ by $f_j = (1 + z_j) f_{j,\mathrm{obs}}$, where z_j is the redshift between the observer and the source. The metric perturbation produced by the entire population of sources at time t and location $\vec{x}$ is then given by:

$$\begin{aligned} h_{ab}(t, \vec{x}) &= \sum_{j=1}^{N} \sum_{A} h_j^A(t - \hat{\Omega}_j \cdot \vec{x}) \epsilon_{ab}^A(\hat{\Omega}_j) \\ &= \sum_{j=1}^{N} \frac{\mathcal{A}_j}{2} e^{i[2\pi f_j (t - \hat{\Omega}_j \cdot \vec{x}) + \phi_j]} \epsilon_{ab}(\Omega_j) + \text{c.c.} \,, \end{aligned} \tag{4}$$

where c.c. indicates the complex conjugate, and we defined the complex polarization tensor $\epsilon_{ab} \equiv \epsilon_{ab}^+ - i\epsilon_{ab}^\times$.

The timing residual for the p-th pulsar induced by this metric perturbation is given by:

$$\begin{aligned} \delta t_p(t) &= \frac{\hat{x}_p^a \hat{x}_p^b}{2} \int_{t-L_p}^{t} dt' h_{ab}[t', (t-t')\hat{x}_p] \\ &= \sum_j \frac{\mathcal{A}_j}{4\pi i f_j} R_{p,j}\, e^{i(2\pi f_j t + \phi_j)} + \text{c.c.} \,, \end{aligned} \tag{5}$$

where $\hat{x}_p$ is the unit vector pointing from Earth to the p-th pulsar, and L_p is the distance from the Earth to the pulsar. In the above equation, we have also defined a complex antenna response function: $R_{p,j} \equiv R_p^+(f_j, \hat{\Omega}_j) - iR_p^\times(f_j, \hat{\Omega}_j)$, where the antenna response function for pulsar p is given by:

$$R_p^A(f, \hat{\Omega}) = \left[1 - e^{-2\pi i f L_p (1 + \hat{\Omega}\cdot\hat{x}_p)}\right] F_p^A(\hat{\Omega}) \,, \tag{6a}$$

$$F_p^A(\hat{\Omega}) = \frac{1}{2} \frac{\hat{x}_p^a \hat{x}_p^b}{1 + \hat{\Omega}\cdot\hat{x}_p} \epsilon_{ab}^A(\hat{\Omega}). \tag{6b}$$

It is common in PTA analyses to decompose the timing residual on a discrete, *two-sided* Fourier basis with sampling frequencies $f_i = i/T$, where T is the total observation time:

$$\delta t_p(t) = \sum_i \tilde{a}_i^p e^{2\pi i f_i t} \,, \tag{7}$$

where the summation is usually truncated at some frequency between the largest GWB-dominated frequency and the inverse of the observational cadence. Then, for our population model, the Fourier coefficient for the i-th frequency mode is

$$\tilde{a}_i^p = \sum_j \frac{\mathcal{A}_j}{4\pi i f_j} \left[w_j^- e^{i\phi_j} R_{p,j} - w_j^+ e^{-i\phi_j} R_{p,j}^* \right] \,, \tag{8}$$

where we have introduced the window function, $w_j^\pm \equiv w(f_i \pm f_j)$. For an observation carried out for a finite time T, the window function takes the form $w_j^\pm = \mathrm{sinc}[\pi(f_i \pm f_j)T]$. However, in a realistic PTA observation, the window function will be modified by the marginalization over the timing model parameters. In this work, we will derive analytical results without specifying the form of the window function, but only assuming that they are real functions. Therefore, our final results can be applied to different forms of the window function.

In PTA analyses, the Fourier coefficients of the timing residuals are typically assumed to be a Gaussian-distributed complex random variable, reflecting the idea that the GWB itself is Gaussian. In the next sections, we will find expressions for the higher moments of $\tilde{a}_i^p$ to test this Gaussian assumption.

III. MOMENTS OF THE FINITE-POPULATION GRAVITATIONAL WAVE BACKGROUND

In this section, we derive expressions for the moments of the timing-residual Fourier coefficients, $\tilde{a}_i^p$, generated by the finite population model discussed in the previous section.

We start by computing $\langle \tilde{a}_i^p \rangle$, where $\langle \cdot \rangle$ denote the ensemble average. In our model, the ensemble average consists of averaging over phases, ϕ_j, source positions, $\hat{\Omega}_j$, and the number of sources. We start by computing the average over phases, which is defined as:

$$\langle Q \rangle_\phi \equiv \int_0^{2\pi} \frac{d\phi_1}{2\pi} \ldots \int_0^{2\pi} \frac{d\phi_N}{2\pi}\, Q(\phi_1, \ldots, \phi_N)\,. \tag{9}$$

Using the standard result, $\langle e^{i\phi} \rangle_\phi = 0$, we immediately find that

$$\langle \tilde{a}_i^p \rangle_\phi = 0\,, \tag{10}$$

In fact, all odd moments of $\tilde{a}_i^p$ are zero mean because they have an odd number of $e^{i\phi_j}$ components. Since $\langle \tilde{a}_i^p \rangle_\phi$ does not depend on the source positions or properties, the phase average coincides with the ensemble average: $\langle \tilde{a}_i^p \rangle_\phi = \langle \tilde{a}_i^p \rangle$.

We now move our attention to the second moment of timing residuals' Fourier components. We start by computing the cross-pulsar product for pulsars p and q,

$$\begin{aligned} \tilde{a}_i^{*p} \tilde{a}_i^q = \sum_j \frac{\mathcal{A}_j^2}{16\pi^2 f_j^2} \big[w_j^{--} R_{pj}^* R_{qj} + w_j^{++} R_{pj} R_{qj}^* \\ - w_j^{-+} \left(R_{pj} R_{qj} e^{2i\phi_j} + \text{c.c.} \right) \big] \\ + \sum_{j \neq k} \frac{\mathcal{A}_j \mathcal{A}_k}{16\pi^2 f_j f_k} \left(w_j^- R_{pj}^* e^{-i\phi_j} - w_j^+ R_{pj} e^{i\phi_j} \right) \\ \times \left(w_k^- R_{qk} e^{i\phi_k} - w_k^+ R_{qk}^* e^{-i\phi_k} \right), \end{aligned} \tag{11}$$

where j, k are indices over the sources, and we have defined $w_j^{\pm\pm} = w(f_j \pm f_i) w(f_j \pm f_i)$. Taking the auto-pulsar product as the case where $p = q$, we find the phase average as

$$\left\langle |\tilde{a}_i^p|^2 \right\rangle_\phi = \sum_j \frac{\mathcal{A}_j^2}{16\pi^2 f_j^2} |R_{pj}|^2 \left(w_j^{--} + w_j^{++} \right)\,. \tag{12}$$

The above phase average still depends on the source positions via the $|R_{pj}|^2$ term. Therefore, to find the ensemble average, we still have to perform an average over all possible source directions. For PTAs, the sky average of the product of two response functions is related to the well-known Hellings-&-Downs curve, $\Gamma(\gamma)$, [52]:

$$\Gamma_{pq} \equiv \Gamma(\gamma_{pq}) = \frac{3}{2} \left\langle R_{pj}^* R_{qj} \right\rangle_{\hat{\Omega}}\,, \tag{13}$$

where γ_{pq} is the angular separation between pulsar p and pulsar q, and the normalization factor of $3/2$ is chosen such that the auto-correlation $\Gamma_{pp} = 1$. From this it follows that:

$$\left\langle |\tilde{a}_i^p|^2 \right\rangle_{\phi,\hat{\Omega}} = \sum_j \frac{\mathcal{A}_j^2}{24\pi^2 f_j^2} \left(w_j^{--} + w_j^{++} \right)\,. \tag{14}$$

While we have only reported the second moment for the $p = q$ case, the cross-pulsar second moment is very similar, and it can be obtained by multiplying the right-hand side of Equation 14 by Γ_{pq}.

Finally, we will average over fluctuations in the number of sources. To do this, we divide the parameter space of the SMBHB model into equal-size bins, and replace the sum over individual sources in Equation 14 with a sum over these bins:

$$\sum_{j=1}^N \frac{\mathcal{A}_j^2}{f_j^2} \quad \Rightarrow \quad \sum_J N_J \frac{\mathcal{A}_J^2}{f_J^2}\,, \tag{15}$$

where N_J is the number of sources in the J-th bin, and A_J, f_J are the amplitude and frequency of the sources in that bin. The number of sources in each bin is a Poisson random variable, with a mean value $\bar{N}_J$. This mean value of sources can be derived, for a given set of astrophysical parameters, using semi-analytical models (SAM) or numerical simulations (for more details on SAMs, see Appendix B). Therefore, after averaging over N_J, we get:

$$\left\langle |\tilde{a}_i^p|^2 \right\rangle_{\phi,\hat{\Omega},N} = \sum_J \bar{N}_J \frac{\mathcal{A}_J^2}{24\pi^2 f_J^2} \left(w_J^{--} + w_J^{++} \right)\,. \tag{16}$$

Equation 16 is equivalent to SAMs of GWBs used in astrophysical inference of PTA data [7, 40], with the addition of window functions from finite observation times.

Following an analogous procedure (i.e., averaging over phases, sky positions, and number of sources), we find that the four-point function is given by (for a detailed derivation, see Appendix A):

$$\left\langle|\tilde{a}_i^p|^4\right\rangle_{\phi,\hat{\Omega},N} = 2\left\langle|\tilde{a}_i^p|^2\right\rangle^2_{\phi,\hat{\Omega},N} + \sum_J \bar{N}_J \frac{\mathcal{A}_J^4}{24^2\pi^4 f_J^4}\left[\frac{7}{10}(w_J^{--}+w_J^{++})^2 + \frac{7}{5}w_J^{--}w_J^{++}\right] + 4\left(\sum_J \bar{N}_J \frac{\mathcal{A}_J^2}{24\pi^2 f_J^2} w_J^{-+}\right)^2 . \quad (17)$$

Equation 17 is one of the main results of this work, so let us discuss it in more detail. First, we immediately notice that $\tilde{a}_i^p$ is non-Gaussian. This non-Gaussianity, which arises from a combination of finite-source effects and Poisson fluctuations in the total number of GW sources, can be quantified by computing the kurtosis, $\kappa \equiv \langle|\tilde{a}_i^p|^4\rangle/\langle|\tilde{a}_i^p|^2\rangle$. For a circularly complex Gaussian random variable, $\kappa = 2$ (see section C for more details), while from Equation 17 we find that the excess kurtosis, $\bar{\kappa}$, is

$$\bar{\kappa} \simeq \frac{(7/10)\sum_J \bar{N}_J \mathcal{A}_J^4 w_J^{--}/f_J^4 + 4\left(\sum_J \bar{N}_J \mathcal{A}_J^2 w_J^{-+}/f_J^2\right)^2}{\left(\sum_J \bar{N}_J \mathcal{A}_J^2 w_J^{--}/f_J^2\right)^2} \quad (18)$$

where, assuming that $w_J^{++} \ll 1$, we have ignored the w_J^{++} terms in the square bracket of Equation 17. This is a justified assumption if $w_J^{\pm} = \mathrm{sinc}[(f_i - f_J)T]$, but it might not be true for more general window functions. Often, in astrophysical inference analyses, the sources are assumed to be emitting only at frequencies that are integer multiples of the observing time. In this case, $w^{--} = 1$ while $w^{++,+-} = 0$, and the excess kurtosis reduces to:

$$\bar{\kappa} \simeq \frac{(7/10)\sum_J \bar{N}_J \mathcal{A}_J^4 w_J^{--}/f_J^4}{\left(\sum_J \bar{N}_J \mathcal{A}_J^2 w_J^{--}/f_J^2\right)^2} . \quad (19)$$

A. Single-source limit

It is also instructive to study the form of our results in two opposite limits: a single source and many sources. If a single source, with amplitude $\mathcal{A}$ and frequency f, dominates the GWB in a specific frequency bin, the two- and four-point functions take the form

$$\left\langle|\tilde{a}_i^p|^2\right\rangle = \frac{\mathcal{A}^2}{24\pi^2 f^2} w^{--} , \quad (20)$$

$$\left\langle|\tilde{a}_i^p|^4\right\rangle \simeq 2\left\langle|\tilde{a}_i^p|^2\right\rangle + \frac{7}{10}\left\langle|\tilde{a}_i^p|^2\right\rangle , \quad (21)$$

where we have ignored all terms containing a w^{++} or w^{+-}, since we are working under the assumption that they are much smaller compared to w^{--}. So it is easy to see that, in this limit, the excess kurtosis is $\bar{\kappa} \simeq 7/10$; a result that we will explicitly check with the numerical simulations discussed in section IV. This value for the excess kurtosis is derived under the assumption of circular, face-on SMBHBs. However, generalizing the SMBHB model to allow eccentricity and inclined binaries will result in larger coefficients in each term of the moments of $\tilde{a}_i^p$, and hence increase the factor of 7/10 in the equivalent excess kurtosis equation [29]. Computing this factor for a generalized SMBHB model is left as work for a future publication.

B. Many-sources limit

In the opposite limit, when many sources contribute to the GWB such that $N_J \gg 1$ for the majority of the parameter-space bin, the four-point function reads

$$\left\langle|\tilde{a}_i^p|^4\right\rangle \simeq 2\left\langle|\tilde{a}_i^p|^2\right\rangle + 4\left(\sum_J \bar{N}_J \frac{\mathcal{A}_J^2}{24\pi^2 f_J^2} w^{-+}\right)^2 , \quad (22)$$

where we have ignored the second term on the right-hand side of Equation 17 since, for $N_J \gg 1$,

$$\sum_J \bar{N}_J \mathcal{A}_J^4 \ll \left(\sum_J \bar{N}_J \mathcal{A}_J^2\right)^2 . \quad (23)$$

From this equation, we can see that, even in the many-sources limit, there is a residual non-Gaussian term. As we will see in the next section, this residual non-Gaussianity derives from the covariance between the real and imaginary parts of $\tilde{a}_i^p$ introduced by the finite observation time. From Figure 8, we see that the 'cross-window' term $w^{+-} \ll w^{--}$ and can usually be ignored. See Appendix D for further discussion.

Very often, the large number of expected sources is the justification for modeling gravitational wave backgrounds as zero-mean Gaussian processes, fully described by the two-point function of the $\tilde{h}_A(f,\hat{\Omega})$ coefficients in Equation 2. The second and fourth moments of the Fourier coefficients in this model are

$$\left\langle|\tilde{a}_i^p|^2\right\rangle_{\hat{\Omega}} = \int_{-\infty}^{\infty} \frac{S_h(f)}{24\pi^2 f^2}\left(w^{--} + w^{++}\right) df , \quad (24a)$$

$$\left\langle|\tilde{a}_i^p|^4\right\rangle_{\hat{\Omega}} = 2\left\langle|a_i^p|^2\right\rangle^2_{\hat{\Omega}} + \left(\int_{-\infty}^{\infty} \frac{S_h(f)}{24\pi^2 f^2} w^{-+} df\right)^2 . \quad (24b)$$

Note that the second term in Equation 24b is comparable to the second term in Equation 22 – it is a covariance term between the real and imaginary components of a Gaussian random variable. The difference of a factor of 4 is a result of our integration limits and definition of the (two-sided) GWB strain PSD $S_h(f)$. The integration limits of Equation 22 are between 0 and ∞ because discrete sources have positive frequencies. If we change the integration limits in the second term in Equation 24b such that the PSD is one-sided, it would introduce a factor of 4. Hence, the many-source limit of the astrophysical GWB can be approximated as a Gaussian GWB.

C. Argument of the finite-population Fourier coefficient

Finally, we analyze the argument of the Fourier coefficient using Equation C5. Taking the real and imaginary components of Equation 8, using trigonometric identities, assume that $f_j L_p \gg 1$ such that $R_{pj} \approx F_{pj}$, and substituting $\phi'_j = \phi_j - \pi/2$, we find that

$$\mathrm{Re}\,[\tilde{a}_i^p] = \sum_j \frac{\mathcal{A}_j}{4\pi f_j}\left(w_j^- + w_j^+\right)\left[F_{pj}^+ \cos\phi'_j + F_{pj}^\times \cos\left(\phi'_j - \frac{\pi}{2}\right)\right] \tag{25}$$

$$\mathrm{Im}\,[\tilde{a}_i^p] = \sum_j \frac{\mathcal{A}_j}{4\pi f_j}\left(w_j^- - w_j^+\right)\left[F_{pj}^+ \sin\phi'_j + F_{pj}^\times \sin\left(\phi'_j - \frac{\pi}{2}\right)\right] \tag{26}$$

Hence, the argument φ_i^p *of a single realization* of the Fourier coefficient measured by pulsar p at observation frequency f_i is,

$$\tan\varphi_i^p = \tag{27}$$

$$\frac{\sum_j \frac{\mathcal{A}_j}{f_j}\left(w_j^- - w_j^+\right)\left[F_{pj}^+ \sin\phi'_j + F_{pj}^\times \sin\left(\phi'_j - \frac{\pi}{2}\right)\right]}{\sum_j \frac{\mathcal{A}_j}{f_j}\left(w_j^- + w_j^+\right)\left[F_{pj}^+ \cos\phi'_j + F_{pj}^\times \cos\left(\phi'_j - \frac{\pi}{2}\right)\right]}. \tag{28}$$

Without loss of generality, we locate the pulsar p parallel to the z-plane such that $F^\times(f,\hat{\Omega}) = 0$. In the case where $N = 1$, then,

$$\tan\varphi_i^p = \frac{\left(w_j^- - w_j^+\right)}{\left(w_j^- + w_j^+\right)}\tan\phi'_j, \tag{29}$$

Therefore, if $\tan\phi'_j$ follows a Cauchy distribution with location and scale parameters $x_{0,j}$ and γ_j, then Equation 29 is a Cauchy distribution with location $x'_{0,j} = \beta_j x_{0,j}$ and scale $\gamma'_j = \beta_j\gamma_j$, where

$$\beta_j = \frac{w_j^- + w_j^+}{w_j^- - w_j^+}. \tag{30}$$

Future studies can introduce eccentric single sources, in which case the equal-amplitude assumption of Equation 3 is invalid and the distribution of $\tan\varphi_i^p$ for a single source may no longer be Cauchy. Regardless, in the large source limit, the distribution of $\tan\varphi_i^p$ will tend towards Cauchy. More details on Cauchy random variables may be found in Appendix E.

IV. COMPARING TO NUMERICAL RESULTS

In this section, we use numerical simulations to test the analytical results presented in section III. Specifically, we will use these simulations to estimate the moments of timing residuals Fourier coefficients of a single pulsar, which we will term the *auto-pulsar moments*, as well as the cross-moments of Fourier coefficients of two pulsars, which we will term *cross-pulsar moments.*

For cross-pulsars, the second moment is the ensemble average of the product of two Fourier components, $\langle\tilde{a}_i^{*p}\tilde{a}_i^q\rangle$, while the fourth moment is the absolute square of this product, $\langle|\tilde{a}_i^{*p}\tilde{a}_i^q|^2\rangle$. An important result in the large-source-number limit is that the cross-pulsar fourth moment tends towards twice the second moment of *either* pulsar individually: $\langle|\tilde{a}_i^{*p}\tilde{a}_i^q|^2\rangle \approx 2\langle|\tilde{a}_i^p|^2\rangle^2$.

A. Toy Model

We start by considering the simple model of SMBHB (first described in [28]) where we have a fixed number of sources, N, and the amplitude $\mathcal{A}_j$ of the j^{th} source is given by $\mathcal{A}_j = \mathcal{A}j^{-\frac{1}{3}}$ (where $\mathcal{A}$ is an arbitrary amplitude). Moreover, we assume that all sources emit at frequencies which are integer multiples of the inverse of the observing time, such that the window functions are $w_j^-, w_j^+ = 1, 0$. We then generate 10^5 realizations of this toy model, and for each of these realizations, we use Equation 8 to compute the timing residuals Fourier coefficients. We then use these coefficients to estimate the excess kurtosis.

In Figure 1, we show the excess kurtosis as a function of the number of sources obtained in this way. These results are compared with the analytical predictions obtained in Equation 19, which for this toy model takes the form

$$\bar{\kappa} = \frac{7}{10}\frac{\sum_j j^{-4/3}}{\left(\sum_j j^{-2/3}\right)^2}. \tag{31}$$

We can see that numerical and analytical results show good agreement, and for $N = 1$, the excess kurtosis is $7/10$, as predicted. We also see that as $N \to \infty$, then $\bar{\kappa} \to 0$. This is expected because as $N \to \infty$, the numerator of Equation 31 $\to \zeta\left(\frac{4}{3}\right)$, where ζ is the Riemann zeta function. This is finite. However, the denominator diverges, hence the excess kurtosis goes to zero as $N \to \infty$.[1]

Next, we consider the predictions made for how the second and fourth moments of Fourier coefficients are related. We test the case of a single SMBHB source, in Figure 2, which validates the frequency scaling of Equation 21 and its cross-pulsar generalization. As the base amplitude A is arbitrary, we have rescaled the axes so that the predicted value at the lowest frequency is equal to unity.

---

[1] It is tempting to state that the denominator of Equation 31 is $\zeta\left(\frac{2}{3}\right)^2$ for $N \to \infty$. However, this would not be correct because the Riemann zeta function $\zeta(s)$ is defined for $\mathrm{Re}[s] > 1$, with its *analytic continuation* used to extend the domain.

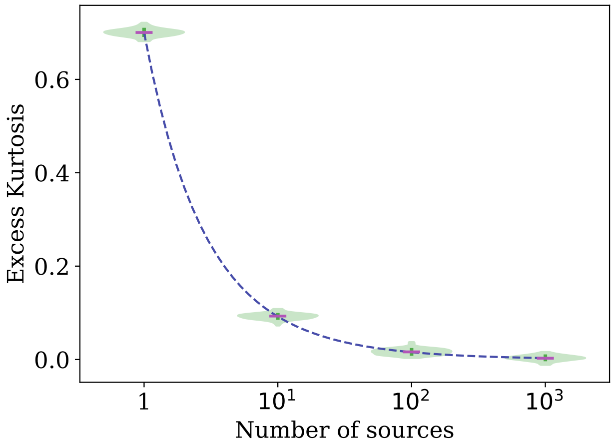

FIG. 1. Comparing excess kurtosis as a function of the number of sources numerically (green violin plots) and analytically (dashed blue curve). The vertical green bar indicates the 10% to 90% quantile range of the distribution of the excess kurtosis estimator across the pulsars in the array; the horizontal pink bar indicates the median. Numerical distributions are all of 100 samples of kurtosis generated from 10^5 realizations of an SMBHB population with $f_i = f_j = 2/T$ and a single pulsar.

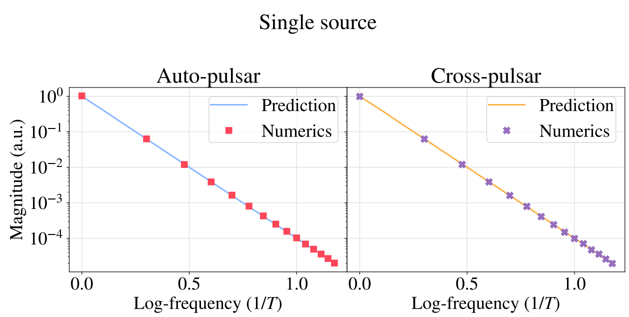

FIG. 2. With a single SMBHB source, fourth moments of the Fourier coefficients match the predictions of Equation 21 and its cross-pulsar generalization. The data shown here are generated from 10^7 realizations for frequencies from $1/T$ to $15/T$. The vertical axes have been rescaled so that the predicted value at $f = 1/T$ is 1 in arbitrary units in both cases.

We test the case of many sources (here, $N = 1000$) in Figure 3 to validate Equation 22 and the cross-pulsar extension, given in Equation A23. In this case, we have no analytic functions to compare to; instead, we compare the variance and the square of the mean directly. As before, the averaging is done over 10^7 realizations, and the normalization of the fourth moment at $f = 1/T$ to unity is done to remove the arbitrary base amplitude.

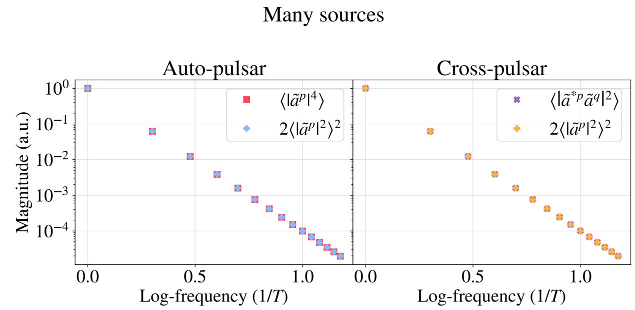

FIG. 3. With many sources, fourth moments of the Fourier coefficients match the predictions of Equation 22,Equation A23, respectively. The data shown here are generated from 10^7 realizations of an SMBHB population of 1000 sources for frequencies from $1/T$ to $15/T$. The vertical axes have been rescaled so that $\langle|\tilde{a}_i^p|^4\rangle = 1$ or $\langle|\tilde{a}_i^{*p}\tilde{a}_i^q|^2\rangle = 1$, as appropriate, in arbitrary units at $f = 1/T$.

B. Astrophysical Model

Let us now consider a more realistic model, which uses data generated for a total observing time of $T = 15$ yr and the 67 pulsars from the NANOGrav 15yr dataset [55], with the 10^4 loudest sources following the SMBHB number density model described in Middleton *et al.* [56] with a normalized merger rate of $\dot{n}_0 = 1.5 \cdot 10^{-3}$, a characteristic mass of $\mathcal{M}_* = 1.8 \cdot 10^8\, \mathcal{M}_\odot$, a mass spectral index of $\alpha = 0$, a characteristic redshift of $z_0 = 1.8$, and a redshift spectral index of $\beta = 2$ [30, 31]. We perform $11,000$ realizations and, for each of those, derive the timing residual Fourier coefficients for $f = n/T$, $n \in [1, 13]$. We additionally perform a simplified timing model subtraction, which analytically removes the linear and quadratic terms from the timing residuals.

Similarly to what we did for the toy model of the previous section, we start by studying the excess kurtosis, calculated for each pulsar and frequency. The distribution by frequency is found in Figure 4. The distributions all tend to be leptokurtic (i.e., more likely to produce outliers than a Gaussian distribution), with lower frequencies– except for the first one– agreeing fairly well with the expectation of $\bar{\kappa} = 0$, valid in the large source limit. For the first frequency, we see significant excess kurtosis, and there is a slight rise in excess kurtosis at higher frequencies, driven in part by a few outliers, e.g., a single significant outlier in the thirteenth frequency.

To understand the origin of this deviation from Gaussianity, it is useful to rewrite the kurtosis for a complex Gaussian variable, $Z = X + iY$, as

$$\kappa = \frac{3\left\langle X^2\right\rangle^2 + 3\left\langle Y^2\right\rangle^2 + 2\left\langle X^2\right\rangle\left\langle Y^2\right\rangle}{\left(\left\langle X^2\right\rangle + \left\langle Y^2\right\rangle\right)^2} + \frac{\bar{\kappa}_X\left\langle X^2\right\rangle^2 + \bar{\kappa}_Y\left\langle Y^2\right\rangle^2 + 4\left\langle XY\right\rangle^2}{\left(\left\langle X^2\right\rangle + \left\langle Y^2\right\rangle\right)^2}, \tag{32}$$

where $\bar{\kappa}_X, \bar{\kappa}_Y$ are the excess kurtoses for X, Y. Consider

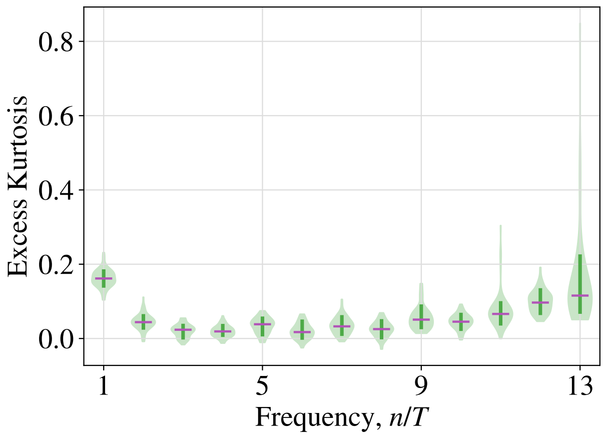

FIG. 4. The distribution of excess kurtosis of all 67 pulsars by frequency. Most frequencies agree fairly well with the prediction of zero in the many-sources limit. At $f = 1/T$, the deviation is mostly driven by the correlation between the real and imaginary parts of the coefficients. Color and bar conventions match Figure 1.

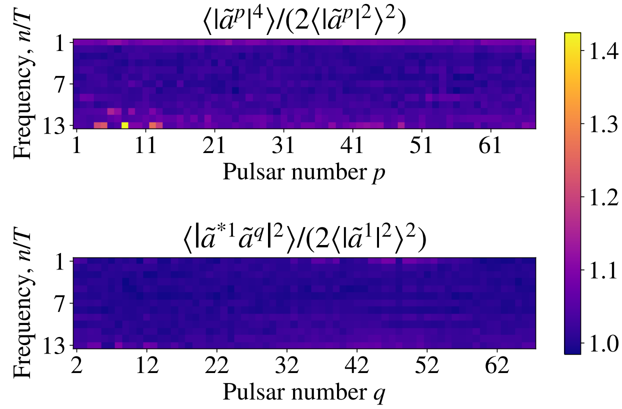

FIG. 5. The ratio of fourth moments to twice the second moments for auto-pulsars (upper) and a fixed-partner subset of cross-pulsars (lower) for an astrophysical model of a background and using the 67 pulsars of the NANOGrav 15yr dataset. At the lowest frequency, a mild disagreement for the auto-pulsar result is likely due to Poisson noise from truncating sources at very low frequencies. The cross-pulsar result matches predictions fairly well across all pairings and frequencies.

now the case of circular Z. Then, $\langle X^2 \rangle = \langle Y^2 \rangle$ and so the first term above simplifies to 2. The second term contains information about the excess kurtoses in X and Y, which should be the same if Z is circular, as well as the covariance between the real and imaginary parts, which should be zero if Z is circular. If Z is additionally Gaussian, the excess kurtoses should be zero, and we recover $\kappa = 2$ as expected. Examining the above equation, we see that any of $\bar{\kappa}_X, \bar{\kappa}_Y \neq 0$, $\langle X^2 \rangle \neq \langle Y^2 \rangle$, or $\langle XY \rangle \neq 0$ can lead to a non-zero excess kurtosis for the complex variable Z.

The results of our simulations show that $\bar{\kappa}_X, \bar{\kappa}_Y \approx 0$ and $\langle X^2 \rangle \approx \langle Y^2 \rangle$ at all frequencies, with a slight increase towards higher frequencies. For all frequencies except $f = 1/T$, X and Y are effectively uncorrelated, but in the lowest frequency we see a correlation of ≈ 0.36. One explanation for this correlation is that at this frequency, the approximation $w^{+-} \ll 1$ is marginal.

Now, consider the moment predictions on the auto- and cross-correlations of the Fourier coefficients. We find the second and fourth moments of all 67 auto-pulsar cases and the 66 cross-pulsar cases obtained by fixing one of the pulsars (without loss of generality, pulsar #1) across $11,000$ realizations. We then calculate the ratio of actual variance to large-source-limit variance and plot them in Figure 5. The auto-pulsar results display good agreement with the prediction for the second through the thirteenth frequencies. In the lowest frequency, a smaller deviation from prediction is likely due to truncating sources emitting below the bin edge of $0.5/T$, introducing Poisson noise. Note the single bright pixel at $f = 13/T$ for pulsar $p = 8$—this outlier is the same pulsar–frequency pair as produced the outlier in Figure 4. The cross-pulsar results display better agreement with the prediction across all pairings and frequencies. Some notable exceptions, again in the first frequency, occur when the pulsars are very nearly aligned with one another (e.g., $q = 49, 50$), that is, $\hat{x}_p \cdot \hat{x}_q \approx 1$. Then, even though $p \neq q$, terms like the averaged χ might behave more like their $p = q$ forms.

Finally, we use this realistic population to test the Cauchy distribution prediction introduced in Equation 27. As a reminder, for a circular Gaussian random variable, we expect its complex argument to be uniformly distributed, and thus the tangent thereof (which is directly related to the ratio of imaginary and real parts of the Fourier coefficients) to be Cauchy-distributed with a location parameter (mean) of zero and a shape parameter of unity. From Figure 5, we expect the distribution of $\tan\varphi$ to best match this prediction for the second through thirteenth frequencies.

To test this, we proceed as follows. We find the best fit parameters, assuming a Cauchy distribution, to the distribution of the $11,000$ realizations of $\tan\varphi_i^p$, indexing across all pulsars p and frequencies f_i. We then combine the data across pulsars and average the best fits to obtain a mean Cauchy location and shape parameter per frequency.[2] At each frequency, we then compare the data distribution to the predicted Cauchy distribution using these averaged parameters, as shown in Figure 6.

We see here similar behavior as in Figure 5. The first frequency has scale and shape parameters $\approx (0.30, 0.79)$,

[2] The weighted sum of Cauchy-distributed independent random variables is Cauchy-distributed, with location and shape parameters given by the identically-weighted sums of the location and shape parameters of the same.

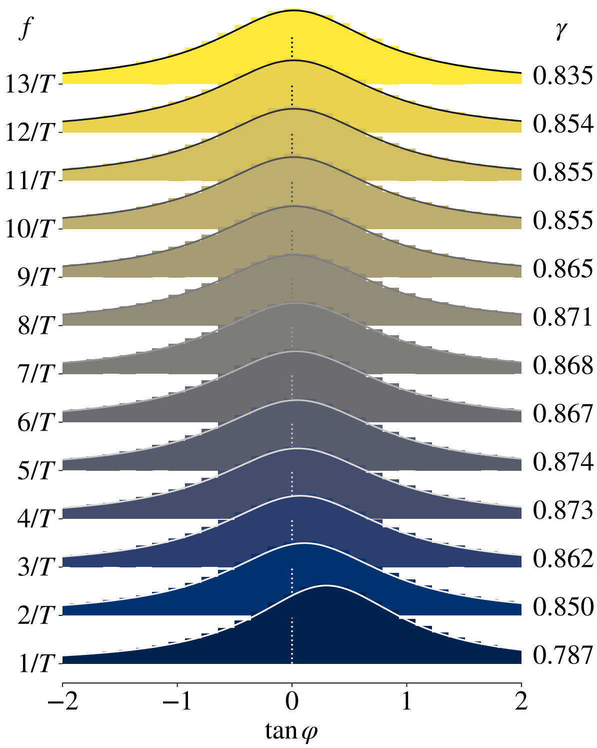

FIG. 6. Histograms of the $\tan\varphi$ for all $11,000$ realizations of all 67 pulsars by frequency, compared to the averaged best-fit Cauchy distributions by frequency. A vertical dashed line indicates $\tan\varphi = 0$. A circular Gaussian distribution in the Fourier coefficients would lead to a mean of zero and shape parameter $\gamma = 1$, but mild correlations and slightly unequal variances in the realistic model's coefficients cause deviations from this idealized case.

far from the predicted (circular Gaussian) result of $(0, 1)$. This suggests that there is a significant correlation between the real and imaginary parts of the Fourier coefficients at the lowest frequency, as well as a milder mismatch in their variances. The second through thirteenth frequencies have location parameters in the range $0.018\ldots0.108$ and shape parameters in the range $0.835\ldots0.874$; combined, this suggests a small to negligible correlation with a ratio of variances comparable to the first-frequency case while still being closer to parity. Thus, we can interpret the mild disagreement in the auto-pulsar results at the lowest frequency, as shown in Figure 5, as being mainly due to correlations between real and imaginary parts of the Fourier coefficients.

V. DISCUSSION

We have developed a general framework for calculating the variance and kurtosis of a single pulsar's timing-residual Fourier coefficients in response to a GWB from a finite population of sources. We applied it to both a toy model and a more astrophysical model of an SMBHB population and compared limits of our model to a Gaussian GWB. For the first time, our work accounts for sources emitting a non-central frequencies via the explicit inclusion of window functions.

We have analytically computed the excess kurtosis for the Fourier coefficients of an astrophysical GWB, and shown that for a finitely-sourced GWB, its always leptokurtic (i.e., more likely to produce outliers than a Gaussian distribution). In the large-source limit, the moments tend towards the moments of a Gaussian random variable.

Complementing the moments on the magnitudes of the Fourier coefficients, we have shown that the distribution of the tangent of their arguments should follow a standard Cauchy distribution. The distributions on magnitudes and arguments combined paves the way for quickly generating the timing-residual Fourier coefficients of a realistic SMBHB background.

Our work highlights the need for developments in current astrophysical inference. Firstly, we have demonstrated the importance of not assuming that all sources emit GW at harmonics of $f_n = n/T$ for $n \in \mathbb{N}^+$. This should be easy to implement with a SAM model assuming that the appropriate window function w is applied.

In PTA data analyses, the Gaussian assumption is achieved by marginalizing the Fourier coefficients in a hierarchical likelihood over a Gaussian prior with variance $S(f)\Delta f$ [57]. The PSD $S(f)$ is modeled as a frequency-dependent spectrum, e.g., a power-law conditioned on hyperparameters to be estimated by MCMC. To deal with non-Gaussianity, future analyses should model the GWB as a combination of a Gaussian GWB and individual continuous-wave sources [26, 42], with further development possible to model non-Gaussianities in regimes with a low number of unresolvable binaries [32].

Additional complexities are introduced with more astrophysical models, where the amplitudes depend on physical parameters such as chirp mass and luminosity distance, and which may account for the inclination, polarization angle, or eccentricity of the source. However, the specific astrophysical model we studied here shows that our results hold even in more realistic scenarios. Including additional astrophysical parameters such as polarization angle and inclination of the binary will modify the prefactors of the 'Poisson term' [29] as well as the distribution of the argument of the resulting Fourier coefficient. In future work, we will generalize the astrophysical model to explore the moments of Fourier coefficients and PSDs for eccentric, inclined, and polarized binaries. Adding eccentricity effects will also introduce inter-frequency covariances and non-stationarity [3, 58], which will need to be modeled with our framework.

Our work highlights the current issues in astrophysical inference of PTA data sets and introduces some potential solutions for accurate, but computationally tractable, analyses. With further development, our work can be extended to any finite population of GW sources across the GW frequency spectrum and with any GW detector, to determine the source, and to characterize the sources, of gravitational wave backgrounds.

ACKNOWLEDGEMENTS

Discussions: We thank Bruce Allen, Alexander Criswell, Celia Fielding, Kyle Gersbach, Jeffrey Hazboun, Thomas Konstandin, Nima Laal, Matthew Miles, Polina Petrov, Kai Schmitz, Levi Schult, and Serena Valtolina for useful conversations. We thank Kayhan Gültekin for early collaboration and extensive comments on a draft of this paper.

Funding/resources: Our work was supported by the NANOGrav NSF Physics Frontier Center awards #2020265 and #1430284. S.R.T acknowledges support from NSF AST-2307719, and an NSF CAREER #2146016. S.R.T is also grateful for support from a Vanderbilt University Chancellor's Faculty Fellowship. E.L.H. acknowledges summer research support from the Dean's Office of the Wentworth School of Sciences & Humanities. S.C.S acknowledges awards from NSF #2309246. J.M.W. acknowledges support from the APS–Simons Travel & Professional Development Awards. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, TN. W.G.L thanks the Fisk-Vanderbilt Bridge Program for their computing support.

Appendix A: Ensemble averaging over realizations of finite populations

To compute the moments of $\tilde{a}_i^p$, we must ensemble average over quantities that describe the source, such as the wave phase ϕ_j, GW propagation direction $\hat{\Omega}_j$, and amplitude parameters $\vec{\theta}_j$, where j indexes the sources. That average is taken over a probability measure $p(X)$; for a generic observable Q,

$$\langle Q \rangle_X = \int p(x) Q(x) dx, \text{ where } \int p(x) dx = 1. \tag{A1}$$

We will derive the fourth moment of $\tilde{a}_i^p$, which requires taking moments of $|\tilde{a}_i^p|^4$,

$$\begin{aligned}
|\tilde{a}_i^p|^4 = &\sum_j \frac{\mathcal{A}_j^4}{256\pi^4 f_j^4} |R_{pj}|^4 \left[\left(w_j^{--} + w_j^{++}\right)^2 + 2 w_j^{-+} w_j^{-+} + a\left(e^{i\phi_j}, e^{-i\phi_j}\right)\right] \\
&+ \sum_{j\neq k} \frac{\mathcal{A}_j^4}{256\pi^4 f_j^2 f_k^2} |R_{pj}|^2 |R_{pk}|^2 \left[\left(w_j^{--} + w_j^{++}\right)\left(w_k^{--} + w_k^{++}\right) + g\left(e^{i\phi_j}, e^{-i\phi_j}\right) b\left(e^{i\phi_k}, e^{-i\phi_k}\right)\right] \\
&+ \sum_{j\neq k}\sum_{l\neq m} \frac{\mathcal{A}_j \mathcal{A}_k \mathcal{A}_l \mathcal{A}_m}{256\pi^4 f_j f_k f_l f_m} \left(w_j^- R_{pj} e^{i\phi_j} - w_j^+ R_{pj}^* e^{-i\phi_j}\right)\left(w_k^- R_{pk}^* e^{-i\phi_j} - w_k^+ R_{pk} e^{i\phi_k}\right) \\
&\qquad \times \left(w_l^- R_{pl}^* e^{-i\phi_l} - w_l^+ R_{pl} e^{i\phi_l}\right)\left(w_m^- R_{pm} e^{i\phi_m} - w_m^+ R_{pm}^* e^{-i\phi_m}\right),
\end{aligned} \tag{A2}$$

where a and b are functions with an odd number $e^{i\phi_j}$ or $e^{i\phi_j}$ terms. These terms do not contribute to the fourth moment after averaging over wave phase $\phi \in [0, 2\pi)$. We note the following identities,

$$\left\langle e^{i\phi_j} \right\rangle_\phi = \int_0^{2\pi} e^{i\phi_j} \frac{d\phi_j}{2\pi} = 0 \tag{A3}$$

$$\left\langle e^{i(\phi_j - \phi_k)} \right\rangle_\phi = \iint_0^{2\pi} e^{i(\phi_j - \phi_k)} \frac{\mathrm{d}\phi_j}{2\pi} \frac{\mathrm{d}\phi_k}{2\pi} = \delta_{jk} \tag{A4}$$

$$\left\langle e^{i(\phi_j - \phi_k - \phi_l + \phi_m)} \right\rangle_\phi = \delta_{jl}\delta_{km} \text{ with } j \neq k, l \neq m\,. \tag{A5}$$

Averaging over phase removes the covariant phase functions and removes the second double summation in the final line of Equation A2. We get,

$$\left\langle |\tilde{a}_i^p|^4 \right\rangle_\phi = \sum_j \frac{\mathcal{A}_j^4}{256\pi^4 f_j^4} |R_{pj}|^4 \left[\left(w_j^{--} + w_j^{++} \right)^2 + 2 w_j^{-+} w_j^{-+} \right] + \sum_{j \neq k} \frac{2\mathcal{A}_j^4}{256\pi^4 f_j^2 f_k^2} |R_{pj}|^2 |R_{pk}|^2 \left(w_j^{--} + w_j^{++} \right) \left(w_k^{--} + w_k^{++} \right) \tag{A6}$$

Next, we average over the GW propagation direction $\hat{\Omega}$. We must find the following two quantities: $\left\langle |R_{pj}|^2 \right\rangle_{\hat{\Omega}}$ and $\left\langle |R_{pj}|^4 \right\rangle_{\hat{\Omega}}$. When we cross-correlate the (complex) antenna response $R_p = R_p^+ - iR_p^\times$ with sources j, k at pulsars p and q, we have the following averaging:

$$\begin{aligned} \left\langle R_p^*(f_j, \hat{\Omega}_j) R_q(f_k, \hat{\Omega}_k) \right\rangle_{\hat{\Omega}} &= \left\langle \left[1 - e^{2\pi i f_j L_p (1 + \hat{\Omega}_j \cdot \vec{x}_p)} \right] \left[1 - e^{-2\pi i f_k L_q (1 + \hat{\Omega}_k \cdot \vec{x}_q)} \right] \sum_A F_p^A(\hat{\Omega}_j) F_q^A(\hat{\Omega}_k) \right\rangle_{\hat{\Omega}} \\ &\approx \left\langle \left[1 - e^{2\pi i f_j L_p (1 + \hat{\Omega}_j \cdot \vec{x}_p)} \right] \left[1 - e^{-2\pi i f_k L_q (1 + \hat{\Omega}_k \cdot \vec{x}_q)} \right] \right\rangle \left\langle \sum_A F_p^A(\hat{\Omega}_j) F_q^A(\hat{\Omega}_k) \right\rangle_{\hat{\Omega}} \end{aligned} \tag{A7}$$

The exponential functions oscillate rapidly as a function of $\hat{\Omega}$, while the expectation of the product of antenna responses F_p^A varies relatively slowly. Hence, we can separate the expectation into two products as shown in the second line of Equation A7. The first expectation is

$$\begin{aligned} \langle \chi_{pq} \rangle_{\hat{\Omega}} &= \left\langle \left[1 - e^{2\pi i f_j L_p (1 + \hat{\Omega}_j \cdot \vec{x}_p)} \right] \left[1 - e^{-2\pi i f_j L_q (1 + \hat{\Omega}_j \cdot \vec{x}_q)} \right] \right\rangle_{\hat{\Omega}} \\ &\approx \begin{cases} 2 & \text{if } p = q \\ 1 & \text{else,} \end{cases} \end{aligned} \tag{A8}$$

This result is an approximation because the cross-terms generate cosine functions. With current PTAs of $T = \mathcal{O}(10\text{yr})$ and $L_p > 1\,\text{kpc}$, this is a rapidly oscillating function and can be ignored. However, in distant future data sets, this oscillation term will become relevant [59].

For a GWB measurement via a pulsar p with antenna response functions $F_p = F_p^+ - iF_p^\times$, we can define the unpolarized, polarized, and cross-polarized responses for a pulsar pair as being proportional to $\text{Re}[F_p^* F_q]$, $\text{Im}[F_p F_q]$, and $\text{Abs}[F_p F_q]$, respectively. The moments of these responses allow us to calculate the moments of the Fourier coefficients.

For an unpolarized GWB, the mean of the product of antenna responses $\mu_u(\gamma_{pq})$ is given by

$$\begin{aligned} \mu_u(\gamma_{pq}) &= \frac{1}{4\pi} \int d\hat{\Omega} \sum_{A \in \{+,\times\}} F_p^A(\hat{\Omega}) F_q^A(\hat{\Omega}) \\ &= \frac{1}{4\pi} \int d\hat{\Omega}\, \text{Re}[F_p^*(\hat{\Omega}) F_q(\hat{\Omega})] \\ &= \frac{1}{4} + \frac{1}{12} \cos\gamma_{pq} + \frac{1}{2} \Psi(\gamma_{pq}), \end{aligned} \tag{A9}$$

$$\Psi(\gamma_{pq}) = (1 - \cos\gamma_{pq}) \log\left(\frac{1 - \cos\gamma_{pq}}{2} \right). \tag{A10}$$

This is called the Hellings-&-Downs (HD) correlation [52], and it is the signature correlation of a GWB.

Now, Equation A7 simplifies to [28]

$$\left\langle R_p^*(f_j, \hat{\Omega}_j) R_q(f_j, \hat{\Omega}_j) \right\rangle_{\hat{\Omega}} \equiv \left\langle R_{pj}^* R_{qj} \right\rangle = \langle \chi_{pq} \rangle \mu_u(\gamma_{pq}) \tag{A11}$$

for a single source j, where we used the condensed notation $R_p(f_j, \hat{\Omega}_j) \equiv R_{pj}$.

Finally, we choose to normalize the inter-pulsar correlation such that the auto-correlation term is equal to 1 when pulsars $p = q$, as is common in the literature. This allows us to write cross-correlations as a fraction of the autocorrelation,

$$\Gamma_{pq} = \frac{3}{2} \mu_u(\gamma_{pq}) \tag{A12a}$$

(A12b)

We now turn our attention to the ensemble average of $|R_{pj}|^4$. We start by looking at $\left\langle |R_{pj}|^2 |R_{qj}|^2 \right\rangle_{\hat{\Omega}}$ by applying results from Allen [28]. We can rewrite this four-point function in terms of real and imaginary components,

$$\left\langle |R_{pj}|^2 \, |R_{qj}|^2 \right\rangle_{\hat{\Omega}} = \left\langle \left|R^*_{pj} R_{qj}\right|^2 \right\rangle_{\hat{\Omega}} \tag{A13}$$
$$\approx \left\langle |\chi_{pq}|^2 \right\rangle \left(\left\langle \mathrm{Re}\left[F^*_{pj} F_{qj}\right]^2 \right\rangle_{\hat{\Omega}} + \left\langle \mathrm{Im}\left[F^*_{pj} F_{qj}\right]^2 \right\rangle_{\hat{\Omega}} \right).$$

The real term represents the second moment of the "unpolarized" Hellings-&-Downs correlation, while the imaginary term represents the second moment of the "polarized" Hellings-&-Downs correlation. We will deal with each quantity separately. When we find higher moments of a point source, we require the second moment of χ_{pq},

$$\left\langle |\chi_{pq}|^2 \right\rangle_{\hat{\Omega}} = \left\langle \chi^*_{pq} \chi_{pq} \right\rangle_{\hat{\Omega}} \approx \begin{cases} 6 & \text{if } p = q \\ 4 & \text{else.} \end{cases} \tag{A14}$$

In this case, the cross-terms that we ignored because they rapidly oscillated in the mean now cross with their conjugates, and contribute towards the average.

The second moment of the unpolarized Hellings-&-Downs correlation is,

$$\left\langle \mathrm{Re}\left[F^*_{pj} F_{qj}\right]^2 \right\rangle_{\hat{\Omega}} = \mu_u^2(\gamma_{pq}) + \sigma_u^2(\gamma_{pq}), \tag{A15}$$

where $\sigma_u^2(\gamma_{pq})$ is the variance of the unpolarized Hellings-&-Downs correlation, which is derived in Appendix D of Allen [28]. We give its form here; however, to be consistent with the normalization of Equation A12a, we also normalize the variance.

$$\left(\frac{3}{2}\right)^2 \sigma_u^2(\gamma_{pq}) = \left(\frac{873}{320} + \frac{3}{32}\cos(\gamma_{pq}) - \frac{839}{320}\cos^2\gamma_{pq} \right) + \frac{3}{16}\left(18 - 10\cos\gamma_{pq} - 3\Psi(\gamma_{pq})\right)\Psi(\gamma_{pq}) \tag{A16}$$

Similarly, the second moment of the polarized Hellings-&-Downs correlation is,

$$\left\langle \mathrm{Im}\left[F^*_{pj} F_{qj}\right]^2 \right\rangle_{\hat{\Omega}} = \mu_p^2(\gamma_{pq}) + \sigma_p^2(\gamma_{pq}). \tag{A17}$$

The mean of polarized sources $\mu_p(\gamma_{pq}) = 0 \, \forall \, \gamma_{pq}$, which can be understood by example: one origin of polarization might be the inclination of the binary's orbital plane with respect to the observation line, but assuming a uniformly-distributed inclination angle, on average the plane is not inclined and thus the wave is not polarized. The variance $\sigma_p^2(\gamma_{pq})$ of the polarized Hellings-&-Downs correlation is derived in Appendix E of Allen [28]. For consistency, we also normalize this quantity,

$$\sigma_p^2(\gamma_{pq}) = \frac{21}{8}\left(\cos^2\gamma_{pq} - 1\right) + \frac{9}{16}\left(3\cos\gamma_{pq} - 7\right)\Psi(\gamma_{pq}). \tag{A18}$$

Bringing everything together, we find that,

$$\left\langle \left(\frac{3}{2}\right)^2 |R_{pj}|^2 |R_{qj}|^2 \right\rangle_{\Omega} = \left\langle |\chi_{pq}|^2 \right\rangle \left(\Gamma_{pq}^2 + \sigma_u^2(\gamma_{pq}) + \sigma_p^2(\gamma_{pq}) \right) = \left\langle |\chi_{pq}|^2 \right\rangle \sigma_c^2(\gamma_{pq}), \tag{A19}$$

where $\sigma_c^2(\gamma_{pq})$ is the variance of the cross-polarized Hellings-&-Downs correlation, which is derived in Appendix F of Allen [28]. We have shown that this quantity is the sum of the variances of the real and imaginary parts of the cross-correlation of the antenna response functions.

$$\sigma_c^2(\gamma_{pq}) = \frac{39}{160} + \frac{3}{16}\cos\gamma_{pq} + \frac{3}{160}\cos^2\gamma_{pq}\,. \tag{A20}$$

We plot these various quantities as a function of γ_{pq} in Figure 7. Finally, to ensure that we have normalized our quantities correctly, we must substitute

$$\left\langle R^*_{pj} R_{qj} \right\rangle \to \frac{2}{3}\left\langle \frac{3}{2} R^*_{pj} R_{qj} \right\rangle, \tag{A21}$$

in any ensemble average that we compute.

Therefore, substituting these quantities into Equation A6 and setting $p = q$ (i.e., $\gamma_{pq} = 0$),

$$\left\langle |\tilde{a}_i^p|^4 \right\rangle_{\phi,\hat{\Omega}} = 2\left\langle |\tilde{a}_i^p|^2 \right\rangle^2_{\phi,\hat{\Omega}} + \sum_j \frac{\mathcal{A}_j^4}{24^2\pi^4 f_j^4}\left[\frac{7}{10}(w_j^{--} + w_j^{++})^2 + \frac{27}{5} w_j^{--} w_j^{++} \right] + 4\sum_{j\neq k} \frac{\mathcal{A}_j^2 \mathcal{A}_k^2}{24^2\pi^4 f_j^2 f_k^2} w_j^{-+} w_k^{-+}. \tag{A22}$$

The final step finds the ensemble average over astrophysical parameters $\vec{\theta}$. We detail this process in section B, and the final fourth moment is given by Equation 17.

By way of example of the utility of these relationships, we present below the fourth cross-pulsar moment of the Fourier coefficients, averaged over sky position and phase:

$$\left\langle|\tilde{a}_i^{*p}\tilde{a}_i^q|^2\right\rangle_{\phi,\hat{\Omega}} = \sum_j \frac{A_j^4}{24^2\pi^4 f^4}\left[\left\langle\chi_{pp}^2\right\rangle\sigma_c^2(\gamma_{pq})\left(w_j^{--}+w_j^{++}\right)^2 + \left\langle\chi_{pq}^2\right\rangle\left(4\Gamma_{pq}^2 + 2\left(2\sigma_u^2(\gamma_{pq}) - \sigma_c^2(\gamma_{pq})\right)\right)w_j^{--}w_j^{++}\right]$$
$$+ \sum_{j\neq k}\frac{A_j^2A_k^2}{24^2\pi^4 f_j^2 f_k^2}\left(\left(1+\Gamma_{pq}^2\right)\left(w_j^{--}+w_j^{++}\right)\left(w_k^{--}+w_k^{++}\right) + \left(\Gamma_{pq}^2-1\right)\left(w_j^{-}w_k^{+}+w_j^{+}w_k^{-}\right)^2\right) \quad \text{(A23)}$$

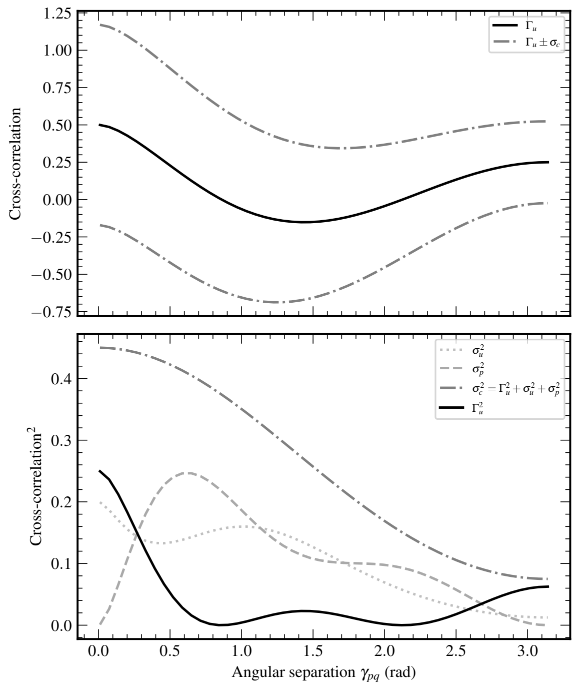

FIG. 7. Top: The Hellings-&-Downs curve as a function of angular separation between pulsars p, q (solid line) plus/minus a standard deviation (dot-dashed). Bottom: A figure of $(\Gamma_{pq})^2$ (solid), σ_u^2 (dotted), σ_p^2 (dashed), and σ_c^2 (dot-dashed) as a function of angular separation γ_{pq}.

In the limit $p = q$, this reproduces Equation A22. It is useful to have when determining the origin of a number of terms and prefactors of interest.

Appendix B: Semi-Analytic Models

There are two quantities in our model that are dependent on $\vec{\theta}$: the GW strain amplitude $A(f_j, \vec{\theta}_j) \equiv A_j$. The *mean* number of GW sources $\Delta\bar{N}(f, \hat{\Omega}, \vec{\theta})$ that emit within a frequency bin Δf, sky-pixel $\hat{\Omega}$, and astrophysical parameter bins $\Delta\vec{\theta}$. In this case, we can replace the summation over $\bar{N}$ binaries within a given parameter bin with a factor of dN:

$$\sum_j^{\Delta\bar{N}(f,\hat{\Omega},\vec{\theta})} Q(f_j, \hat{\Omega}_j, \vec{\theta}_j) \approx \Delta\bar{N}(f, \hat{\Omega}, \vec{\theta})Q(f, \hat{\Omega}, \vec{\theta})$$
$$= \frac{\Delta\bar{N}}{\Delta f\Delta\hat{\Omega}\Delta\vec{\theta}}Q(f, \hat{\Omega}, \vec{\theta})\Delta f\Delta\hat{\Omega}\Delta\vec{\theta}. \quad \text{(B1)}$$

In the limits $\Delta f, \Delta\hat{\Omega}, \Delta\vec{\theta} \to 0$, all $\Delta\bar{N}(f, \hat{\Omega}, \vec{\theta}) \to d\bar{N}(f, \hat{\Omega}, \vec{\theta})$ binaries within that infinitesimal bin emit at the same frequency $f_j = f$, the same GW propagation $\hat{\Omega}_j = \hat{\Omega}$, and the same strain amplitude parameters $\vec{\theta}_j = \vec{\theta}$. The infinitessimal number density $d\bar{N}/dfd\hat{\Omega}d\vec{\theta}$ models the distribution of the number of GW sources as a function of frequency and strain amplitude parameters, which becomes our probability measure that we use to average over $\vec{\theta}$ and $\hat{\Omega}$. In PTA data analyses, this quantity is often described as $d\bar{N}/df_r d\hat{\Omega} d\log\mathcal{M}dz$, where $\mathcal{M}$ is the chirp mass of an SMBHB and z is the redshift to the SMBHB [31, 40, 56], and evaluated at the rest-frame GW frequency $f_r = f(1+z)$.

Hence, the summation becomes an integration,

$$\sum_j^{\bar{N}} Q(f_j, \hat{\Omega}_j, \vec{\theta}_j) \to \int \frac{d\bar{N}}{dfd\hat{\Omega}d\vec{\theta}}Q(f, \hat{\Omega}, \vec{\theta})dfd\hat{\Omega}d\vec{\theta}. \quad \text{(B2)}$$

It is important to note that, in the limit $\Delta f \to 0$, $\Delta\hat{\Omega} \to 0$ and $\Delta\vec{\theta} \to 0$ for a finite population, there are either 0 or 1 sources emitting at exactly frequency f with GW propagation direction $\hat{\Omega}$ and astrophysical parameters $\vec{\theta}$. Therefore, the square of the number of sources is also 0 or 1. This is the definition of shot noise, or Poisson variance [30, 31]. The number of GW sources N in a realization of a GWB is Poisson-distributed with mean $\bar{N}$ and it is conditioned on model hyperparameters $\vec{\eta}$. The hyperparameters $\vec{\eta}$ include parameters that limit the number of binaries above certain chirp masses and redshifts, hence, changing the distribution of sources along each astrophysical parameter, modifying the number of binaries and their amplitudes. Therefore, the mean number density of GW sources per unit frequency, GW propagation unit angle, and per astrophysical parameter is $d\bar{N}(f, \hat{\Omega}, \vec{\theta}|\eta)/dfd\hat{\Omega}d\vec{\theta}$ is simply the *ensemble averaged number density* of sources for a given population model and hyperparameters η. [23, 31]. This is the semi-analytical model (SAM) of the astrophysical GWB.

To compute variances, we note the following relations to aid our computation for an arbitrary function Q(x):

$$\left(\sum_j Q(x_j)\right)^2 = \sum_j Q(x_j)^2 + \sum_{j\neq k} Q(x_j)Q(x_k) \quad \text{(B3a)}$$

$$\left(\int_x Q(x)dx\right)^2 = \iint_{x,x'} Q(x)Q(x')(1-\delta(x-x'))dxdx' + \int_x Q(x)^2 dx \quad \text{(B3b)}$$

In short, Equation B3a and Equation B3b are expressions that the square of a sum can be decomposed into its cross-terms and its diagonal.

Appendix C: Complex random variables

When we say that the $\tilde{a}_i^p$ are Gaussian, we mean that they follow a *circular complex Gaussian distribution*. “Complex” here means that the random variable is complex; circularity is a property of complex random variables which says (loosely) that the statistics are invariant under arbitrary rotation. Formally, for variable Z and a fixed but arbitrary $\phi \in [-\pi, \pi)$,

$$\langle Z \rangle = \left\langle e^{i\phi} Z \right\rangle = e^{i\phi} \langle Z \rangle \, , \quad \text{(C1a)}$$

$$\langle ZZ \rangle = \left\langle e^{i2\phi} ZZ \right\rangle = e^{i2\phi} \langle ZZ \rangle \, . \quad \text{(C1b)}$$

Because ϕ is arbitrary, it must be that $\langle Z \rangle = 0$ and $\langle ZZ \rangle = 0$. The first quantity is properly the mean, but $\langle ZZ \rangle$ is called the *pseudovariance*, and has no direct analogy for real random variables. Therefore, a circular complex Gaussian has zero mean and zero pseudocovariance matrix.

The definitions of statistics for complex random variables require a more careful treatment. For a complex random variable Z, $\langle Z \rangle$ and $\langle (Z^* - \langle Z^* \rangle)(Z - \langle Z \rangle) \rangle$ are agreed upon as the mean and variance, respectively (and note that $\langle Z^* \rangle = \langle Z \rangle^*$). Similarly, with a second complex random variable W, $\langle (Z - \langle Z \rangle)(W^* - \langle W^* \rangle) \rangle$ is the covariance (and so note $\mathrm{Cov}(Z, W) = \mathrm{Cov}(W, Z)^*$). Skewness and kurtosis, which we are interested in studying, do not have unambiguous extensions. Consider, for example, $\left\langle |Z^* Z|^3 \right\rangle$ as a skewness measure: since this discards phase information, it cannot determine the “direction of the tilt” of the distribution in the complex plane; moreover, it would return a non-zero value for a circular complex Gaussian distribution, which manifestly should have zero skewness.

To compute the (excess) kurtosis of the timing residual Fourier components, we extend the idea of the kurtosis as a ratio between a fourth central moment and the variance squared and write

$$\kappa = \frac{\left\langle (Z^* - \langle Z^* \rangle)^2 (Z - \langle Z \rangle)^2 \right\rangle}{\sigma^4} \, , \quad \text{(C2)}$$

where σ^2 is the variance as defined above. This measure is real-valued and thus preserves the notion of kurtosis as a measure of tailedness of the distribution.

In the main text, we did not analytically calculate the odd moments of the $\tilde{a}_i^p$ such as skewness, which should be zero. When we perform checks of simulated $\tilde{a}_i^p$, we will use

$$\gamma_1 = \frac{\left\langle (Z - \langle Z \rangle)^3 \right\rangle}{\sigma^{3/2}} \, . \quad \text{(C3)}$$

This measure keeps full phase information and will properly return zero for a circular complex Gaussian.

As a final note, Isserlis’ theorem [60] works for complex variables just as well as for real variables. We can use it to show that the kurtosis of a circular complex Gaussian is *not* 3, as it is for a real univariate Gaussian. Rather, with

$$\langle ZZZ^*Z^* \rangle = \langle ZZ \rangle \langle Z^*Z^* \rangle + 2 \langle Z^*Z \rangle \quad \text{(C4)}$$

and recalling $\langle Z \rangle = 0$, $\langle ZZ \rangle = 0$ if Z is circular, then $\kappa = \langle ZZZ^*Z^* \rangle / \langle Z^*Z \rangle = 2$. Thus, the excess kurtosis for a complex random variable is the kurtosis minus two.

The above moments can be thought of as moments of the magnitude $|Z|$ of the complex number Z. However, complex numbers are uniquely defined by their magnitude and their argument $\varphi = \arg Z$.[3] The tangent of the argument is given by,

$$\tan\varphi = \frac{\mathrm{Im}[Z]}{\mathrm{Re}[Z]}. \quad \text{(C5)}$$

By analyzing the distribution of the argument, we can gather more information about the distribution of the complex number, as discussed further in Appendix E.

Appendix D: Combinations of window functions

The product of coefficients gives rise to a product of window functions. For compactness, we adopt a further reduced notation,

$$w^{\pm\pm} = w^{\pm} w^{\pm} = w(f \pm f_i) w(f \pm f_i) \, . \quad \text{(D1)}$$

In this work, we also take products of products of window functions. In this case, we are always free to redistribute the four signs as we see fit (e.g., $w^{+-}w^{+-} = w^{--}w^{++}$; note that $w^{+-} = w^{-+}$ follows from the definition). Assuming that $w^{\pm} \to \delta(f \pm f_i)$ as $T \to \infty$, then ‘crossed’ window terms will go to zero. This is because (for $f_j > 0$),

$$w^{-+} \to \delta(f - f_i)\delta(f + f_i) = 0 \, , \quad \text{(D2a)}$$

$$w^{--}w^{++} \to \delta(f - f_i)^2 \delta(f + f_i)^2 = 0 \, . \quad \text{(D2b)}$$

[3] This is also called the *phase* of the complex number, but to avoid confusion with the wave phase of a GW source, we will use the term *argument*.

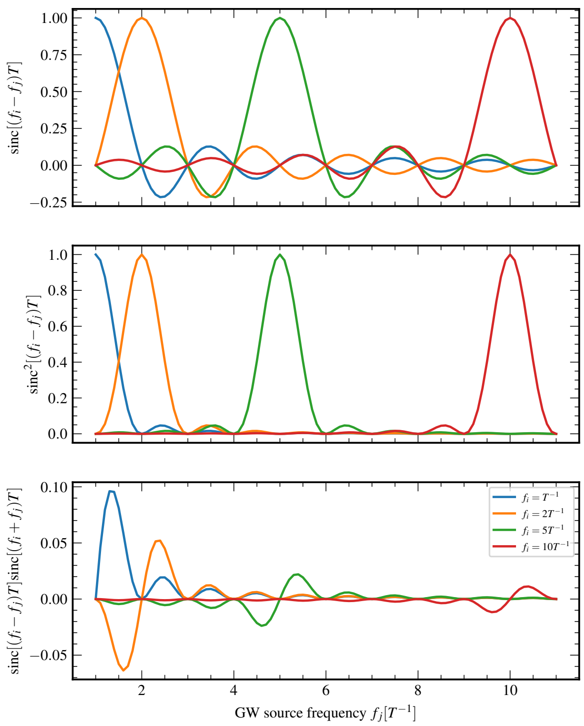

FIG. 8. An illustration of w^- (top), w^{--} (middle), and w^{+-} (bottom) for various measured frequencies as a function of GW source frequency. The first two functions strongly select at the measured frequency, but are subject to leakage from nearby frequencies; the "crossed" function suppresses at all frequencies but less strongly near the measured frequency.

We will also often consider the *one-sided* Fourier transform, which changes our range of integration in frequency from $\mathbb{R}$ to $\mathbb{R}_{\geq 0}$. Under the assumption that the window function is symmetric around $f_i = 0$, the one-sided power spectrum is simply twice the two-sided power spectrum. Similarly, the one-sided fourth moment of $\tilde{a}_i^p$ is four times the two-sided quantity.

Next, consider w^{+-} for various different observed GWB frequency f and sample frequency f_i. When $f \approx f_i$, $w^+ \ll w^-$ and $w^{+-} \ll w^{--}$. When f and f_i are well-separated (but positive), w^+ and w^- are of comparable (small) magnitude. When computing the variance, we will often be able to neglect terms like w^{+-} in comparison to terms like w^{--}, sometimes independently of their precise subscripts. This is the assumption often made in SAMs.

Finally, when we are discussing the finite population model, we modify our notation to include the index j, $w_j^\pm \equiv w(f_j \pm f_i)$. Figure 8 shows how different combinations of sinc functions compare as a function of GW source frequency.

Appendix E: Cauchy random variables

This appendix serves as a brief introduction to the most relevant properties of Cauchy distributions. Most of these definitions and results can be found in a standard probability text, but are presented here for convenience.

The Cauchy distribution has two parameters: the location parameter ($x_0 \in \mathbb{R}$) and the shape parameter ($\gamma \in \mathbb{R}^+$), and PDF

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma} \frac{1}{1 + \left(\frac{x - x_0}{\gamma}\right)^2} . \tag{E1}$$

It is sometimes called the *Breit–Wigner* distribution in high-energy particle physics. We will denote a Cauchy-distributed random variable X by $X \sim \mathrm{Cy}(x_0, \gamma)$. Its mean, variance, and all higher moments are undefined, but its median and mode are both simply equal to x_0. A last general property to note is that the (weighted) sum of Cauchy random variables is itself a Cauchy random variable: with $X_i \sim \mathrm{Cy}(x_{0,i}, \gamma_i)$ for $i = 1 \dots n$ and all n Xs independent,

$$\bar{X} \equiv \sum_{i=1}^{n} \beta_i X_i \sim \mathrm{Cy}(\bar{x}_0, \bar{\gamma}) \tag{E2}$$

$$\text{where} \quad \bar{x}_0 = \sum_{i=1}^{n} \beta_i x_{0,i} \, , \; \bar{\gamma} = \sum_{i=1}^{n} \beta_i \gamma_i \, .$$

Importantly, this means that the central limit theorem applied to Cauchy random variables predicts that the sum is still a Cauchy random variable, and not Gaussian. This is related to the Cauchy distribution being a stable distribution, and helps explain why we expect $\tan\varphi_i^p$ to still be Cauchy-distributed, even for $N > 1$.

A construction of a Cauchy random variable, particularly relevant to this work, is as follows. Take two independent standard normal random variables: $X, Y \sim \mathcal{N}(0, 1)$. Then, their ratio follows a standard Cauchy distribution: $W \equiv Y/X \sim \mathrm{Cy}(0, 1)$. This can be extended to *correlated* normal random variables with unequal variance as follows. First, note that if T is a Cauchy random variable, then so too is $W \equiv (aT + b)/(cT + d)$ where $a, b, c, d \in \mathbb{R}$. Specifically, defining a complex Cauchy parameter $\psi = x_0 + i\gamma$ and saying $T \sim \mathrm{Cy}(\psi_t)$, then $W \sim \mathrm{Cy}(\psi_w)$ where $\psi_w = (a\psi_t + b)/(c\psi_t + d)$.

Now, let U, V be independent standard normal random variables and thus $T \equiv U/V$ is standard Cauchy, with $\psi_t = i$. Then W, by the above construction, is also Cauchy, with

$$\psi_w = \frac{ac + bd + i(ad - bc)}{c^2 + d^2} \tag{E3}$$

$$\rightarrow \quad x_{0w} = \frac{ac + bd}{c^2 + d^2} \, , \; \gamma_w = \frac{ad - bc}{c^2 + d^2} \, . \tag{E4}$$

We can rearrange the definition of W in terms of T to be

$$W \equiv \frac{aU + bV}{cU + dV} = \frac{Y}{X} \, , \tag{E5}$$

where $Y \sim \mathcal{N}(0, \sigma_y^2 = a^2 + b^2)$, $X \sim \mathcal{N}(0, \sigma_x^2 = c^2 + d^2)$ are normal random variables with $\mathrm{Cov}[X, Y] = ac + bd$. Then noting that the correlation of X and Y is defined as $\rho = \mathrm{Cov}[X, Y]/\sigma_x \sigma_y$, we can write

$$x_{0w} = \frac{\sigma_y}{\sigma_x}\rho\,,\ \gamma_w = \frac{\sigma_y}{\sigma_x}\sqrt{1-\rho^2}\,. \tag{E6}$$

Thus, with a complex Gaussian random variable $Z = X + iY$ and $\tan\varphi_i^p = \mathrm{Im}[Z]/\mathrm{Re}[Z] \leftrightarrow W = Y/X$, a nonzero location parameter indicates nonzero correlations between the real and imaginary parts of the Fourier coefficients, as seen in Figure 6. When the location parameter is (effectively) zero but the shape parameter is non-unity, this indicates that the real and imaginary parts have unequal variance. But if Z is a circular complex Gaussian, then its argument will follow a standard Cauchy distribution. Deviations from standard Cauchy are signs of non-circularity, as discussed in section IV B.

[1] G. Agazie *et al.* (NANOGrav), The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background, Astrophys. J. Lett. **951**, L8 (2023), arXiv:2306.16213 [astro-ph.HE].

[2] J. Antoniadis *et al.* (EPTA, InPTA:), The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals, Astron. Astrophys. **678**, A50 (2023), arXiv:2306.16214 [astro-ph.HE].

[3] D. J. Reardon *et al.*, Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array, Astrophys. J. Lett. **951**, L6 (2023), arXiv:2306.16215 [astro-ph.HE].

[4] H. Xu *et al.*, Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I, Res. Astron. Astrophys. **23**, 075024 (2023), arXiv:2306.16216 [astro-ph.HE].

[5] M. T. Miles, R. M. Shannon, D. J. Reardon, M. Bailes, D. J. Champion, M. Geyer, P. Gitika, K. Grunthal, M. J. Keith, M. Kramer, *et al.*, The meerkat pulsar timing array: The first search for gravitational waves with the meerkat radio telescope, Monthly Notices of the Royal Astronomical Society **536**, 1489 (2025).

[6] R. Manchester, The international pulsar timing array, Classical and Quantum Gravity **30**, 224010 (2013).

[7] G. Agazie, A. Anumarlapudi, A. M. Archibald, P. T. Baker, B. Bécsy, L. Blecha, A. Bonilla, A. Brazier, P. R. Brook, S. Burke-Spolaor, *et al.*, The nanograv 15 yr data set: Constraints on supermassive black hole binaries from the gravitational-wave background, The Astrophysical Journal Letters **952**, L37 (2023).

[8] A. Afzal, G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Bécsy, J. J. Blanco-Pillado, L. Blecha, K. K. Boddy, *et al.*, The nanograv 15 yr data set: Search for signals from new physics, The Astrophysical Journal Letters **951**, L11 (2023).

[9] G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Bécsy, L. Blecha, A. Brazier, P. R. Brook, S. Burke-Spolaor, *et al.*, The nanograv 15 yr data set: Search for anisotropy in the gravitational-wave background, The Astrophysical Journal Letters **956**, L3 (2023).

[10] G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Becsy, L. Blecha, A. Brazier, P. R. Brook, S. Burke-Spolaor, *et al.*, The nanograv 15 yr data set: Bayesian limits on gravitational waves from individual supermassive black hole binaries, The Astrophysical Journal Letters **951**, L50 (2023).

[11] J. Antoniadis, P. Arumugam, S. Arumugam, S. Babak, M. Bagchi, A.-S. B. Nielsen, C. Bassa, A. Bathula, A. Berthereau, M. Bonetti, *et al.*, The second data release from the european pulsar timing array-iv. implications for massive black holes, dark matter, and the early universe, Astronomy & Astrophysics **685**, A94 (2024).

[12] J. Antoniadis, P. Arumugam, S. Arumugam, S. Babak, M. Bagchi, A.-S. B. Nielsen, C. Bassa, A. Bathula, A. Berthereau, M. Bonetti, *et al.*, The second data release from the european pulsar timing array-v. search for continuous gravitational wave signals, Astronomy & Astrophysics **690**, A118 (2024).

[13] K. Grunthal, R. S. Nathan, E. Thrane, D. J. Champion, M. T. Miles, R. M. Shannon, A. D. Kulkarni, F. Abbate, S. Buchner, A. D. Cameron, *et al.*, The meerkat pulsar timing array: Maps of the gravitational wave sky with the 4.5-yr data release, Monthly Notices of the Royal Astronomical Society **536**, 1501 (2025).

[14] G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, J. G. Baier, P. T. Baker, B. Bécsy, L. Blecha, A. Brazier, P. R. Brook, *et al.*, The nanograv 15 yr data set: Looking for signs of discreteness in the gravitational-wave background, The Astrophysical Journal **978**, 31 (2024).

[15] J. Ellis, M. Fairbairn, G. Hütsi, M. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, Prospects for future binary black hole gravitational wave studies in light of pta measurements, Astronomy & Astrophysics **676**, A38 (2023).

[16] C. Caprini and D. G. Figueroa, Cosmological backgrounds of gravitational waves, Classical and Quantum Gravity **35**, 163001 (2018).

[17] A. Afzal, G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Bécsy, J. J. Blanco-Pillado, L. Blecha, K. K. Boddy, A. Brazier, P. R. Brook, S. Burke-Spolaor, R. Burnette, R. Case, M. Charisi, S. Chatterjee, K. Chatziioannou, B. D. Cheeseboro, S. Chen, T. Cohen, J. M. Cordes, N. J. Cornish, F. Crawford, H. T. Cromartie, K. Crowter, C. J. Cutler, M. E. DeCesar, D. DeGan, P. B. Demorest, H. Deng, T. Dolch, B. Drachler, R. von Eckardstein, E. C. Ferrara, W. Fiore, E. Fonseca, G. E. Freedman, N. Garver-Daniels, P. A. Gentile, K. A. Gersbach, J. Glaser, D. C. Good, L. Guertin, K. Gültekin, J. S. Hazboun, S. Hourihane, K. Islo, R. J. Jennings, A. D. Johnson, M. L. Jones, A. R. Kaiser, D. L. Kaplan, L. Z. Kelley, M. Kerr, J. S. Key, N. Laal, M. T. Lam, W. G. Lamb, T. J. W. Lazio, V. S. H. Lee, N. Lewandowska, R. R. L. dos Santos, T. B. Littenberg, T. Liu, D. R. Lorimer, J. Luo, R. S. Lynch, C.-P. Ma, D. R. Madison, A. McEwen, J. W. McKee, M. A. McLaughlin, N. McMann, B. W.

Meyers, P. M. Meyers, C. M. F. Mingarelli, A. Mitridate, J. Nay, P. Natarajan, C. Ng, D. J. Nice, S. K. Ocker, K. D. Olum, T. T. Pennucci, B. B. P. Perera, P. Petrov, N. S. Pol, H. A. Radovan, S. M. Ransom, P. S. Ray, J. D. Romano, S. C. Sardesai, A. Schmiedekamp, C. Schmiedekamp, K. Schmitz, T. Schröder, L. Schult, B. J. Shapiro-Albert, X. Siemens, J. Simon, M. S. Siwek, I. H. Stairs, D. R. Stinebring, K. Stovall, P. Stratmann, J. P. Sun, A. Susobhanan, J. K. Swiggum, J. Taylor, S. R. Taylor, T. Trickle, J. E. Turner, C. Unal, M. Vallisneri, S. Verma, S. J. Vigeland, H. M. Wahl, Q. Wang, C. A. Witt, D. Wright, O. Young, and K. M. Zurek (NANOGrav), The NANOGrav 15-year Data Set: Search for Signals from New Physics, Astrophys. J. Lett. **951**, L11 (2023), arXiv:2306.16219 [astro-ph.HE].

[18] M. Rajagopal and R. W. Romani, Ultra-low frequency gravitational radiation from massive black hole binaries, arXiv preprint astro-ph/9412038 (1994).

[19] A. H. Jaffe and D. C. Backer, Gravitational waves probe the coalescence rate of massive black hole binaries, The Astrophysical Journal **583**, 616 (2003).

[20] J. S. B. Wyithe and A. Loeb, Low-frequency gravitational waves from massive black hole binaries: predictions for lisa and pulsar timing arrays, The Astrophysical Journal **590**, 691 (2003).

[21] A. Sesana, F. Haardt, P. Madau, and M. Volonteri, Low-frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies, The Astrophysical Journal **611**, 623 (2004).

[22] S. Burke-Spolaor, S. R. Taylor, M. Charisi, T. Dolch, J. S. Hazboun, A. M. Holgado, L. Z. Kelley, T. J. W. Lazio, D. R. Madison, N. McMann, *et al.*, The astrophysics of nanohertz gravitational waves, The Astronomy and astrophysics review **27**, 5 (2019).

[23] A. Sesana, A. Vecchio, and M. Volonteri, Gravitational waves from resolvable massive black hole binary systems and observations with pulsar timing arrays, Monthly Notices of the Royal Astronomical Society **394**, 2255 (2009).

[24] P. A. Rosado, A. Sesana, and J. Gair, Expected properties of the first gravitational wave signal detected with pulsar timing arrays, Monthly Notices of the Royal Astronomical Society **451**, 2417 (2015).

[25] L. Z. Kelley, L. Blecha, L. Hernquist, A. Sesana, and S. R. Taylor, Single sources in the low-frequency gravitational wave sky: properties and time to detection by pulsar timing arrays, Monthly Notices of the Royal Astronomical Society **477**, 964 (2018).

[26] B. Bécsy, N. J. Cornish, and L. Z. Kelley, Exploring realistic nanohertz gravitational-wave backgrounds, The Astrophysical Journal **941**, 119 (2022).

[27] J. A. Marck and J. P. Lasota, eds., *Relativistic gravitation and gravitational radiation. Proceedings, School of Physics, Les Houches, France, September 26-October 6, 1995* (Cambridge Univ.Pr., Cambridge, UK, 1997).

[28] B. Allen, Variance of the hellings-downs correlation, Physical Review D **107**, 043018 (2023).

[29] B. Allen and S. Valtolina, Pulsar timing array source ensembles, Phys. Rev. D **109**, 083038 (2024).

[30] W. G. Lamb and S. R. Taylor, Spectral Variance in a Stochastic Gravitational Wave Background From a Binary Population, arXiv e-prints , arXiv:2407.06270 (2024), arXiv:2407.06270 [gr-qc].

[31] G. Sato-Polito and M. Kamionkowski, Exploring the spectrum of stochastic gravitational-wave anisotropies with pulsar timing arrays, Physical Review D **109**, 123544 (2024).

[32] N. J. Cornish and J. D. Romano, When is a gravitational-wave signal stochastic?, Physical Review D **92**, 042001 (2015).

[33] B. Allen, D. Agarwal, J. D. Romano, and S. Valtolina, Source anisotropies and pulsar timing arrays, Physical Review D **110**, 123507 (2024).

[34] G. Sato-Polito and M. Zaldarriaga, Distribution of the gravitational-wave background from supermassive black holes, Physical Review D **111**, 023043 (2025).

[35] G. Sato-Polito, M. Zaldarriaga, and E. Quataert, Where are the supermassive black holes measured by ptas?, Physical Review D **110**, 063020 (2024).

[36] A. Toubiana, L. Sberna, M. Volonteri, E. Barausse, S. Babak, R. Enficiaud, D. Izquierdo-Villalba, J. Gair, J. Greene, and H. Q. Leclere, Reconciling pta and jwst, and preparing for lisa with pompoco: a parametrisation of the massive black hole population for comparison to observations, Astronomy & Astrophysics **700**, A135 (2025).

[37] M. Crisostomi, R. van Haasteren, P. M. Meyers, and M. Vallisneri, Beyond diagonal approximations: improved covariance modeling for pulsar timing array data analysis, arXiv preprint arXiv:2506.13866 (2025).

[38] V. Di Marco, A. Zic, R. M. Shannon, and E. Thrane, Systematic errors in searches for nanohertz gravitational waves, Monthly Notices of the Royal Astronomical Society **532**, 4026 (2024).

[39] V. Di Marco, A. Zic, R. M. Shannon, E. Thrane, and A. D. Kulkarni, Choosing suitable noise models for nanohertz gravitational-wave astrophysics, arXiv preprint arXiv:2502.04653 (2025).

[40] A. Sesana, A. Vecchio, and C. N. Colacino, The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays, Mon. Not. Roy. Astron. Soc. **390**, 192 (2008), arXiv:0804.4476 [astro-ph].

[41] B. Bécsy, N. J. Cornish, P. M. Meyers, L. Z. Kelley, G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, L. Blecha, *et al.*, How to detect an astrophysical nanohertz gravitational wave background, The Astrophysical Journal **959**, 9 (2023).

[42] B. Bécsy and N. J. Cornish, Joint search for isolated sources and an unresolved confusion background in pulsar timing array data, Classical and Quantum Gravity **37**, 135011 (2020).

[43] S. C. Sardesai, J. Simon, and S. J. Vigeland, Characterizing the nanohertz gravitational wave background using a t-process power spectral density, The Astrophysical Journal **976**, 212 (2024).

[44] L. Lentati, P. Alexander, M. P. Hobson, S. Taylor, J. Gair, S. T. Balan, and R. van Haasteren, Hyper-efficient model-independent bayesian method for the analysis of pulsar timing data, Physical Review D—Particles, Fields, Gravitation, and Cosmology **87**, 104021 (2013).

[45] X. Xue, Z. Pan, and L. Dai, Non-gaussian statistics of nanohertz stochastic gravitational waves, Physical Review D **111**, 043022 (2025).

[46] C. J. Moore and A. Vecchio, Ultra-low-frequency gravitational waves from cosmological and astrophysical processes, Nature Astronomy **5**, 1268 (2021).

[47] W. G. Lamb, S. R. Taylor, and R. van Haasteren, Rapid

refitting techniques for bayesian spectral characterization of the gravitational wave background using pulsar timing arrays, Physical Review D **108**, 103019 (2023).

[48] H. Quelquejay Leclere, P. Auclair, S. Babak, A. Chalumeau, D. A. Steer, J. Antoniadis, A.-S. B. Nielsen, C. Bassa, A. Berthereau, M. Bonetti, *et al.*, Practical approaches to analyzing pta data: Cosmic strings with six pulsars, Physical Review D **108**, 123527 (2023).

[49] J. Ellis, M. Fairbairn, G. Hütsi, J. Raidal, J. Urrutia, V. Vaskonen, and H. Veermäe, Gravitational waves from supermassive black hole binaries in light of the nanograv 15-year data, Physical Review D **109**, L021302 (2024).

[50] E. Phinney, A practical theorem on gravitational wave backgrounds, arXiv preprint astro-ph/0108028 (2001).

[51] M. Falxa and A. Sesana, Modeling non-gaussianities in pulsar timing array data analysis using gaussian mixture models, arXiv preprint arXiv:2508.08365 (2025).

[52] R. Hellings and G. Downs, Upper limits on the isotropic gravitational radiation background from pulsar timing analysis, Astrophysical Journal, Part 2-Letters to the Editor, vol. 265, Feb. 15, 1983, p. L39-L42. **265**, L39 (1983).

[53] B. Allen and J. D. Romano, Hellings and Downs correlation of an arbitrary set of pulsars, Phys. Rev. D **108**, 043026 (2023), arXiv:2208.07230 [gr-qc].

[54] B. Allen and J. D. Romano, Optimal Reconstruction of the Hellings and Downs Correlation, Phys. Rev. Lett. **134**, 031401 (2025), arXiv:2407.10968 [gr-qc].

[55] G. Agazie, M. F. Alam, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, L. Blecha, V. Bonidie, A. Brazier, P. R. Brook, *et al.*, The nanograv 15 yr data set: observations and timing of 68 millisecond pulsars, The Astrophysical Journal Letters **951**, L9 (2023).

[56] H. Middleton, W. Del Pozzo, W. M. Farr, A. Sesana, and A. Vecchio, Astrophysical constraints on massive black hole binary evolution from pulsar timing arrays, Monthly Notices of the Royal Astronomical Society: Letters **455**, L72 (2015).

[57] R. van Haasteren and M. Vallisneri, Low-rank approximations for large stationary covariance matrices, as used in the bayesian and generalized-least-squares analysis of pulsar-timing data, Monthly Notices of the Royal Astronomical Society **446**, 1170 (2015).

[58] M. Falxa, J. Antoniadis, D. Champion, I. Cognard, G. Desvignes, L. Guillemot, H. Hu, G. Janssen, J. Jawor, R. Karuppusamy, *et al.*, Modeling nonstationary noise in pulsar timing array data analysis, Physical Review D **109**, 123010 (2024).

[59] C. M. Mingarelli and T. Sidery, Effect of small interpulsar distances in stochastic gravitational wave background searches with pulsar timing arrays, Physical Review D **90**, 062011 (2014).

[60] L. Isserlis, On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, Biometrika **12**, 134 (1918), https://academic.oup.com/biomet/article-pdf/12/1-2/134/481266/12-1-2-134.pdf.