Non-Gaussian statistics of nanohertz stochastic gravitational waves

Xiao Xue[1,2,3,*] Zhen Pan[4,5,†] and Liang Dai[6,‡]

[1]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology,*
Campus UAB, 08193 Bellaterra (Barcelona), Spain
[2]*II. Institute of Theoretical Physics, Universität Hamburg, 22761 Hamburg, Germany*
[3]*Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*
[4]*Tsung-Dao Lee Institute, Shanghai Jiao-Tong University, Shanghai,*
520 Shengrong Road, 201210, People's Republic of China
[5]*School of Physics and Astronomy, Shanghai Jiao-Tong University, Shanghai,*
800 Dongchuan Road, 200240, People's Republic of China
[6]*Department of Physics, University of California,*
366 Physics North MC 7300, Berkeley, California 94720, USA

Multiple pulsar timing arrays (PTAs) have recently reported evidence for nHz stochastic gravitational wave background (SGWB), stimulating intensive discussions about its physical origin. In principle, the sources may be either supermassive black hole binaries (SMBHBs) or processes in the early Universe. One key difference between the two lies in the statistics of the SGWB frequency power spectrum. In particular, the often assumed Gaussian random SGWB does not accurately describe the distribution of the collective SMBHB emission. This work presents a semianalytical framework for calculating the non-Gaussian statistics of SGWB power expected from SMBHBs. We find that (a) wave interference between individual SMBHBs with indistinguishable observed frequencies and (b) the Poisson fluctuation of the source numbers, together shape the non-Gaussian statistics. Implementing the non-Gaussian statistics developed in this work, we investigate the sensitivity of current and future PTA datasets in distinguishing the origin of the SGWB through non-Gaussian information. Additionally, we find an interesting approximation of the non-Gaussian statistics, which has implications for accurately and practically treating non-Gaussianity in PTA Bayesian analyses.

I. INTRODUCTION

The recent detection of nanohertz (nHz) stochastic gravitational wave background (SGWB) by pulsar timing arrays (PTAs) [1–8] has inspired intensive discussions of its astrophysical implications. Supermassive black hole binaries (SMBHBs) are a promising source of SGWB, from which we may infer the abundance and evolution history of the cosmic SMBHB population [9–16]. On the other hand, various early-Universe processes have been speculated to be alternative nHz SGWB sources (see [17] for a summary).

In principle, SMBHBs and early Universe sources can be distinguished by the different statistical properties of the SGWB they produce. SMBHBs at redshifts $z < 1$ contribute significantly to the SGWB at nanohertz frequencies, characterized by strong signal strength due to their proximity, and they emit mostly monochromatic gravitational waves (GWs) over decades of observation. In contrast, early Universe sources, produced at high redshift, are heavily redshifted and contribute minimally due to their larger distance. These sources usually emit over a broad frequency range. A much larger number of sources would be required for early Universe sources to contribute to the same signal power as SMBHBs at the same observed frequency. Thus, it is usually assumed that the latter produces a Gaussian and isotropic SGWB, while the former is certainly non-Gaussian due to Poissonian fluctuations of a finite number of SMBHBs and may show random power anisotropy if there are nearby loud SMBHBs. Along this line, there have been some efforts towards measuring the spectral variance beyond Gaussian fluctuations [18–24], or anisotropies of the nHz SGWB [25–32]. In these analyses, spectral variance is usually quantified or parametrized based on SMBHB population synthesis. Specially, Refs. [33,34] derived the probability distribution of the

---
*Contact author: xxue@ifae.es
†Contact author: zhpan@sjtu.edu.cn
‡Contact author: liangdai@berkeley.edu

characteristic strain $P(h_c^2)$ given a SMBHB population model, which enabled a quantitative constraint on SMBHB population models using the PTA data [34]. For the study of anisotropy, forecasts are made based on SMBHB population simulations, and measurements from the PTA data are performed under the simplifying assumption of a Gaussian random SGWB. Spectral variance and anisotropy are two different manifestations of Poisson fluctuations of a finite number of SMBHBs, in the intrinsic (e.g., GW amplitude and frequency) and extrinsic (e.g., sky localization) parameter spaces, respectively. Focusing on a different aspect, Allen *et al.* [35,36] calculated the variance of the Hellings-Downs correlation [37] for a simple population of SMBHBs, taking into account interference between GW sources of overlapping frequencies. They demonstrated two sources of variance: one arising from the finite number of pulsar pairs being used and another cosmic variance due to interference effects.

As summarized above, different aspects of nHz SGWB sourced by a finite number of SMBHBs are related. Both spectral variance and anisotropy originate from Poissonian fluctuation in the number of SMBHBs and are affected by interference, and deviation from the Hellings-Downs correlation can be result from interference [35,36,38,39] or spatial anisotropy [40–42]. In previous works, each aspect has only been investigated separately, assuming other aspects are known or independent. Here, we consider a unified framework for dealing with Poissonian fluctuations in the number of SMBHBs and GW interference. The major challenge is efficiently computing the non-Gaussian probability distribution $P(\delta z)$ for general SMBHB population models, where δz is observed photon redshift from a given single pulsar in the network.

This work demonstrates an efficient method for computing $P(\delta z)$ that accounts for the Poisson statistics and interference. The key is to use the cumulant generating function for δz (see Sec. II for details). As a result, we find that interference is the major source of variance in $|\delta z|^2$. At the same time, the non-Gaussianity originates from the Poisson fluctuation in the source number (see Fig. 4). With the non-Gaussian statistics $P(\delta z)$ developed in this work, we estimate the likelihood ratio against the conventionally used Gaussian statistics and find that the current PTA dataset is capable of finding substantial evidence of non-Gaussian statistics (Fig. 5). However, we find narrowing down the population model parameters challenging due to the significant statistical variance caused by interference (Fig. 6) given the current PTA sensitivity. As we will see, a more extended observation period and a lower timing noise level expected from future PTA data will enable substantially improved constraints on population model parameters. Under current PTA sensitivity, we find the Gaussian statistics are still a good approximation in inferring the strain power spectrum $h_c^2(f)$ of the SGWB from SMBHBs (Fig. 7). As PTA data of lower timing noise accumulates in

the foreseeable future, we find the Gaussian statistics will eventually bias the inference of $h_c^2(f)$ (Fig. 7).

This paper is organized as follows. In Sec. II, we briefly introduce compound Poisson statistics and the corresponding cumulant generating function (CGF) are key to computing the exact PTA signal distribution $P(\delta z)$ for SMBHB population models. In Sec. III, we introduce a parametrized form of the SMBHB population model to be considered in this work and provide details about calculating the redshift distribution $P(\delta z)$. Building on these, we perform a likelihood ratio test comparing non-Gaussian and Gaussian statistics. Concluding remarks will be given in Sec. IV. Throughout this work, we adopt a flat Λ cold dark matter cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km s^{-1} Mpc^{-1}.

II. COMPOUND POISSON STATISTICS

In this section, we develop the mathematical formalism for computing the probability density function (PDF) of the PTA observable, $P(\delta z)$. The basic idea was also explored in Ref. [34]. Assuming that individual gravitational wave sources are monochromatic over the observational period of PTAs, each source can be characterized by its current observed frequency f_{GW} and another set of M parameters collectively referred to as Θ. The following differential distribution can describe a general source population:

$$\frac{d^{1+M}\bar{N}}{d^M\Theta d\ln f}.\qquad(1)$$

The average source number in a logarithmic frequency interval $\Delta \ln f$ and in a multidimensional parameter-space volume element $\Delta^M\Theta$ is thus given by $\Delta\bar{N}$.

The actual source number ΔN is expected to be a random number that fluctuates around its expectation value $\langle\Delta N\rangle = \Delta\bar{N}$. It follows Poisson statistics, i.e.,

$$\Delta N \sim \text{Pois}(\Delta\bar{N}).\qquad(2)$$

Source numbers in nonoverlapping volumes of the parameter space are independent random numbers.

For any signal $s(\Theta, f)$ that obeys the superposition principle, the population-summed signal $\mathcal{S}$ is a weighted sum over the entire source parameter space:

$$\mathcal{S} = \sum_n s_n \Delta N_n,\qquad(3)$$

where we divide the source parameter space into individual blocks indexed by n, and $s_n = s(\Theta_n, f_{GW}^n)$ is the corresponding weight. Linearity of Eq. (3) suggests that the total signal $\mathcal{S}$ is drawn from a compound Poisson distribution. In the special case where the weight $s(\Theta, f)$ is a constant, the total signal also follows a Poisson distribution. Unfortunately, we are not aware of an analytical expression

for the PDF of $\mathcal{S}$, $P(\mathcal{S})$, for the case of general weights $s(\Theta, f)$.

Instead, we pursue a numerical method to evaluate the PDF $P(\mathcal{S})$ that makes use of the CGF for $\mathcal{S}$, defined as

$$K_{\mathcal{S}}(t) = \ln \langle \exp(i\mathcal{S}t) \rangle, \tag{4}$$

where t is the variable conjugate to $\mathcal{S}$ and is defined on the real axis from $-\infty$ to ∞, and the notation $\langle \cdots \rangle$ stands for the statistical average. Since ΔN_n's are statistically independent of each other, Eq. (4) can be evaluated as follows:

$$K_{\mathcal{S}}(t) = \sum_n \ln \langle \exp\left(is_n \Delta N_n t\right) \rangle = \sum_n \Delta \bar{N}_n(e^{is_n t} - 1), \tag{5}$$

where the last step follows from the CGF of Poisson distribution. Equation (5) then allows us to numerically evaluate the PDF of $\mathcal{S}$ through an inverse Fourier transformation,

$$P(\mathcal{S}|\Theta, f) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt \exp\left(i\mathcal{S}t + [K_{\mathcal{S}}(t)]^*\right), \tag{6}$$

where $[\cdots]^*$ stands for complex conjugation. The PDF evaluated in this way is correctly normalized.

If the source parameter space is continuously parametrized by (Θ, f), then Eq. (5) can be revised to the appropriate continuous limit, where summation is replaced by integration

$$\sum_n \Delta N_n \longrightarrow \int_V d \ln f d^M \Theta \frac{d^{1+M}\bar{N}}{d^M \Theta d \ln f}. \tag{7}$$

The CGF converges in this limit, giving a unique answer for the PDF of the signal $\mathcal{S}$.

It is straightforward to generalize the above framework to the case where a set of $\mathcal{N}$ different signals ($\mathcal{N} > 1$), collectively denoted as $\mathcal{S}$, are measured. The corresponding CGF depends on a set of $\mathcal{N}$ conjugate variables, which we collectively refer to as t. The CGF is given by

$$K_{\mathcal{S}}(t) = \sum_n \ln \langle \exp\left(is_n \cdot t \Delta N_n\right) \rangle = \sum_n \Delta \bar{N}_n(e^{is_n \cdot t} - 1), \tag{8}$$

and the multivariate PDF, in principle, can be found from the multidimensional Fourier transformation,

$$P(\mathcal{S}|\Theta, f) = \frac{1}{(2\pi)^{\mathcal{N}}} \int d^{\mathcal{N}}t \exp\left(i\mathcal{S} \cdot t + [K_{\mathcal{S}}(t)]^*\right). \tag{9}$$

If an observable $\mathcal{S}$ is also contaminated by noise n, and if the noise is statistically uncorrelated with the signal, then the total measurement $\mathcal{S} + n$ has a CGF

$$K_{\mathcal{S}+n}(t) = K_{\mathcal{S}}(t) + K_n(t), \tag{10}$$

where $K_n(t)$ is the CGF for the noise n. The corresponding PDF is essentially the likelihood function for obtaining data $\mathcal{S} + n$ given a model with model parameters $\{\Theta, \varsigma\}$:

$$P(\mathcal{S} + n | \Theta, \varsigma), \tag{11}$$

where ς is a set of parameters that characterize the noise. Based on the above general results following the probability theory, we will develop a framework to perform Bayesian inference of GW source population parameters.

III. NON-GAUSSIAN STATISTICS

In Sec. III A, we will first introduce a convenient parametrized form for the SMBHB population following [11],

$$\frac{d^3 \bar{N}}{d \ln f d \log_{10} \mathcal{M} dz}(\Theta). \tag{12}$$

This gives the SMBHB number density in the phase space defined jointly by the GW frequency f, the binary chirp mass $\mathcal{M}$, and the cosmological redshift z. We then introduce some basics of the PTA observable, redshift of photons δz propagating from a pulsar to the earth under the influence of GWs (Sec. III B), and the details for computing the non-Gaussian redshift PDF, $P(\delta z)$, given a SMBHB population model (Sec. III C). With the redshift PDF, we compare the non-Gaussian statistics with the conventionally used Gaussian statistics in Sec. III D.

A. SMBHB population

Supermassive black holes (SMBHs) are commonly found at the centers of their host galaxies. Galaxy mergers inevitably bring multiple SMBHs into the same postmerger galaxy. SMBHBs are the most promising sources for the nHz SGWB observed at PTAs, though the detailed processes for the formation and mergers of SMBHBs still need to be fully understood. The mass of the central SMBH is empirically known to correlate strongly with the velocity dispersion in the host galaxy bulge. This is the famous $M_{\rm BH} - \sigma$ relation [43–45]. We model the mean $M_{\rm BH} - \sigma$ relation, together with the expected scatter around it, with the following log-normal probability distribution

$$P(\log_{10} M_{\rm BH} | \sigma) = \frac{1}{\sqrt{2\pi} \epsilon_0} \exp \left\{ -\frac{1}{2\epsilon_0^2} \left[\log_{10} \frac{M_{\rm BH}}{M_\odot} - \log_{10} 10^{a_\bullet} \left(\frac{\sigma}{200 \, {\rm km \, s^{-1}}} \right)^{b_\bullet} \right]^2 \right\}, \tag{13}$$

where $a_\bullet$ and $b_\bullet$ are constants, and ϵ_0 quantifies the intrinsic scatter. On the other hand, the velocity dispersion function (VDF) of the galaxy bulge is parametrized as

$$\phi(\sigma) = \frac{\phi_* \beta}{\sigma_*} \left[\frac{\sigma}{\sigma_*}\right]^{\alpha-1} \frac{e^{-(\sigma/\sigma_*)^\beta}}{\Gamma(\alpha/\beta)}. \qquad (14)$$

Here α, β are dimensionless constants, σ_* characterizes the turnover of the velocity dispersion, and ϕ_* is a normalization constant, which also sets the total number of the galaxies per Mpc^3. Taking these into account, the SMBH mass function is

$$\frac{dn}{d \log_{10} M_{BH}} = \int d\sigma P(\log_{10} M_{BH} | \sigma) \phi(\sigma), \qquad (15)$$

where n is the total number of black holes per Mpc^3.

We assume that all SMBHBs in the PTA band are circular inspirals. This assumption is expected to be true when GW radiation dominates the hardening of the binary orbit, as GW emission tends to circularize the orbit. For circular inspirals, the chirp mass $\mathcal{M}$ and source redshift z determine the observed GW amplitude. Chirp mass $\mathcal{M}$ is defined as

$$\mathcal{M} = \eta^{3/5} M_{BH}, \qquad \eta \equiv \frac{q}{(1+q)^2}, \qquad (16)$$

where $q < 1$ is the binary mass ratio. We express the population model as follows:

$$\frac{dn}{d\log_{10}\mathcal{M}} = \int_{q_{min}}^{1} dq \, p_q(q) \frac{dn}{d\log_{10}M_{BH}}, \qquad (17)$$

$p_q(q) = \mathcal{N}_q q^\delta$ and $\mathcal{N}_q$ are the normalization constants. The above expression assumes one merger per galaxy.

Only SMBHBs located on our past light cone are observed through the SGWB. The relevant comoving volume is given by dV_c/dt,

$$\frac{dV_c}{dt_r} = \frac{dV_c/dz}{dt_r/dz} = 4\pi c d_L(z)^2 (1+z), \qquad (18)$$

where t_r is the proper time of a comoving observer in Friedmann-Robertson-Walker cosmology, and $d_L(z) = (c(1+z)/H_0) \int_0^z dz'/E(z')$ is the luminosity distance to redshift z, where $E(z) = \sqrt{(1+z)^3 \Omega_m + \Omega_\Lambda}$. Assuming the redshift dependence is separable, the SMBHB population can be written as [46]

$$\frac{d^3\bar{N}}{d\log_{10}\mathcal{M}dz d\ln f_r} = p_z(z) \frac{dn}{d\log_{10}\mathcal{M}} \frac{dt_r}{d\ln f_r} \frac{dV_c}{dt_r}, \qquad (19)$$

where $f_r = (1+z)f$ is GW frequency in the rest frame of the SMBHB barycenter, $p_z(z) = \mathcal{N}_z z^\gamma e^{-z/z_*}$, $dt_r/d\ln f_r$ is determined by the binary orbital hardening rate, z_* is the turnover redshift, and $\mathcal{N}_z$ is a normalization constant. We adopt the simplifying assumption that the spectral shape of the mass function is independent of redshift z. We introduce

ξ to the frequency dependence

$$\frac{dt_r}{d\ln f_r} = \frac{5}{96(G\mathcal{M}/c^3)^{5/3}(\pi f_{yr})^{8/3}(f_r/f_{yr})^{8/3-\xi}}, \qquad (20)$$

especially $\xi = 0$ if the hardening is driven by GW emission only. To summarize, in our parametrized SMBHB population model, $\{\epsilon_0, a_\bullet, b_\bullet\}$ characterize the $M_{BH} - \sigma$ relation, $\{\phi_*, \sigma_*, \alpha, \beta\}$ characterize the VDF, and $\{\delta, \gamma, z_*, \xi\}$ characterize the SMBHB distribution with respect to binary mass ratio, redshift, and frequency.

Following Ref. [11], we set fiducial values for the population parameters: $\sigma_* = 159.6 \, km s^{-1}$, $\{\alpha, \beta,; a_\bullet, b_\bullet; \gamma, z_*; \delta; \xi\} = \{0.41, 2.59; 8.32, 5.64; 0.3, 0.5; -1; 0\}$. We leave ϕ_* and ϵ_0 as free parameters. These choices are justified by studies on the velocity distribution function [47], on the $M - \sigma$ relation [45], and SMBHB population simulations [48]. Furthermore, in this work, we consider an expanded population model that accounts for source dependence on extra binary parameters in addition to $\mathcal{M}$ and z,

$$\Lambda = \{\log_{10}\mathcal{M}, z\} \cup \{\cos\theta, \phi; \cos\iota, \psi; \varphi\}, \qquad (21)$$

where ι is the binary orbit inclination, ψ is the polarization angle, θ and ϕ parametrize the source's sky location, and φ is a binary orbital phase constant. It is reasonable to assume that binaries have no preference regarding spatial location and orbital orientation. Therefore, the extended population model is

$$\frac{d^8\bar{N}}{d\ln f_r d^7\Lambda} = \frac{1}{32\pi^3} \frac{d^3\bar{N}}{d\ln f_r d\log_{10}\mathcal{M}dz}. \qquad (22)$$

Next, we will look into the intensity of GWs from the circular SMBHBs. For a slowly-evolving binary with chirp mass $\mathcal{M}$ and redshift z, the GW tensor at the observer $\boldsymbol{x} = 0$ can be written as

$$h_{ab}(t, 0) = h_+(t, 0)\epsilon_{ab}^+ + h_\times(t, 0)\epsilon_{ab}^\times, \qquad (23)$$

where the polarization tensor $\epsilon_{ab}^{+,\times}(\theta, \phi, \psi)$ depends on the source sky location (θ, ϕ) and the source polarization angle ψ [49]. The two polarization states have amplitudes

$$h_+(t, 0) = h_0 \frac{1 + \cos^2\iota}{2}\cos(2\pi ft + \varphi),$$
$$h_\times(t, 0) = h_0 \cos\iota \sin(2\pi ft + \varphi), \qquad (24)$$

where the dimensionless amplitude is

$$h_0(f, \mathcal{M}, z) = \frac{4c(\pi f_{yr})^{2/3}(f_r/f_{yr})^{2/3}(G\mathcal{M}/c^3)^{5/3}}{d_L(z)}. \qquad (25)$$

A commonly used quantity is the characteristic strain of the GWs, the expected value of which from a SMBHB
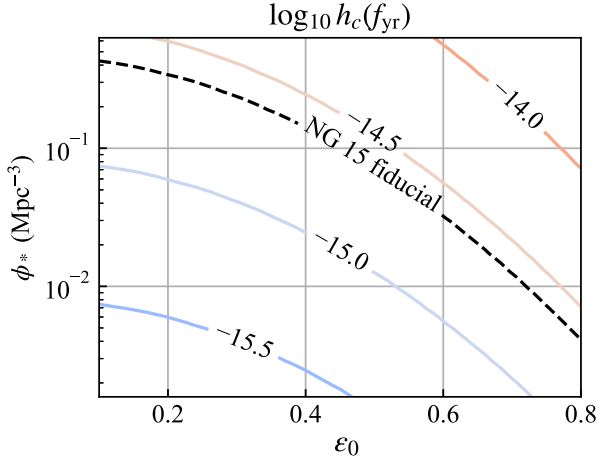
FIG. 1.   The contours of $\log_{10}(h_c)(f_{\rm yr})$ on $\phi_* - \epsilon_0$ plane. The parameters consistent with the 15-year NANOGrav data are marked as the black dashed line. We fix all the other parameters at their fiducial values (given in the text).

population can be written as [46,50]

$$
h_c^2(f) = \int_0^{+\infty} dz \int_{-\infty}^{+\infty} d\log_{10}\mathcal{M} \langle h_+^2 + h_\times^2 \rangle_{\cos\iota,\varphi}
$$
$$
\times \frac{d^3\bar{N}}{d\log_{10}\mathcal{M}dz d\ln f_r}, \tag{26}
$$

where

$$
\langle h_+^2 + h_\times^2 \rangle_{\cos\iota,\varphi} = \frac{2}{5} h_0^2 \tag{27}
$$

is the GW amplitude squared averaged over random binary orbital inclination and phase constant. In Fig. 1, we show contours of $h_c(f_{\rm yr})$ on the ϕ_*–ϵ_0 plane, as well as the specific contour that corresponds to the best-fit value from the 15-year NANOGrav data [1], $h_c(f_{\rm yr}) = 2.4 \times 10^{-15}$.

B. Pulsar redshift

As radio photons travel from a pulsar to the Earth, the arrival time is perturbed by the SGWB. For a planar GW propagating in the direction $\hat{\mathbf{\Omega}}$ (with $-\hat{\mathbf{\Omega}}$ being the direction pointing from the Earth to the GW source), the GW tensor is generally written as

$$
h_{ab}(t,\mathbf{x}) = h_+(t - \mathbf{x} \cdot \hat{\mathbf{\Omega}})\epsilon_{ab}^+ + h_\times(t - \mathbf{x} \cdot \hat{\mathbf{\Omega}})\epsilon_{ab}^\times. \tag{28}
$$

Considering a pulsar with a sky location $\hat{p}$ and at a distance L from the earth, the observed photon redshift (or the fractional change in apparent pulsar spin period) for the observer at the coordinate origin $\mathbf{x} = 0$ is [37]

$$
\delta z(t) = \delta z_{\rm E} - \delta z_{\rm P}
$$
$$
= \frac{1}{2}\frac{\hat{p}^a \hat{p}^b}{1 + \hat{\mathbf{\Omega}} \cdot \hat{p}}[h_{ab}(t,0) - h_{ab}(t - L, L\hat{p})]. \tag{29}
$$

Here, the subscripts E and P stand for the Earth and the pulsar terms. For a slowly evolving SMBHB where the GW frequency evolution during the PTA observational period T is negligible (with $2\pi \dot{f}T < T^{-1}$), the two polarization components in the Earth term are in the form

$$
h_+(t,0) = h_0 \frac{1 + \cos^2\iota}{2} \cos(2\pi f_{\rm E}t + \varphi),
$$
$$
h_\times(t,0) = h_0 \cos\iota \sin(2\pi f_{\rm E}t + \varphi), \tag{30}
$$

where the GW amplitude and the frequency $f_{\rm E}$ are taken as a constant during the observational period $(0, T)$. Similarly, we have for the pulsar term,

$$
h_+(t - L, L\hat{p}) = h_0 \frac{1 + \cos^2\iota}{2}\cos(2\pi f_{\rm P}t + \varphi - \Delta),
$$
$$
h_\times(t - L, L\hat{p}) = h_0 \cos\iota \sin(2\pi f_{\rm P}t + \varphi - \Delta), \tag{31}
$$

where the phase difference $\Delta = c^{-1}\int_0^L 2\pi f(t)dt(1 + \hat{\mathbf{\Omega}} \cdot \hat{p}) \approx \pi(f_{\rm E} + f_{\rm P})L(1 + \hat{\mathbf{\Omega}} \cdot \hat{p})/c$, and $f_{\rm P}$ is also taken as a constant during the observational period $(0, T)$. We note that for a given pulsar, the frequency difference between the Earth term and the pulsar term, $f_{\rm E} - f_{\rm P}$, can be significant and measurable if $2\pi \dot{f}L/c > T^{-1}$. For SMBHB at nHz frequency, $f_{\rm E}$ and $f_{\rm P}$ are indistinguishable.

In practice, the PTA data analysis is more straightforward in the frequency domain, where the redshifts $\delta z(f)$ in different frequency bins are statistically uncorrelated. We first divide the relevant frequency range into bins, with central frequencies f_n ($n = 1, \ldots, n_{\max}$) and equal bin sizes T^{-1}. The GW amplitude from a SMBHB in a frequency bin of central frequency f_n is

$$
h_{ab}(f_n; f, \mathbf{\Theta}) = TW(f - f_n)h_0(f, \mathcal{M}, z)e^{i\varphi}
$$
$$
\times \left[i\frac{1 + \cos^2\iota}{2}\epsilon_{ab}^+ + \cos\iota\epsilon_{ab}^\times\right], \tag{32}
$$

where f is the observed GW frequency from the binary, and W is the top-hat window function,

$$
W(f - f_n) = \begin{cases} 1, & f_n - 1/2T < f < f_n + 1/2T \\ 0, & \text{otherwise.} \end{cases} \tag{33}
$$

As Eq. (29) shows, two terms contribute to the redshift of a photon emitted from a pulsar and received on the Earth. However, it is widely accepted that only the Earth term exhibits interpulsar correlation, namely the Hellings-Downs curve [51]. Here, we assume that we successfully subtracted the Earth term. Taking the Earth term,

$$\delta z_{\rm E}(f) = -\frac{1}{2T}\frac{p_a h_{ab}(x_{\rm E}, f) p_b}{1 + \hat{\bf \Omega} \cdot \hat{\bf p}}. \tag{34}$$

The expression above is dimensionless by introducing the prefactor T^{-1}. Without loss of generality, we take $\hat{\bf p} = (0, 0, 1)$ for simplicity, then the real and imaginary components of $\delta z_{\rm E}$ are

$$\delta z_{\rm E,Re}^{\rm circ} = h_0 \lambda_z(\theta, \iota, \psi, \varphi),$$
$$\delta z_{\rm E,Im}^{\rm circ} = h_0 \lambda_z(\theta, \iota, \psi, \varphi + \pi/2), \tag{35}$$

where

$$\lambda_z(\theta, \iota, \psi, \varphi) = \sin^2\frac{\theta}{2}\left[\left(\frac{1 + \cos\iota^2}{2}\right)\cos(2\psi)\sin\varphi\right.$$
$$\left. - \cos\iota \sin(2\psi)\cos\varphi\right]. \tag{36}$$

For a population of SMBHBs, the total redshift is therefore

$$\delta z_{\rm E} = \sum_n \Delta N_n (\delta z_{\rm E}^{\rm circ})_n, \tag{37}$$

where we divide the source parameter space into individual blocks indexed by n. The distribution $P(\lambda_z)$ is evaluated using the Monte Carlo method by uniformly sampling $\cos\theta$, $\cos\iota$, ψ, and φ, and we store numerical values for further use. The mean value of any arbitrary function $F(\lambda_z)$ is equivalent to the following integration:

$$\langle F(\lambda_z)\rangle := \int d\lambda_z P(\lambda_z) F(\lambda_z)$$
$$= \frac{1}{32\pi^3}\int_{-1}^1 d\cos\theta \int_{-1}^1 d\cos\iota \int_0^{2\pi} d\phi \int_0^{2\pi} d\psi$$
$$\times \int_0^{2\pi} d\varphi F(\lambda_z(\theta, \iota, \psi, \varphi)). \tag{38}$$

Then it is straightforward to find $\langle\lambda_z\rangle = 0$ and $\langle\lambda_z^2\rangle = 1/15$.

If a SMBHB is slowly evolving, with the Earth term and the pulsar term contributing to the same frequency bin, then we find

$$\delta z^{\rm circ} = \delta z_{\rm E}^{\rm circ} - \delta z_{\rm P}^{\rm circ}$$
$$= h_0\left[\tilde{\lambda}_z(\theta, \iota, \psi, \varphi, \Delta) + i\tilde{\lambda}_z(\theta, \iota, \psi, \varphi + \pi/2, \Delta)\right], \tag{39}$$

where

$$\tilde{\lambda}_z(\theta, \iota, \psi, \varphi, \Delta) = 2\sin(\Delta/2)\lambda_z(\theta, \iota, \psi, \varphi - \Delta/2). \tag{40}$$

It is clear that the pulsar term further increases the variance of the redshift. The calculation of the redshift distribution, in this case, is completely parallel to the previous case with

the replacement $\lambda_z(\theta, \iota, \psi, \varphi) \to \tilde{\lambda}_z(\theta, \iota, \psi, \varphi, \Delta)$. In the next subsection, we will focus on the PDF of the redshift $P(\delta z_{\rm E}(f))$.

C. Distribution of pulsar redshift

Using the method developed in Sec. II, we can now calculate the PDF of the redshift $P(\delta z_{\rm E}(f))$ given a SMBHB population model. We first assemble the real and imaginary parts into a vector $s = (\delta z_{\rm E,Re}^{\rm circ}, \delta z_{\rm E,Im}^{\rm circ})$. The total signal is a superposition of many individual sources and the conjugate variables also into a vector $t = (t_{\rm Re}, t_{\rm Im})$, then

$$s \cdot t = \delta z_{\rm E,Re}^{\rm circ} t_{\rm Re} + \delta z_{\rm E,Im}^{\rm circ} t_{\rm Im} = t_\rho h_0 \lambda_z(\theta, \iota, \psi, \varphi - t_\varphi), \tag{41}$$

where $t_{\rm Re}$ and $t_{\rm Im}$ are real variables conjugate to $\delta z_{\rm E,Re}^{\rm circ}$ and $\delta z_{\rm E,Im}^{\rm circ}$, $t_\varphi = \arctan(t_{\rm Im}/t_{\rm Re})$, $t_\rho = \sqrt{t_{\rm Re}^2 + t_{\rm Im}^2}$. Because of the uniform distribution of φ, the CGF does not depend on the variable t_φ. The two-dimensional (2D) CGF turns out to be

$$K_{\delta z}(t_\rho) = \Delta \ln f \iiint d\log_{10}\mathcal{M} dz d\lambda_z P(\lambda_z)$$
$$\times [\cos(t_\rho h_0 \lambda_z) - 1]\frac{d^3\bar{N}}{d\log_{10}\mathcal{M} dz d\ln f_r}. \tag{42}$$

We find that $K_{\delta z}(t_{\rm Re}, t_{\rm Im}) = K_{\delta z}(t_\rho)$ is a real function because $P(\lambda_z)$ is even, and $K_{\delta z}(t_{\rm Re}, t_{\rm Im})$ only depends on the magnitude of the conjugate vector t_ρ.

The cumulants of the redshift can be inferred from the CGF $K_{\delta z}$. First, it is easy to verify that all odd order of cumulants vanish because $P(\lambda_z)$ is even, and the higher and even order of cumulants, namely $\langle\delta z_{\rm E,Re}^{2k}\rangle$ with $k > 1$, do not exist because of the divergence at the nearby end (cosmology redshift $z \to 0$) of the SMBHB population, which implies that the corresponding PDF of δz is a heavy-tailed distribution. The only nonzero and finite cumulant equals the variance of the redshift $\delta z_{\rm E,Re}$ and $\delta z_{\rm E,Im}$, with

$$\langle\delta z_{\rm E,Re}^2\rangle = \langle\delta z_{\rm E,Im}^2\rangle = -K_{\delta z}''(0). \tag{43}$$

In particular, $K_{\delta z}''(t_\rho)$ is free of divergence at the nearby end of the SMBHB population:

$$K_{\delta z}''(t_\rho) = -\Delta \ln f \iiint d\log_{10}\mathcal{M} dz d\lambda_z P(\lambda_z)$$
$$\times h_0^2 \lambda_z^2 \cos(t_\rho h_0 \lambda_z)\frac{d^3\bar{N}}{d\log_{10}\mathcal{M} dz d\ln f_r}, \tag{44}$$

where it is straightforward to find $0 \le |K_{\delta z}''(t_\rho)/K_{\delta z}''(0)| \le 1$ using Cauchy-Schwarz inequality, and $K''(t_\rho \to +\infty) = 0$ due to the fast oscillatory integrand.
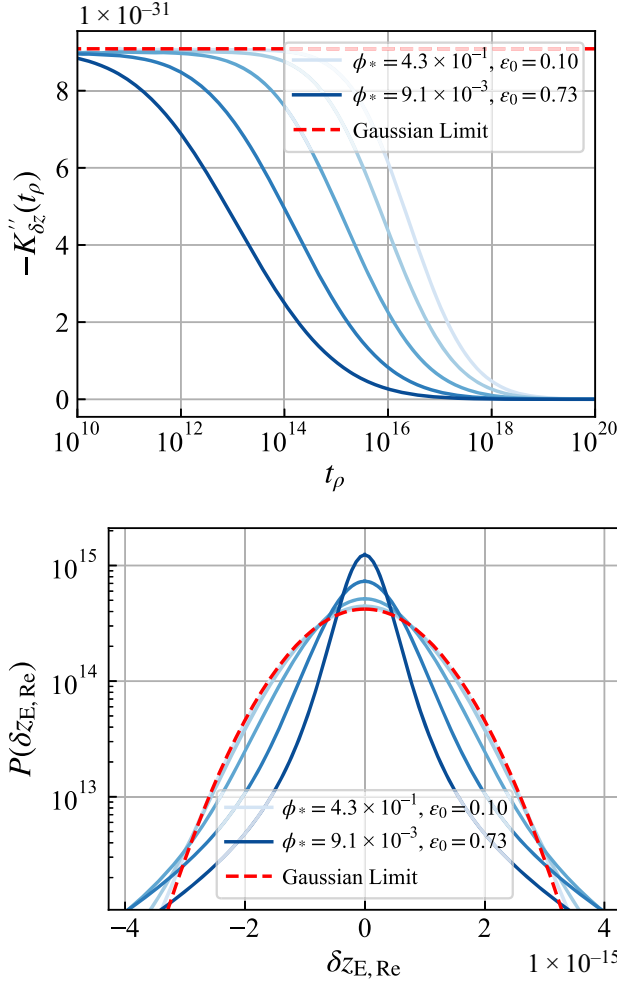
FIG. 2. Top: the second derivative of the CGF. Bottom: the PDF of the observed redshift $\delta z_{E,Re}$. Note that all PDFs in this plot share the same variance, $\langle \delta z_{E,Re}^2 \rangle$. We consider $f = 10$ nHz with a bandwidth $\Delta f = 2$ nHz. Curves of lighter colors correspond to SMBHB populations with more SMBHBs but lower black hole masses. As the number of individual SMBHB increases, the CGF and the PDF approach the Gaussian limit.

In comparison, in the Gaussian limit $K_{\delta z}^{Gaus}(t_\rho) \propto -t_\rho^2/2$, meaning $K_{\delta z}^{Gaus\prime\prime}(t_\rho)$ equals a constant. To recover the Gaussian limit, we can artificially rescale $d\bar{N} \to x d\bar{N}$ and $h^2 \to h^2/x$ and take the limit of $x \to +\infty$, which corresponds to the limit of an infinite number of arbitrarily weak GW sources. That is to say,

$$K_{\delta z}^{Gaus\prime\prime}(t_\rho) = K_{\delta z}^{\prime\prime}(0) = -\frac{1}{6}h_c^2 \Delta \ln f, \qquad (45)$$

where we have used Eqs. (26), (27), and (38).

In the left panel of Fig. 2, we show the numerical results of $K''(t_\rho)$ for a range of SMBHB populations with different population parameters (ϕ_*, ϵ_0) that are consistent with 15-year NANOGrav data (see the black-dashed line in Fig. 1), all having the same variance $-K_{\delta z}''(0)$

[Eqs. (43) and (45)]. As expected, $K''(t_\rho)$ is closer to the Gaussian limit for a larger number of GW sources (higher ϕ_*). The comparison demonstrates the origin of the non-Gaussian nature of the SGWB from SMBHBs.

The 2D probability distribution, properly normalized, can be expressed in terms of an integral

$$
\begin{aligned}
P(\delta z_{E,Re}, \delta z_{E,Im}) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} dt_{Re} \\
&\times \int_{-\infty}^{\infty} dt_{Im} e^{[K_{\delta z}(t_\rho)]^* + it_{Re}\delta z_{E,Re} + it_{Im}\delta z_{E,Im}} \\
&= \frac{1}{2\pi} \int_0^{+\infty} dt_\rho t_\rho e^{[K_{\delta z}(t_\rho)]^*} J_0(t_\rho |\delta z_E|) \\
&:= f(|\delta z_E|).
\end{aligned}
\qquad (46)
$$

In practice, we define $t_\rho = e^\tau$ and numerically implement the integral using the τ variable instead. It is important to notice that $\delta z_{E,Re}$ and $\delta z_{E,Im}$ are generally not statistically independent, i.e., $P(\delta z_{E,Re}, \delta z_{E,Im}) \neq P(\delta z_{E,Re})P(\delta z_{E,Im})$. The marginalized distribution of $\delta z_{E,Re}$ is

$$P(\delta z_{E,Re}) = 2 \int_0^{+\infty} f(|\delta z_E|) d\sqrt{|\delta z_E|^2 - \delta z_{E,Re}^2}. \qquad (47)$$

On the right panel of Fig. 2, we show $P(\delta z_{E,Re})$ for several $\phi_* - \epsilon_0$ pairs (same as those in the left panel). We note that the true distribution $P(\delta z_{E,Re})$ is both heavy tailed and strong peaked.

The PDF of $|\delta z_E|^2$ is

$$P(|\delta z_E|^2) = \pi f(|\delta z_E|), \qquad (48)$$

and is shown in Fig. 3.

So far, the calculation in this section is based on Eq. (37), which includes interference between individual sources. Here, we consider another commonly studied quantity—the signal power summed incoherently over sources:

$$|\delta z_E|_{NI}^2 \equiv \sum_n \Delta N_n \langle |\delta z_E^{circ}|_n^2 \rangle_{\cos\theta,\phi,\cos\iota,\psi,\varphi}, \qquad (49)$$

where the source parameter space is defined only by $\mathcal{M}$, z, and f_r, and $\langle |\delta z_E^{circ}|_n^2 \rangle_{\cos\theta,\phi,\cos\iota,\psi,\varphi} = \frac{2}{15}h_0(f_n, \mathcal{M}_n, z_n)^2$, the subscript "NI" stands for "No Interference." It is easy to prove that $\langle |\delta z_E|_{NI}^2 \rangle = \langle |\delta z_E|^2 \rangle$. However, comparing to the true $|\delta z_E|^2$, cross terms such as $(\delta z_E^{circ})_n (\delta z_E^{circ})_{n'}$ with $n \neq n'$, which correspond to interference between sources, are absent in the NI counterpart. The corresponding CGF is

$$
\begin{aligned}
\tilde{K}_{|\delta z|^2}(t) &\equiv \Delta \ln f \iint d\log_{10}\mathcal{M} dz \left[\exp\left(\frac{2}{15}ith_0^2\right) - 1 \right] \\
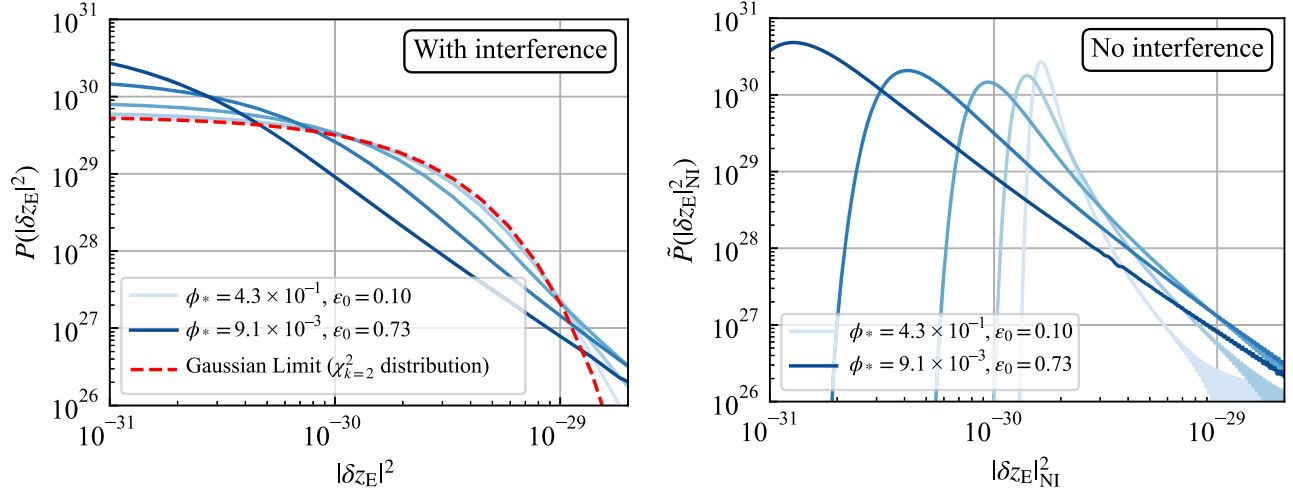&\times \frac{d^3\bar{N}}{d\log_{10}\mathcal{M} dz d\ln f_r},
\end{aligned}
\qquad (50)
$$

FIG. 3. The PDF of $|\delta z_E|^2$ and $|\delta z_E|^2_{NI}$. We consider $f = 10$ nHz in both subplots, with a bandwidth $\Delta f = 2$ nHz. Curves of lighter colors correspond to SMBHB populations with more SMBHBs but lighter black hole masses. Note that on the left panel, the "Gaussian Limit" corresponds to the χ^2_k distribution with $k = 2$, which is also an exponential distribution.

and the PDF is

$$\tilde{P}(|\delta z_E|^2_{NI}) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dt\, e^{[\tilde{K}_{|\delta z|^2}(t)]^* + it|\delta z_E|^2_{NI}}. \quad (51)$$

In Fig. 3, we find that the tails of $P(|\delta z_E|^2)$ and $\tilde{P}(|\delta z_E|^2_{NI})$ can be very well approximated by power law distributions. This result is in agreement with the findings of Refs. [12,24,33]. The major difference is that in the no interference case, $\tilde{P}(|\delta z_E|^2_{NI})$, is strong peaked at some positive value while the peak is smeared and always located at zero in the presence of interference.

In Fig. 4, we show the PDF of $\log_{10}|\delta z_E|^2$,

$$P(\log_{10}|\delta z_E|^2) = \ln 10 |\delta z_E|^2 P(|\delta z_E|^2), \quad (52)$$

and the no interference counterpart $\tilde{P}(\log_{10}|\delta z_E|^2_{NI})$ at different frequency bins for two SMBHB populations and three total observation times. The true PDF is much more diffused than the no interference case. The comparison shows that interference dominates the variance in $|\delta z_E|^2$, while Poisson fluctuations play a minor role. In the no interference case, the expectation value of $|\delta z_E|^2_{NI}$ is consistently higher than the median of $|\delta z_E|^2_{NI}$, especially at higher frequencies. This is because $\tilde{P}(\log_{10}|\delta z_E|^2_{NI})$ is a skewed distribution that decreases to zero rather slowly at large $|\delta z_E|^2_{NI}$, contrary to rapid falloff of the low $|\delta z_E|^2_{NI}$ end (Fig. 3). This observation is consistent with the Monte Carlo results by Refs. [46,48]. The difference between the median and mean values is more prominent for combinations between smaller ϕ_* with larger ϵ_0, which correspond to stronger non-Gaussianity. We also find that the true PDF with the interference effect is well approximated by a Gaussian convolution of the no interference PDF,

$$\tilde{P}(\delta z_{E,Re}, \delta z_{E,Im}) = \int_0^{+\infty} d|\delta z_E|^2_{NI}\tilde{P}(|\delta z_E|^2_{NI})$$
$$\times \frac{e^{-(\delta z^2_{E,Re}+\delta z^2_{E,Im})/|\delta z_E|^2_{NI}}}{\pi|\delta z_E|^2_{NI}}, \quad (53)$$

where the numerical comparison between $P(\delta z_{E,Re}, \delta z_{E,Im})$ and $\tilde{P}(\delta z_{E,Re}, \delta z_{E,Im})$ is shown in Fig. 4. Numerically, the two distributions are nearly indistinguishable from the realistic SMBHB population models we consider. However, in the Appendix, we demonstrate that they are not mathematically identical.

Equation (53) as an excellent approximate PDF implies that non-Gaussianity originates from the Poisson fluctuation of the source numbers. The interference effect is well approximated by randomly drawing the real and imaginary parts of δz_E from a 2D Gaussian distribution. It justifies the population inference done in [34] at the practical level in the following sense: NANOGrav had first derived the likelihoods for the frequency-binned power $|\delta z_E|^2_{NI}$ from the redshift residuals of individual pulsars under the assumption that the SGWB is a Gaussian random field [9]. Reference [34] then used these likelihoods as summary data and performed Bayesian population inference using the correct non-Gaussian PDF $\tilde{P}(|\delta z_E|^2_{NI})$ for the total power (hence without interference). These two analyses combined are equivalent to using Eq. (53) to perform Bayesian inference at the redshift observable level.

In principle, one would like to perform Bayesian inference for SMBHB population models by applying a multivariate non-Gaussian PDF $P(\{\delta z^{(I)}_{E,Re}, \delta z^{(I)}_{E,Im}\})$ for the redshifts of all N_p pulsars $\delta z^{(I)}_E, I = 1, 2, ..., N_p$ to direct PTA observables. Unfortunately, writing down an
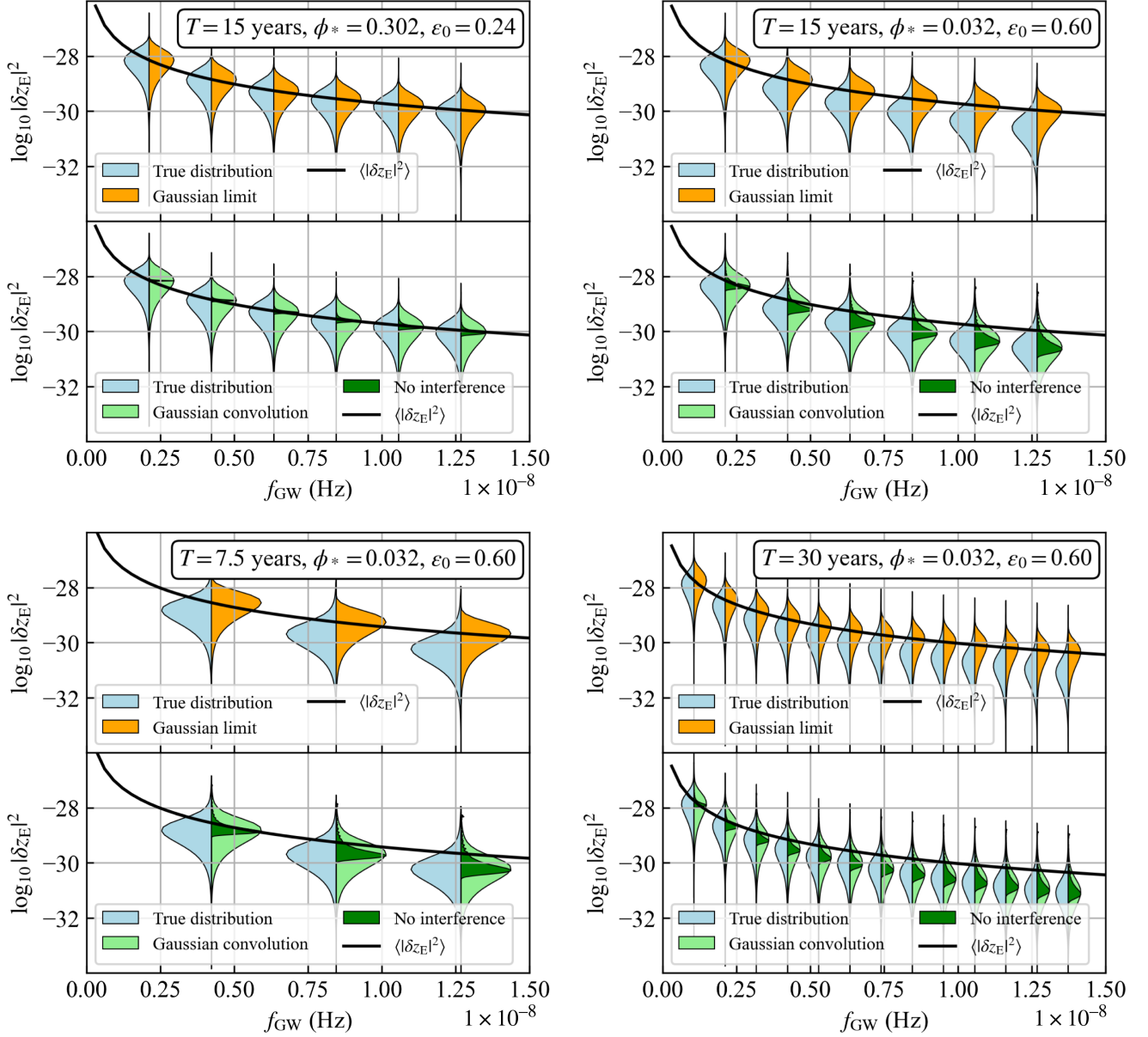
FIG. 4.   Violin plot for $P(\log_{10}|\delta z_E|^2)$ and $\tilde{P}(\log_{10}|\delta z_E|^2_{NI})$ in discrete frequencies binned at a width $\Delta f = 1/T$. All violins at the same frequency in each panel share the same expectation value of $\langle|\delta z_E|^2\rangle$, which equals to $\langle|\delta z_E|^2_{NI}\rangle$, shown by the black curve.

exact expression for this PDF seems prohibitive,[1] as redshift signals in different pulsars are correlated in the Earth term.

However, the numerical success of the approximate PDF in Eq. (53) guides us to conjecture that if one convolves $\tilde{P}(|\delta z_E|^2_{NI})$ with a multivariate Gaussian distribution for the redshifts of all pulsars, with the correct covariance set up between pulsars,

$$
\tilde{P}\left(\left\{\delta z_{E,Re}^{(I)}, \delta z_{E,Im}^{(I)}\right\}\right) = \int_0^{+\infty} d|\delta z_E|^2_{NI} \tilde{P}(|\delta z_E|^2_{NI})
$$

$$
\times \frac{e^{-\delta z_E^\dagger C^{-1}(|\delta z_E|^2_{NI})\delta z_E}}{\pi^{N_p} \det C(|\delta z_E|^2_{NI})}, \tag{54}
$$

then the result may very well approximate the exact joint PDF. In Eq. (54), the vector δz_E collectively represents redshifts of all pulsars, and $C(|\delta z_E|^2_{NI})$ stands for the multipulsar covariance matrix that would result from an isotropic Gaussian SGWB with a given full-sky average power $|\delta z_E|^2_{NI}$. We note that applying Eq. (54) to Bayesian inference at the level of individual pulsar redshifts will be mathematically equivalent to the combination of two steps:

---

[1]Even though a straightforward generalization of the formalism we present in this work will allow the calculation for the corresponding multivariate characteristic function, inverse Fourier transform in high dimensions would be impractical.

first deriving the likelihoods for the mean frequency-binned powers assuming a Gaussian random SGWB, and then inferring the SMBHB population by applying $\tilde{P}(|\delta z_{\rm E}|_{\rm NI}^2)$ to the derived likelihoods as summary data.

If true, such an approach will be a practical path toward full Bayesian treatment, including non-Gaussianity and at the level of pulsar redshifts. However, how this idea may be quantitatively validated is an open question.

D. Sensitivity estimation

With the non-Gaussian statistics developed above, we are in a position to estimate the PTA sensitivities in measuring the non-Gaussianity in the nHz SGWB and constraining the SMBHB population parameters and to test possible bias in inferring the population properties if the Gaussian statistics are applied instead.

1. Likelihood ratio between non-Gaussian and Gaussian model

We first implement a simple likelihood ratio test by applying both the non-Gaussian and the Gaussian statistics to mock data generated from a SMBHB population model. We assume that measurement noises for $\delta z_{\rm Re}$ and $\delta z_{\rm Im}$ follow a Gaussian distribution with the variance of ς. In the presence of noise, the new PDF $P(\delta z_{\rm E,Re}, \delta z_{\rm E,Im}|\varsigma)$ is calculated in the similar way, see Eq. (46), with the CGF modified as

$$K_{\delta z}(t_\rho|\varsigma) = K_{\delta z}(t_\rho) - \frac{\varsigma^2 t_\rho^2}{2}. \qquad (55)$$

For comparison, the CGF for a Gaussian PDF $P_{\rm Gaus}(\delta z_{\rm E,Re}, \delta z_{\rm E,Im}|\varsigma)$ with the same variance is

$$K_{\delta z}^{\rm Gaus}(t_\rho|\varsigma) = \frac{t_\rho^2 K_{\delta z}''(0|\varsigma)}{2}. \qquad (56)$$

We benchmark two observational noise levels: one with a noise level $\varsigma^k = 7 \times 10^{-15}$ over $T = 15$ years, and another with a noise level $\varsigma^k = 1 \times 10^{-15}$ over $T = 30$ years, where k labels the kth frequency bin. The first benchmark case is consistent with the current PTA sensitivity, where the SGWB signal stands out of noise in five frequency bins.

We define the likelihood ratio test statistic as follows:

$$\lambda_{\rm LR} = 2 \ln \frac{P(\delta z_{\rm E,Re}, \delta z_{\rm E,Im}|\varsigma)}{P_{\rm Gaus}(\delta z_{\rm E,Re}, \delta z_{\rm E,Im}|\varsigma)}, \qquad (57)$$

where $P(\delta z_{\rm E,Re}, \delta z_{\rm E,Im}|\varsigma) = \prod_{k=1}^{k_{\max}} P(\delta z_{\rm E,Re}^k, \delta z_{\rm E,Im}^k|\varsigma^k)$ and $\delta z_{\rm E,Re}^k$ and $\delta z_{\rm E,Im}^k$ are the real and imaginary components of the redshift data in the kth frequency bin. We ensure that the Gaussian distribution, $P_{\rm Gaus}(\delta z_{\rm E,Re}^k, \delta z_{\rm E,Im}^k|\varsigma^k)$, has the same variance as the non-Gaussian distribution
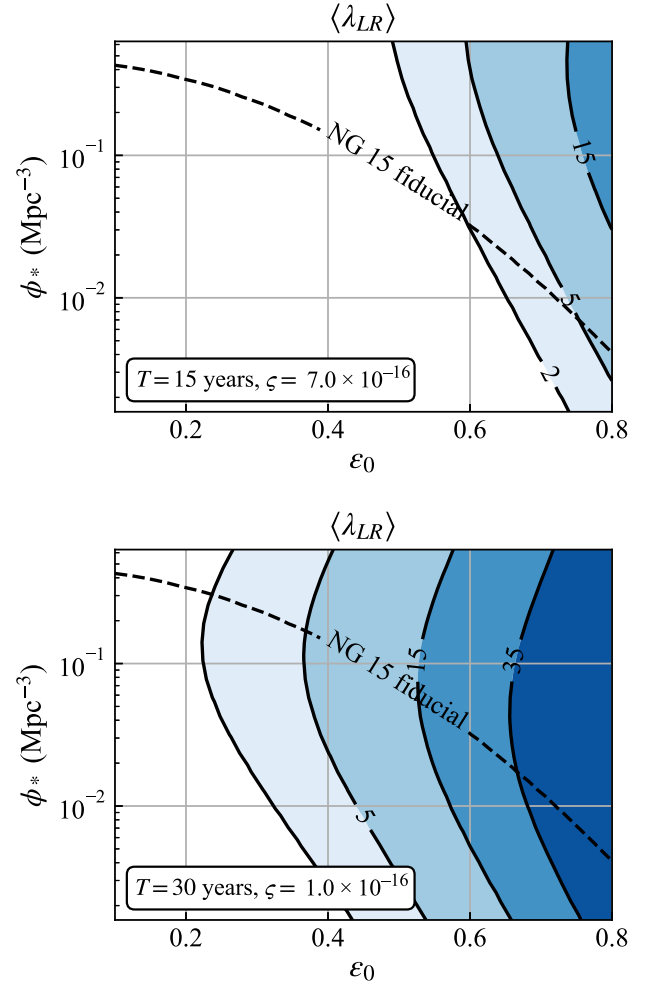

FIG. 5. The contours of the expectation value of $\lambda_{\rm LR}$, $\langle\lambda_{\rm LR}\rangle$, on the $\phi_* - \epsilon_0$ plane. The top panel corresponds to the current PTA sensitivity. The statistics $\lambda_{\rm LR}$, defined in Eq. (57), characterize the likelihood ratio between a non-Gaussian PDF and the Gaussian PDF, assuming the data is drawn from the non-Gaussian PDF.

$P(\delta z_{\rm E,Re}^k, \delta z_{\rm E,Im}^k|\varsigma^k)$. We then evaluate the expectation value $\langle\lambda_{\rm LR}\rangle$ using the following formula:

$$\langle\lambda_{\rm LR}\rangle = \int d\delta z_{\rm E,Re} d\delta z_{\rm E,Im} P(\delta z_{\rm E,Re}, \delta z_{\rm E,Im}|\varsigma)\lambda_{\rm LR}. \qquad (58)$$

We choose $k_{\max} = 14$ and $k_{\max} = 28$ for the two benchmark cases described above. The integral can be simplified using the fact that all frequency bins are independent of each other and that at each bin the PDFs only have dependence on the modulus of $\delta z_{\rm E}^k$.

The results are shown in Fig. 5. With current PTA sensitivity (which has five frequency bins above the noise floor), we expect to find $\langle\lambda_{\rm LR}\rangle > 2$ for $\epsilon_0 > 0.6$. This may increase to $\langle\lambda_{\rm LR}\rangle > 15$ for the same parameter range in the more optimistic situation.

2. Parameter estimation of the SMBHB population model

We also perform a Monte Carlo simulation, in which we generate random data points drawing from the non-Gaussian PDF for a given combination of ϕ_* and ϵ_0.

We then define the likelihood function as follows:

$$\mathcal{L}\big(\log_{10}\phi_*, \epsilon_0 | \delta z_{\mathrm{E,Re}}^{\mathrm{mock}}, \delta z_{\mathrm{E,Im}}^{\mathrm{mock}}\big) = P\big(\delta z_{\mathrm{E,Re}}^{\mathrm{mock}}, \delta z_{\mathrm{E,Im}}^{\mathrm{mock}} | \varsigma, \phi_*, \epsilon_0\big), \tag{59}$$

where $P\big(\delta z_{\mathrm{E,Re}}^{\mathrm{mock}}, \delta z_{\mathrm{E,Im}}^{\mathrm{mock}} | \varsigma, \phi_*, \epsilon_0\big)$ is the calculated with the population model parameters ϕ_*, ϵ_0, while other parameters are fixed at their fiducial values.

Next, we choose the second benchmark and use the population model parameters $(\phi_*, \epsilon_0) = (0.032, 0.61)$ to generate mock data $(\delta z_{\mathrm{E,Re}}^{\mathrm{mock}}, \delta z_{\mathrm{E,Im}}^{\mathrm{mock}})$ 10,000 times. Then for each realization, we calculate the likelihood values on the $\phi_* - \epsilon_0$ plane and find the maximal-likelihood value points. We then collect the best-fit points and investigate their distribution. The results are shown in Fig. 6.

3. Strain power spectrum inference

PTA data have been used in inferring the strain power spectrum of the nHz SGWB, assuming the strain is a Gaussian random field with a power-law frequency dependence,

$$h_c^2(f) = h_c^2(f_{\mathrm{yr}})(f/f_{\mathrm{yr}})^{2\alpha}. \tag{60}$$

In this subsection, we apply the Gaussian statistics to mock data generated from a SMBHB population model and

quantify the resulted bias. We construct the log-likelihood function as follows:

$$\ln \mathcal{L}\big(\delta z_{\mathrm{E,Re}}^{\mathrm{mock}}, \delta z_{\mathrm{E,Im}}^{\mathrm{mock}} | \log_{10} h_c(f_{\mathrm{yr}}), \gamma\big)$$
$$= \sum_k [-|\delta z_{\mathrm{E}}^k|^2 / \langle |\delta z_{\mathrm{E}}^k|^2 \rangle - \ln \pi \langle |\delta z_{\mathrm{E}}^k|^2 \rangle], \tag{61}$$

where

$$\langle |\delta z_{\mathrm{E}}^k|^2 \rangle = \frac{1}{3} h_c^2(f_k) \Delta \ln f_k + 2(\varsigma^k)^2. \tag{62}$$

In accordance with the convention of PTA community, we have introduced the notation γ, which is the power index of the time delay power spectrum and is related to α by $\alpha = (3-\gamma)/2$. Similarly, we choose the earlier defined two benchmarks and simulate 200,000 random realizations for given population model parameters, then collect the best-fit

FIG. 7.   The contours of the best-fit parameters using the classical power-law model. We assume two different population model parameters and simulate 200,000 realizations of mock data for each. We then calculate the likelihood values for each realization using the power-law model and investigate the distributions of the best-fit points. The solid and dashed contours indicate 1σ and 2σ confidence regions, respectively.
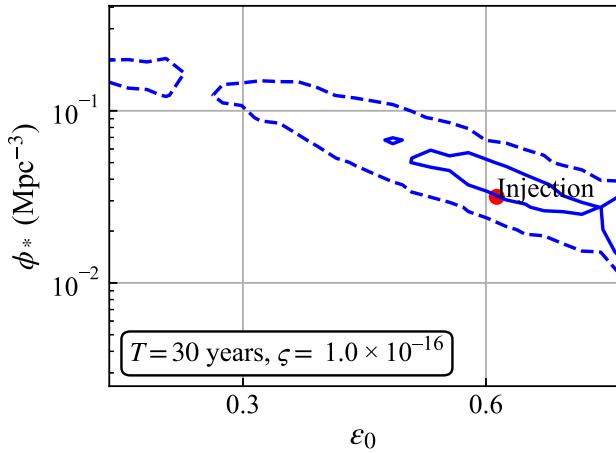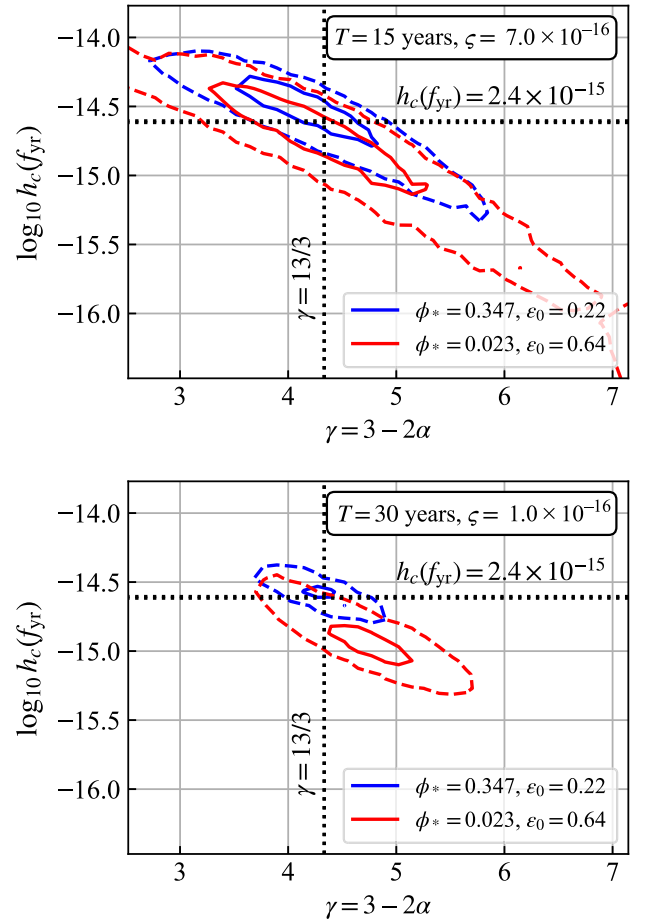

FIG. 6.   The contour of the best-fit parameters using the non-Gaussian PDF we derived in this work. We choose one pair of the population model parameters (ϕ_*, ϵ_0) and simulate 10,000 realizations of mock data. The true parameters are marked as the red point. We then calculate the likelihood values for each realization using the non-Gaussian PDF and investigate the distributions of the best-fit points. The solid and dashed contours indicate 1σ and 2σ confidence regions, respectively.

$(\log_{10} h_c(f_{yr}), \gamma)$ points, and plot the distribution. We show the results in Fig. 7. Comparing the blue and red contours, we find that as the non-Gaussianity increases, the characteristic strain $h_c(f_{yr})$ tends to be underestimated. At the same time, the power-law index γ is overestimated, consistent with what is implied in Fig. 4. This observation is in agreement with findings in [52]. The Gaussian statistics are still a good approximation in inferring the strain power spectrum with current PTA sensitivity and will bias the inference in the foreseeable future.

In conclusion, with the current sensitivity, PTAs cannot determine SMBHB population model parameters with a decent precision. This is fundamentally limited by the large variance arising from interference and the limited number of frequency bins above the noise level. However, as the observation period increases and pulsar timing quality improves, it will become possible to constrain these parameters more precisely in the foreseeable future.

IV. CONCLUSIONS

Since the PTA collaborations dropped the first evidence for the nanohertz SGWB, there has been an ongoing debate over whether it originates from astrophysical sources, such as SMBHBs, or has a primordial origin. These two broad hypotheses suggest different models for constructing the SGWB observed today: fewer sources with stronger individual signals or more sources with weaker individual signals. In this work, we have developed a semianalytic mathematical framework for computing the non-Gaussian PDF of the redshift $P(\delta z)$ for a SMBHB population model, where we have accounted for both Poissonian fluctuations in the number of SMBHBs and GW interference.

To quantify the significance of this distinction, we have numerically calculated the exact PDF of the GW strain power in the frequency domain as a function of population model parameters. With current PTA sensitivities, evidence of non-Gaussianity may be detected in some areas of the population parameter space with fewer and individually louder SMBHBs. As PTA sensitivity improves over time, we expect more robust evidence for non-Gaussianity. We also find the Gaussian statistics are still a good approximation in inferring the strain power spectrum $h_c^2(f)$ with current PTA sensitivity, though it will bias the inference as PTA data of lower noise accumulates in the foreseeable future.

We have proposed an approximated formula to calculate the PDF incorporating data from many pulsars across the sky. We have shed light on why the approximation is numerically very close to the exact answer while mathematically not the same. The correction of the approximated PDF will naturally lead to modifications in two-point correlation functions [21,36,38,42,53,54]. Additionally, developing a numerically efficient method to calculate the analytical PDFs for data analysis will be beneficial. These topics will be addressed in future work.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX: TOY MODEL

To see if the approximation in Eq. (53) works, we perform Fourier transform of $\tilde{P}(\delta z_{E,Re}, \delta z_{E,Im})$ defined in Eq. (53) to obtain the conjugate of the characteristic function (CF),

$$
\begin{aligned}
\tilde{\Phi}^*(t_\rho) &= \int_{-\infty}^{+\infty} d\delta z_{Re} \int_{-\infty}^{+\infty} d\delta z_{Im} \\
&\quad \times \tilde{P}(\delta z_{E,Re}, \delta z_{E,Im}) e^{-it_{Re}\delta z_{Re} - it_{Im}\delta z_{Im}} \\
&= \int_0^{+\infty} d|\delta z_E|^2 \tilde{P}(|\delta z_E|_{NI}^2) \exp\left(-t_\rho^2 |\delta z_E|_{NI}^2/4\right),
\end{aligned}
\tag{A1}
$$

where CGF is the logarithm of the CF

$$
\tilde{\Phi}(t_\rho) \equiv \exp \tilde{K}_{\delta z}(t_\rho), \qquad \tilde{K}_{\delta z}(t_\rho) \equiv \ln \tilde{\Phi}(t_\rho). \tag{A2}
$$

As a simple example, we consider a binary population where the binary number density is a delta function of $\log_{10} \mathcal{M}$, z and $\ln f_r$,

$$
\frac{d^3\bar{N}}{d\log_{10}\mathcal{M} dz d\ln f_r} = \bar{N}\delta(\log_{10}\mathcal{M} - \log_{10}\mathcal{M}_0)\delta(z - z_0)
$$
$$
\times \delta(\ln f_r - \ln f_{r,0}). \tag{A3}
$$

Inserting the above population model into Eqs. (50) and (51) we find the PDF as follows:

$$
\tilde{P}(|\delta z_E|_{NI}^2) = e^{-\bar{N}} \sum_{k=0}^{+\infty} \frac{\bar{N}^k}{k!} \delta\left(|\delta z_E|_{NI}^2 - \frac{2}{15}kh_0^2\right), \tag{A4}
$$

here k are integers, $h_0 = h_0(f_0, \mathcal{M}_0, z_0)$, and $f_0 = f_{r,0}/(1 + z_0)$. Inserting the above expression in Eq. (A1),

we easily find the result for the CF,

$$\tilde{\Phi}^*(t_\rho) = e^{-\bar{N}} \sum_{k=0}^{+\infty} \frac{\bar{N}^k}{k!} \exp\left(-\frac{1}{30} k t_\rho^2 h_0^2\right)$$

$$= \exp\left\{\left[\exp\left(-\frac{1}{30} t_\rho^2 h_0^2\right) - 1\right]\bar{N}\right\}. \quad (A5)$$

The CGF of the approximated PDF is

$$\tilde{K}_{\delta z}(t_\rho) = \ln \tilde{\Phi}(t_\rho) = \left[\exp\left(-\frac{1}{30} t_\rho^2 h_0^2\right) - 1\right]\bar{N}. \quad (A6)$$

In comparison, the CGF of the exact PDF is

$$K_{\delta z}(t_\rho) = \int d\lambda_z P(\lambda_z)[\cos(t_\rho h_0 \lambda_z) - 1]\bar{N}. \quad (A7)$$

Using Eq. (38), we easily find

$$K''_{\delta z}(0) = \tilde{K}''_{\delta z}(0) = -\frac{h_0^2 \bar{N}}{15}. \quad (A8)$$

The difference emerges at the fourth-order cumulant (Kurtosis),

$$K^{(4)}_{\delta z}(0) = \frac{3 h_0^4 \bar{N}}{175}, \qquad \tilde{K}^{(4)}_{\delta z}(0) = \frac{h_0^4 \bar{N}}{75}. \quad (A9)$$
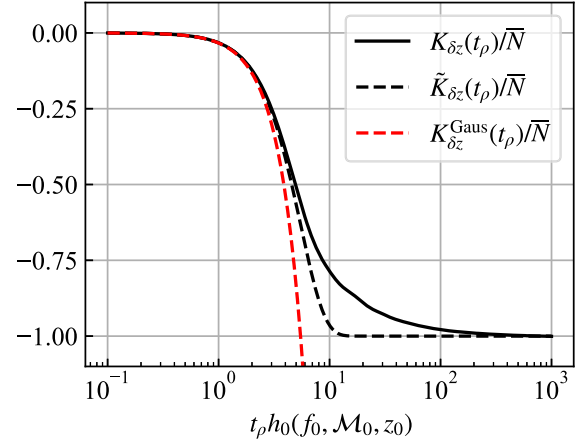

FIG. 8. We compare $K_{\delta z}(t_\rho)/\bar{N}$, $\tilde{K}_{\delta z}(t_\rho)/\bar{N}$ and $K^{\text{Gaus}}_{\delta z}(t_\rho)/\bar{N}$. We calculate the CGFs by assuming all sources have the same chirp mass $\mathcal{M}_0$ and located at the same redshift z_0.

In Fig. 8, we compare $\tilde{K}_{\delta z}(t_\rho)$ and $K_{\delta z}(t_\rho)$ in Eqs. (A6) and (A7) with the CDF of the Gaussian distribution, $K^{\text{Gaus}}_{\delta z}(t_\rho) = -t_\rho^2 h_0^2 \bar{N}/30$. The "No interference + Gaussian convolution" approximation correctly captures the asymptotic behavior of the true CGF at $t_\rho \to 0$ and $t_\rho \to +\infty$. However, in the intermediate region, the approximation deviates from the true PDF and underestimates the non-Gaussianity since it is closer to the Gaussian limit.

[1] Gabriella Agazie *et al.* (NANOGrav Collaboration), The NANOGrav 15 yr data set: Evidence for a gravitational-wave background, Astrophys. J. Lett. **951**, L8 (2023).

[2] Gabriella Agazie *et al.* (NANOGrav Collaboration), The NANOGrav 15 yr data set: Observations and timing of 68 millisecond pulsars, Astrophys. J. Lett. **951**, L9 (2023).

[3] G. Agazie, J. Antoniadis, A. Anumarlapudi, A. M. Archibald, P. Arumugam, S. Arumugam, Z. Arzoumanian, J. Askew, S. Babak, M. Bagchi, M. Bailes, A.-S. Bak Nielsen, P. T. Baker, C. G. Bassa *et al.* (The International Pulsar Timing Array Collaboration), Comparing recent PTA results on the nano-hertz stochastic gravitational wave background, Astrophys. J. **966**, 105 (2024).

[4] Daniel J. Reardon *et al.*, Search for an isotropic gravitational-wave background with the Parkes Pulsar Timing Array, Astrophys. J. Lett. **951**, L6 (2023).

[5] Andrew Zic *et al.*, The Parkes Pulsar Timing Array third data release, Publ. Astron. Soc. Aust. **40**, e049 (2023).

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

[7] J. Antoniadis *et al.* (EPTA Collaboration), The second data release from the European Pulsar Timing Array I. The dataset and timing analysis, Astron. Astrophys. **678**, A48 (2023).

[8] Heng 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).

[9] Gabriella Agazie *et al.* (NANOGrav Collaboration), The NANOGrav 15 yr data set: Constraints on supermassive black hole binaries from the gravitational-wave background, Astrophys. J. Lett. **952**, L37 (2023).

[10] Yan-Chen Bi, Yu-Mei Wu, Zu-Cheng Chen, and Qing-Guo Huang, Implications for the supermassive black hole binaries from the NANOGrav 15-year data set, Sci. China Phys. Mech. Astron. **66**, 120402 (2023).

[11] Gabriela Sato-Polito, Matias Zaldarriaga, and Eliot Quataert, Where are NANOGrav's big black holes?, Phys. Rev. D **110**, 063020 (2024).

[12] John Ellis, Malcolm Fairbairn, Gert Hütsi, Juhan Raidal, Juan Urrutia, Ville Vaskonen, and Hardi Veermäe, Gravitational waves from supermassive black hole binaries in

light of the NANOGrav 15-year data, Phys. Rev. D **109,** L021302 (2024).

[13] John Ellis, Malcolm Fairbairn, Juan Urrutia, and Ville Vaskonen, In search of the biggest bangs since the big bang, Int. J. Mod. Phys. **33,** 2440004 (2024).

[14] John Ellis, Malcolm Fairbairn, Gert Hütsi, Juan Urrutia, Ville Vaskonen, and Hardi Veermäe, Consistency of JWST black hole observations with NANOGrav gravitational wave measurements, Astron. Astrophys. **691,** A270 (2024).

[15] Hamsa Padmanabhan and Abraham Loeb, Constraints on supermassive black hole binaries from JWST and NANOGrav, Astron. Astrophys. **684,** L15 (2024).

[16] Emily R. Liepold and Chung-Pei Ma, Big galaxies and big black holes: The massive ends of the local stellar and black hole mass functions and the implications for nanohertz gravitational waves, Astrophys. J. Lett. **971,** L29 (2024).

[17] Adeela Afzal *et al.* (NANOGrav Collaboration), The NANOGrav 15 yr data set: Search for signals from new physics, Astrophys. J. Lett. **951,** L11 (2023); Astrophys. J. Lett. **971,** L27(E) (2024).

[18] Elinore Roebber, Gilbert Holder, Daniel E. Holz, and Michael Warren, Cosmic variance in the nanohertz gravitational wave background, Astrophys. J. **819,** 163 (2016).

[19] John Ellis, Malcolm Fairbairn, Gabriele Franciolini, Gert Hütsi, Antonio Iovino, Marek Lewicki, Martti Raidal, Juan Urrutia, Ville Vaskonen, and Hardi Veermäe, What is the source of the PTA GW signal?, Phys. Rev. D **109,** 023522 (2024).

[20] William G. Lamb and Stephen R. Taylor, Spectral variance in a stochastic gravitational-wave background from a binary population, Astrophys. J. Lett. **971,** L10 (2024).

[21] Reginald Christian Bernardo, Stephen Appleby, and Kin-Wang Ng, Toward a test of Gaussianity of a gravitational wave background, arXiv:2407.17987.

[22] Gabriella Agazie *et al.*, The NANOGrav 15 yr data set: Looking for signs of discreteness in the gravitational-wave background, Astrophys. J. **978,** 31 (2025).

[23] Shashwat C. Sardesai, Joseph Simon, and Sarah J. Vigeland, Characterizing the nanohertz gravitational wave background using a *t*-process power spectral density, Astrophys. J. **976,** 212 (2024).

[24] Juhan Raidal, Juan Urrutia, Ville Vaskonen, and Hardi Veermäe, Eccentricity effects on the SMBH GW background, Astron. Astrophys. **691,** A212 (2024).

[25] Selim C. Hotinli, Marc Kamionkowski, and Andrew H. Jaffe, The search for anisotropy in the gravitational-wave background with pulsar-timing arrays, Open J. Astrophys. **2,** 8 (2019).

[26] Stephen R. Taylor, Rutger van Haasteren, and Alberto Sesana, From bright binaries to bumpy backgrounds: Mapping realistic gravitational wave skies with pulsar-timing arrays, Phys. Rev. D **102,** 084039 (2020).

[27] Nihan Pol, Stephen R. Taylor, and Joseph D. Romano, Forecasting pulsar timing array sensitivity to anisotropy in the stochastic gravitational wave background, Astrophys. J. **940,** 173 (2022).

[28] Gabriella Agazie *et al.* (NANOGrav Collaboration), The NANOGrav 15 yr data set: Search for anisotropy in the gravitational-wave background, Astrophys. J. Lett. **956,** L3 (2023).

[29] Anna-Malin Lemke, Andrea Mitridate, and Kyle A. Gersbach, Detecting gravitational wave anisotropies from supermassive black hole binaries, arXiv:2407.08705.

[30] Paul Frederik Depta, Valerie Domcke, Gabriele Franciolini, and Mauro Pieroni, Pulsar timing array sensitivity to anisotropies in the gravitational wave background, arXiv:2407.14460.

[31] Emiko C. Gardiner, Luke Zoltan Kelley, Anna-Malin Lemke, and Andrea Mitridate, Beyond the background: Gravitational-wave anisotropy and continuous waves from supermassive black hole binaries, Astrophys. J. **965,** 164 (2024).

[32] Qing Yang, Xiao Guo, Zhoujian Cao, Xiaoyun Shao, and Xi Yuan, Anisotropy of nanohertz gravitational wave background and individual sources from supermassive binary black holes: Probe of cosmic large scale structure, arXiv:2408.05043.

[33] John Ellis, Malcolm Fairbairn, Gert Hütsi, Martti Raidal, Juan Urrutia, Ville Vaskonen, and Hardi Veermäe, Prospects for future binary black hole gravitational wave studies in light of PTA measurements, Astron. Astrophys. **676,** A38 (2023).

[34] Gabriela Sato-Polito and Matias Zaldarriaga, The distribution of the gravitational-wave background from supermassive black holes, arXiv:2406.17010.

[35] Bruce Allen, Variance of the Hellings-Downs correlation, Phys. Rev. D **107,** 043018 (2023).

[36] Bruce Allen and Serena Valtolina, Pulsar timing array source ensembles, Phys. Rev. D **109,** 083038 (2024).

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

[38] Yu-Mei Wu, Yan-Chen Bi, and Qing-Guo Huang, The spatial correlations between pulsars for interfering sources in pulsar timing array and evidence for gravitational-wave background in NANOGrav 15-year data set, arXiv:2407.07319.

[39] Joseph D. Romano and Bruce Allen, Answers to frequently asked questions about the pulsar timing array Hellings and Downs curve, Classical Quantum Gravity **41,** 175008 (2024).

[40] Neil J. Cornish and A. Sesana, Pulsar timing array analysis for black hole backgrounds, Classical Quantum Gravity **30,** 224005 (2013).

[41] Stephen R. Taylor and Jonathan R. Gair, Searching for anisotropic gravitational-wave backgrounds using pulsar timing arrays, Phys. Rev. D **88,** 084001 (2013).

[42] Thomas Konstandin, Anna-Malin Lemke, Andrea Mitridate, and Enrico Perboni, The impact of cosmic variance on PTAs anisotropy searches, arXiv:2408.07741.

[43] Laura Ferrarese and David Merritt, A fundamental relation between supermassive black holes and their host galaxies, Astrophys. J. Lett. **539,** L9 (2000).

[44] John Kormendy and Luis C. Ho, Coevolution (or not) of supermassive black holes and host galaxies, Annu. Rev. Astron. Astrophys. **51,** 511 (2013).

[45] Nicholas J. McConnell and Chung-Pei Ma, Revisiting the scaling relations of black hole masses and host galaxy properties, Astrophys. J. **764,** 184 (2013).

[46] Alberto Sesana, Alberto Vecchio, and Carlo Nicola Colacino, The stochastic gravitational-wave background

from massive black hole binary systems: Implications for observations with pulsar timing arrays, Mon. Not. R. Astron. Soc. **390**, 192 (2008).

[47] Mariangela Bernardi, Francesco Shankar, J. B. Hyde, Simona Mei, Federico Marulli, and Ravi K. Sheth, Galaxy luminosities, stellar masses, sizes, velocity dispersions as a function of morphological type, Mon. Not. R. Astron. Soc. **404**, 2087 (2010).

[48] Luke Zoltan Kelley, Laura Blecha, and Lars Hernquist, Massive black hole binary mergers in dynamical galactic environments, Mon. Not. R. Astron. Soc. **464**, 3131 (2017).

[49] Zaven Arzoumanian, Paul T. Baker, Laura Blecha, Harsha Blumer, Adam Brazier, Paul R. Brook, Sarah Burke-Spolaor, Bence Bécsy, J. Andrew Casey-Clyde, Maria Charisi *et al.*, The NANOGrav 12.5 yr data set: Bayesian

limits on gravitational waves from individual supermassive black hole binaries, Astrophys. J. Lett. **951**, L28 (2023).

[50] E. S. Phinney, A practical theorem on gravitational wave backgrounds, arXiv:astro-ph/0108028.

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

[52] Bence Bécsy *et al.*, How to detect an astrophysical nanohertz gravitational wave background, Astrophys. J. **959**, 9 (2023).

[53] Bruce Allen, Deepali Agarwal, Joseph D. Romano, and Serena Valtolina, Source anisotropies and pulsar timing arrays, Phys. Rev. D **110**, 123507 (2024).

[54] Reginald Christian Bernardo and Kin-Wang Ng, Charting the nanohertz gravitational wave sky with pulsar timing arrays, arXiv:2409.07955.