The Astrophysical Journal Letters, 971:L10 (7pp), 2024 August 10 https://doi.org/10.3847/2041-8213/ad654a

Spectral Variance in a Stochastic Gravitational-wave Background from a Binary Population

William G. Lamb and Stephen R. Taylor
Department of Physics & Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA; william.g.lamb@vanderbilt.edu, stephen.r.taylor@vanderbilt.edu
Received 2024 July 8; revised 2024 July 17; accepted 2024 July 18; published 2024 August 5

Abstract

A population of compact object binaries emitting gravitational waves that are not individually resolvable will form a stochastic gravitational-wave signal. While the expected spectrum over population realizations is well known from Phinney, its higher-order moments have not been fully studied before or computed in the case of arbitrary binary evolution. We calculate analytic scaling relationships as a function of gravitational-wave frequency for the statistical variance, skewness, and kurtosis of a stochastic gravitational-wave signal over population realizations due to finite source effects. If the time derivative of the binary orbital frequency can be expressed as a power law in frequency, we find that these moment quantities also take the form of power-law relationships. We also develop a numerical population synthesis framework against which we compare our analytic results, finding excellent agreement. These new scaling relationships provide physical context to understanding spectral fluctuations in a gravitational-wave background signal and may provide additional information that can aid in explaining the origin of the nanohertz-frequency signal observed by pulsar timing array campaigns.

Unified Astronomy Thesaurus concepts: Gravitational wave astronomy (675); Gravitational wave sources (677); Gravitational waves (678); General relativity (641); Compact objects (288); Astrophysical black holes (98); Supermassive black holes (1663); LIGO (920); Astrostatistics (1882); Astronomical simulations (1857); Neutron stars (1108)

1. Introduction

With strong evidence for a nanohertz-frequency gravitational-wave (GW) background (GWB) signal recently presented by international collaborations of pulsar timing arrays (PTAs; Agazie et al. 2023a; Antoniadis et al. 2023a; Reardon et al. 2023; Xu et al. 2023), attention is now shifting to the elusive nature of its origin. While a population of GW-radiating supermassive black hole binaries (SMBHBs) seems the most likely, many models of early-Universe processes can be formed to explain the PTA data (e.g., Agazie et al. 2023b; Antoniadis et al. 2023b; Afzal et al. 2023; Smarra et al. 2023).

In the simplest model of an astrophysical GWB one usually assumes a statistically isotropic distribution of SMBHBs across the sky, where the binary orbital evolution at PTA frequencies is driven entirely by the emission of GWs, resulting in a power-law characteristic strain spectrum with $h_c^2(f) \propto f^{-4/3}$ (Phinney 2001). However, this relation is in fact an average over many realizations of an SMBHB population, and several studies have shown that fluctuations in this spectrum are expected (e.g., Sesana et al. 2008; Roebber et al. 2016).

Recent work has attempted to account for binary population variance in the spectrum and to distinguish between astrophysical and cosmological signal origins. This has included comparing the spectra of cosmological models against a power law (Kaiser et al. 2022; Afzal et al. 2023), fitting Gaussian processes or neural networks that are trained on many realizations of simulated astrophysical GWBs to PTA data (Taylor et al. 2017; Afzal et al. 2023; Bonetti et al. 2024), searching for anisotropy (Agazie et al. 2023c) that may be frequency dependent (Gardiner et al. 2024), quantifying the discreteness of the GWB at high frequencies (Agazie et al. 2024), and testing the Gaussian ensemble model given a finite population of GW sources (Allen & Valtolina 2024).

However, there has been little work to compute the expected variance or higher moments of the GWB signal for a finite population of GW sources. This could be applied to model selection and parameter estimation of the GWB to improve upon the simplistic power-law model for a binary population and provide additional leverage against which other signal models can be compared. There is a brief discussion of this in Jaffe & Backer (2003) in terms of the moment generating function of $h_c^2(f)$, while Sato-Polito & Zaldarriaga (2024)[1] re-analyzed the NANOGrav 15 yr data set (Agazie et al. 2023a, 2023b) with a GWB spectral model that considers deviations from a power law due to a finite population of SMBHBs. However, explicit scaling relationships for variance, skewness, and kurtosis of the distribution of characteristic strain as a function of frequency were not provided, and the calculations assumed that binaries were only evolving through GW emission in both works. Additionally, Renzini & Golomb (2024) investigated the GWB detection capabilities of LIGO-Virgo-KAGRA (LVK) given uncertainties in the source population from single event analyses. In this Letter, we compute the variance and higher-order moments of the GWB for a general finite population of GW sources that create an unresolved GWB. While here we consider only population finiteness effects, we note that GW signal interference will in general also contribute to the variance and higher moments of the GWB. We provide explicit frequency-scaling relationships

[1] Sato-Polito & Zaldarriaga (2024) use techniques similar to Ginat et al. (2020) and Ginat et al. (2024).

for arbitrary binary evolution and compare directly to many numerically synthesized populations. These results are relevant not only for SMBHBs in the PTA frequency regime but for other sources of a GWB across the GW frequency spectrum, including LVK (e.g., Abbott et al. 2021) and LISA (Digman & Cornish 2022; Babak et al. 2023; European Space Agency 2024). One can regard this work as a generalization of Phinney (2001) to higher moments of the GWB strain distribution.

This Letter is laid out as follows. In Section 2 we introduce a framework to model a finite population of GW sources, which we then use to compute the mean, variance, skewness, and kurtosis of the GWB spectrum across realizations of the population. In Section 3 we develop a population synthesis model derived from Sato-Polito & Kamionkowski (2024, hereafter SK23). The comparison between our analytical and numerical results is shown in Section 4, followed in Section 5 by a discussion of how our new results could aid in answering the origin question for the PTA signal.

2. GW Strain Spectrum from a Binary Population

The cosmological energy density in GWs, as a fraction of closure density, and as a function of frequency, is given by

$$\Omega_{\rm GW}(f) = \frac{1}{\rho_{\rm c}}\frac{d\rho_{\rm GW}}{d\ln f} = \frac{2\pi^2}{3H_0^2} f^2 h_{\rm c}^2(f), \tag{1}$$

where $\rho_{\rm GW}$ is the energy density of GWs, $\rho_{\rm c}$ is the closure density, f is the observed GW frequency, H_0 is the Hubble constant, and $h_{\rm c}^2(f)$ is the squared characteristic strain. Within a logarithmic frequency bin $\left[\ln f - \frac{\Delta \ln f}{2}, \ln f + \frac{\Delta \ln f}{2}\right]$, $\Omega_{\rm GW}(f)$ is a quadrature sum of GW strain amplitude, h, from each source (Sesana et al. 2008):

$$\Omega_{\rm GW}(f) = \left(\frac{2\pi^2}{3H_0^2}\right) f^2 \int_{\Delta \ln f} \frac{d\ln f'}{\Delta \ln f} \times \int_{\vec{\theta}} d\vec{\theta}' \frac{dN}{{\rm d}\vec{\theta}' d\ln f_r'} h^2(f_r', \vec{\theta}'), \tag{2}$$

where N is the number of sources, h is the GW strain amplitude from a single GW source, the subscript r denotes the rest frame of the source such that $f_r = (1+z)f$, and $\vec{\theta}$ are parameters of the source. For example, for populations of SMBHBs, $\vec{\theta}$ includes the chirp mass $\mathcal{M}$ and redshift z of each SMBHB. We denote the binning of sources within a logarithmic frequency bin as an integral over frequency divided by $\Delta \ln f$, which effectively averages $\Omega_{\rm GW}(f)$ across the log-frequency bin.

We compute the number of binaries per bin of GW frequency and source parameters $\vec{\theta}$ by

$$\frac{dN}{{\rm d}\vec{\theta} d\ln f} = \frac{dn}{d\vec{\theta}} \frac{dt}{d\ln f_r} \frac{dz}{dt} \frac{dV_{\rm c}}{dz}, \tag{3}$$

where n is the number of SMBHBs per unit comoving volume, $V_{\rm c}$. The quantity df/dt describes the dynamical evolution of the SMBHB and is related to the typical amount of time that a binary spends emitting at a given frequency (also known as the residence time), $f/(df/dt)$. For a circular SMBHB evolving entirely due to the emission of GWs, $df/dt \propto f^{11/3}$ (Peters 1964). Other binary hardening drivers, such as eccentricity of the binary orbit, three-body interactions with stars in the stellar loss cone, or interactions with a circumbinary gas disk, will result in different spectral indices for the power-law relation (Kocsis & Sesana 2011; Sampson et al. 2015; Burke-Spolaor et al. 2019). For example, for three-body stellar scattering, $df/dt \propto f^{1/3}$, while different models of gas interactions can result in $df/dt \propto f^{10/3}$ or $\propto f^{4/3}$. Hence, we model the frequency evolution of the binary as $df/dt \propto f^{\lambda}$, where λ is a free parameter. With $\lambda = 11/3$ for GW emission alone, the GWB spectrum has the following power-law form:

$$\Omega_{\rm GW}(f) = \frac{2\pi^2}{3H_0^2} A_{\rm ref}^2 f^2 \left(\frac{f}{f_{\rm ref}}\right)^{2\alpha}, \tag{4}$$

where $A_{\rm ref}$ is the strain amplitude measured at a reference frequency $f_{\rm ref}$, and $\alpha = (7/3 - \lambda)/2 = -2/3$.

The original calculation from Phinney (2001) can be further generalized by considering the distribution of SMBHB source positions,

$$\begin{aligned}\Omega_{\rm GW}(f, \vec{\theta}) &= \int_{\hat{\Omega}} \Omega_{\rm GW}(f, \vec{\theta}, \hat{\Omega}') {\rm d}\hat{\Omega}' \\ &= \frac{2\pi^2}{3H_0^2} f^2 \int_{\Delta \ln f} \frac{d\ln f'}{\Delta \ln f} \int_{\vec{\theta}} d\vec{\theta}' \\ &\quad \times \int_{\hat{\Omega}} \frac{d\hat{\Omega}'}{4\pi} \frac{dN}{{\rm d}\vec{\theta}' d\ln f_r' {\rm d}\hat{\Omega}'} h^2(f_r', \vec{\theta}'),\end{aligned} \tag{5}$$

where $\hat{\Omega}$ is the direction of propagation of a GW.

SK23 adopt this form and discretize the GW source populations as follows. The sky is divided up into infinitesimal cells with unit vector $\hat{\Omega}_i$ pointing toward the center of cell i. In each cell, there is exactly zero or one GW source within the frequency bin $\left[\ln f - \frac{\Delta \ln f}{2}, \ln f + \frac{\Delta \ln f}{2}\right]$. If N_i is the number of binaries in cell i, then the expected number of binaries in each cell is

$$\langle N_i(f_i, \vec{\theta}_i)\rangle d\vec{\theta}_i = \frac{1}{4\pi \Delta \ln f_i} \int_{\Delta \ln f_i} d\ln f_i' \frac{dN}{d\vec{\theta}_i' d\ln f_{r,i}'} d\vec{\theta}_i', \tag{6}$$

where $\langle\cdot\rangle$ denotes expectation over sky position. Note that, by definition, $N_i^2 = N_i$. Cells i and j should be uncorrelated; thus for a frequency bin $\left[\ln f - \frac{\Delta \ln f}{2}, \ln f + \frac{\Delta \ln f}{2}\right]$,

$$\langle N_i N_j\rangle = \begin{cases} \langle N_i\rangle \langle N_j\rangle & \text{for } i \neq j \\ \langle N_i^2\rangle \delta_{\vec{\theta}_i \vec{\theta}_j} = \langle N_i\rangle \delta_{\vec{\theta}_i \vec{\theta}_j} & \text{for } i = j, \end{cases} \tag{7}$$

where δ is the Kronecker delta. For a cell i, the expected energy density of the GWB is $\langle \Omega_{\rm GW}(f_i, \vec{\theta}_i, \hat{\Omega}_i)\rangle \propto f^2 \langle h_{\rm c}^2(f_i, \vec{\theta}_i, \hat{\Omega}_i)\rangle = \langle N_i(f_i, \vec{\theta}_i)\rangle h^2(f_i, \vec{\theta}_i)$, and the expected energy density across the sky is given by

$$\begin{aligned}\langle \Omega_{\rm GW}(f_i)\rangle &= \tfrac{2\pi^2}{3H_0^2} f^2 \int d\hat{\Omega}_i' \int_{\vec{\theta}_i} d\vec{\theta}_i' \langle N_i(f_i, \vec{\theta}_i')\rangle h^2(f_i, \vec{\theta}_i') \\ &= \tfrac{2\pi^2}{3H_0^2} f^2 \int_{\vec{\theta}_i} d\vec{\theta}_i' \langle N(f_i, \vec{\theta}_i')\rangle h^2(f_i, \vec{\theta}_i'),\end{aligned} \tag{8}$$

where $\langle N\rangle = \int d\hat{\Omega}_i \langle N_i\rangle = 4\pi \langle N_i\rangle$, the expected number of binaries in a given realization. This agrees with Equation (2).

Table 1
A Table of Relationships between Different Quantities Relative to the Fractional Energy Density $\Omega_{\rm GW}$

	Scaling	Mean	Variance
$\Omega_{\rm GW}(f)$	...	$f^{13/3-\lambda}$	$f^{26/3-\lambda}$
$h_c^2(f)$	$f^{-2}\Omega_{\rm GW}$	$f^{7/3-\lambda}$	$f^{14/3-\lambda}$
$S_h(f)$	$f^{-3}\Omega_{\rm GW}$	$f^{4/3-\lambda}$	$f^{8/3-\lambda}$
$S_{\delta t}(f)$	$f^{-5}\Omega_{\rm GW}$	$f^{-2/3-\lambda}$	$f^{-4/3-\lambda}$

Note. Assumes an SMBHB frequency evolution of $df/dt \propto f^{\lambda}$. Skewness and kurtosis scale as $f^{\lambda/2}$ and f^{λ}, respectively, for all quantities. Featured quantities include: squared characteristic strain h_c^2; the one-sided power spectral density of the GWB Fourier modes S_h; and the one-sided power spectral density of pulsar timing residuals induced by a GWB $S_{\delta t}$.

For the second noncentral moment of $h_c^2(f)$,

$$\begin{aligned}
&\langle h_c^2(f_i, \vec{\theta}_i, \hat{\Omega}_i) h_c^2(f_j, \vec{\theta}_j, \hat{\Omega}_j)\rangle \\
&= \delta_{f_i f_j}\langle N_i(f_i, \vec{\theta}_i) N_j(f_j, \vec{\theta}_j)\rangle h^2(f_i, \vec{\theta}_i) h^2(f_j, \vec{\theta}_j) \\
&= \delta_{f_i f_j}\langle N_i(f_i, \vec{\theta}_i)\rangle \langle N_j(f_j, \vec{\theta}_j)\rangle h^2(f_i, \vec{\theta}_i) h^2(f_j, \vec{\theta}_j) \\
&\quad + \delta_{ij}\delta_{f_i f_j}\delta_{\vec{\theta}_i \vec{\theta}_j}\langle N_i(f_i, \vec{\theta}_i)\rangle h^4(f_i, \vec{\theta}_i). \qquad (9)
\end{aligned}$$

Integrating over sky positions $\hat{\Omega}_i$ and $\hat{\Omega}_j$ and binary parameters $\vec{\theta}_i$, $\vec{\theta}_j$, and substituting Equation (7) gives

$$\begin{aligned}
\langle h_c^2(f_i) h_c^2(f_j)\rangle &= \delta_{f_i f_j}\left(\int_{\vec{\theta}} d\vec{\theta}' \langle N(f_i, \vec{\theta}')\rangle h^2(f_i, \vec{\theta}')\right)^2 \\
&\quad + \delta_{f_i f_j}\int_{\vec{\theta}} d\vec{\theta}' \langle N(f_i, \vec{\theta}')\rangle h^4(f_i, \vec{\theta}'). \qquad (10)
\end{aligned}$$

Integrating over frequency bin $\Delta \ln f_j$, we can then compute the variance on $\Omega_{\rm GW}(f)$:

$$\begin{aligned}
{\rm Var}[\Omega_{\rm GW}(f)] &= \left(\frac{2\pi^2 f^2}{3H_0^2}\right)^2 [\langle (h_c^2(f))^2\rangle - \langle h_c^2(f)\rangle^2] \\
&= \frac{4\pi^4}{9H_0^4}\frac{f^4}{\Delta \ln f}\int d\vec{\theta}\, \langle N(f, \vec{\theta})\rangle h^4(f, \vec{\theta}). \qquad (11)
\end{aligned}$$

Since $\langle N\rangle \propto dN/d\ln f \propto f dt/df_r \propto f^{1-\lambda}$, $\Delta \ln f \propto 1/f$, and $h(f) \propto f^{2/3}$, we get the following relation:

$${\rm Var}[\Omega_{\rm GW}(f)] \propto f^4 \cdot f \cdot f^{1-\lambda} \cdot f^{\frac{8}{3}} \propto f^{\frac{26}{3}-\lambda}. \qquad (12)$$

We derive the mean and variance of various useful parameters related to $\Omega_{\rm GW}(f)$ in Table 1, using the standard transformation ${\rm Var}[aX + b] = a^2{\rm Var}[X]$ for random variable X and constants a, b. Note that our results agree with the frequency relation in Equation (29) of Mingarelli et al. (2013), where they compute $\sigma(\Omega_{\rm GW}(f))/\Omega_{\rm GW}(f) \propto f^{1/2}/f^{-4/3} = f^{11/6}$, where $\sigma^2 = {\rm Var}$. We also see that upon taking the ratio of the variance to the mean, the frequency scaling no longer depends on the binary evolution parameter λ, thus creating a clean probe of the Poissonian statistics of the binary population without the need to specify the dynamical mechanisms of binary decay.

The calculations for skewness and excess kurtosis on $\Omega_{\rm GW}(f)$ proceed in much the same way. The skewness is a measure of the asymmetry of a distribution about its mean, where a positive skewness indicates that most of the mass of the distribution is to the right of the mean, and vice versa. The excess kurtosis is a measure of the extremities of deviations from the mean (i.e., outliers) relative to a Gaussian distribution, where a positive kurtosis indicates more extreme outliers resulting in wider tails, and a negative kurtosis indicates fewer extreme outliers and a thinner tail, compared to a Gaussian distribution. Note that ${\rm Skew}[aX + b] = {\rm Skew}[X]$ and ${\rm Kurt}[aX + b] = {\rm Kurt}[X]$; therefore, the skewness and kurtosis are equal between linear transforms of $\Omega_{\rm GW}$.

$${\rm Skew}(f) = \left(\frac{1}{\Delta \ln f}\right)^{\frac{1}{2}} \frac{\int d\vec{\theta}\, \langle N(f, \vec{\theta})\rangle h^6(f, \vec{\theta})}{\left(\int d\vec{\theta}\, \langle N(f, \vec{\theta})\rangle h^4(f, \vec{\theta})\right)^{\frac{3}{2}}} \qquad (13)$$

$${\rm Kurt}(f) = \frac{1}{\Delta \ln f}\frac{\int d\vec{\theta}\, \langle N(f, \vec{\theta})\rangle h^8(f, \vec{\theta})}{\left(\int d\vec{\theta}\, \langle N(f, \vec{\theta})\rangle h^4(f, \vec{\theta})\right)^2} \qquad (14)$$

The skewness and kurtosis are always positive. We derive the following frequency relations, which are relevant for all linear transformations of $\Omega_{\rm GW}$:

$${\rm Skew}(f) \propto f^{\frac{1}{2}} \cdot f^{1-\lambda} \cdot f^4 \cdot (f^{1-\lambda} \cdot f^{\frac{8}{3}})^{-\frac{3}{2}} \propto f^{\frac{\lambda}{2}} \qquad (15)$$

$${\rm Kurt}(f) \propto f \cdot f^{1-\lambda} \cdot f^{\frac{16}{3}} \cdot (f^{1-\lambda} \cdot f^{\frac{8}{3}})^{-2} \propto f^{\lambda}. \qquad (16)$$

We now see that taking the ratio between the kurtosis and the squared skewness removes all frequency dependence, providing a clear prediction for a stochastic GW signal generated by a Poisson-distributed binary population.

3. Population Synthesis

We simulate 10^7 realizations of a population of SMBHBs. We use a Schechter-based model to describe the number density of SMBHBs per unit redshift and chirp mass (Middleton et al. 2015; see also Sesana et al. 2008; Agazie et al. 2023b; SK23):

$$\begin{aligned}
\frac{d^2 n}{dz d\log \mathcal{M}} &= \dot{n}_0 \left[\left(\frac{\mathcal{M}}{10^7 M_\odot}\right)^{-\alpha} e^{-\mathcal{M}/\mathcal{M}_*}\right] \\
&\quad \times [(1+z)^{\beta} e^{-z/z_0}]\frac{dt_r}{dz}, \qquad (17)
\end{aligned}$$

where $\dot{n}_0$ is the normalized merger rate, $\mathcal{M}_*$ is the characteristic mass, α is the mass spectral index, z_0 is the characteristic redshift, and β is the redshift spectral index. (See Middleton et al. 2015 for details.) To create the population of SMBHBs, we Poisson sample the number of sources in each frequency, chirp mass, and redshift bin such that (Sesana et al. 2008; Middleton et al. 2015; Agazie et al. 2023b)

$$\begin{aligned}
h_c^2(f) &= \sum_{\log \mathcal{M}, z} \mathcal{P}\left(\frac{d^3 N}{d\log \mathcal{M} dz d\ln f_r}\Delta \log \mathcal{M} \Delta z \Delta \ln f\right) \\
&\quad \times \frac{h^2(f_r, \log \mathcal{M}, z)}{\Delta \ln f}. \qquad (18)
\end{aligned}$$

In our analyses, we implement the five different models used in SK23, calibrating $\dot{n}_0$ such that $A_{1\ {\rm yr}^{-1}} \approx 2 \times 10^{-15}$ at $f = 1\ {\rm yr}^{-1}$, as was found in Agazie et al. (2023a). We model a circular binary whose evolution is purely driven by GW

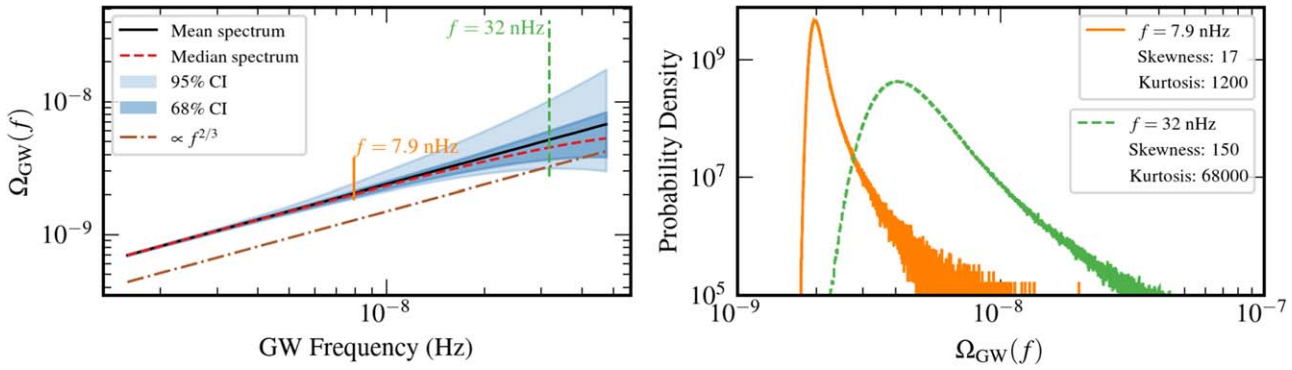

Figure 1. Gravitational-wave background from 10^7 realizations of a binary population described by Model 3 SK23. Left: we plot the mean and median cosmological energy density as a function of GW frequency. The blue contours represent the 68% and 95% confidence intervals of the spectra, and the orange and green lines represent the 99% confidence interval at two selected frequencies. For comparison, we also plot a $\propto f^{2/3}$ spectrum as expected from a population of circular SMBHB, inspiraling due to GW emission only. Right: the GWB distribution across 10^7 realizations for two selected frequencies, noted in the left plot. As frequency increases, the distribution is wider; hence the variance, skewness, and kurtosis increase.

emission, and hence $\lambda = 11/3$. Our population synthesis code is available on GitHub.[2]

4. Results

We now compare the 10^7 realizations of synthesized SMBHB populations against our analytical results. Given that $\lambda = 11/3$ for circular SMBHBs hardened by GW emission, from Table 1 we expect the following frequency relationships in our results: $\Omega_{\rm GW}(f) \propto f^{2/3}$, variance $\mathrm{Var}[\Omega_{\rm GW}(f)] \propto f^5$; $\mathrm{Skew}(f) \propto f^{11/6}$; $\mathrm{Kurt}(f) \propto f^{11/3}$.

In Figure 1, we show the mean, median, and 68% and 95% confidence intervals of simulated GWB spectra across 10^7 realizations of Model 3 from SK23 in terms of $\Omega_{\rm GW}(f)$. We see that the mean follows the expected $\Omega_{\rm GW}(f) \propto f^{2/3}$ relation that we would expect from a GWB sourced from a finite population of circular SMBHBs, while the spread in the spectrum increases as frequency increases. We show the 99% confidence interval for two frequencies, $f = 7.9$ nHz and $f = 32$ nHz, as vertical lines to further show the increase in the width of the distribution as frequency increases. In the right-hand panel, we show the shape of the distribution of $\Omega_{\rm GW}(f)$ across the 10^7 realizations at these two frequencies. We see that as frequency increases, the distribution widens, and hence the variance increases. The mass of the distributions is also to the left of the mean, and they become more asymmetrical and feature more extreme outliers (as shown by the wider tail) as frequency increases. This suggests positive skewness and positive excess kurtosis, which increases with increasing frequency, which agrees with our interpretation of Equations (13) and (14).

Figure 2 shows the mean, variance, skewness, and kurtosis of our 10^7 GWB realizations as a function of frequency for the five models defined in SK23. Since we defined $\dot{n}_0$ in each model such that $A_{1\ \mathrm{yr}^{-1}} \approx 2 \times 10^{-15}$ (Agazie et al. 2023a), it is expected, and observed, that the mean spectrum of each model is identical. Comparing the computed variance of the GWB to the guiding lines of $\propto f^5$ shows excellent consistency. As we get to higher-order moments, the effect of having a finite number of samples to approximate and analyze the distribution becomes apparent. For Models 1 and 2, the skewness and kurtosis follow the expected relation very well, despite some instability at higher frequencies. The skewness of Model 3 follows the frequency relation well across all frequencies too; however, for high frequencies, the kurtosis slightly deviates from the expected $f^{11/6}$ relation. For Models 4 and 5, the skewness and kurtosis follow the correct frequency relation at lower frequencies but deviate from this relation at higher frequencies. We attribute this simply to a limited sample size. Some of these models have very long tails dominated by rare massive binaries, and sampling a population realization where the GWB at this frequency deviates so far from the mass of the distribution to explore these tails is extremely unlikely. Hence, we underestimate the skewness and kurtosis at these frequencies. To create 10^7 realizations, we ran our code for 12 hr across 100 CPUs for each model. Running at least 10^{8-9} realizations could begin to ameliorate this; however, this would be computationally expensive. We also note that in Model 5 at $\sim$6 nHz there is a deviation from the expected skewness and kurtosis; this is due to a single extreme outlier in one realization. Removing this point removes this deviation, and the measured skewness and kurtosis are consistent with our expectations.

The magnitude (and numerical stability) of the variance, skewness, and kurtosis at high frequencies are model dependent. The populations that result in the largest variance and instability in the spectra are models that have a larger characteristic mass $\mathcal{M}_*$, which controls the exponential cutoff to the number of high-mass binaries in Equation (17). In these models where very high-mass SMBHBs are more likely, their large strain amplitudes cause them to dominate the GWB spectrum such that the expected number of SMBHBs in these frequency bins is $\mathcal{O}(1)$. Therefore, some realizations will not have a binary in these bins at all, while others will, resulting in a large variance (and skewness and kurtosis) of the GWB spectrum. To investigate the effect of the characteristic mass on the distribution of the spectrum, we create two sets of 10^7 realizations of a Model 3 population, where we change the characteristic mass to $\log \mathcal{M}_* = 7$ and $\log \mathcal{M}_* = 9$ (and normalize $\dot{n}_0$ as required to maintain $A_{1\ \mathrm{yr}^{-1}}$).

We show the results of these simulations in Figure 3. Increasing the characteristic mass increases the variance of the GWB, which strongly follows the f^5 dependence from Table 1. For the model with $\log \mathcal{M}_* = 9$, we see in the left panel that the mean spectrum across realizations is outside of the 68% confidence interval, suggesting strong influence by a small number of realizations with large $\Omega_{\rm GW}$ at high frequencies, corresponding to a small number of SMBHBs in those

[2] https://github.com/astrolamb/pop_synth

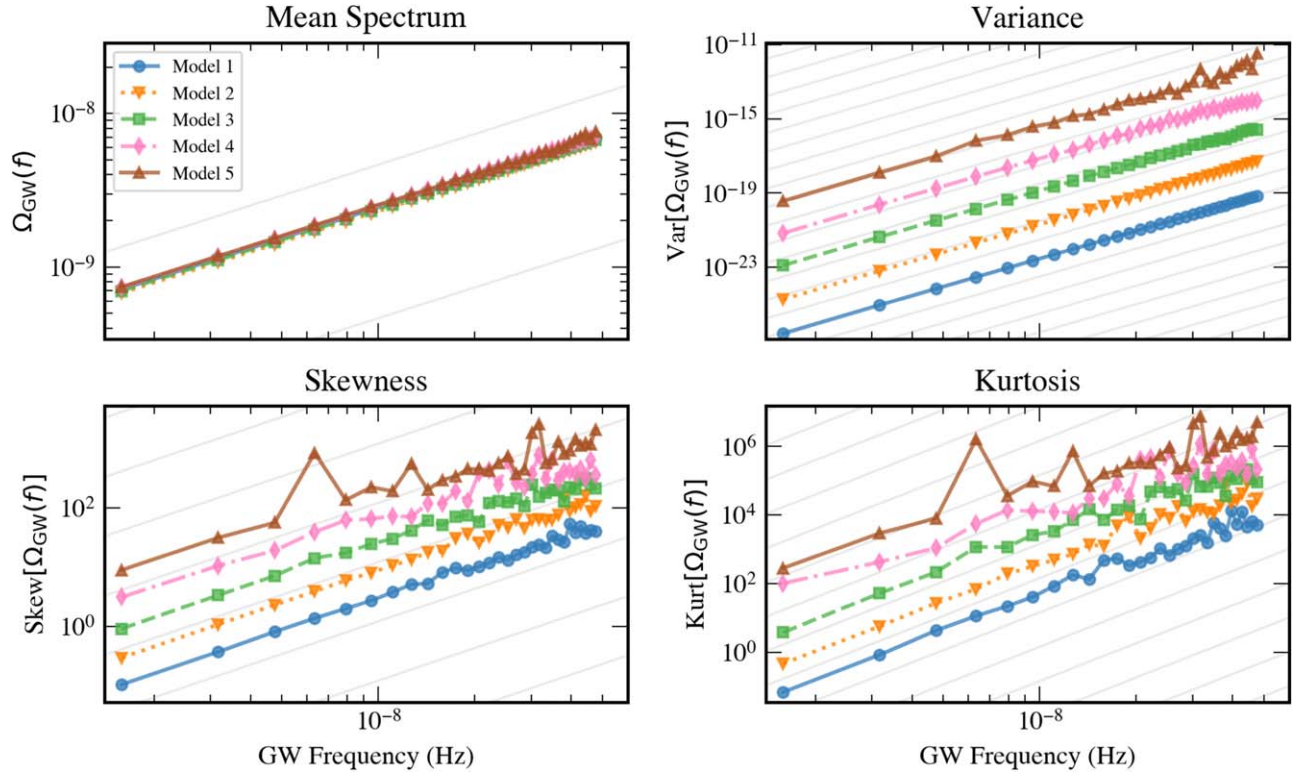

Figure 2. Comparing the mean, variance, skewness, and kurtosis of 10^7 realizations of the GWB as a function of GW frequency for the five models defined in SK23. The gray lines represent the general frequency scaling relation for the given statistic in the case of $\lambda = 11/3$, which are $\Omega_{GW}(f) \propto f^{2/3}$, $\mathrm{Var}[\Omega_{GW}] \propto f^5$, Skew $\propto f^{11/6}$, and Kurt $\propto f^{11/3}$. For all models, we see strong agreement with the relations found in Table 1.

realizations. However, for the majority of realizations there are no binaries and hence no GWB in these frequency bins, causing the median spectrum to depart from the mean.

Decreasing $\log \mathcal{M}_*$ results in decreasing variance across the spectrum because of the exponential cutoff at a relatively low mass, which leads to a large number of low-mass (hence low-strain-amplitude) binaries to create a GWB consistent with $A_{1\ \mathrm{yr}^{-1}} = 2 \times 10^{-15}$. The 68% confidence interval for $\log \mathcal{M}_* = 7$ is very narrow and cannot be seen in Figure 3. In the right panel of Figure 3, we compare the measured variance for the three choices of $\log \mathcal{M}_*$ (shown by markers) against the analytically derived variance in Equation (11) (plotted lines), showing excellent consistency.

5. Discussion

We derived analytical scaling relationships for the variance, skewness, and kurtosis of the GWB spectrum across realizations of a finite population of GW sources for the first time. For a binary population hardened by GW emission, we find that the mean, variance, skewness, and kurtosis of $\Omega_{GW}(f)$ scale as $f^{2/3}$, f^5, $f^{11/6}$, and $f^{11/3}$, respectively. We created simulated populations of SMBHBs and compared the spectral moments to our analytical forms, finding excellent agreement.

In addition, our approach can compute the moment statistics of the GWB spectrum for any df/dt. For example, we expect that SMBHBs interact with their galaxy environments at larger separations (hence lower GW frequencies), resulting in accelerated GW evolution in this frequency range. Assuming that each mechanism being considered follows a power-law relation f^{λ_i} for each mechanism i, we can model df/dt as

$$\frac{df}{dt} = g(f|\{c_i, \lambda_i\}) = \sum_i c_i f^{\lambda_i}, \tag{19}$$

for coefficients c_i, which depend on other binary and environmental properties. Furthermore, while this Letter considers only circular binary populations, one could extend our formalism to populations with eccentric orbits.

We have shown our analytical results for linear transformations of the GW cosmological energy density, $\Omega_{GW}(f)$. However, other useful quantities involve nonlinear transformations of $\Omega_{GW}(f)$, the most notable of which is the rms characteristic strain $h_c(f)$. Transforming these variables requires an expansion of the variable about the mean, and applying the statistic to that expansion. To leading order this gives $\mathrm{Var}[h_c(f)] \propto f^{7/3}$ for GW-hardened binaries, but this assumes that fluctuations about the mean are small. We have tested this against our numerical simulations, observing good agreement at low frequencies but poor performance at higher frequencies for models in which large strain outliers are more probable, e.g., low characteristic redshift and high characteristic chirp mass.

These results could be used in exploring the origin of GWB signals. For example, in PTA science the standard astrophysical GWB model is a power-law spectrum. However, a single population realization's spectrum will deviate from a power law. Hence, including information on higher-order moment statistics could lead to more robust analyses to account for spectral excursions. One could directly fit SMBHB population models to PTA data with this information (Sato-Polito & Zaldarriaga 2024), bypassing computationally expensive

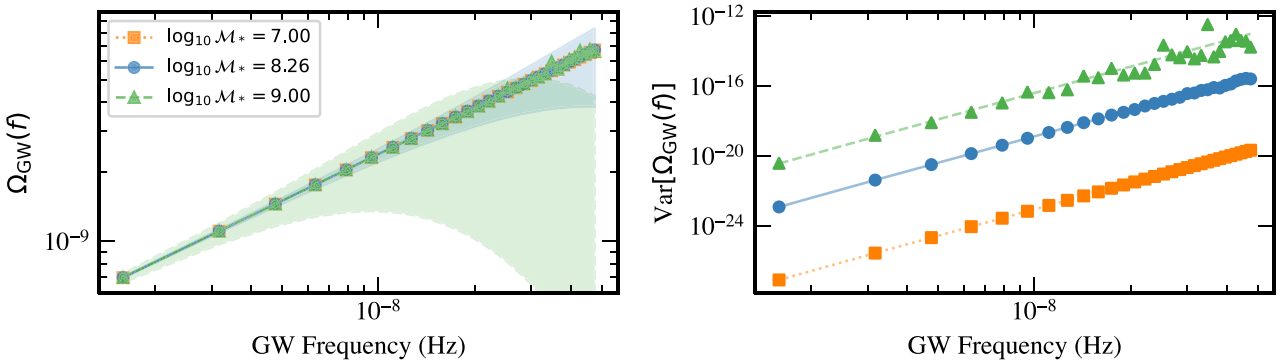

Figure 3. The characteristic mass $\mathcal{M}_*$ has a strong effect on the variance of the GWB. We use "Model 3" of SK23, but change $\mathcal{M}_*$ (and $\dot{n}_0$). Left: the mean $\Omega_{\rm GW}(f)$ as a function of frequency for the three submodels, along with the 68% confidence interval (CI) of the GWB (shaded contours). As $\mathcal{M}_*$ increases, the width of the 68% CI increases. Note that the mean spectrum for $\log \mathcal{M}_* = 9.00$ is outside of the 68% CI at high frequencies. Right: the variance of the GWB as a function of frequency for the three submodels. The markers are the measured variance from 10^7 realizations. The plotted lines are the analytically computed variance, given by Equation (11), for each model. The measured variance and analytical variance are strongly consistent.

simulations and machine learning (Taylor et al. 2017; Agazie et al. 2023b; Shih et al. 2024) using the methods developed in Lamb et al. (2023). This will enable more refined GWB and noise estimation in PTA data if the GWB is assumed to be of astrophysical origin. Indeed by adopting a more general dynamical model like Equation (19) in our calculations, one could leverage higher statistical moment information to probe the dynamics of SMBHB evolution in the subparsec regime.

PTA spectral characterization is likely too insensitive at present to accurately measure skewness and kurtosis, but these higher-order moments could be useful in the future. Meanwhile, the future LISA mission will be sensitive to several stochastic signals, such as cosmological stochastic backgrounds from first-order phase transitions, cosmic strings, and primordial black holes, a stochastic foreground from Galactic white dwarf binaries, and potentially backgrounds from extragalactic stellar-mass black hole binaries and extragalactic white dwarf binaries (Farmer & Phinney 2003; Babak et al. 2023; Colpi et al. 2024). Incorporating information on the statistics of spectral deviations from simple power-law models could help LISA analyses distinguish astrophysical from cosmological GWB signals.

Finally, beyond the population finiteness effects studied here, the additional influence of GW signal interference between sources will contribute as a source of (cosmic) variance (Allen 2023; Allen & Valtolina 2024) and potentially to higher statistical moments of the GWB spectrum. Frequency-scaling relationships for signal interference effects, and the relative importance of population finiteness and signal interference, will be explored in future work.

Acknowledgments

The authors wish to thank Kai Schmitz for fruitful discussions about this work. The authors are members of the NANOGrav collaboration, which receives support from NSF Physics Frontiers Center award No. 1430284 and 2020265. S. R.T. acknowledges support from an NSF CAREER No. 2146016, NSF AST-2007993, and NSF AST-2307719. W.G. L. acknowledges support from the Vanderbilt Physics & Astronomy McMinn Summer Research Award. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, TN.

Software: numpy (Harris et al. 2020), matplotlib (Hunter 2007).

ORCID iDs

William G. Lamb https://orcid.org/0000-0003-1096-4156
Stephen R. Taylor https://orcid.org/0000-0003-0264-1453

References

Abbott, R., Abbott, T., Abraham, S., et al. 2021, PhRvD, 104, 022004
Afzal, A., Agazie, G., Anumarlapudi, A., et al. 2023, ApJL, 951, L11
Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023a, ApJL, 951, L8
Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023b, ApJL, 952, L37
Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023c, ApJL, 956, L3
Agazie, G., Baker, P. T., Bécsy, B., et al. 2024, arXiv:2404.07020
Allen, B. 2023, PhRvD, 107, 043018
Allen, B., & Valtolina, S. 2024, PhRvD, 109, 083038
Antoniadis, J., Arumugam, P., Arumugam, S., et al. 2023a, A&A, 678, A50
Antoniadis, J., Arumugam, P., Arumugam, S., et al. 2023b, A&A, 685, A94
Babak, S., Caprini, C., Figueroa, D. G., et al. 2023, JCAP, 2023, 034
Bonetti, M., Franchini, A., Galuzzi, B. G., & Sesana, A. 2024, A&A, 687, A42
Burke-Spolaor, S., Taylor, S. R., Charisi, M., et al. 2019, A&ARv, 27, 1
Colpi, M., Danzmann, K., Hewitson, M., et al. 2024, arXiv:2402.07571
Digman, M. C., & Cornish, N. J. 2022, ApJ, 940, 10
Farmer, A. J., & Phinney, E. S. 2003, MNRAS, 346, 1197
Gardiner, E. C., Kelley, L. Z., Lemke, A.-M., & Mitridate, A. 2024, ApJ, 965, 164
Ginat, Y. B., Desjacques, V., Reischke, R., & Perets, H. B. 2020, PhRvD, 102, 083501
Ginat, Y. B., Reischke, R., Rapoport, I., & Desjacques, V. 2024, PhRvD, 109, 083526
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Hunter, J. D. 2007, CSE, 9, 90
Jaffe, A. H., & Backer, D. C. 2003, ApJ, 583, 616
Kaiser, A. R., Pol, N. S., McLaughlin, M. A., et al. 2022, ApJ, 938, 115
Kocsis, B., & Sesana, A. 2011, MNRAS, 411, 1467
Lamb, W. G., Taylor, S. R., & van Haasteren, R. 2023, PhRvD, 108, 103019
Middleton, H., Del Pozzo, W., Farr, W. M., Sesana, A., & Vecchio, A. 2015, MNRAS: Letters, 455, L72
Mingarelli, C. M., Sidery, T., Mandel, I., & Vecchio, A. 2013, PhRvD, 88, 062005
Peters, P. C. 1964, PhRv, 136, 1224
Phinney, E. 2001, arXiv:astro-ph/0108028
Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, ApJL, 951, L6
Renzini, A. I., & Golomb, J. 2024, arXiv:2407.03742
Roebber, E., Holder, G., Holz, D. E., & Warren, M. 2016, ApJ, 819, 163
Sampson, L., Cornish, N. J., & McWilliams, S. T. 2015, PhRvD, 91, 084055
Sato-Polito, G., & Kamionkowski, M. 2024, PhRvD, 109, 123544
Sato-Polito, G., & Zaldarriaga, M. 2024, arXiv:2406.17010
Sesana, A., Vecchio, A., & Colacino, C. N. 2008, MNRAS, 390, 192

Shih, D., Freytsis, M., Taylor, S. R., Dror, J. A., & Smyth, N. 2024, PhRvL, 133, 011402
Smarra, C., Goncharov, B., Barausse, E., et al. 2023, PhRvL, 131, 131
Taylor, S. R., Simon, J., & Sampson, L. 2017, PhRvL, 118, 181102
Xu, H., Chen, S., Guo, Y., et al. 2023, RAA, 23, 075024