

Probability distribution of astrophysical gravitational-wave background fluctuations


Yonadav Barry Ginat^{1,*}, Vincent Desjacques^{1,2,†}, Robert Reischke^{1,3} and Hagai B. Perets^{1,2,4}

¹*Physics department, Technion, Haifa 3200003, Israel*

²*Asher Space Science Institute, Technion, Haifa 3200003, Israel*

³*Department of Natural Sciences, The Open University of Israel, 1 University Road, P.O. Box 808, Ra'anana 4353701, Israel*

⁴*Theoretical Astrophysics 130-33, California Institute of Technology, Pasadena, California 91125, USA*

 (Received 13 October 2019; revised 3 July 2020; accepted 10 August 2020; published 1 October 2020)

The coalescence of compact binary stars is expected to produce a stochastic background of gravitational waves (GWs) observable with future GW detectors. Such backgrounds are usually characterized by their power spectrum as a function of the frequency. Here, we present a method to calculate the full one-point distribution of strain fluctuations. We focus on time series data, but our approach generalizes to the frequency domain. We illustrate how this probability distribution can be evaluated numerically. In addition, we derive accurate analytical asymptotic expressions for the large strain tail, which demonstrate that it is dominated by the nearest source. As an application, we calculate the distribution of strain fluctuations for the astrophysical GW background produced by binary mergers of compact stars in the Universe, and the distribution of the observed confusion background obtained upon subtracting bright, resolved sources from the signal. We quantify the extent to which they deviate from a Gaussian distribution. Our approach could be useful for the spectral shape reconstruction of stochastic GW backgrounds.

DOI: [10.1103/PhysRevD.102.083501](https://doi.org/10.1103/PhysRevD.102.083501)

I. INTRODUCTION

The direct discovery of gravitational waves (GWs) from binary black-hole mergers in 2015 [1] sparked a new interest in gravitational waves, which constitute a new window to the Universe. Because of the relative weakness of gravity, the amplitudes of gravitational waves are rather small, and many GW-emitting processes pass under our noses undetected. Their cumulative effect amounts to a gravitational-wave background that bathes the detectors (for a recent review, see [2]) and may, when investigated, reveal details of its physical origin (see, for instance, [3–13] for astrophysical backgrounds). This background shares a lot of properties with other cosmic backgrounds, such as the cosmic microwave background (e.g., their stochastic nature), but it is also unique, in ways we will explore here.

That the stochastic gravitation-wave background (SGWB) from cosmological sources (such the primordial GWs produced during inflation) is a Gaussian random field is well known [14]. The situation is less clear for backgrounds of astrophysical origin, although, *prima facie*, the central limit theorem suggests that the distribution of observed strain should converge toward a Gaussian when the number of sources N tends to infinity. As opposed to

primordial GWs, which are generated by inherently random quantum fluctuations during inflation, astrophysical sources are purely deterministic, insofar as their position, separation, masses, etc., specify the waveforms fully. Hence, while in the former case the wave amplitudes themselves are random, in the latter—which is our main concern in this paper—the only randomness is in the spatiotemporal location of the source on the past light cone of the observer (as well as its position in the relevant parameter space). Therefore, one expects that a superposition of signals from such sources need not be a Gaussian random field, unless N becomes very large.

So far, most of the literature on SGWBs has focused on the quantity

$$\Omega_{\text{gw}}(f) = \frac{1}{\rho_{\text{crit}}} \frac{d\rho_{\text{gw}}}{d \ln f}, \quad (1)$$

where

$$\rho_{\text{gw}} = \frac{c^2}{32\pi G} \langle \dot{h}_{ij} \dot{h}^{ij} \rangle \quad (2)$$

is the energy density of the observed GW strain h at the (observer-frame) frequency f . The density parameter Ω_{gw} integrated over all frequencies is then the *mean* SGWB energy density signature. In this paper, we aim to provide a

*ginat@campus.technion.ac.il
†dvince@physics.technion.ac.il

tractable approach for determining the full (one-point) probability distribution function (PDF) of h around this mean energy. For simplicity, we will focus on strain time series $h(t)$ analogous to those simulated in, e.g., [15–17], but we emphasize that our approach is capable of resolving the frequency dependence of the full distribution function (see Sec. VI for a brief discussion of this point).

In order to carry out this task, we exploit the fact that GWs propagate over distances much larger than the typical clustering length r_ξ of the sources. Therefore, Poisson clustering should provide a good approximation when the bulk of the sources lie at distances $r \gg r_\xi$ from the detector. This simplification enables us to calculate the PDF of the observed strain from the knowledge of the characteristic functions of individual sources solely. In particular, we are interested in the asymptotic behavior of the strain distribution in the large strain limit, the impact of interferences, and the validity of the Gaussian approximation. Furthermore, as a demonstration of the applicability of our methods, we will apply them to the characterization of the PDF of the SGWB produced by binary mergers of compact stars.

The paper is organized as follows: We lay down our assumptions and spell out our approach in Sec. II. We then move on to exemplify our method with a toy model in Sec. III, which we subsequently expand in Sec. IV in order to estimate the probability distribution of the SGWB produced by binary mergers of compact stars in the Universe. In Sec. V we consider the removal of bright sources from the background estimation. We discuss our results, along with a number of possible extensions, in Sec. VI before concluding in Sec. VII. Throughout, we assume that space-time is described by a flat Friedman-Lemaître-Robertson-Walker (FLRW) metric, with cosmological parameters $\Omega_m = 0.32$, $\Omega_{\text{rad}} = 9.187 \times 10^{-5}$, $\Omega_\Lambda = 1 - \Omega_m - \Omega_{\text{rad}}$, $h = 0.674$ [18].

II. GENERAL CONSIDERATIONS

Suppose that a gravitational-wave detector located at the origin of the coordinate system receives signals from N sources distributed uniformly within a sphere of (comoving) radius R . The sources radiate signals—each of which is characterized by its own relevant physical processes—which travel to the detector, where they interfere to produce a total signal [19]

$$s(t) = h(t) + n(t) \quad \text{with} \quad h(t) \equiv \sum_{\text{source } i} h_i(t). \quad (3)$$

Here, $n(t)$ is the detector noise [which is uncorrelated with $h(t)$], and $h_i(t)$ is a linear combination of the two different polarizations h_+ and h_\times of the waves [20]. Presupposing that the detector noise can be mitigated (using advanced interferometers like LISA and/or cross-correlation among multiple detectors), the strain time series $h(t)$ (sampled into

small time intervals)—which contains the largest amount of information about the GW background (down to the residual noise level)—could be extracted from the data (possibly sampled into small time intervals).

Our goal is to find a way to calculate the one-point PDF $P(h)$, which gives the probability of measuring a strain $[h, h + dh]$ at the detector, and to comprehend its basic properties.

For this purpose, we make the following assumptions:

- (1) Space-time is described by a flat FLRW metric in comoving coordinates (t, \mathbf{r}) , with cosmological parameters as specified in Sec. I. Henceforth, t will stand for the cosmic time, while η will denote the conformal time. The sources and the detector are idealized comoving frames.
- (2) The GW sources are independent of each other and are all instances of the same type of sources (say, compact binaries), but may have varying physical parameters (say, chirp masses).
- (3) The sources are distributed homogeneously in comoving space according to an isotropic Poisson process. Therefore, the number N of sources within a sphere of comoving radius R is Poisson distributed with a mean count $\lambda(R)$: $N \sim \text{Pois}(\lambda(R))$.
- (4) Each source has a probability density $R_*(t)$ of turning on at a given cosmic time t .
- (5) A source which turned on at $t = t_*$ at a comoving distance $r = |\mathbf{r}|$ from the detector produces a strain at $r = 0$ and time $t_0 > t_*$, which is given by a known function g :

$$h(t) = \frac{d}{d_L(r)} e^{-r/r_0} g \left(\int_0^{\eta_{0,\text{ret}}} d\eta a(\eta), \varphi \right), \quad (4)$$

$\eta_{0,\text{ret}} = \eta_0 - r/c$ is the retarded conformal time, η_0 is the conformal time measured by the detector (i.e., the conformal age of the Universe today), d is an arbitrary distance chosen to normalize the strain (to make g dimensionless), and φ is a random phase uniformly distributed in the interval $[0, 2\pi]$. We include an exponential decay with characteristic length scale r_0 for generality. This could represent a physical damping caused by an anisotropic dark matter energy-momentum tensor [21,22], for instance.

- (6) The wavelength of the GWs is considerably smaller than the radius of curvature of the Universe. As a result, the polarizations decouple and propagate along null geodesics ([19] Sec. IV.1.4).

Let us comment briefly on some of these assumptions: In general, the properties of the GW signal g depends on a set of model parameters (which include, e.g., the chirp mass of a binary star), $\xi = (\xi_1, \xi_2, \dots)$. As mentioned earlier, one expects that the distant astrophysical sources be clustered spatially (a phenomenon which also leads to anisotropies

on large angular scales, as recently discussed in the literature by, e.g., [3,8,10,23–26]). However, since the GWs propagate over large distances without being attenuated significantly (any viable r_0 is much greater than r_ξ), it is reasonable to assume that the sources follow Poisson clustering. When r_0 is large but finite, it is convenient to work with the mean number N_0 of sources inside a sphere of radius r_0 . As the source counts are also homogeneous along the radial direction, one has $N_0 = \lambda(R)(r_0/R)^3$.

In practice, h is a superposition of independent contributions, so it is natural to calculate the Fourier transform of its PDF $P(h)$, which decomposes into a product of the characteristic functions of the waves emitted by the individual sources [27–29]. For a spatial Poisson process, the probability to find a *single* source at position \mathbf{r} with parameters ξ and phase φ is

$$p(\mathbf{r}, \xi, \varphi) = \frac{4\pi r^2 dr}{\frac{4}{3}\pi R^3} \cdot \phi(\xi) d\xi \cdot \frac{d\varphi}{2\pi}. \quad (5)$$

Here and henceforth, $\phi(\xi)$ denotes the measure on the source parameter space [30].

Let $\psi(q)$ be the corresponding characteristic function—the PDF’s Fourier transform—

$$\psi(q) = \frac{3}{4\pi R^3} \int_0^R dr r^2 \int_{S^2} d\Omega \int d\xi \tilde{p}(q; \xi; \mathbf{r}), \quad (6)$$

where

$$\tilde{p}(q; \xi; \mathbf{r}) = \int_0^{2\pi} \frac{d\varphi}{2\pi} \exp \left[\frac{iqd e^{-\frac{r}{r_0}}}{d_L(r)} g \left(\int_0^{\eta_{0,\text{ret}}} d\eta a(\eta), \varphi \right) \right] \quad (7)$$

is the phase average of $\exp(iqh(t))$ for a single source.

Since the sources are identical, they share the same characteristic function $\psi(q)$, whence the characteristic function $\Xi(q)$ for Poisson distributed sources in a sphere of radius R is the Poisson mixture

$$\begin{aligned} \Xi(q) &= \sum_{k=0}^{\infty} \left(\frac{\lambda^k}{k!} e^{-\lambda} \right) \cdot \psi(q)^k \\ &= e^{\lambda\psi(q) - \lambda}. \end{aligned} \quad (8)$$

(One could start the sum at $k = 1$ and end up with an irrelevant additive constant term which can be dropped.) Consequently, the PDF of the observed strain produced by Poisson distributed sources is given by the inverse Fourier transform

$$P \left(\sum h_k = h \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dq e^{-iqh} e^{\lambda\psi(q) - \lambda}. \quad (9)$$

The main question is how to obtain a meaningful expression for $\psi(q)$. As we will show below, the asymptotic

properties of $P(h)$ can be gleaned from Eq. (9) without explicit knowledge of $\psi(q)$, as they rely only on universal properties of the physical system. For example, provided that each source has an equal probability of emitting $+h_k$ as $-h_k$, $\psi(q)$ is an even function of q and, therefore, $P(h)$ is also an even function of h . This implies that $P'(h = 0) = 0$; i.e., $P(h)$ flattens for small values of $|h|$.

An analytical derivation of the form of $P(h)$ at large h is somewhat more challenging and is expounded below. The results are reminiscent of cosmological results obtained with the theory of large deviations (see [31] for a review) when applied in the context of the large scale structure (see, for instance, [32–38]). At this point, it is worthwhile to draw an analogy between the physical system considered here and random walks. Since the observed strain is a superposition of independent but identical waves, its time evolution is analogous to a random walk in the complex plane: The position of the walker after n steps is of the form $\sum_{k=1}^n A_k \exp(i\varphi_k)$, which is precisely a sum of n random waves, with φ_k representing the phase of the k th wave as it reaches the detector (see [39,40]). Here, however, the amplitudes of the waves have a unique property—they all obey the law of gravity (they decay as $1/r$), but *ipso facto* one will be able to deduce general properties of the PDF.

To exemplify our method and illustrate the key properties of $P(h)$, it is instructive to consider a simplified scenario first. This is the focus of Sec. III. A more realistic calculation will be carried out in Sec. VI.

III. A SIMPLIFIED MODEL

In order to understand the features of $P(h)$, we consider a simple test case in which the GWs emitted are all described by a pure sine wave with a constant frequency ω , i.e., $g(t) = A \cos(\omega t + \varphi)$. Furthermore, we ignore any cosmological effects, but assume there is a finite attenuation length r_0 . Picking up $d \equiv r_0$, the strain produced by a single source at the detector is

$$g(t, \varphi) = \frac{Ar_0 e^{-r/r_0}}{r} \cos(\omega t + \varphi). \quad (10)$$

The source parameters are $\xi = (A, \omega)$ in this case. We leave the distributions of A and ω , $\phi(A)$ and $\phi(\omega)$, undetermined. While this simplified model serves its purpose of demonstrating the salient points of the mathematical technique used in this paper, it does not, however, constitute a simplified version of the physical case discussed below in Sec. IV.

A. The characteristic function

The phase average of $\exp(iqh(t))$ needed for the computation of the single source characteristic function is

$$\begin{aligned}\tilde{p}(q; A, \omega; r) &= \int_0^{2\pi} \frac{d\varphi}{2\pi} \exp \left[i q \frac{A r_0 e^{-r/r_0}}{r} \cos(\omega t + \varphi) \right] \\ &= J_0 \left(q \frac{A r_0 e^{-r/r_0}}{r} \right),\end{aligned}\quad (11)$$

where J_0 is the zeroth (cylindrical) Bessel function.

The characteristic function $\psi(q)$ is obtained upon a further average over the volume of a sphere of radius R (corresponding to the maximum distance at which sources can form, regardless of the detector sensitivity), the distribution of the amplitude A , and, in principle, the frequency ω (although the latter is immaterial in the one-point PDF as \tilde{p} is independent of it). Explicitly,

$$\psi(q) = \int d\omega \phi(\omega) \int dA \phi(A) \frac{3}{R^3} \int_0^R dr r^2 \tilde{p}(q; A, \omega; r),\quad (12)$$

or, upon substituting the variables $\tau \equiv r/r_0$, $s \equiv q h_c$, and $b \equiv A/h_c$,

$$\psi(s) = \frac{3N_0}{\lambda} \int db \phi(b) \int_0^{R/r_0} d\tau \tau^2 J_0 \left(s \frac{b e^{-\tau}}{\tau} \right). \quad (13)$$

The normalization strain h_c is defined through

$$\begin{aligned}\lambda h_c^2 &\equiv 3N_0 \int dA \phi(A) \int_0^{R/r_0} d\tau \tau^2 \cdot \frac{1}{2} \left(\frac{A e^{-\tau}}{\tau} \right)^2 \\ &= \frac{3}{4} N_0 \langle A^2 \rangle (1 - e^{-2R/r_0}) \\ &\approx \frac{3}{4} N_0 \langle A^2 \rangle.\end{aligned}\quad (14)$$

The last approximation is valid only if $R \gg r_0$. A multiplicative factor of λ is taken out of h_c^2 , so that the latter is defined for a mean count of unity. In fact, $\lambda h_c^2 \equiv \langle h^2 \rangle$ is exactly the variance of strain fluctuations, as is readily visible from computing the second moment of $P(h)$ from the general relation

$$\langle h^k \rangle = (-i)^k \frac{d^k}{dq^k} [e^{\lambda \psi(q) - \lambda}] \Big|_{q=0}. \quad (15)$$

The transformation $q \mapsto s$ implies that h is now measured in units of h_c . All one has to do to obtain a PDF for h itself is to divide by h_c , where needed.

It is evident from Eq. (13) that the quantity we are interested in, $\lambda \psi(s) - \lambda$, depends linearly on N_0 . Therefore, it is easier to define $N_0 G(s) = \lambda \psi(s) - \lambda$, so that $G(s)$ is independent of N_0 . Then,

$$G(s) = 3 \int db \phi(b) \int_0^{R/r_0} d\tau \tau^2 \left[J_0 \left(s \frac{b e^{-\tau}}{\tau} \right) - 1 \right]. \quad (16)$$

Note that $G(s)$ is nothing but the cumulant generating function for $N_0 = 1$. It satisfies $G(0) = G'(0) = 0$, while $G''(0) < 0$, it reaches its global maximum at $s = 0$.

B. The shape of the distribution

The large- h asymptotic behavior of $P(h)$ can be obtained for any N_0 , *vide infra* and Appendix A. Suppose first that N_0 is small. Taking $h > 0$ [without loss of generality since $P(h)$ is even in h], one may expand the characteristic function $e^{N_0 G(s)} \approx 1 + N_0 G(s)$. Ignoring the average over b for the moment, one obtains

$$\begin{aligned}P(h) &\sim \frac{1}{2\pi} \int_{-\infty}^{\infty} ds e^{-ish} (1 + N_0 G(s)) \\ &= \frac{3N_0}{2\pi} \int_0^{R/r_0} d\tau \tau^2 \int_{-\infty}^{\infty} ds J_0 \left(\frac{b s e^{-\tau}}{\tau} \right) e^{-ish}.\end{aligned}\quad (17)$$

The Fourier transform may now be performed easily, viz.,

$$P(h) \sim \frac{3N_0}{\pi} \int_0^{R/r_0} d\tau \begin{cases} \frac{\tau^2}{\sqrt{b^2 e^{-2\tau}/\tau^2 - h^2}}, & \text{if } |h| < \frac{b e^{-\tau}}{\tau}, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

If $u = \tau e^\tau$, then, provided that $b/|h| \leq R/r_0 e^{R/r_0}$ (which ought to be the case for sufficiently large $|h|$),

$$P(h) \sim \frac{3N_0}{\pi} \int_0^{b/|h|} \frac{du W^3(u)}{b(1+W(u))\sqrt{1-h^2 u^2/b^2}}, \quad (19)$$

where $W(u)$ is Lambert's W function. If $|h| \gg b$, then the integration domain contains solely small values of u , for which $W^3(u) \approx u^3$, whence, to leading order,

$$P(h) \sim \frac{3N_0}{\pi} \frac{b^3}{h^4} \int_0^1 \frac{dx x^3}{\sqrt{1-x^2}}, \quad (20)$$

where $x = hu/b$. Upon performing the final integral, the first order asymptotic expansion of $P(h)$ turns out to be

$$P(h) \sim \frac{2N_0 \langle b^3 \rangle}{\pi h^4}. \quad (21)$$

This meshes well with the numerical evaluation of $P(h)$ accurately in the large- h limit, as can be seen from the top panel of Fig. 1.

We contend that this asymptotic relation still holds even when N_0 is not small. We give a simple argument for this here and present a full derivation in Appendix A. Away from the origin, $G(s)$ is a smooth (C^∞) function, and, consequently, its Fourier transform decays faster than any power law. Indeed, if we divide up the real line into three intervals $I_0 = (-\delta, \delta)$, $I_- = (-\infty, -\delta + \varepsilon)$, $I_+ = (\delta - \varepsilon, +\infty)$, where $0 < \varepsilon < \delta \ll 1$, and take a partition of unity $\{\eta_0, \eta_+, \eta_-\}$ subordinate to this division, then we can

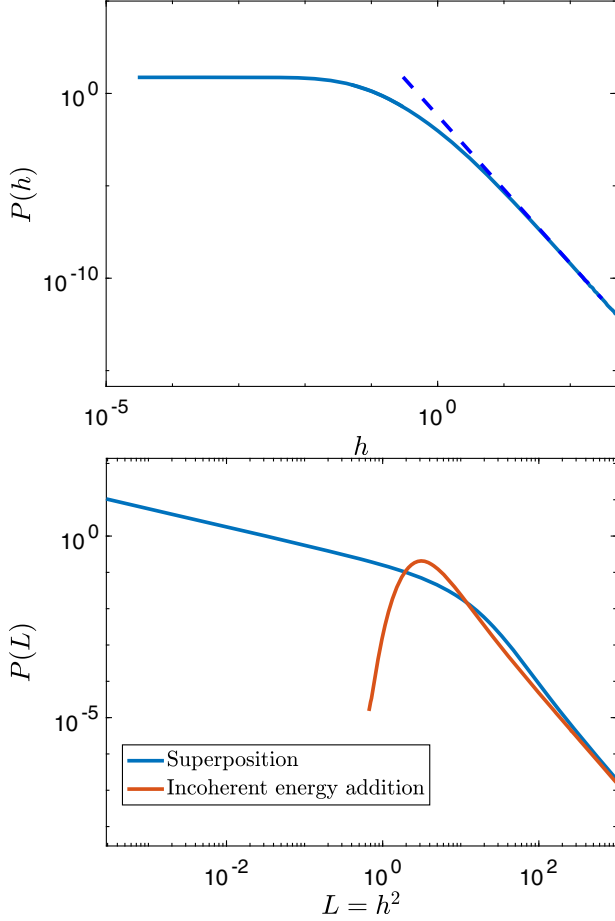


FIG. 1. Top: Numerical evaluation of the Fourier transform of $\exp(\lambda\psi - \lambda)$ (solid) along with the leading, large- h asymptotic result (21) for the illustrative model considered in Sec. III. Bottom: Comparison between the probability distribution for h^2 obtained via Eq. (9) and that for L obtained by summing up the intensities of individual source (see text). N_0 is set to 0.1 for the top panel, and 10 for the bottom panel. This leads to $\langle L \rangle = \langle h^2 \rangle = 7.5$ in the right panel. A distribution $\phi(A)$ peaked sharply at $A = 1$ is assumed throughout.

divide $P(h)$ accordingly, too. More precisely, let $P_i = \frac{1}{2\pi} \int_{\mathbb{R}} \eta_i e^{-ish} e^G ds$, so that $P = \sum_{i \in \{0, \pm\}} P_i$. For $i \neq 0$, P_i is the Fourier transform of a smooth function, and therefore decays exponentially with h . Hence, a power-law decay must originate from P_0 , if there is one at all [41]. Thus, $P(h) \sim P_0(h)$ as $h \rightarrow \infty$. The advantage of adding $\eta_0(s)$ is that it vanishes outside I_0 , and one thus may expand the exponential as before. After expanding, η_0 may be removed and the integration limits restored to $\pm\infty$, accruing only exponentially small errors. This argument is similar to that presented in [42], in the context of the method of stationary phase.

C. The importance of interferences

Equation (21) implies that the probability of observing a large value of h is dominated by the nearest neighbor.

To see this, assume that the individual strains $h_i(t)$ are fully incoherent. As a result, cross terms vanish, and the total intensity L is the sum of the individual intensities h_i^2 ,

$$L \equiv \sum_{\text{source } i} \bar{h}_i^2, \quad (22)$$

as one would expect for incoherent electromagnetic radiation. Here, \bar{h}_i^2 is the average of $h_i^2(t)$ over the duration of the experiment. In analogy with the ionizing background produced by quasars, we have (e.g., [29])

$$P(L) \stackrel{L \rightarrow \infty}{\sim} L^{-5/2}. \quad (23)$$

This power-law behavior is known to arise from the nearest neighbor [43]. Similarly, bearing in mind that $P(h^2) = P(h)/h$, the large- h asymptotic scaling (21) implies

$$P(h^2) = P(h)/h \stackrel{h \rightarrow \infty}{\sim} h^{-5} \sim (h^2)^{-5/2}, \quad (24)$$

which demonstrates our assertion.

The distribution $P(L)$ can be computed using the technique outlined above (cf. [29]) upon substituting

$$g(t) = \frac{A^2 r_0^2 e^{-2r/r_0}}{2r^2} \quad (25)$$

in the expression of $\tilde{p}(q; A, \omega; r)$. We included a factor of $\frac{1}{2}$ so that Eq. (25) is precisely the time average of Eq. (10). Consequently, the dependence on the random phase φ is trivial, and one is left with

$$G(s) = 3 \int db \phi(b) \int_0^{R/r_0} d\tau \tau^2 \left[\exp\left(i s b \frac{e^{-2\tau}}{2\tau^2} \right) - 1 \right], \quad (26)$$

where $s \equiv q/L_c$ and $b \equiv A^2/L_c$. We choose the normalization to be $L_c \equiv h_c^2$, such that $\lambda L_c \equiv \langle L \rangle$. The function $G(s)$ gives $P(L)$ upon a Fourier transformation.

The bottom panel of Fig. 1 displays $P(L)$ along with $P(h^2)$ for the comparison. While the two distributions exhibit the same power-law behavior at high L , as explained above, there is a stark difference at low L . This emphasizes the crucial role that interference, both constructive and destructive, plays in the determination of the observed time series. Note that the means of both distributions are equal, $\lambda L_c = \lambda h_c^2$, reflecting the fact that the variance of independent random variables is additive. In other words, a measurement of $\langle h^2 \rangle$ (that is, ρ_{gw} for a realistic GW background) does not provide any information about the (in)coherence of the signal. As a rule of thumb, the distribution is well approximated by a Gaussian for $|h| \ll \sqrt{\lambda} h_c$ (many sources contribute to the observed

strain) and by power-law tails for $|h| \gg \sqrt{\lambda} h_c$ (a single, nearby source dominates the signal).

IV. STOCHASTIC BACKGROUND FROM BINARIES

Having understood the salient features of the PDF of the observed strain in a simplified case, we now turn our attention to a more realistic source of the SGWB: binary mergers of any combination of white dwarfs, neutron stars, or black holes. All but the final stages of such systems may be described by a Keplerian orbit perturbed by the gravitational radiation reaction [19].

A. Basic relations

We approximate the sources as Keplerian throughout, and consider circular orbits—an assumption justified by the circularizing effect of gravitational-wave emission. As a result, the detector measures a certain linear superposition of the two polarizations $h = F_+ h_+ + F_\times h_\times$ with waveforms given by [19]

$$F_+ h_+(\mathbf{t}) = h_0(\mathbf{t}) \frac{1 + \cos^2 i}{2} \cos(2\theta) \cos \phi(\mathbf{t}), \quad (27)$$

$$F_\times h_\times(\mathbf{t}) = h_0(\mathbf{t}) \cos i \sin(2\theta) \sin \phi(\mathbf{t}), \quad (28)$$

where θ, i describe the orientation of the binary relative to the detector, $\mathbf{t} = t_{\text{coal}} - t$ is the time to coalescence as measured by the observer (or at the source, as it is comoving), and

$$\phi(\mathbf{t}) = -2 \left(\frac{5G(1+z)M_c}{c^3} \right)^{-5/8} \mathbf{t}^{5/8} + \varphi, \quad (29)$$

$$h_0(\mathbf{t}) = \frac{4}{d_L(z)} \left(\frac{G(1+z)M_c}{c^2} \right)^{5/3} \left(\frac{\pi f_{\text{gw}}^{\text{obs}}(\mathbf{t})}{c} \right)^{2/3}, \quad (30)$$

$$f_{\text{gw}}^{\text{obs}}(\mathbf{t}) = \frac{1}{\pi} \left(\frac{5}{256\mathbf{t}} \right)^{3/8} \left(\frac{G(1+z)M_c}{c^3} \right)^{-5/8}. \quad (31)$$

Here,

$$M_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \quad (32)$$

is the chirp mass, φ is a (random) phase, and z is the source redshift. Furthermore, the time to coalescence reads

$$\mathbf{t} = \tau_0(M_c, T) - (t_{0,\text{ret}}(r) - t_*), \quad (33)$$

where t_* is the formation time of the binary,

$$t_{0,\text{ret}}(r) = \int_0^{n_0 - r/c} a(\eta) d\eta \quad (34)$$

is the retarded age of the Universe, and $\tau_0(M_c, T)$ is the lifetime of a binary—the time it would take until both members collide; assuming gravitational-wave emission solely, we have $\tau_0(M_c, T) = 5c^5 T^{8/3} (GM_c)^{-5/3}$ ([19] p. 171). We set the present-day scale factor to unity.

Since the phase φ is random, the average of $\exp(iq\hbar(\mathbf{t}))$ over φ returns a Bessel function for \tilde{p} , as before. Including an exponential decay with attenuation length r_0 (to which d is set), this implies

$$\tilde{p}(q; \xi; \mathbf{r}) = J_0 \left(\sqrt{F_+^2 + F_\times^2} h_0(\mathbf{t}, z(r)) q e^{-r/r_0} \right). \quad (35)$$

The redshift dependence induces a dependence on the comoving distance r through the Friedmann equations.

Next, \tilde{p} must be further averaged over the sphere \mathcal{S}^2 as well as all the model parameters ξ in order to get ψ . But first of let us stress that the integral over t_{coal} is equivalent to an integral over the “starting time” and, as such, it must be weighted by the density of binary progenitors (which we take to be proportional to the star formation rate).

Finally, the large- h asymptotics of $P(h)$ can be obtained upon separating the argument of the Bessel function into an “amplitude” part $A(r)$ multiplied by a “propagation” part $\exp(-r/r_0)/d_L = \exp(-r/r_0)a(r)/r$. Following the procedure outlined in Sec. III, one obtains the same asymptotic expansion at large h . Namely (see also Appendix A),

$$P(h) \sim \frac{2N_0 \langle b(\tau=0)^3 \rangle}{\pi h^4}, \quad (36)$$

where $b(\tau=0)$ is the normalized amplitude evaluated at $\tau=0$. The coefficient b is evaluated at $\tau=0$, because, in passing through Eq. (18), b is now a function of τ , which we Taylor expand about 0. The terms above zeroth order contribute higher powers of h in the denominator and are thus neglected.

B. Probability measure for the source parameter space

Before delving into the calculation of the SGWB characteristic function, let us remark on the probability measure we adopt for the model parameters $\xi = (t_*, M_c, T, M_{\text{gal}})$. While we attempt to use physically plausible rates, our main goal is to illustrate how our formalism can be used to calculate the PDF of the SGWB produced by binary mergers. Therefore, even though the rates and parameters chosen here were picked so as to simplify this task, we stress that the procedure laid down in this section may be used for any choice of rate functions, such as more accurate estimates based on stellar evolution simulations and past GW merger events [44,45].

To demonstrate our technique, we adopt the so-called “reference model” of Cusin *et al.* [46], which we summarize below for the readers’ convenience. Let us approximate

the probability density that a binary of compact stars forms at cosmic time t_* with semimajor axis a and chirp mass M_c as the product distribution

$$\phi(t_*, M_c, T, M_{\text{gal}}) = R_*(t_*)\phi(M_{\text{gal}}|t_*)\phi(M_c|t_*, M_{\text{gal}}) \times \frac{\ln(a_{\text{max}}/a_{\text{min}})}{a}. \quad (37)$$

$R_*(t_*)$ is a probability distribution which accounts for the redshift evolution of the star formation rate, which we model as

$$R_*(z) = (1+z)e^{-z^2/(2\sigma^2)} \left(\sigma^2 + \sqrt{\frac{\pi}{2}}\sigma \right)^{-1}. \quad (38)$$

Here, $\sigma = \sqrt{6}$, so that R_* peaks at redshift $z = 2$ [47]. $R_*(z)$ can be converted into $R_*(t_*)$ using the redshift-to-cosmic-time relation. The conditional probabilities $\phi(M_{\text{gal}}|t_*)$ and $\phi(M_c|t_*, M_{\text{gal}})$ will be discussed shortly. Finally, the $1/a$ probability density for the initial orbital period is derived from Öpik's law [48], which is a reasonable approximation to the observed Galactic period distribution over a fairly large range of periods [49]. We choose the bounds $a_{\text{min}} = 0.014$ AU and $a_{\text{max}} = 4000$ AU as in [46]. They translate into limits on the initial period T

(by Kepler's third law) using the masses of the binary components.

The knowledge of T allows us to compute the binary lifetime τ_0 . Then, $G(s)$ is obtained from an integration over t_* , M_c , T , M_{gal} . The nontrivial integration limits are as follows: T lies between T_{max} and $\max\{T_{\text{min}}, \tau_0^{-1}(t_{0,\text{ret}} - t_*)\}$, while t_* runs from 0 to $t_{0,\text{ret}}(r)$. The reason for this choice of integration limits is twofold: Binaries which merge such that the signal from the merger event reaches the detector (origin) before t_0 do not contribute to the gravitational-wave signal at t_0 and neither do binaries which form at (t_*, r) such that the signal sent at their birth does not reach the detector in time.

To model the chirp-mass distribution $\phi(M_c|t_*, M_{\text{gal}})$, suppose that the initial masses m_1 and m_2 of the two binary members have both broken power-law densities, $\phi(m) = Cm^{-\alpha}$ with α dependent on m . We choose a Kroupa mass function [50,51] (in the mass range we consider $\alpha = 2.7$) rather than a Salpeter mass function as in [46]. This is the only difference with their model.

The SGWB is mainly produced by remnants of stellar evolution, whose final mass is related to m_1 , m_2 by the so-called ‘‘initial-to-final-mass function’’ $\mu(m, Z)$, which depends on the metallicity Z . Following [46], we adopt the functional form of the delayed model presented in Fryer *et al.* [52] (here, all masses are in solar mass units):

$$\mu(m, Z) = \begin{cases} 1.3, & \text{if } m \leq 11, \\ 1.1 + 0.2e^{(m-11)/4} - (2+Z)e^{2(m-26)/5}, & \text{if } 11 < m \leq 30, \\ \min\{33.35 + (4.75 + 1.25Z)(m-34), m - \sqrt{Z}(1.3m - 18.35)\}, & \text{otherwise.} \end{cases} \quad (39)$$

The metallicity depends on the cosmic time at which the system formed, a dependence which we model (following again [46]) using the fit of Ma *et al.* [53] for the gas density (from which the stars formed):

$$\log_{10}\left(\frac{Z(z, M_{\text{gal}})}{Z_{\odot}}\right) = 0.35 \left[\log_{10}\left(\frac{M_{\text{gal}}}{M_{\odot}}\right) - 10 \right] + 0.93e^{-0.43z} - 1.05. \quad (40)$$

The redshift is z then, again, converted to cosmic time t_* using the cosmic-time-to-redshift relation.

All this leads to a chirp-mass distribution

$$\phi(M_c|t_*, M_{\text{gal}}) \equiv \phi(M_c|Z(t_*, M_{\text{gal}})) \quad (41)$$

given by

$$\phi(M_c|Z) = \iint dm_1 dm_2 \phi(m_1)\phi(m_2) \times \delta_{\text{D}}\left(M_c - \frac{(\mu(m_1, Z)\mu(m_2, Z))^{3/5}}{(\mu(m_1, Z) + \mu(m_2, Z))^{1/5}}\right), \quad (42)$$

where $\delta_{\text{D}}(x)$ is the Dirac delta function.

The last ingredient is $\phi(M_{\text{gal}}|t_*)$, which we model using the halo mass function of Tinker *et al.* [54] and under the assumption that the total stellar mass in a galaxy is proportional to its halo mass.

The knowledge of the functions $R_*(t_*)$, $\phi(M_{\text{gal}}|t_*)$, $\phi(M_c|t_*, M_{\text{gal}})$, and $\phi(a) \propto 1/a$ fully specifies the PDF of $\phi(\xi)$, up to an overall rate encapsulated in the value of N_0 which we are free to specify.

C. Determination of N_0

N_0 is the average number of overlapping sources in the time-domain data, whose frequencies are in the detector's observing band. While N_0 could be as large as, say, 10^6 – 10^9 if we were observing the stochastic background across all frequencies, it is of order $\mathcal{O}(10)$ only if frequencies are restricted to $f_{\text{gw}}^{\text{obs}} \gtrsim 1$ Hz [17]. Here and henceforth, we will match our predictions to the LIGO's frequency band. This restricts the source parameters via Eq. (31) given by $f_{\text{min}} \leq f_{\text{gw}}^{\text{obs}} \leq f_{\text{max}}$, where $f_{\text{min}} = 5$ Hz and $f_{\text{max}} = 5000$ Hz. Imposing such a constraint on $f_{\text{gw}}^{\text{obs}}$

through an additional Heaviside function in the definition of $G(s)$ (see below) affects the determination of N_0 .

To determine N_0 [and, thereby, the absolute normalization of the star formation rate $R_*(z)$ in our model], we use the approximate relation

$$N_0 \sim R_{\text{merger}} \cdot V \cdot \langle t(f_{\text{min}}) \rangle, \quad (43)$$

where V is the comoving volume of the observable Universe (because we assume the detector to be noiseless in the observing frequency band), R_{merger} is the present-day merger rate (inferred from the data), and $\langle t(f_{\text{min}}) \rangle$ is the average amount of time spent by a binary in the detector's observing band (which, in addition to f_{min} and f_{max} , sensitively depends on the physics of GWs).

For $R_{\text{merger}} \sim 2000 \text{ Gpc}^{-3} \text{ yr}^{-1}$, $V \sim 10 \text{ Gpc}^3$, and $\langle t(f_{\text{min}}) \rangle \sim 200 \text{ s}$, we obtain $N_0 \sim 10$, in agreement with the value of $N_0 = 15$ calculated by Ref. [17] in the LIGO frequency band [55]. We shall adopt $N_0 = 15$ as our fiducial value.

D. Distribution of observed strain fluctuations

All that remains is to evaluate $G(s)$, i.e., to compute

$$\begin{aligned} G(s) &= 3 \int_{r_{\text{min}}/r_0}^{R/r_0} d\tau \tau^2 \int_{S^2} d\Omega \int dM_c \int dT \int dt_* \\ &\times \int dM_{\text{gal}} \varphi(t_*, M_c, T, M_{\text{gal}}) (\tilde{p}(s; \xi; \tau, \theta, \phi) - 1) \\ &\times \Theta[(f_{\text{gw}}^{\text{obs}} - f_{\text{min}})(f_{\text{max}} - f_{\text{gw}}^{\text{obs}})], \end{aligned} \quad (44)$$

where Θ is Heaviside's function. We perform this integral numerically assuming $R = 14 \text{ Gpc}$ and $r_0 = 10 \text{ Gpc}$ (see Appendix B for details). For our fiducial cosmology, 10 and 14 Gpc correspond to the comoving radial distance to redshift $z \simeq 12$ and to the last scattering surface, respectively. The result is plotted in Fig. 2.

The distribution $P(h)$ computed from an inverse Fourier transform of $\exp(N_0 G)$ is shown in Fig. 3 assuming $N_0 = 15$. We find a strain normalization of $h_c = 2.06 \times 10^{-25}$. Here again, there is an excellent agreement between the exact numerical result (solid curve) and the asymptotic expression (dashed line) in the large- h limit. We have also overlaid a Gaussian distribution (dotted-dashed curve) with variance $\langle h^2 \rangle$ equal to that of the full distribution. This emphasizes that a Gaussian is a bad approximation over the range of strain values considered here owing to the low number of sources. As N_0 increases, the transition to the h^{-4} power-law tail moves to larger values of h , so that the Gaussian approximation improves (at fixed value of h). In Fig. 4, we show the effect of varying N_0 on the probability $P(h)$ of measuring a squared strain h .

To conclude this section, recall that the variance of observed strain fluctuations is related to the density parameter $\Omega_{\text{gw}}(f)$ through

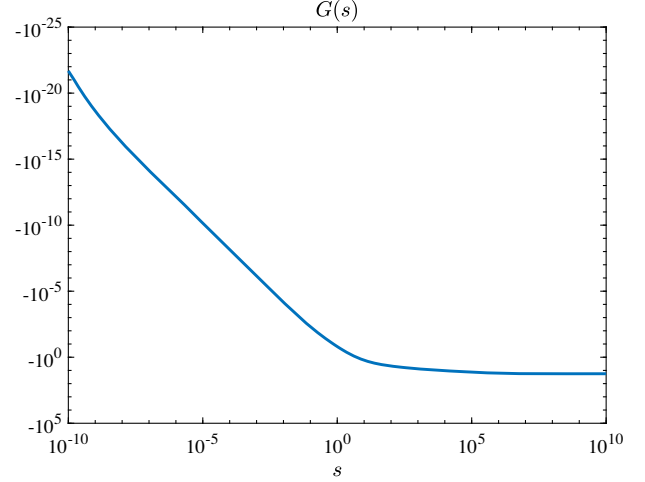


FIG. 2. The value of $G(s)$ [Eqs. (35) and (44)] for the parameters specified in Table I.

$$\langle h^2(t) \rangle = \frac{8G}{\pi c^2} \rho_{\text{crit}} \int_0^\infty d \ln f f^{-2} \Omega_{\text{gw}}(f). \quad (45)$$

The unequal time correlator $\langle h(t)h(t') \rangle$ encodes additional information on, e.g., whether the SGWB produced by binary mergers of compact stars falls in the ‘‘continuous,’’ ‘‘shot noise,’’ or ‘‘popcorn’’ regime [16]. This depends on the ratio between the duration of events and time interval between successive events. At the distribution level, quantifying the correlation structure of time series data would amount to calculating the two-point PDF $P(h(t), h(t'))$ (and higher order statistics). This ought to be relatively straightforward, albeit beyond the scope of this paper.

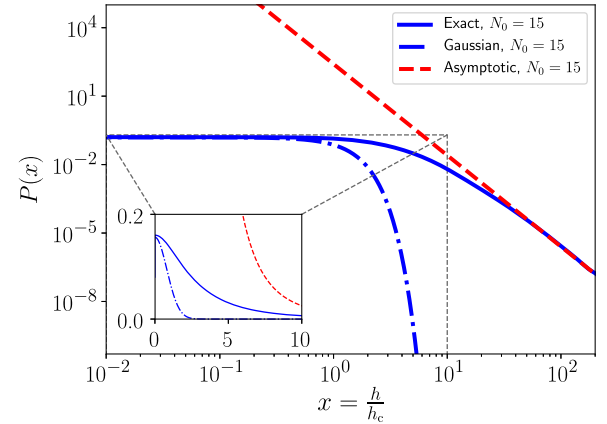


FIG. 3. The probability density function of the cumulative SGWB strain from binaries for a mean number of sources $N_0 = 15$ (solid blue) within one attenuation volume (see text for details). There is excellent agreement with the large- h asymptotic prediction in Eq. (36) (red dashed). The Gaussian approximation of the probability density function given by the second order expansion of $G(s)$ is also shown for comparison (blue dotted dashed).

V. THE CONFUSION BACKGROUND

So far, we have computed the distribution $P(h)$ of time-domain strain fluctuations measured in the frequency domain $f_{\min} \leq f_{\text{gw}}^{\text{obs}} \leq f_{\max}$ by an idealized, noiseless detector without any signal postprocessing (i.e., source identification, etc). In this section, we will model the outcome of a realistic analysis in which bright sources significantly above the detector sensitivity are subtracted from the time-domain series. The resulting distribution, which we denote $P_r(h)$, will characterize the so-called ‘‘confusion background’’ of unresolved sources [57,58].

A. Bright-source subtraction

When a binary system has a large apparent brightness— as is the case of nearby sources—its signal can rise well above the residual noise rms variance $\sqrt{\langle n^2 \rangle}$. As a result, it can be resolved as an individual event and subtracted from the SGWB time series if it does not significantly interfere with other signals. In this case, it ceases to contribute to the SGWB as defined above. In practice, the identification, modeling, and subsequent removal of bright sources are quite complicated. Therefore, we shall consider here the following simplified implementation.

A black-hole binary produces waves with frequency rising in time, until its components coalesce. The detector observes the merging, while the frequency $f_{\text{gw}}^{\text{obs}}$ of the gravitational waves lies between f_{\min} and f_{\max} . The signal-to-noise ratio of the detector is [19]

$$\begin{aligned} \left(\frac{S}{N}\right)^2 &= 4 \int_{f_{\min}}^{f_{\max}} df \frac{|\tilde{h}(f)|^2}{S_n(f)} \\ &= \frac{1}{3\pi^{4/3}} \frac{c^2}{d_L(z)^2} \left(\frac{GM_c(1+z)}{c^3}\right)^{5/3} \\ &\quad \times \int_{f_{\min}}^{\min\{2f_{\text{ISCO}}, f_{\max}\}} df \frac{f^{-1/3}}{f^2 S_n(f)}, \end{aligned} \quad (46)$$

where $f_{\text{ISCO}} = 2200 \text{ Hz} \times \frac{M_{\odot}}{m_1 + m_2}$, and $S_n(f)$ is the detector’s noise spectral density [59]. For convenience, let us introduce

$$I = \int_{f_{\min}}^{440 \text{ Hz}} df \frac{f^{-1/3}}{f^2 S_n(f)}. \quad (47)$$

Because of the $f^{-7/3}$ power in the integrand, changing the upper limit of I from the maximum possible value of $2f_{\text{ISCO}}$ in the model used here, $\sim 440 \text{ Hz}$, to its minimum, has a small effect on its value. Thus, requiring the source SNR to be more than a certain value $n\sigma$ amounts to requiring that

$$\frac{(M_c(z+1))^{5/3}}{d_L^2} \geq \frac{3\pi^{4/3} c^3 k^2}{G^{5/3} I} \approx 3.3n^2 \times 10^{-7} M_{\odot}^{5/3} \text{ Mpc}^{-2}. \quad (48)$$

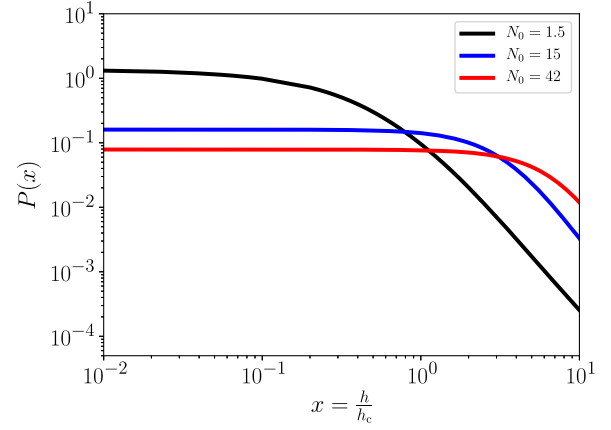


FIG. 4. The probability density function $P(h)$ of the strain h for different values of N_0 . The black line corresponds to our fiducial choice of $N_0 = 15$, while red and blue are $N_0 = 1.5$ and $N_0 = 42$, respectively. Note that N_0 always refers to the attenuation volume, which we give in Appendix B.

If this inequality is satisfied, the source is deemed *bright*, and its signal is removed from the data provided that its time to coalescence t is smaller than 5 y (so that it merges during a 5 y observational run). In practice, the bound on the SNR may be formally expressed as a bound on a function of the source parameters. The latter is then inserted as a Heaviside function into the integrand of Eq. (44), thereby ensuring that the condition (48) is not satisfied by the sources making up the confusion background.

B. Distribution

The condition Eq. (48) introduces another strain scale h_{cutoff} in addition to h_c . Above this scale, typically much larger than h_c , one expects an exponential falloff of $P_r(h)$ (cf. Appendix C). For an experiment with LIGO characteristics, we obtain indeed $h_{\text{cutoff}} \sim 10^4 h_c \gg h_c$. In this case, we expect three different regimes as shown in Fig. 5: $h \ll N_0 h_c$, where $P_r(h)$ is roughly constant, $h_{\text{cutoff}} \gg h \gg N_0 h_c$, where $P_r(h)$ behaves like a power law, and $h > h_{\text{cutoff}}$, above which $P_r(h)$ decreases exponentially. This emphasizes that the distribution $P_r(h)$ of the confusion background $P(h)$ is very close to $P(h)$ when $h_{\text{cutoff}} \gg h_c$. Overall, any cutoff h_{cutoff} will regularize the moments of $P(h)$ (which are all infinite starting from the third), and give rise to an exponential decline. The variance $\langle h^2 \rangle$ remains weakly sensitive to the value of h_{cutoff} (i.e., it varies at most by $\sim 10\%$) so long as $h_{\text{cutoff}} \gg h_c$.

To illustrate the impact of bright-source subtraction when h_{cutoff} is comparable to h_c , we consider a hypothetical detector with sensitivity comparable to h_c (and, thus, much better than LIGO). Upon removing sources which are k times brighter than h_c , i.e., $\text{SNR} > kh_c$, with $k = 1$ or 10, we obtain the distribution $P_r(h)$ of the confusion background displayed in Fig. 6 (the evaluation of the Fourier

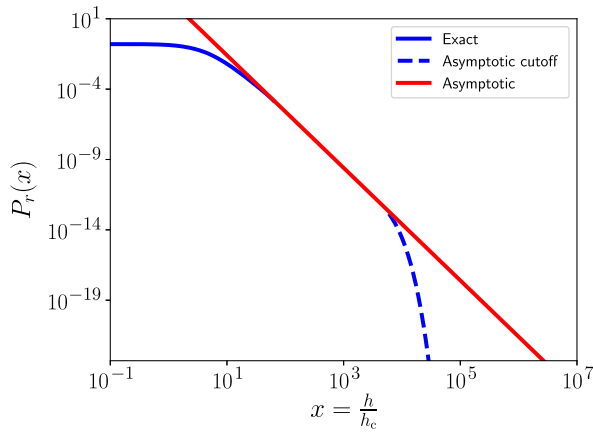


FIG. 5. Full shape of $P_r(x)$ with realistic h_{cutoff} : power law and exponential for $x(h) \rightarrow \infty$. Details of the computation can be found in Appendix C.

transform becomes computationally challenging for large values of x). One can spot a cutoff around $x \sim k$ at which the slope of the distribution changes. This is particularly clear for the case $k = 10$ because the intermediate h^{-4} power-law regime is still visible. For the case $k = 1$, this intermediate regime is not apparent because $P_r(h)$ quickly transitions to the steeper falloff, rendering the full shape of $P_r(h)$ Gaussian-like. This agrees with the numerical findings of [57], who found that $P_r(h)$ becomes closer to a Gaussian when bright sources are removed. Clearly however, $P_r(h)$ depends on the details of the source subtraction procedure. Note also that, by definition, a given detector will be sensitive to values of h not significantly below the cutoff h_{cutoff} . However, hypothetically speaking, one could decide to deem a source as bright only if it is detected at 5σ , but settle for finding the SGWB at 3σ , in which case one could still measure the flat and/or power-law behavior of $P_r(h)$.

To conclude this discussion, we emphasize that cross-correlations between different detectors can help reduce the experimental noise [19]. In this case, however, it is the probability density $P_r(h^2)$ which is the quantity being measured [60]. This can help probe the distribution $P_r(h)$ of the confusion background for values of h much smaller than the sensitivity of a single detector.

VI. DISCUSSION AND GENERAL PROPERTIES

The salient features of $P(h)$ that we outlined in this paper remain valid in a more general setting—when $g = A(t) \exp[i(\omega(t)t + \varphi)]$ —provided that A and ω are slowly varying functions of t . As in Secs. III and IV, averaging over the phase φ gives a Bessel function:

$$\tilde{p} = J_0 \left(qA \frac{de^{-r/r_0}}{d_L(r)} \right). \quad (49)$$

For $h_{\text{cutoff}} = \infty$, the dominant contribution arises again from the $|s|^3$ term in $G(s)$ and generates an h^{-4} power law.

The -4 exponent actually reflects the specific r dependence of the GW luminosity distance in general relativity (GR). Obviously, this exponent may change if the flux does not satisfy the familiar $1/r^2$ law (as is the case in some extensions of GR [63,64]). For $h_{\text{cutoff}} < \infty$, the methods of Sec. V and Appendix C apply, *mutatis mutandis*.

A. Isotropy

One critical assumption made throughout this paper is that of the isotropy of the source spatial distribution (the mean number density is allowed to vary along the radial direction), so that our $\langle h^2 \rangle$ actually corresponds to the shot noise term discussed in [25], rather than the angular power spectra C_ℓ 's calculated in [8,10] (see also [65–67] for similar calculations in the context of cosmological backgrounds).

While isotropy is a reasonable assumption for the main extra-Galactic background sources of LIGO and Virgo [25], there is a significant contribution from Galactic white-dwarf binaries to the background that should be observed by LISA [19]. The formalism considered here can be extended to include generic clustering (along the lines of, e.g., [43]) and projection effects induced by inhomogeneities (peculiar velocities, gravitational redshift, etc.) into the calculation of $G(s)$ (see, for instance, [26]).

In the specific case of contributions from our Galaxy, Eq. (5) will have to be amended to include a number density which depends on the sky direction in accordance with the Galactic density profile (approximately a disk; see [13]). The SGWB will then be the superposition of an extra-Galactic part h_{ext} (studied in this paper) and a Galactic part h_{Gal} . Since both are independent, the characteristic function of $h = h_{\text{extra}} + h_{\text{Gal}}$ is the product of their characteristic functions. One can still denote the characteristic function of h_{Gal} by $\exp(N_{\text{Gal}} G_{\text{Gal}}(s))$, with N_{Gal} measuring the mean number of Galactic sources across the sky (even if this distribution is neither isotropic nor Poissonian). Therefore, the function $G(s)$ should also separate into two parts: $N_{\text{tot}} G(s) = N_0 G_{\text{ext}}(s) + N_{\text{Gal}} G_{\text{Gal}}(s)$. If $N_{\text{Gal}} \gg 1$, Laplace's method (this time treating N_{Gal} as the big parameter, not h) ensures that, for all values of h but the very largest, where the approximations of Appendix C apply, the dominant contribution to G_{Gal} is its $O(s^2)$ term in its Maclaurin series. Thus, one expects that for large N_{Gal} , h is well approximated by a sum of h_{ext} , whose distribution we have described here, and a Gaussian random variable, provided that h is not too large.

B. Frequency domain

All of the above is true for the entire gravitational-wave amplitude integrated, so to speak, over all frequencies. However, we can resolve the probability distribution for each frequency. A signal processing in the frequency

domain is also desirable [especially when $\langle n^2 \rangle$ is of order λh_c^2 or larger] because the detector noise usually has a very specific frequency dependence.

Let $\tilde{h}(f) = \int dt e^{2\pi i f t} h(t)$ be the Fourier amplitudes. For N sources, we have

$$\tilde{h}(f) = \sum_{k=1}^N \tilde{h}_k(f). \quad (50)$$

The individual Fourier modes [68] $\tilde{h}_k(f)$ also are independent by the assumptions stated in Sec. II, with the caveat that they are complex random variables. Treating $\tilde{h}(f)$ as a complex variable, the single source characteristic function $\psi(q)$ should be defined as

$$\psi(q) = \langle \exp[i\Re(q\tilde{g}(f))] \rangle, \quad (51)$$

where q is now complex, $\tilde{g}(f)$ denotes the Fourier transform of $g(t)$, and $\Re(z)$ is the real part. Alternatively, since $\tilde{h}(f)$ must satisfy the reality condition [because $h(t)$ is real], we can restrict ourselves to the statistics of the real part $\Re(\tilde{h}(f))$ of the Fourier amplitudes without losing any information. In this case, the methods of Sec. II apply exactly provided that $g(t)$ is replaced by

$$\tilde{g}(f) = \frac{1}{T} \int_0^T dt \cos(2\pi f t) g(t). \quad (52)$$

(T expresses the finite duration T of the experiment; frequencies $f < 1/T$ are poorly sampled by the data.)

As an illustration, consider the simplified model of Sec. III. For a general $T < \infty$, the Fourier transform g yields

$$\begin{aligned} \tilde{g}(f) = \frac{A}{2T} \frac{r_0 e^{-r/r_0}}{r} & \left[\left(\frac{\cos \omega_+}{\omega_+} + \frac{\cos \omega_-}{\omega_-} \right) \sin \varphi \right. \\ & \left. + \left(\frac{\sin \omega_+}{\omega_+} + \frac{\sin \omega_-}{\omega_-} \right) \cos \varphi \right], \end{aligned} \quad (53)$$

where $\omega_{\pm} = \omega \pm 2\pi f$. Integrating over the random phase φ , $\tilde{p}(q, f; \xi; \mathbf{r})$ becomes, after some manipulations,

$$\tilde{p}(q, f; \xi; \mathbf{r}) = J_0 \left(q \frac{A r_0 e^{-r/r_0}}{r} \alpha(\omega, f, T) \right), \quad (54)$$

provided that $\alpha(\omega, f, T)$ is defined as

$$\alpha(\omega, f, T) = \frac{\sqrt{\omega^2 + 4\pi^2 f^2 + (\omega^2 - 4\pi^2 f^2) \cos(4\pi f T)}}{\sqrt{2} |\omega^2 - 4\pi^2 f^2| T}. \quad (55)$$

Equation (54) is identical to Eq. (11) except for a frequency-dependent factor of $\alpha(\omega, f, T)$. Our previous

arguments remain valid, so the PDF $P(\Re(\tilde{h}(f)) = h)$ will be dominated by the nearest neighbor at large h . Its asymptotic form thus exhibits the power-law behavior $\sim h^{-4}$, although its overall amplitude is now modulated in accordance with $\alpha(\omega, f, T)$.

C. Central limit theorem

The reader might wonder whether central limit theorems (CLTs) hold here, especially the *classical* CLT that guarantees the *pointwise* convergence of a sum of identically distributed variables with *finite* variance. In our case, however, the variance of a single source is infinite.

To see this, consider the simplified model of Sec. III, where the emitted amplitude b is constant, and let us evaluate $\mathbb{E}\langle h^2 \rangle$ ($\mathbb{E}h = 0$), where $\langle \cdot \rangle$ is a phase average. The tail formula for the expectation gives

$$\begin{aligned} \mathbb{E}\langle h^2 \rangle &= \int_0^\infty dt P(\langle h^2 \rangle > t) = \int_0^\infty dt P\left(\frac{d^2 b^2}{2r^2} > t\right) \\ &= \int_0^\infty dt P\left(r < \frac{db}{\sqrt{2t}}\right) = \int_0^\infty dt \int_0^{\frac{db}{\sqrt{2t}}} dr \frac{3r^2}{R^3} \\ &= \int_0^\infty dt \frac{d^3 b^3}{R^3 (2t)^{3/2}}. \end{aligned} \quad (56)$$

This diverges, of course, at $t = 0$. Therefore, the SGWB does not satisfy the *classical* CLT. However, we argue that $P(h)$ still converges (nonuniformly) toward a Gaussian, despite the infinite variance of the single source distribution, in agreement with the existence of CLTs for random variables with infinite variance (e.g., [69–71]).

To demonstrate this point, let us scrutinize Eq. (A7). If the series in the argument of the exponential would involve equal powers of N_0 , $\langle b^2 \rangle$ and s , i.e., $-N_0 \langle b^2 \rangle s^2 + (N_0 \langle b^2 \rangle)^{3/2} |s|^3 + \dots$ [we omit $\mathcal{O}(1)$ series coefficients for clarity], then any change in the value of N_0 or $\langle b^2 \rangle$ could be absorbed into a redefinition of s . Therefore, the shape of $P(h)$ for large $N_0 \gg 1$ would be approximately the same as that for $N_0 \sim 1$, that is, highly non-Gaussian. The linearity of the series expansion in N_0 implies that $P(h)$ eventually converges to a Gaussian in the limit $N_0 \rightarrow \infty$. However, the rate of convergence strongly depends on the value of h and the statistical properties of the source distribution. In the random walk analogy, the slow convergence manifests itself as anomalous diffusion on a longer timescale than expected from the classical CLT.

Comparing the s^2 and $|s|^3$ terms in Eq. (A7) shows that the closer $P(h)$ is to a Gaussian, the smaller is the ratio

$$\frac{\langle b^3 \rangle}{\sqrt{N_0} \langle b^2 \rangle^{3/2}} \equiv \frac{s_3(b)}{\sqrt{N_0}}, \quad (57)$$

where $s_3(b)$ is the skewness of the b distribution only if $\langle b \rangle = 0$. In other words, source distributions with a large s_3 lead to a slower convergence rate, in agreement with the

Berry-Esseen theorem [69,70]. For illustration, let us evaluate s_3 for the initial, power-law mass distribution $\phi(m)$. Assuming $m_{\min} < m < m_{\max}$ and $2 < \alpha < 3$, we find

$$s_3(m) \approx m_{\max}^{(\alpha-1)/2} \quad \text{for } m_{\max} \gg m_{\min}; \quad (58)$$

i.e., $s_3(m)$ grows nearly linearly with m_{\max} for $\alpha = 2.7$. As a result, taking a value of m_{\max} 10 times larger implies that N_0 should be 100 times bigger to achieve a similar level of convergence.

Overall, the nonuniformity of the convergence rate is reflected by Eq. (36), which we have shown to be valid for any value of N_0 . Indeed, we find that for $h \lesssim N_0 h_c$, the Gaussian approximation holds, while for larger values of h , the power-law asymptotic is recovered. When there is not bright-source removal, for any *finite* N_0 , $P(h)$ therefore assumes the shape of a power law at sufficiently large values of h . As N_0 increases, the power-law regime is pushed further and further to higher values of h .

VII. CONCLUSION

In this paper, we exemplified a formalism to calculate the PDF of a stochastic gravitational-wave background produced by Poisson-clustered compact binaries. A similar approach was considered in the context of the high-redshift ionizing UV background but, to the extent of our knowledge, it is the first time that it is applied to stochastic GW backgrounds, where interference is taken into account in a consistent manner.

In contrast with earlier works on the topic, we provided expressions for the full distribution function $P(h)$, not only its variance (or power spectrum). We also demonstrated how our approach can be extended to calculate the distribution $P_r(h)$ of the confusion background obtained after the subtraction of bright sources. Our formalism thus has the advantage of being able to model large deviations, as well as typical fluctuations. We evaluated the PDF of the stochastic gravitational strain numerically and derived accurate asymptotic expressions for the large strain tail. The latter turns out to be dominated by the nearest active source. The resulting h^{-4} scaling is, in fact, a universal phenomenon, insofar as it is independent of the particular properties and characteristics of the sources—so long as they are deterministic, point sources—and arises solely from the nature of gravity, the four-dimensional nature of space-time, and the inverse-square law decay of the wave amplitude. Furthermore, we demonstrated that both $P(h)$ and $P_r(h)$ are generically non-Gaussian as they depend on the *moments* (rather than the cumulants) of the source parameter distribution owing to the projection along the line of sight to the observer. As a corollary, one must be careful with the application of central limit theorems because the source variance is infinite. We expect these conclusions to hold also for the distributions of Fourier amplitudes.

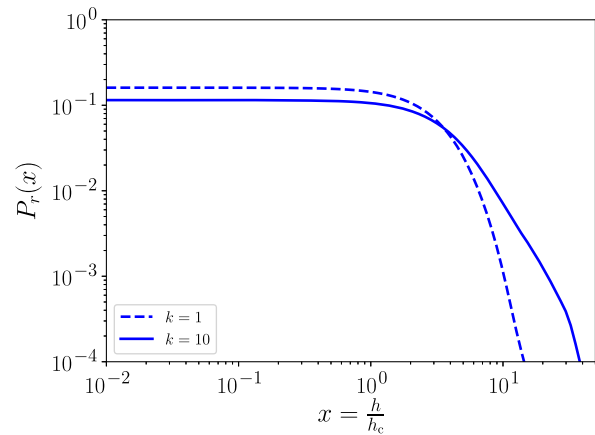


FIG. 6. The distribution of the confusion background $P_r(h)$ for two different cutoff SNRs $> kh_c$, with $k = 1$ and 10, for the bright-source subtraction.

Section IV described a calculation of the stochastic background produced by binary mergers of compact stars, assuming Newtonian orbits throughout the binary’s evolution in conjunction with other simplifying assumptions on the rates, mass functions, and spatial clustering of the sources. These, however, do not reflect a limitation of the formalism, but rather the authors’ wish to simplify the numerics. In fact, $G(s)$ may be calculated much more precisely in an analogous way, using rates and mass function derived from simulations directly and taking into account inhomogeneities in the spatial distribution of the sources. Finally, our results can be extended in the Fourier domain and, therefore, could be useful for the frequency reconstruction of stochastic GW backgrounds (see, e.g., [72]). We defer a thorough study of all these issues to future work.

ACKNOWLEDGMENTS

We would like to thank Adi Nusser for helpful discussions and the anonymous referee for a helpful criticism of our paper. Y. B. G. acknowledges support from the Technion Jacobs scholarship. V. D. and R. R. acknowledge funding from the Israel Science Foundation (Grants No. 1395/16 and No. 255/18). V. D. would also like to thank the Munich Institute for Astro- and Particle Physics of the DFG cluster of excellence “Origin and Structure of the Universe,” and the Institut d’Astrophysique de Paris for their hospitality during the completion of this work. H. B. P. acknowledges support from the Merle Kingsley fund at Caltech.

APPENDIX A: ASYMPTOTIC EXPANSION FOR ANY N_0

We demonstrate that the power-law asymptotic behavior in Eq. (21) is true even when N_0 is not small. As we are

interested in the large- $|h|$ limit, the dominant contribution comes from small values of s in $G(s)$, so it is important to expand Eq. (16) around $s = 0$. We do this using a Mellin transform method (see Appendix A.2 of [39]): let

$$\bar{G}(\mu) = \int_0^\infty ds G(s) s^{\mu-1} \quad (\text{A1})$$

be the Mellin transform of G . The integral converges for $0 > \Re(\mu) > -2$, and may be performed analytically [73] to give

$$\bar{G}(\mu) = 3 \int db \phi(b) \int_0^{R/r_0} d\tau \tau^2 \left(\frac{be^{-\tau}}{\tau} \right)^{-\mu} \frac{\operatorname{cosec}(\frac{\mu\pi}{2})}{2^{1-\mu}\Gamma^2(1-\frac{\mu}{2})}. \quad (\text{A2})$$

Integrating further over τ , we arrive at

$$\bar{G}(\mu) = 3 \int db \phi(b) b^{-\mu} (-\mu)^{-(3+\mu)} \frac{\operatorname{cosec}(\frac{\mu\pi}{2})\Gamma(3+\mu)}{2^{1-\mu}\Gamma^2(1-\frac{\mu}{2})}. \quad (\text{A3})$$

This formula can be analytically continued to any complex μ , except for $\mu = 0, \pm 2, \pm 4, \pm 6, \dots$ (the cosecant's poles) and $\mu = -3, -4, -5, \dots$ (the numerator gamma function's poles). The Mellin transform inversion theorem implies that, for $c = \Re(\mu)$ between 0 and -2 , we can write

$$G(s) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} d\mu s^{-\mu} \bar{G}(\mu). \quad (\text{A4})$$

$$P(h) \sim Q(h) \equiv \frac{1}{\pi} \int_0^\infty ds \cos(hs) \exp\left(-\frac{3N_0\langle b^2 \rangle}{8\pi} s^2 + \frac{N_0\langle b^3 \rangle}{3\pi} s^3 + \frac{24\gamma_E - 27 + 12 \ln 2}{64\pi} N_0\langle b^4 \rangle s^4\right). \quad (\text{A7})$$

Taking the Mellin transform of this equation gives

$$\bar{Q}(\mu) = \int_0^\infty dh \int_0^\infty ds h^{\mu-1} \cos(hs) \exp\left(-\frac{3N_0\langle b^2 \rangle}{8\pi} s^2 + \frac{N_0\langle b^3 \rangle}{3\pi} s^3 + \frac{24\gamma_E - 27 + 12 \ln 2}{64\pi} N_0\langle b^4 \rangle s^4\right). \quad (\text{A8})$$

The h integral is just

$$\int_0^\infty dh h^{\mu-1} \cos(hs) = s^{-\mu} \Gamma(\mu) \cos\left(\frac{\pi\mu}{2}\right) \quad (\text{A9})$$

when $0 < \Re\mu < 1$. The s integral may be written as

$$\int_0^\infty ds s^{-\mu} \exp\left(-\frac{3N_0\langle b^2 \rangle}{8\pi} s^2\right) \sum_{n=0}^\infty \frac{1}{n!} \left(\frac{N_0\langle b^3 \rangle}{3\pi} s^3\right)^n \sum_{k=0}^\infty \frac{1}{k!} \left(\frac{24\gamma_E - 27 + 12 \ln 2}{64\pi} N_0\langle b^4 \rangle s^4\right)^k. \quad (\text{A10})$$

Each integral is of the form

$$\int_0^\infty ds s^{-\nu} \exp\left(-\frac{3N_0\langle b^2 \rangle}{8\pi} s^2\right) = \frac{1}{2} \left(\frac{3N_0\langle b^2 \rangle}{8\pi}\right)^{(\nu-1)/2} \Gamma\left(\frac{1-\nu}{2}\right), \quad (\text{A11})$$

Now, using the analytical continuation of Eq. (A3) to shift the integration contour all the way to the left, all the negative integer poles of Eq. (A3), starting at -2 , contribute to the small- s expansion of $G(s)$. The first three terms come from $\mu = -2, -3, -4$, viz.,

$$G(s) \sim -\frac{3\langle b^2 \rangle}{8\pi} s^2 + \frac{\langle b^3 \rangle}{3\pi} s^3 + \frac{24\gamma_E - 27 + 12 \ln 2}{64\pi} \langle b^4 \rangle s^4 + \dots \quad (\text{A5})$$

where γ_E is Euler's constant.

The Mellin transform is only affected by positive values of s , so the above equation is correct only for small, positive s . For negative values, recall that $G(s)$ is an even function of s , and extend Eq. (A5) to negative values evenly by adding an absolute value to all odd powers of s :

$$G(s) \sim -\frac{3\langle b^2 \rangle}{8\pi} s^2 + \frac{\langle b^3 \rangle}{3\pi} |s|^3 + \frac{24\gamma_E - 27 + 12 \ln 2}{64\pi} \langle b^4 \rangle s^4 + \dots \quad (\text{A6})$$

This shows that the cumulants of $P(h)$ beyond the variance do not vanish (they depend on the *moments* of the source parameter distribution), and, consequently, $P(h)$ does not *uniformly* converge to a Gaussian (an h^{-4} tail is always present for any finite $N_0 \gg 1$).

We evaluate the asymptotic expansion for $P(h)$ at large (positive) h using a Mellin transform argument, again. Equation (A6) implies [bear in mind that $G(s)$ is even] that

where $\nu = \mu - 3n - 4k$. The Mellin transform inversion theorem implies (again) that for $0 < c = \Re(\mu) < 1$,

$$Q(h) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} d\mu h^{-\mu} \bar{Q}(\mu). \quad (\text{A12})$$

The expression for \bar{Q} has poles at $\mu = 0, -2, -4, \dots$ due to $\Gamma(\mu) \cos(\mu\pi/2)$, as well as poles at $\nu = 1, 3, 5, \dots$. The former set of poles does not affect the large- h expansion, as one needs to consider poles with $\Re(\mu) \geq 1$. The cosine cancels any ν pole unless μ is even; whence n has to be odd for there to be a nonzero residue. As μ increases with n , the low values of n contribute lower powers of $1/h$ to the expansion of Q at $h \gg 1$. The lowest such μ for which there is a nonzero residue therefore dominates. It comes from $n = 1, k = 0$, and is given by $\mu = 4$ (the residues from $\mu = 1, 2$, and 3 are zero), which corresponds to $\nu = 1$. This yields

$$Q(h) = -\left(-\frac{12}{2}\right) \frac{N_0 \langle b^3 \rangle}{3\pi} h^{-4} \quad (\text{A13})$$

on using

$$\text{Res}\left(\cos\left(\frac{\pi z}{2}\right) \Gamma(z) \Gamma\left(\frac{1}{2} - \frac{1}{2}(-4k - 3n + z)\right), \{z = 4\}\right) = -12. \quad (\text{A14})$$

An additional minus sign comes from the fact that, when the line integral in the Mellin transform inversion formula [Eq. (A12)] is shifted all the way to the right, it translates into clockwise contour integrals around the poles. Together with Eq. (A7) this yields

$$P(h) \sim \frac{2N_0 \langle b(\tau = 0)^3 \rangle}{\pi h^4}, \quad (\text{A15})$$

which is precisely Eq. (36) (the $\tau = 0$ dependence is added as in Sec. IV).

It is clear that powers of s higher than s^3 do not affect this scaling as they contribute only larger values of μ and, therefore, are subdominant.

APPENDIX B: NUMERICAL INTEGRATION AND PARAMETERS

As explained in Sec. IV, we evaluate the eight-dimensional integral Eq. (44), as well as the various normalization constants numerically, before performing the inverse Fourier transform to obtain $P(h)$. For the integration, we use Monte Carlo integration. In particular, we employ importance sampling, i.e., the VEGAS algorithm. Since the integrand is well behaved, convergence can be reached very quickly, outperforming nested integration in this particular case. The integration in the T variable is carried out logarithmically. For the final plots we used 10^8 function

TABLE I. Summary of the integration boundaries, the corresponding units, and additional parameters appearing in the numerically evaluated integral, Eq. (44).

Parameter	Value	Units	Description
R	14	Gpc	Comoving size of the Universe
r_{\min}	0.01	Mpc	UV cutoff on comoving distance from detector
m_{\max}	45	M_{\odot}	Maximum mass of binary member
m_{\min}	8	M_{\odot}	Minimum mass of binary member
a_{\max}	4000	AU	Maximum initial semimajor axis of binary
a_{\min}	0.014	AU	Minimum initial semimajor axis of binary
r_0	10	Gpc	Attenuation length

evaluations to sample the integral. As a sanity check, we compared it to nested integration and obtained excellent agreement. We tabulate the values for $G(s)$ and interpolate logarithmically to perform the final, one-dimensional inverse Fourier transformation. At high values of h , the inverse Fourier transform can pick up modes from the numerical noise of the Monte Carlo integration, but by increasing the number of function evaluations the value of h where this starts to happen can be pushed to ever higher values. For all integrations, we used the routines provided by the GSL [74].

All the parameters entering in the evaluation of the integral in Eq. (44) are summarized in Table I. A short distance cutoff $r_{\min} = 0.01$ Mpc was introduced only to avoid possible numerical issues. It does not affect our conclusions in any way.

APPENDIX C: SHAPE OF EXPONENTIAL DECLINE IN SEC. V

By introducing an effective cutoff in $G(s)$ due to bright sources, as described in Sec. V, $G(s)$ is made into an analytic function, and the $|s|^3$ irregular term is regularized. This implies that, by the Paley-Wiener theorem, its Fourier transform $P_r(h)$ must decline faster than any power law at large values of h ; i.e., the power-law approximation ceases to hold for sufficiently large h . The purpose of this appendix is to find the asymptotic limit of $P_r(h)$ in the case where $G(s)$ is analytic.

The integral we need to approximate is given by Eq. (9), where $\psi(z)$ is an even function of z , a task we perform using the steepest descents method [75]. This method relies on finding the saddle points of the exponent, and then deforming the integration contour in the complex plane so that it coincides with the steepest descents contour (on which the imaginary part of the exponent is constant) that passes through the relevant saddle points. The remaining parts of the integral would give exponentially small errors. In our case, the exponent is $\zeta(z) = -ihz + G(z)$, whence the saddle points are those where $G'(z) = ih$. The relevant

steepest descents contour coincides with the imaginary axis, for G is even. Therefore, we set $z = -iy$ so that $\zeta = -hy + G(-iy)$. Moreover, since $G(0) = 0$ and G is an entire function, then z has to be large for G' to be as large as ih . This justifies the replacement of $G(-iy)$ by its asymptotic expansion as $|y| \rightarrow \infty$. In the particular case at hand, it is of the form

$$-hy + c \exp(ay)y^{-b}, \quad (\text{C1})$$

where $a, b, c > 0$ are independent of y .

The reason for Eq. (C1) is as follows: \tilde{p} , as given by Eq. (35), is $J_0(\text{const} \times s)$ (where the constant is not a function of s), so $G(iy)$ is the expectation value of $J_0(\text{const} \times iy) = I_0(\text{const} \times y)$. When y is large (as is necessary for a saddle point), $I_0(Ay) \sim \exp(Ay)(2\pi Ay)^{-1/2}$ [73]. One needs to calculate the expectation value over A , which, in the case of Sec. IV, is a multidimensional integral. As y is large, this integral is well approximated by Laplace's method: One may change variables so that A is one of the integration variables; the value of A is maximized at the boundary, so Laplace's method implies that this integration contributes one negative power of y to the result. Therefore, $b = \frac{3}{2}$. Furthermore, a is the value of

A at the maximum. c is the value of the expected value evaluated at a and integrated over all other parameters.

The critical points of $\zeta(y) = -hy + ce^{ay}/y^{3/2}$ are the solutions of

$$h \sim \frac{ace^{ay}}{y^{1.5}}, \quad (\text{C2})$$

bearing in mind that y is large and retaining only the dominant terms. This equation is solved by the Lambert W function, which is approximated by $ay \sim \ln h$ [73]. Applying Laplace's method yet another time to the y integration, the steepest descents method yields

$$P_r(h) \sim \frac{1}{\sqrt{2\pi ch}} \frac{(\ln h)^{3/4}}{a^{7/4}} \exp \left[-\frac{h \ln h}{a} + \frac{a^{3/2} ch}{(\ln h)^{3/2}} \right]. \quad (\text{C3})$$

In Fig. 5 we show the three regimes of $P_r(x)$ ($x = h/h_c$), that is, the exact calculation, the intermediate power-law regime, and the exponential cutoff at very large x . The latter is described by Eq. (C3). The exponential curve and the power law x^{-4} are connected smoothly at some reference value x_0 upon adjusting the free constants a and c in Eq. (C3).

-
- [1] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [2] C. Caprini and D. G. Figueroa, *Classical Quantum Gravity* **35**, 163001 (2018).
- [3] D. I. Kosenko and K. A. Postnov, *Astron. Astrophys.* **355**, 1209 (2000).
- [4] A. J. Farmer and E. S. Phinney, *Mon. Not. R. Astron. Soc.* **346**, 1197 (2003).
- [5] D. Falta and R. Fisher, *Phys. Rev. D* **84**, 124062 (2011).
- [6] E. Thrane, *Phys. Rev. D* **87**, 043009 (2013).
- [7] T. Callister, L. Sammut, S. Qiu, I. Mandel, and E. Thrane, *Phys. Rev. X* **6**, 031018 (2016).
- [8] G. Cusin, C. Pitrou, and J.-P. Uzan, *Phys. Rev. D* **96**, 103019 (2017).
- [9] R. Brito, S. Ghosh, E. Barausse, E. Berti, V. Cardoso, I. Dvorkin, A. Klein, and P. Pani, *Phys. Rev. Lett.* **119**, 131101 (2017).
- [10] A. C. Jenkins and M. Sakellariadou, *Phys. Rev. D* **98**, 063509 (2018).
- [11] E. Barausse, R. Brito, V. Cardoso, I. Dvorkin, and P. Pani, *Classical Quantum Gravity* **35**, 20LT01 (2018).
- [12] N. Christensen, *Rep. Prog. Phys.* **82**, 016903 (2019).
- [13] C. Conneely, A. H. Jaffe, and C. M. F. Mingarelli, *Mon. Not. R. Astron. Soc.* **487**, 562 (2019).
- [14] M. Maggiore, *Astrophysics and Cosmology, Gravitational Waves*, Vol. 2 (Oxford University Press, Oxford, 2018).
- [15] D. M. Coward, R. R. Burman, and D. G. Blair, *Mon. Not. R. Astron. Soc.* **329**, 411 (2002).
- [16] T. Regimbau, *Res. Astron. Astrophys.* **11**, 369 (2011).
- [17] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **120**, 091101 (2018).
- [18] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak *et al.* (Planck Collaboration), *arXiv*: 1807.06209.
- [19] M. Maggiore, *Theory and Experiments, Gravitational Waves*, Vol. 1 (Oxford University Press, Oxford, 2008).
- [20] We assume that the total observational time is much larger than the periods of the waves received by the detector.
- [21] R. Flauger and S. Weinberg, *Phys. Rev. D* **97**, 123506 (2018).
- [22] R. Flauger and S. Weinberg, *Phys. Rev. D* **99**, 123030 (2019).
- [23] G. Cusin, I. Dvorkin, C. Pitrou, and J.-P. Uzan, *Phys. Rev. Lett.* **120**, 231101 (2018).
- [24] A. C. Jenkins and M. Sakellariadou, *Phys. Rev. D* **100**, 063508 (2019).
- [25] A. C. Jenkins, J. D. Romano, and M. Sakellariadou, *Phys. Rev. D* **100**, 083501 (2019).

- [26] D. Bertacca, A. Ricciardone, N. Bellomo, A. C. Jenkins, S. Matarrese, A. Raccanelli, T. Regimbau, and M. Sakellariadou, *Phys. Rev. D* **101**, 103513 (2020).
- [27] L. Zuo, *Mon. Not. R. Astron. Soc.* **258**, 36 (1992).
- [28] M. A. Fardal and J. M. Shull, *Astrophys. J.* **415**, 524 (1993).
- [29] A. Meiksin and M. White, *Mon. Not. R. Astron. Soc.* **342**, 1205 (2003).
- [30] Throughout the paper, we use $\phi(\cdot)$ to denote the probability measure on any subset of the model parameters ξ .
- [31] H. Touchette, *Phys. Rep.* **478**, 1 (2009).
- [32] R. Balian and R. Schaeffer, *Astron. Astrophys.* **220**, 1 (1989).
- [33] F. Bernardeau and R. Schaeffer, *Astron. Astrophys.* **349**, 697 (1999).
- [34] P. Valageas, *Astron. Astrophys.* **382**, 450 (2002).
- [35] P. Valageas and D. Munshi, *Mon. Not. R. Astron. Soc.* **354**, 1146 (2004).
- [36] F. Bernardeau, C. Pichon, and S. Codis, *Phys. Rev. D* **90**, 103519 (2014).
- [37] C. Uhlemann, S. Codis, O. Hahn, C. Pichon, and F. Bernardeau, *Mon. Not. R. Astron. Soc.* **469**, 2481 (2017).
- [38] A. Barthelemy, S. Codis, C. Uhlemann, F. Bernardeau, and R. Gavazzi, *Mon. Not. R. Astron. Soc.* **492**, 3420 (2020).
- [39] B. D. Hughes, *Random Walks and Random Environments* (Clarendon Press, Oxford, 1995), Vol. 1, pp. xxii, 631.
- [40] L. Hui, S. T. McWilliams, and I.-S. Yang, *Phys. Rev. D* **87**, 084009 (2013).
- [41] The function $G(s)$ itself is not smooth, but it contains an $|s|^3$ term in its series expansion near the origin, as may be visible from a Mellin transform argument; see Appendix A.
- [42] D. S. Jones, *J. Inst. Mat Appl.* **2**, 197 (1966).
- [43] V. Desjacques, A. M. Dizgah, and M. Biagetti, *Mon. Not. R. Astron. Soc.* **444**, 2793 (2014).
- [44] L. Boco, A. Lapi, S. Goswami, F. Perrotta, C. Baccigalupi, and L. Danese, *Astrophys. J.* **881**, 157 (2019).
- [45] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **882**, L24 (2019).
- [46] G. Cusin, I. Dvorkin, C. Pitrou, and J.-P. Uzan, *Phys. Rev. D* **100**, 063004 (2019).
- [47] R. Smit, R. J. Bouwens, M. Franx, G. D. Illingworth, I. Labbé, P. A. Oesch, and P. G. van Dokkum, *Astrophys. J.* **756**, 14 (2012).
- [48] E. Öpik, *Publ. Tartu Astrofizica Obs.* **25**, 1 (1924).
- [49] G. Duchêne and A. Kraus, *Annu. Rev. Astron. Astrophys.* **51**, 269 (2013).
- [50] P. Kroupa, *Science* **295**, 82 (2002).
- [51] J. Binney and S. Tremaine, *Galactic Dynamics*, 2nd ed., edited by J. Binney and S. Tremaine (Princeton University Press, Princeton, NJ, 2008).
- [52] C. L. Fryer, K. Belczynski, G. Wiktorowicz, M. Dominik, V. Kalogera, and D. E. Holz, *Astrophys. J.* **749**, 91 (2012).
- [53] X. Ma, P. F. Hopkins, C.-A. Faucher-Giguère, N. Zolman, A. L. Muratov, D. Kereš, and E. Quataert, *Mon. Not. R. Astron. Soc.* **456**, 2140 (2016).
- [54] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz, *Astrophys. J.* **688**, 709 (2008).
- [55] The rates in Ref. [17] are slightly lower than those in the more recent Refs. [45,56], but consistent with them, and the value of N_0 changes only by an order-unity constant, which does not alter any of our conclusions.
- [56] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
- [57] S. E. Timpano, L. J. Rubbo, and N. J. Cornish, *Phys. Rev. D* **73**, 122001 (2006).
- [58] T. Regimbau and S. A. Hughes, *Phys. Rev. D* **79**, 062002 (2009).
- [59] L. Barsotti, P. Fritschel, M. Evans, and S. Gras, LIGO Report No. LIGO-T1800044-v5, 2018 [<https://dcc.ligo.org/LIGO-T1800044/public>].
- [60] The overlap reduction function is used to take into account the fact that the detectors are not spatially coincident [61,62]. However, the interpretation of the cross-correlation signal may be less straightforward. We thank the anonymous referee and Amos Ori for pointing this out to us.
- [61] E. E. Flanagan, *Phys. Rev. D* **48**, 2389 (1993).
- [62] B. Allen and J. D. Romano, *Phys. Rev. D* **59**, 102001 (1999).
- [63] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, *Phys. Rep.* **513**, 1 (2012).
- [64] M. Ishak, *Living Rev. Relativity* **22**, 1 (2019).
- [65] M. Geller, A. Hook, R. Sundrum, and Y. Tsai, *Phys. Rev. Lett.* **121**, 201303 (2018).
- [66] N. Bartolo, D. Bertacca, S. Matarrese, M. Peloso, A. Ricciardone, A. Riotto, and G. Tasinato, *Phys. Rev. D* **100**, 121501 (2019).
- [67] N. Bartolo, D. Bertacca, V. De Luca, G. Franciolini, S. Matarrese, M. Peloso, A. Ricciardone, A. Riotto, and G. Tasinato, *J. Cosmol. Astropart. Phys.* **02** (2020) 028.
- [68] One could also consider a wavelet decomposition.
- [69] A. C. Berry, *Trans. Am. Math. Soc.* **49**, 122 (1941).
- [70] C.-G. Esseen, *Ark. Mat. Astron. Fys.* **28**, 1 (1942).
- [71] C. Börgers and C. Greengard, *C. R. Math.* **356**, 679 (2018).
- [72] C. Caprini, D. G. Figueroa, R. Flauger, G. Nardini, M. Peloso, M. Pieroni, A. Ricciardone, and G. Tasinato, *J. Cosmol. Astropart. Phys.* **11** (2019) 017.
- [73] NIST Digital Library of Mathematical Functions, Release 1.0.23 of June 15, 2019, edited by f. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders, <http://dlmf.nist.gov/>.
- [74] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, F. Rossi, and R. Ulerich, GNU Scientific Library Release 2.6, 2019.
- [75] F. W. J. Olver, *Asymptotics and Special Functions*, Computer Science and Applied Mathematics (Academic Press, New York, 1974).