PHYSICAL REVIEW D VOLUME 48, NUMBER 6 15 SEPTEMBER 1993

ARTICLES

Sensitivity of the Laser Interferometer Gravitational Wave Observatory to a stochastic background, and its dependence on the detector orientations

Eanna E. Flanagan
Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125
(Received 17 May 1993)

We analyze the sensitivity of a network of interferometer gravitational-wave detectors to the gravitational-wave stochastic background, and derive the dependence of this sensitivity on the orientations of the detector arms. We build on and extend the recent work of Christensen, but our conclusion for the optimal choice of orientations of a pair of detectors differs from his. For a pair of detectors (such as LIGO) that subtends an angle at the center of the Earth of $\lesssim 70°$, we find that the optimal configuration is for each detector to have its arms make an angle of 45° (modulo 90°) with the arc of the great circle that joins them. For detectors that are farther separated, each detector should instead have one arm aligned with this arc. We show that the broadband sensitivity to the stochastic background of a detector pair which are $\lesssim 3000$ km apart is essentially determined by their relative rotation. Their average rotation with respect to the arc joining them is unimportant. We also describe in detail the optimal data-analysis algorithm for searching for the stochastic background with a detector network, which is implicit in earlier work of Michelson. The LIGO pair of detectors will be separated by ~ 3000 km. The minimum detectable stochastic energy density for these detectors with their currently planned orientations is $\sim 3\%$ greater than what it would be if the orientations were optimal, and ~ 4 times what it would be if their separation were $\lesssim$ a few kilometers. (The detectors are chosen to be far apart so that their sources of noise will be uncorrelated, and in order to improve the angular resolution of the determinations of positions of burst sources.)

PACS number(s): 04.30.+x, 04.80.+z, 95.55.Ym, 98.80.Es

I. INTRODUCTION AND SUMMARY

A. Background and motivation

Construction will begin soon on the American Laser Interferometer Gravitational Wave Observatory (LIGO) [1], and on its French-Italian counterpart, VIRGO [2]. Early in the next century there will likely be in operation a worldwide network of detectors, with sites in America, Europe, and possibly Japan and Australia [3]. It is important at this stage for physicists to look ahead and identify the types of science that the community might focus on using this network when it reaches a mature stage, perhaps a decade after the first gravitational waves are detected. One of the reasons for doing so is that some properties and parameters of the network, which ultimately will constrain what it can accomplish in the future, are being finalized today. The orientation of the detector arms is one example.

One of the long term aims of this detector network will be to place upper limits on (or perhaps detect) the energy density of a stochastic background (SB) of gravitational waves. This background would be analogous to the relic 3 K electromagnetic background, except that its spectrum is not expected to be thermal. The spectrum is usually characterized by a quantity $\Omega_g(f)$ which is the gravitational-wave energy density per unit logarithmic frequency, divided by the critical energy density ρ_c to close the Universe:

$$\Omega_g(f) = \frac{1}{\rho_c}\frac{dE}{d^3x\, d(\ln f)}. \tag{1.1}$$

Some possible sources of a SB include (i) random superposition of many weak signals from binary-star systems [4], (ii) decaying cosmic strings [5] and first-order phase transitions [6] in the early Universe, and (iii) parametric amplification of quantum mechanical zero-point fluctuations in the metric tensor during inflation [7–9]. See Refs. [10–12] for reviews. The predicted wave strengths from all of these stochastic sources are highly uncertain, reflecting our relative ignorance of the relevant physics and/or astrophysics. Hence, detecting or placing upper limits on the SB can bring us valuable information, particularly about the very early Universe.

Relic gravitational waves produced during inflation are particularly interesting, because, as Grischuk has shown [13], the energy spectrum for these waves contains a unique imprint of the time evolution of the Universe's scale factor $a(t)$. We now discuss what is known about the magnitude of the contribution to $\Omega_g(f)$ from these waves, at frequencies relevant to LIGO and/or VIRGO. The predictions for Ω_g from cosmological models are not very firm: they can vary between ~ 1 and $\sim 10^{-14}$ or less. However, observational upper bounds on $\Omega_g(f)$ in various frequency bands give interesting constraints on the models [8]. In turn these constraints can be used, within the context of particular inflation models, to place upper bounds on the contribution of relic gravitons to the value

of $\Omega_g(f)$ in LIGO wave band, see, e.g., Ref. [14]. The strongest such upper bound comes from matching the normalization of the scalar and tensor fluctuations produced during inflation to the recent Cosmic Background Explorer (COBE) measurement of the microwave background anisotropy [9, 15]. The result of this matching is somewhat discouraging: standard exponential inflation models predict that $\Omega_g(f) \leq 3 \times 10^{-14}$ at the 95% confidence level [9], far too small to be detected [cf. Eq. (6.6) below]. While it is far from certain that exponential inflation is correct, it seems unlikely that the expansion during the inflationary era was so much faster than exponential as to tilt the gravitational wave spectrum enough to give a detectable signal at high, LIGO frequencies. It is also possible, of course, that the observed microwave anisotropy was caused by physical processes of structure formation other than inflation.

Despite these pessimistic prospects for the detection of relic gravitons by LIGO and/or VIRGO, it is certainly possible that there will be a detectable signal from other sources such as cosmic strings [5]. Hence, it is important to determine how the detector arm orientations, which will not be changeable in the future, affect the sensitivity of the detector network to the SB. To do so is the first of the two principal purposes of this paper. This issue was first considered by Michelson [16], and has been extensively discussed by Christensen [12, 17]. Essentially we build on and extend slightly their analyses. Our conclusions are also slightly different from those of Christensen.

The second principal purpose of this paper is to spell out the optimal data-processing procedure for searching for the SB with a network of detectors. The algorithm for two detectors is implicit in Michelson [16] (and is incorrectly treated in Ref. [17]); we give a more detailed description and a generalization to a network of detectors, taking into account the possible effect of correlated sources of noise. We now turn to a description of our results and an explanation of how they relate to earlier work.

B. Detection of the stochastic background

The effect of the SB on a gravitational-wave detector is essentially to produce a small contribution to the random, Gaussian noise in its output. For one detector this contribution will be swamped by the detector's own sources of noise, unless the SB strength is implausibly large ($\Omega_g \sim 10^{-6}$; see Sec. II). For two detectors which have no common sources of noise, however, the only contribution to the correlated fluctuations in the detector outputs will be the SB. By cross correlating the outputs of the detectors, the SB can in principle be measured. If one had identical, LIGO-type detectors at the same site, oriented in same way so that they respond in exactly the same way to the SB, and with levels of intrinsic noise corresponding to the "advanced detectors" of Ref. [1], then cross correlating would give a sensitivity to Ω_g of the order of 10^{-10} in the frequency band $10\,\mathrm{Hz} \lesssim f \lesssim 1000\,\mathrm{Hz}$ [10].

For a pair of separated, nonaligned detectors, two new physical effects complicate the analysis [12, 16, 17]. First, if the detectors are not aligned the same way, they will respond to different polarization components of the SB. Orthogonal polarization components of the SB are expected to be statistically independent, and so the cross correlation will be reduced. Second, for each mode there will be a time lag between exciting the first detector and the second detector, and hence phase lags in the cross correlation. For the LIGO detectors separated by $\sim 3000\,\mathrm{km}$ and having maximum sensitivity at a frequency of $f \sim 70\,\mathrm{Hz}$, a typical phase lag will be of order unity. Hence, there will be some destructive interference between parts of the cross correlation that are due to modes which propagate in different directions. Thus we expect a reduction in the sensitivity of the detector pair to the SB.

To analyze this reduction, it is necessary to (i) determine how to *optimally process* the output from the detectors, and (ii) find the signal-to-noise ratio (SNR) that results from this method of filtering. We also want to (iii) determine how the optimal SNR depends on the detector orientations, and (iv) find those orientations that maximize the SNR. Steps (i) and (ii) were analyzed in Refs. [16, 17]. They found that, when optimal signal processing is used, the square of the signal-to-noise ratio for a broadband measurement of the SB is [18]

$$\frac{S^2}{N^2} = \left(\frac{4G\rho_c}{5\pi c^2}\right)^2 2\hat{\tau} \int_0^\infty df\, \frac{\Omega_g(f)^2 \gamma(f)^2}{f^6 S_n(f)^2}. \tag{1.2}$$

Here $\hat{\tau}$ is the duration of the measurement, and $S_n(f)$ is the spectral noise density in either detector. The key quantity appearing in this equation is the dimensionless function $\gamma(f)$, which we call the *overlap reduction function*. It characterizes the reduction in sensitivity to the SB of the detector pair at frequency f that is due to their separation and nonoptimal orientations, and its value is unity for coincident, aligned detectors. In Refs. [16, 17] a formula for the overlap reduction function was derived, which expresses it as an integral over all solid angles of the complex phase lag between the detectors, weighted by combinations of the detector beam pattern functions [cf. Eq. (2.13) below]. Christensen [12, 17] numerically calculated this function for various detector configurations and discussed some of its properties. However its dependence on the detector orientations was not apparent.

In this paper we derive an analytic formula for the overlap reduction function. Using this formula we are able to carry through steps (iii) and (iv) outlined above. We also determine how good are the choices that have been made for the orientations of the detectors in LIGO, VIRGO, and GEO (an as-yet-unfunded British/German Collaboration); i.e., we determine how their sensitivity to the SB compares to the sensitivity they would have if they were optimally oriented.

C. Effect of detector orientations

Our results are as follows. Call σ_1 the angle between the bisector of the arms of the first detector and the arc of the great circle that joins the detectors, and similarly define σ_2 for the second detector. Let $\delta = (\sigma_1 - \sigma_2)/2$ and $\Delta = (\sigma_1 + \sigma_2)/2$, so that δ describes the relative ro-

tation of the detector pair, and Δ describes their average rotation with respect to the line joining them. Then in Sec. V below we show that the optimum SNR (1.2) is given by

$$\frac{S^2}{N^2} = A\cos^2(4\delta) + 2B\cos(4\delta)\cos(4\Delta) + C\cos^2(4\Delta), \tag{1.3}$$

where the quantities A, B, and C are independent of δ and Δ, and A and C are positive. There are thus two possibilities for the optimum detector orientation, depending on the sign of B: (I) $\cos(4\delta) = -\cos(4\Delta) = \pm 1$, corresponding to each detector having an arm along the line joining them, and (II) $\cos(4\delta) = \cos(4\Delta) = \pm 1$, corresponding to each detector arm being at an angle of 45° (mod 90°) to this line. In Fig. 1 we plot the SNR for both of these choices of orientation, as a function of the angle β subtended between the detectors at the center of the earth. Configuration II is optimal for $\beta \lesssim 70°$, while configuration I is optimal for detectors which are further apart. Figure 1 shows that detectors which are close together are the most sensitive; the sensitivity of LIGO is roughly $\sim 25\%$ of what it would be if its detectors were coincident. It also shows that detectors whose planes are roughly perpendicular ($\beta \sim 90°$) have poor sensitivity, as we would expect.

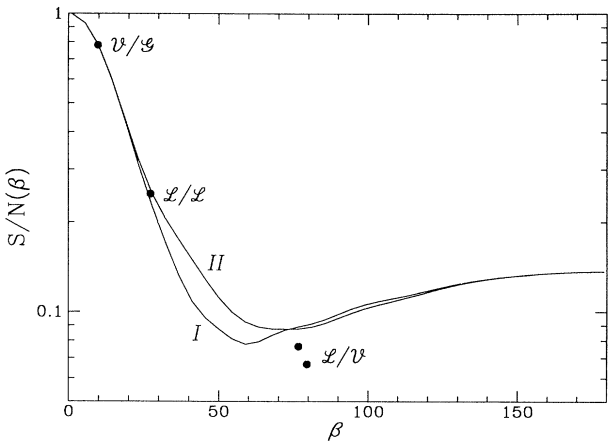

FIG. 1. The broadband signal-to-noise ratio for a pair of detectors as a function of the angle β subtended between them at the center of the Earth, normalized to unity for coincident detectors. Curve I corresponds to each detector having an arm along the arc of the great circle that joins them, and curve II corresponds to each one having an arm at 45° to this arc. The optimal configuration is II for $\beta \lesssim 70°$ (except very close to $\beta = 0$), and I for larger values of β. The point $\mathcal{L}/\mathcal{L}$ and the two points marked $\mathcal{L}/\mathcal{V}$ show the expected sensitivities of the LIGO detector pair and of both LIGO-VIRGO detector pairs, with their current orientations. The point $\mathcal{V}/\mathcal{G}$ shows the sensitivity the detector pair VIRGO-GEO would have if the orientations were chosen optimally for the stochastic background (which will probably not be the case, since optimization for the stochastic background implies sensitivity to only one of the waves' two polarizations, and a corresponding loss of information when studying nonstochastic waves).

We also show in Sec. V that the narrow-band sensitivity of the detector pair near a given frequency f, which is proportional to $|\gamma(f)|$, is also always optimized at either configuration I or II. For example, at very low frequencies, $|\gamma(f)|$ becomes essentially the overlap of the polarization tensors of the two detectors [cf. Eq. (B5) below with $\rho_2(0) = \rho_3(0) = 0$], which is maximized in configuration I. This low frequency limit was previously derived by Christensen [17]. Motivated by this, he suggested that configuration I was always the best orientation to choose. Figure 1 shows that, though this is not true for some values of the separation angle β, the amount lost by choosing I rather than II is never more than a few percent in SNR.

We now discuss the orientations that have been chosen for the detector systems that are under construction or that have been proposed. The relative rotation angle δ for a pair of detectors essentially determines whether the detectors respond to different polarization components, or to the same polarization component, of the gravitational wave field. The advantage in responding to different components ($2\delta \sim 45°$) is that more information can be extracted from incoming burst signals. On the other hand, if the detectors respond to the same component ($\delta \sim 0°$) then the detection signal to noise threshold is reduced, i.e., the fact that the same wave form is seen in both detectors means that one can be more confident that a candidate event is not due to detector noise. These considerations guided the choices of the presently planned values of δ for the detector pairs LIGO-LIGO (there will be two LIGO detectors) and VIRGO-GEO, which are $\delta \approx 0°$ and $2\delta \approx 45°$, respectively [19].

Within the context of these constraints, a key issue that we wanted to understand was the following: given the above values of δ, how much does the broadband sensitivity of the detector pair depend on Δ, i.e., by how much can the SNR be reduced if Δ is chosen arbitrarily instead of being optimally chosen? The answer we obtain (Secs. IV and V below) is that for relatively close detectors with $\beta \lesssim 30°$, the dependence on Δ is very weak; but the dependence is strong for detectors on different continents. Hence for LIGO's parallel detectors, the sensitivity will be close to optimal irrespective of the value of Δ; the present choice of $\Delta = 28.2°$ implies that the SNR for LIGO is $\sim 97\%$ of the sensitivity at optimal orientation. The VIRGO-GEO orthogonal detector pair, however, will have a SNR of less than 10^{-3} times the optimal value, irrespective of the value of Δ, if 2δ is chosen to be 45° (as is planned, so as to optimize the information obtainable from burst sources).

Finally, we estimate the 90% upper confidence limit that can be placed on $\Omega_g(f)$ by the so-called "advanced detectors" in LIGO [1], in a broadband measurement using one-third of a year integration time, to be $\sim 5\times 10^{-10}$ in the frequency band $20\,\mathrm{Hz} \lesssim f \lesssim 70\,\mathrm{Hz}$ [cf. Eq. (6.5) below]. This is a little worse than earlier estimates which assume that the detectors are coincident and aligned [10, 11].

D. Organization of this paper

The layout of this paper is as follows. In Sec. II we define the cross correlation matrix for a network of de-

tectors and give the formula for the contribution to this quantity from the SB. In Sec. III we describe the general optimal data-processing strategy, the justification of which is given in Appendix A. In this appendix we also derive the generalization of the signal-to-noise formula for optimal signal processing (1.2), to a network of more than two detectors, but more importantly to a network that has more than one interferometer per site: we show how to take into account the effect of correlated noise in detectors at the same site by introducing the concept of the *effective spectral noise density* of a detector site. In Appendix B we derive the formula for the overlap reduction function, and we describe some of its properties in Sec. IV.

Next, in Sec. V, we show how to optimize the orientations of a pair of detectors, both for narrow-band and for broadband measurements of the SB. Section VI describes the implications of our results for the LIGO, VIRGO, and GEO detector facilities. Finally in Sec. VII we summarize our main results.

We use units throughout in which the speed of light c and Newton's gravitational constant G are unity.

II. DETECTOR CROSS CORRELATION MATRIX

The effect of a stochastic background on a detector network will be essentially to produce statistical correlations between the outputs of the various detectors. A key result which we will need is an expression for these correlations in terms of the spectrum $\Omega_g(f)$ of the gravitational background. This was first given by Christensen [12, 17], although it is implicit in the work of Michelson [16]. We briefly describe the derivation in this section, and we lay the foundations for our analysis of Appendix A and Sec. III.

A detector network with N detectors will have outputs

$$h_a(t) = h_a^{\rm signal}(t) + n_a(t) + s_a(t), \tag{2.1}$$

for $1 \le a \le N$. Here h_a is the strain amplitude that we read out from the ath detector; it consists of an intrinsic detector noise n_a, a contribution from the SB s_a, and possibly a contribution $h_a^{\rm signal}$ from nonstochastic gravitational waves (bursts and periodic waves). The noise n_a and the SB induced strain s_a are independent random processes, which we assume to be Gaussian and stationary.

The detector correlations can be described by the correlation matrix

$$C_h(\tau)_{ab} = \langle h_a(t+\tau) h_b(t) \rangle - \langle h_a(t+\tau) \rangle \langle h_b(t) \rangle, \tag{2.2}$$

where angular brackets mean an ensemble average or a time average. The Fourier transform of the correlation matrix, multiplied by two, is the power spectral density matrix:

$$S_h(f)_{ab} = 2 \int_{-\infty}^{\infty} d\tau \, e^{2\pi i f \tau} C_h(\tau)_{ab}. \tag{2.3}$$

This is a positive definite Hermitian matrix which satisfies the equations

$$\langle \tilde{h}_a(f) \tilde{h}_b(f')^* \rangle = \frac{1}{2} \delta(f - f') S_h(f)_{ab}, \tag{2.4}$$

and

$$\left\langle e^{i \int dt \, w_a(t) h_a(t)} \right\rangle = \exp\left\{ -\frac{1}{2} \int_0^{\infty} df \tilde{\mathbf{w}}^{\dagger} \cdot \mathbf{S}_h \cdot \tilde{\mathbf{w}} \right\}, \tag{2.5}$$

for any functions $w_a(t)$. Here tildes denote Fourier transforms, according to the convention that

$$\tilde{h}(f) = \int e^{2\pi i f t} h(t) dt.$$

Since the random processes $n_a(t)$ and $s_a(t)$ are uncorrelated, the spectral density matrix of the detector outputs is just the sum of those for the detector noise and for the background:

$$\mathbf{S}_h(f) = \mathbf{S}_n(f) + \mathbf{S}_s(f). \tag{2.6}$$

To derive an expression for $\mathbf{S}_s(f)$, two key ingredients are needed. The first is a mode expansion for the metric perturbation for an isotropic, stationary SB. Expressed in frequency space, this is

$${}^{(\rm SB)}\tilde{h}_{ij}(\mathbf{x}, f) = \int d^2\Omega_n \sum_{A=+,\times} s_{A,\mathbf{n}}(f) \, e^{2\pi i f \mathbf{n}\cdot\mathbf{x}} \, e_{ij}^{A,\mathbf{n}} \tag{2.7}$$

for $f \ge 0$, and ${}^{(\rm SB)}\tilde{h}_{ij}(\mathbf{x}, f) = {}^{(\rm SB)}\tilde{h}_{ij}(\mathbf{x}, -f)^*$ for $f < 0$. Here the tensors $\mathbf{e}^{A,\mathbf{n}}$ are the usual transverse traceless polarization tensors, normalized according to $e_{ij}^{A,\mathbf{n}} e_{ij}^{B,\mathbf{n}} = 2\delta_{AB}$, and $\int d^2\Omega_n$ denotes the integral over solid angles parametrized by the unit vector $\mathbf{n}$. The coefficients $s_{A,\mathbf{n}}$ are random processes which satisfy [20]

$$\langle s_{A,\mathbf{n}}(f) \, s_{B,\mathbf{m}}(f')^* \rangle = \delta_{AB} \, \delta^2(\mathbf{n}, \mathbf{m}) \times \delta(f - f') \, \frac{\rho_c}{4\pi f^3} \, \Omega_g(f) \tag{2.8}$$

and

$$\langle s_{A,\mathbf{n}}(f) \, s_{B,\mathbf{m}}(f') \rangle = 0, \tag{2.9}$$

for $f, f' \ge 0$ [21]. Here $\delta^2(\mathbf{n}, \mathbf{m})$ is the δ function on the unit sphere.

The second ingredient is the expression for the response of the ath detector to the background. This is

$$s_a(t) = \mathbf{d}_a : {}^{(\rm SB)}\mathbf{h}(\mathbf{x}_a, t), \tag{2.10}$$

where $\mathbf{x}_a$ is the position of the detector, the colon denotes a double contraction, and $\mathbf{d}_a$ is a symmetric tensor that characterizes the detector's orientation (its polarization tensor). If the arms of the detector are in the directions of the unit vectors $\boldsymbol{l}$ and $\mathbf{m}$, then $\mathbf{d}_a = (\boldsymbol{l} \otimes \boldsymbol{l} - \mathbf{m} \otimes \mathbf{m})/2$ [22]. From Eqs. (2.7) and (2.10) we obtain that

$$\tilde{s}_a(f) = \sum_A \int d^2\Omega_n \, F_a^A(\mathbf{n}) \, s_{A,\mathbf{n}} \, e^{2\pi i f \mathbf{n}\cdot\mathbf{x}_a}, \tag{2.11}$$

where $F_a^A(\mathbf{n}) \equiv \mathbf{d}_a : \mathbf{e}^{A,\mathbf{n}}$ are the detector beam pattern functions. Inserting this response function into an equation analogous to Eq. (2.4), and using Eqs. (2.8) and (2.9), we obtain

$$S_s(f)_{ab} = \frac{4\rho_c \Omega_g(f)}{5\pi f^3} \, \gamma_{ab}(f), \tag{2.12}$$

where [23]

$$\gamma_{ab}(f) = \frac{5}{8\pi}\int d^2\Omega_n\,(F_a^+F_b^+ + F_a^\times F_b^\times) \times \exp\left[2\pi i f\mathbf{n}\cdot(\mathbf{x}_a - \mathbf{x}_b)\right]. \quad (2.13)$$

The functions γ_{ab} are the overlap reduction functions discussed in Sec. I. It can be seen that their value is unity for coincident, aligned detectors. Below when considering a single pair of detectors we shall write $\gamma_{ab}(f)$ simply as $\gamma(f)$.

III. THE OPTIMAL PROCESSING STRATEGY

In this section we describe the optimal method for filtering the detector outputs when searching for the SB. We discuss the two-detector case in Sec. III A; the filtering method in this subsection is implicit in the formulas of Ref. [16]. A detailed proof that the method is optimal is given in Appendix A. In Sec. III B we discuss the effects of correlated sources of noise, and explain why correlation measurements between detectors at widely separated sites yield much better upper bounds on $\Omega_g(f)$ than correlation measurements between detectors at one site. Next we describe the modifications to the filtering method necessitated by correlated noise, in Sec. III C.

A. General description

To measure the stochastic background one really needs to measure the spectral density matrix of the detector outputs $\mathbf{S}_h(f)$. Now there is no way in principle to separate out the portions of $\mathbf{S}_h(f)$ due to detector noise $\mathbf{S}_n(f)$, and due to the SB $\mathbf{S}_s(f)$. If we had only one detector, we could only conclude that $S_s(f) \le S_h(f)$. From Eq. (2.12) and using $\gamma_{aa}(f) = 1$, this would give an upper bound on Ω_g of

$$\Omega_g(f) \lesssim 2.5\times 10^{-6}\left(\frac{h_n(f)}{10^{-23}}\right)\left(\frac{f}{100\,\mathrm{Hz}}\right)h_{75}^{-2}. \quad (3.1)$$

Here $h_n = \sqrt{fS_h(f)}$, which is projected to be $\gtrsim 10^{-23}$ for LIGO, even at an advanced stage [1]. The quantity h_{75} is the Hubble constant scaled to the value of 75 km $\mathrm{sec}^{-1}\,\mathrm{Mpc}^{-1}$.

It is unlikely that Ω_g will be as large as the value in Eq. (3.1). However, if it does happen that $\Omega_g \gtrsim 10^{-5}$ in the LIGO wave band, then the SB-induced noise may dominate over the other sources of detector noise at some frequencies, and may ultimately constrain the amount of information that we can extract from burst gravitational waves. In this paper we shall from now on assume that Ω_g is small, and restrict attention to measurements made using two or more detectors.

With two or more detectors, one takes advantage of the fact that the sources of noise in each detector will be independent. This will be the case for detectors at widely separated sites, because sources of noise that are correlated between the detectors on time scales of the order of the light travel time between them are expected to be insignificant, or if not they can be monitored and compensated for [17]. When correlated noise is unimportant, then the off-diagonal elements of $\mathbf{S}_n(f)$ will be very small, so that

$$S_s(f)_{ab} \approx S_h(f)_{ab} \quad \text{for } a \neq b. \quad (3.2)$$

By measuring these components we can gain information about the SB. One does this by cross correlating the two output streams [10]. One takes each detector output $h_a(t)$, $a = 1, 2$, and constructs, using an optimizing linear filter $K(t)$, the quantity $\tilde{H}_a(f) = \tilde{K}(f)\tilde{h}_a(f)$. The purpose of this filter is essentially to suppress the signal at those frequencies at which the detector noise is strong, and it is given by [cf. Eq. (A51) below]

$$\tilde{K}(f) = \frac{1}{f^{3/2}S_h(f)}. \quad (3.3)$$

For coincident, aligned detectors the next step is simply to integrate $H_1(t)$ against $H_2(t)$, see, e.g., Ref. [10]. For noncoincident detectors, however, a different strategy is necessary. One first constructs the correlation with time delay τ,

$$Y(\tau) = \int_{-\hat{\tau}/2}^{\hat{\tau}/2} dt\, H_1(t+\tau)H_2(t), \quad (3.4)$$

where $\hat{\tau}$ is the observation time, typically of the order of 1 yr. Then one calculates the weighted average

$$Y = \int_{-\tau_1}^{\tau_1} d\tau\, L(\tau)Y(\tau), \quad (3.5)$$

where τ_1 is the light travel time between the detectors, and $L(\tau)$ is a weighting function which must be carefully chosen for each detector pair in order to maximize the sensitivity. Roughly speaking, this smearing of the cross correlation compensates in some measure for the phase lags between the detectors which were discussed in Sec. I. The quantity Y will then have a truncated Gaussian distribution (i.e., Gaussian but restricted to positive values) with signal to noise ratio given by Eq. (1.2) [18].

In fact the sliding delay function $L(\tau)$ is just the Fourier transform of the overlap reduction function [see Eqs. (A50) and (A51)]. In Appendix B we give an analytic formula for $L(\tau)$, and we show in Fig. 2 the sliding delay function that will need to be used for the LIGO pair of detectors.

In order of magnitude, the 90% confidence upper limit that can be placed on Ω_g by cross correlating between the detectors is [10]

$$\Omega_g^{\max} \approx \frac{\Omega_0}{\sqrt{\hat{\tau}\Delta f}}. \quad (3.6)$$

Here $\Omega_0 \sim 10^{-6}$ is the upper bound (3.1) obtainable from one detector, and Δf is the bandwidth of the measurement. [If no bandpass filtering of the data is carried out, Δf will be roughly the width of the peak of the function $1/(f^3S_n)$; and if filtering with a bandwidth Δf is used, then the domain of integration in Eq. (1.2) must be suitably restricted.]

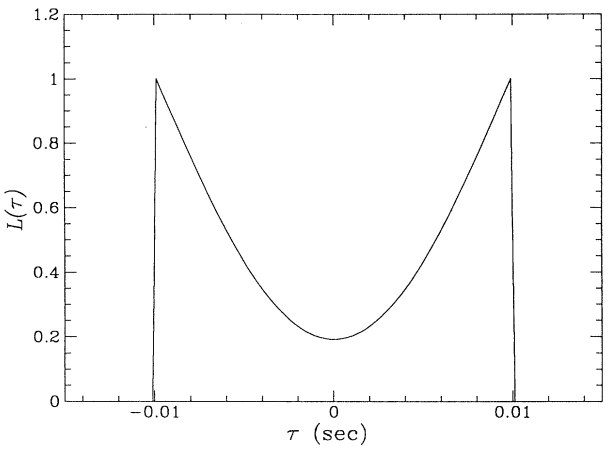

FIG. 2. The "sliding delay function" $L(\tau)$ for the LIGO pair of interferometers. To maximize the broadband sensitivity to the stochastic background, the cross correlation with a time delay τ between the detector pair must be integrated against this function; see text.

B. Effects of correlated noise

Up to this point we have assumed that the intrinsic detector noise is uncorrelated between different detectors, i.e., that the matrix $\mathbf{S}_n(f)$ is diagonal. We now relax this assumption and consider the effects of correlated noise. It is planned for LIGO and VIRGO to ultimately have two or three detectors per detector site, perhaps optimized for different types of gravitational wave sources. Thus, there will be two possible types of correlation measurements: *intrasite* measurements at one site, and *intersite* measurements between detectors at different sites. As previously mentioned, it seems very unlikely that there will be any significant correlated noise in intersite measurements, and so it will only be important for intrasite measurements.

We now consider what information can be extracted from intrasite correlations. This issue has been previously discussed by Christensen [17], who indicates that intrasite upper bounds on the SB will be competitive with those from intersite measurements. As we now discuss, we disagree with this conclusion. The quantity we can measure is the sum $\mathbf{S}_n(f) + \mathbf{S}_s(f)$, the two terms of which are in principle indistinguishable [24]. Hence, we can draw inferences about $\mathbf{S}_s(f)$ *only* if we have some information about the correlated noise. The most obvious such *a priori* information is the fact that $\mathbf{S}_n$ is a positive Hermitian matrix. It follows that if we measure the total power spectral density matrix to be $\hat{\mathbf{S}}_h$, then $\mathbf{S}_s < \hat{\mathbf{S}}_h$. However, this only tells us that $\Omega_g \gtrsim 10^{-6}$ [cf. Eq. (3.1) above], since the inequality applies to the total matrix, and the off-diagonal elements are small compared to the diagonal elements. In particular, it is not true without further assumptions that

$$S_s(f)_{ab} \leq |\hat{S}_h(f)_{ab}| \qquad \text{for } a \neq b \tag{3.7}$$

(as is implicitly assumed in Ref. [17]), since $S_n(f)_{ab}$ may be negative (or complex). A complex value of $S_n(f)_{ab}$ corresponds physically to sources of noise that excite two detectors with a certain preferred phase lag between them.

In Appendix A, we derive the 90% confidence upper limit that one can place on Ω_g in a bandwidth Δf using intrasite correlations, which generalizes Eq. (3.6). If we define the noise correlation coefficients by

$$\mathcal{C}_{ab} \equiv \frac{S_n(f)_{ab}}{\sqrt{S_n(f)_{aa}\, S_n(f)_{bb}}}, \tag{3.8}$$

then the result depends on (i) the measured value $\hat{\mathcal{C}}$ of $\mathcal{C}$ (determined from $\hat{\mathbf{S}}_h$), and (ii) the assumed *a priori* maximum value $\mathcal{C}_{\max}$ of $|\mathcal{C}|$. In order of magnitude we find [cf. Eqs. (A36) and (A40) below]

$$\Omega_g^{\max} \approx \Omega_0 \left[\hat{\mathcal{C}} + \sqrt{\frac{1}{\hat{\tau}\Delta f} + \mathcal{C}_{\max}^2}\right]. \tag{3.9}$$

Since there will be various unknown sources of weak correlated noise, it is not appropriate to choose a very small value of $\mathcal{C}_{\max}$. Hence, the upper bound (3.9) will be much worse than the bound (3.6) obtained from intersite correlations.

We also consider the assumption that the correlated detector noise always excites different detectors in phase, but can be arbitrarily large in magnitude, so that

$$S_n(f)_{ab} \geq 0 \text{ for } a \neq b. \tag{3.10}$$

In this case intrasite correlation measurements cannot be used to *detect* the SB, but can only be used to place upper bounds on its magnitude. The resulting upper limit [cf. Eq. (A38)] is given by Eq. (3.9) with $\mathcal{C}_{\max} = 0$. Thus, the intersite correlations will still give better bounds unless the actual amount of correlated noise in the detectors satisfies

$$\hat{\mathcal{C}} \lesssim \frac{1}{\sqrt{\hat{\tau}\Delta f}}, \tag{3.11}$$

which is $\sim 10^{-4}$ for a year-long measurement with a bandwidth of ~ 50 Hz. Since it is not clear that either of the conditions (3.10) or (3.11) will be appropriate, we henceforth consider only intersite correlations.

C. Filtering intersite correlation measurements

In Appendix A we show how to optimally filter the intersite data when there is a detector network with several interferometers per detector site. In this case the detector noise matrix $\mathbf{S}_n(f)$ will have a block diagonal form, with each block corresponding to a site. Let $\mathbf{S}_A$ be the subblock corresponding to the Ath site. The strategy is essentially to measure the off-diagonal blocks of $\mathbf{S}_h(f)$ (i.e., the intersite correlations), and to use these to obtain information about the SB. The resulting SNR squared is then given [cf. Eq. (A52) below] by a simple generalization of Eq. (1.2):

$$\frac{S^2}{N^2} = \left(\frac{4\rho_c}{5\pi}\right)^2 2\hat{\tau} \int_0^\infty df\, \frac{\Omega_g(f)^2}{f^6} \sum_{A<B} \frac{\gamma_{AB}(f)^2}{S_A^{(\mathrm{eff})}(f) S_B^{(\mathrm{eff})}(f)}. \tag{3.12}$$

Here the sum is over the pairs of sites, and γ_{AB} is the overlap reduction function for any detector at site A together with any detector at site B. The quantities $S_A^{(\mathrm{eff})}(f)$, which we call the effective site spectral noise densities for correlations between sites, are given by

$$S_A^{(\mathrm{eff})}(f) = \left[\sum_{ab} \left(\mathbf{S}_A^{-1}\right)_{ab}\right]^{-1}. \tag{3.13}$$

They will be real since the matrices $\mathbf{S}_A$ are Hermitian. In the case where the off-diagonal elements of $\mathbf{S}_A$ are all equal to $S'_n(f)$, and the common value of the diagonal elements is $S_n(f)$, then we obtain that

$$S_A^{(\mathrm{eff})}(f) = \frac{1}{N} S_n(f) + \left(1 - \frac{1}{N}\right) S'_n(f), \tag{3.14}$$

where N is the number of detectors. This formula shows that the use of N detectors instead of a single detector at one site reduces that sites effective noise density by a factor of $1/N$, *if* the detectors noise sources are independent ($S'_n \approx 0$). If the noise sources are strongly correlated, however, so that $S'_n \approx S_n$, then there is no significant reduction in the effective noise.

IV. THE OVERLAP REDUCTION FUNCTION

In the preceding sections we have seen that for a pair of interferometric detectors, the overlap reduction function $\gamma(f)$ characterizes the dependence on the detector separation and orientations of both the correlation matrix (2.12), and the broadband sensitivity to the SB (1.2). We now give an analytic formula for this function and discuss some of its properties.

A. The general formula for $\gamma(f)$

We first define the variables which describe the orientations and separation of a pair of detectors. Let L be the line joining the two detectors and P_1 (P_2) be the plane formed by the arms of the first (second) detector; see Fig. 3. The variables we will use are (i) the distance d between the detectors, (ii) the acute angle β_1 between L and P_1, (iii) the angle σ_1 between the projection of L onto P_1 and the bisector of the two arms of the first detector, (iv) the corresponding angles β_2 and σ_2, and (v) the angle χ between L and and the intersection of P_1 and P_2. The directions (clockwise or anticlockwise) in which σ_1 and σ_2 are chosen to be positive are unimportant, as long as the conventions for σ_1 and for σ_2 coincide as $\beta_1, \beta_2 \to 0$. Let $\delta \equiv (\sigma_1 - \sigma_2)/2$ and $\Delta = (\sigma_1 + \sigma_2)/2$. In Sec. I we have called these the *relative rotation* and *total rotation* angles, respectively, since for $\beta_1 = \beta_2 = 0$, the angle δ is half of the relative rotation of the detectors while Δ measures the average rotation with respect to the line joining them [25].

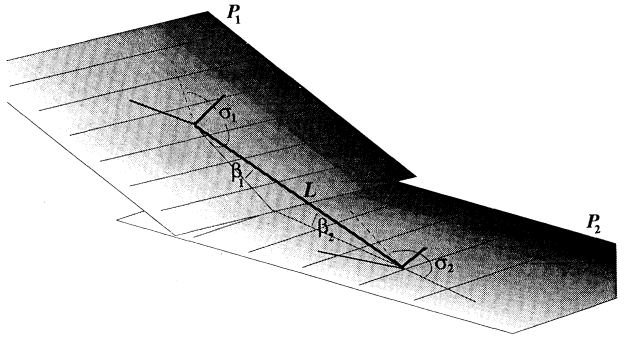

FIG. 3. The angles σ_1, σ_2, β_1, and β_2 formed by a pair of interferometric detectors and the line L which joins them.

In general the overlap reduction function as given in Eq. (2.13) will depend on all of the variables β_1, β_2, δ, Δ, χ, and on the phase lag $\alpha = 2\pi f d$ between the detectors, where f is the frequency. Now for terrestrial detectors $\chi = \pi/2$ and $\beta_1 = \beta_2$ ($= \beta$ say), which is also the angle subtended at the center of the Earth by the detector pair. For this case we derive in Appendix B the formula

$$\gamma(f) = \cos(4\delta)\Theta_1(\alpha,\beta) + \cos(4\Delta)\Theta_2(\alpha,\beta), \tag{4.1}$$

where the functions Θ_1 and Θ_2 are

$$\Theta_1(\alpha,\beta) = \cos^4(\beta/2) g_1(\alpha) \tag{4.2}$$

and

$$\Theta_2(\alpha,\beta) = \cos^4(\beta/2) g_2(\alpha) + g_3(\alpha) - \sin^4(\beta/2)\left[g_2(\alpha) + g_1(\alpha)\right]. \tag{4.3}$$

The functions $g_j(\alpha)$ are given in Eqs. (B11), (B12), and (B13) of Appendix B, and are all linear combinations of the functions $\sin(\alpha)/\alpha^n$ and $\cos(\alpha)/\alpha^n$ where $1 \le n \le 5$.

Several properties of the overlap reduction function $\gamma(f)$, which were discovered by Christensen and others from numerical calculations[12, 16, 17], can be read off the formula (4.1). First, there will always be frequencies f for which $\gamma(f)$ vanishes, and correspondingly near which the narrow-band sensitivity of the detector pair to the SB is very poor. For detectors that are less than a few thousand kilometers apart, the first null frequency is at $f_1 \sim (70\,\mathrm{Hz})(3000\,\mathrm{km}/d)$, irrespective of the detector orientations (see below). Second, the reduction function falls off like $1/f$ when $\alpha \gg 1$, or equivalently when $f \gg f_1$. Hence the 90% confidence limit that we can place on $\Omega_g(f)$ scales like $1/(fd)$ for large d. However for detectors which are $\sim 3000\,\mathrm{km}$ apart, the phase lag α is of order unity for typical detector frequencies, and hence the variation of the sensitivity with distance is more complex than simply scaling like $1/d$ (see Sec. VI below and also Fig. 1 above).

B. Some special cases

To understand the behavior of $\gamma(f)$ as a function of δ, Δ, and β, it is useful to consider some limiting cases. For coplanar, coincident detectors, $d = \beta = 0$, and using

$g_1(0) = 1$ and $g_2(0) = g_3(0) = 0$ we find that $\gamma = \cos(4\delta)$. This is just what we would expect physically. Consider, for instance, detectors which are rotated with respect to one another by 45°. Vertically incident gravitational waves will couple to the two detectors via polarization components that are orthogonal and hence statistically independent; thus these waves give no contribution to the cross correlation. Waves which are incident from nonvertical directions do give rise to correlations between the detectors, but when we average over all incident directions the total correlation vanishes [26]. It is also clear from rotational symmetry that there should be no dependence on the total rotation angle Δ.

Now suppose that the detectors are still coincident but no longer coplanar, so that $\beta \neq 0$. Equivalently, suppose the detectors are separated by a distance d but evaluate $\gamma(f)$ at frequencies $f \ll 1/d$. The result is

$$\gamma(f) = \cos^4(\beta/2)\cos(4\delta) - \sin^4(\beta/2)\cos(4\Delta). \tag{4.4}$$

This equation agrees with Eq. (3.9) of Ref. [17] after making an appropriate change of variables. As we would expect, the dependence on Δ is small when β is small; it is smaller than the dependence on the relative rotation angle δ by a factor of $\tan^4(\beta/2)$ ($\sim 3\times 10^{-3}$ for the planned LIGO detectors; see below).

A third limiting case is when $\beta = 0$ but $d \gtrsim 1/f$, so that the detectors are coplanar but effectively separated. In this case,

$$\gamma(f) = \cos(4\delta)g_1(\alpha) + \cos(4\Delta)\left\{g_2(\alpha) + g_3(\alpha)\right\}. \tag{4.5}$$

Plots of the functions $g_1(\alpha)$ and $g_2(\alpha) + g_3(\alpha)$ are shown in Fig. 4. We see that for $\alpha \lesssim 4$, the $\cos(4\delta)$ term dominates, just as above for coincident detectors for small β. When the phase lag α is large, the functions in Eq. (4.5) can be approximated as $g_0(\alpha) \approx 5\sin\alpha/(16\alpha)$ and $g_2(\alpha) + g_3(\alpha) \approx -g_0(\alpha)$, so we obtain

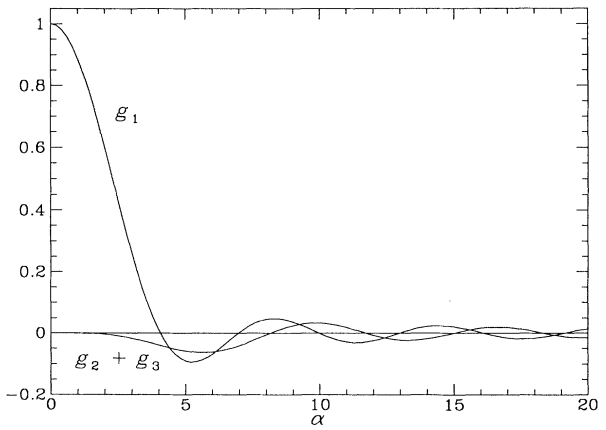

FIG. 4. For coplanar detectors the overlap reduction function is of the form $\gamma(f) = \cos(4\delta)g_1(\alpha) + \cos(4\Delta)\left[g_2(\alpha) + g_3(\alpha)\right]$, where δ and Δ are the relative and total rotation angles, and α is the phase lag between the detectors at frequency f; see text. Here we plot the functions $g_1(\alpha)$ and $g_2(\alpha) + g_3(\alpha)$.

$$\gamma(f) \approx -\frac{5\sin\alpha}{8\alpha}\sin(2\sigma_1)\sin(2\sigma_2). \tag{4.6}$$

What is happening physically here is that the cross correlation is dominated by modes whose propagation vectors are nearly parallel to the line joining the detectors. This can be seen by applying a stationary phase argument to the integral (2.13). Hence it is reasonable that the overlap reduction function should have separate factors for each detector that depend on how they are oriented with respect to the line joining the detectors. In this case the dependence of the overlap reduction function γ on the total rotation angle Δ is strong.

In summary, the dependence on the relative rotation angle δ will dominate over the dependence on the total rotation angle Δ unless two conditions are satisfied: (i) $\beta \gtrsim \pi/4$, and (ii) $\alpha \gtrsim 1$, i.e., $d \gtrsim c/f_n$, where $f_n \sim 100\,\mathrm{Hz}$ is a typical frequency at which the detector is sensitive. In reality of course β and d are not independent; they are related by

$$d = 2\sin(\beta/2)r_E, \tag{4.7}$$

where r_E is the radius of the Earth. Because of the coincidence that the quantity $f_n r_E/c$ is of order unity, conditions (i) and (ii) above become satisfied at approximately the same value of β.

For a pair of terrestrial detectors, as β and d are increased from zero, the following three effects occur: (i) $\gamma(f)$ decreases because d is increasing, as discussed above; (ii) $\gamma(f)$ decreases because β is increasing. From Eq. (4.4) we see that the maximum over all orientations of the value of $\gamma(0)$ is $1 - \sin^2\beta/2$. Hence the effect of increasing β contributes typically a factor of at most 1/2 to the decrease in $\gamma(f)$. (iii) The dependence of $\gamma(f)$ on Δ increases to become as strong as the dependence on δ.

V. OPTIMIZATION OF DETECTOR ORIENTATIONS

We now turn to the issue of how to *optimize* the orientations of a single pair of detectors so as to detect or measure the SB. We assume that the detector's noise curves are the same, so that $S_n(f)_{aa} = S_n(f)_{bb} = S_n(f)$. We consider separately the cases of maximizing the narrow-band sensitivity to the SB in the vicinity of some frequency f, and of maximizing the overall broadband sensitivity.

Now if we rotate a detector through 90°, then its polarization tensor **d** [cf. Eq. (2.10)] will have its sign flipped. Hence $|\gamma(f)|$ is a periodic function of both of the angles σ_1, σ_1 (or δ, Δ) with period 90°. In the following discussion, all values of these angles and all equations involving them should be treated modulo 90°.

A. Narrow-band sensitivity

From Eq. (1.2) we see that the 90% confidence limit that we can place on $\Omega_g(f)$ in a small interval of frequency of width Δf centered about f, using a measurement of duration $\hat{\tau}$, is [18]

$$\Omega_g^{(90\%)}(f;\Delta f) = \left(\frac{5\pi}{4\rho_c}\right)\frac{f^3 S_n(f)}{|\gamma(f)|}\frac{1.65}{\sqrt{2\hat{\tau}\Delta f}}. \qquad (5.1)$$

To minimize this we need to maximize the quantity $|\gamma(f)|$. From Eq. (4.1) it follows that the maximum value over all orientations of $|\gamma(f)|$ is $|\Theta_1(f)| + |\Theta_2(f)|$.

Now if the signs of Θ_1 and Θ_2 are different, then this maximum will be achieved at $\sigma_1 = \sigma_2 = 45°$ (i.e., either $\delta = 45°$, $\Delta = 0°$ or $\delta = 0°$, $\Delta = 45°$). This corresponds to each detector having one arm parallel to the line which joins them. We will call this type of orientation *configuration I*. This will be the optimal orientation at zero frequency, since $\Theta_1(f=0) = \cos^4(\beta/2) > 0$, and $\Theta_2(f=0) = -\sin^4(\beta/2) < 0$. It will also be optimal at high frequencies ($fd \gg 1$), since from Eqs. (4.1) and (B11)–(B14) we see that $\Theta_1(\alpha) \approx -\Theta_2(\alpha)$ for $\alpha \gg 1$.

If, on the other hand, the signs of Θ_1 and Θ_2 are the same, then we see from Eq. (4.1) that $|\gamma(f)|$ will be maximized at $\sigma_1 = \sigma_2 = 0°$ (i.e., either $\delta = \Delta = 0°$ or $\delta = \Delta = 45°$). In this case, which we will call *configuration II*, each detector arm makes an angle of 45° (mod 90°) to the line joining the detectors. This will be optimal, e.g., for $10\,\mathrm{Hz} \lesssim f \lesssim 40\,\mathrm{Hz}$ when $\beta = 29°$ (the LIGO value; see below).

Now suppose that the detector pair is in either configuration I or configuration II. Then for some frequencies the orientations will be optimal, while for some other frequencies they will not be. It is useful to calculate, for a given frequency f, the factor $\mathcal{R}(f)$ by which $|\gamma(f)|$ is reduced if the wrong configuration out of I and II is chosen. Note that choosing between I and II is equivalent to fixing the relative rotation angle δ to be zero, and then choosing Δ to be either 0° or 45°. Hence $\mathcal{R}$ measures the sensitivity of $\gamma(f)$ to the total rotation angle Δ when the detectors are parallel. From Eq. (4.1), we obtain

$$\mathcal{R} = \frac{||\Theta_1| - |\Theta_2||}{|\Theta_1| + |\Theta_2|}. \qquad (5.2)$$

At low frequencies ($\alpha = 2\pi f d \ll 1$) this is approximately

$$\mathcal{R}(f \to 0) \approx \left(1 + \frac{2\nu^4}{1-2\nu^2}\right)^{-1}, \qquad (5.3)$$

where $\nu = \sin(\beta/2)$. Thus for small values of β we have $\mathcal{R}(f \to 0) \approx 1$, and whether configuration I or II is chosen is relatively unimportant.

At high frequencies ($\alpha \gg 1$), however, we find that $\mathcal{R}(f)$ asymptotes to a constant $\mathcal{R}_\infty$ independent of frequency, where

$$\mathcal{R}_\infty \approx \left(\frac{2\nu}{\nu^2+1}\right)^2, \qquad (5.4)$$

which can be quite small. [$\mathcal{R}_\infty \approx 0.2$ for the LIGO detectors; see below.] Hence the choice of Δ will have a large effect on the high frequency sensitivity to the SB, especially for detectors which are close together for which ν is small.

B. Broadband sensitivity

It is quite likely that the SB will be so weak that it will only be detectable (if at all) by integrating over a broad range of frequencies, e.g., from $\sim 20\,\mathrm{Hz}$ to $\sim 70\,\mathrm{Hz}$. Hence it is more important to optimize the overall broadband response (1.2) of the pair of detectors than to optimize the narrow-band sensitivity at some frequency. We now determine how to do this if we assume that the spectrum $\Omega_g(f)$ is approximately constant over the relevant frequency range—the so-called "Zel'dovich spectrum."

By inserting Eq. (4.1) into Eq. (1.2), we find that the SNR squared after optimal filtering is

$$\frac{S^2}{N^2} = A(\beta)\cos^2(4\delta) + 2B(\beta)\cos(4\delta)\cos(4\Delta) + C(\beta)\cos^2(4\Delta), \qquad (5.5)$$

where

$$B(\beta) = \left(\frac{4\rho_c}{5\pi}\right)^2 2\hat{\tau}\int_0^\infty df\,\frac{\Omega_g^2\Theta_1\Theta_2}{f^6 S_n^2}, \qquad (5.6)$$

and A and C are given by similar formulas with $\Theta_1\Theta_2$ replaced by Θ_1^2 and Θ_2^2, respectively. Note that A, B, and C depend on β in two distinct ways, since from Eq. (4.7),

$$\Theta_{1,2}(f) = \Theta_{1,2}[\alpha = 4\pi f r_E \sin(\beta/2), \beta]. \qquad (5.7)$$

The explicit formula (5.5) makes it easy to determine how to optimally orient the detectors at a fixed value of β, and also how the SNR at the optimal orientation varies with β. The SNR (5.5) is maximized at $|\cos(4\delta)| = |\cos(2\Delta)| = 1$, and the optimum orientations are just configurations I and II again, applying when $B < 0$ and $B > 0$, respectively.

VI. IMPLICATIONS FOR LIGO AND VIRGO

The planned LIGO detectors in Hanford, Washington and Livingston, Louisiana will have $\beta = 27.2°$, $\delta = 44.9°$, and $\Delta = 28.2°$ (the last two mod 90°) [27]. A graph of $\gamma(f)$ for this configuration is shown in Fig. 5. In Fig. 6 we show what the functions Θ_1 and Θ_2 look like at the value $\beta = 27.2°$ appropriate for LIGO. For any possible LIGO orientations, $\gamma(f)$ will be a linear combination of these two functions as in Eq. (4.1). From Fig. 6 we can see that there will be a zero of $\gamma(f)$ at $f \sim 70\,\mathrm{Hz}$, whose position

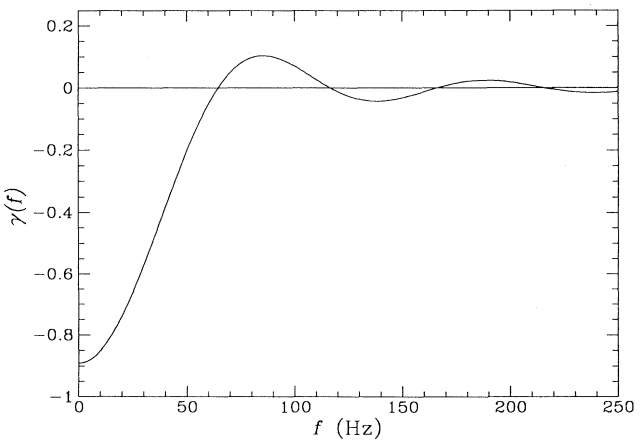

FIG. 5. The overlap reduction function for the LIGO detectors using their currently planned orientations.

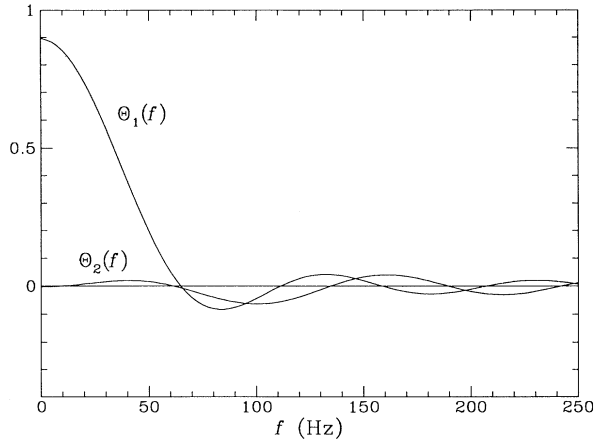

FIG. 6. For *any* orientation angles δ, Δ of the LIGO detectors, the overlap reduction function will be given in terms of two functions of frequency $\Theta_1(f)$ and $\Theta_2(f)$ by $\gamma(f) = \cos(4\delta)\Theta_1(f) + \cos(4\Delta)\Theta_2(f)$. Here we show the functions $\Theta_1(f)$ and $\Theta_2(f)$.

is relatively insensitive to the orientations chosen.

The noise power spectral density in the LIGO detectors after some years of operation might be roughly

$$S_n(f) = \max\left[S_m(f/f_m)^{-4}, S_m(f/f_m)^2\right], \qquad (6.1)$$

where $S_m = 10^{-48}\,\mathrm{Hz}^{-1}$ and $f_m = 70\,\mathrm{Hz}$. For frequencies less than $\sim 10\,\mathrm{Hz}$ the noise will be effectively infinite. Equation (6.1) is a crude analytic fit to the noise curve estimate for the "advanced detectors" given by the LIGO team in Ref. [1]. A more detailed model of the noise is unnecessary because of the uncertainty as to what the actual noise levels might be.

We now estimate what the effective site noise level $S_n^{(\mathrm{eff})}$ [cf. Eq. (3.14)] might be if, after an upgrade of the LIGO facility, there are eventually two or three interferometers at each site, as is planned. The noise spectral density $S_n(f)$ in each detector will be a sum of the form [1, 10]

$$S_n = S_{\mathrm{seismic}} + S_{\mathrm{thermal}} + S_{\mathrm{shot}} + S_{\mathrm{gas}} + \cdots, \qquad (6.2)$$

including contributions due to seismic noise, thermal noise, photon shot noise, and residual gas noise, among others. By contrast, the typical value S_n' of the off-diagonal elements of the spectral density matrix will be approximately a sum of terms that are due to sources of noise which are strongly correlated between different interferometers at the same site:

$$S_n' \lesssim S_{\mathrm{seismic}} + S_{\mathrm{gas}}. \qquad (6.3)$$

Since $S_{\mathrm{seismic}} + S_{\mathrm{gas}} \ll S_n$ when $f \gtrsim 10\,\mathrm{Hz}$ for the advanced detectors of Ref. [1], we find from Eq. (3.13) that to a good approximation for LIGO,

$$S_n^{(\mathrm{eff})}(f)^{-1} \approx \sum_a S_n(f)_{aa}^{-1}. \qquad (6.4)$$

Here the sum is over the different detectors, and $a = 1$ corresponds to the primary, broadband interferometer with noise curve (6.1). Now the second and/or third detectors at each site will most likely be specialized ones such as dual-recycled detectors [29], which have high sensitivity only in some narrow frequency band. Thus in the relevant frequency band of $20\,\mathrm{Hz} \lesssim f \lesssim 70\,\mathrm{Hz}$ [cf. Fig. 9 below], it is likely that $S_n^{(\mathrm{eff})}(f) \approx S_n(f)_{11}$. If, instead, three identical broadband detectors are operated at each site (which is unlikely initially), then $S_n^{(\mathrm{eff})}(f) \approx S_n(f)_{11}/3$, and the 90% confidence upper limit estimated below should be divided by 3.

If we now make the plausible assumption that the SB has approximately constant $\Omega_g(f)$ over the above wave band, then by inserting Eq. (6.1) into Eq. (3.12) we can calculate the 90% confidence upper limit that we can place on Ω_g. Using the planned LIGO positions and orientations we thus obtain

$$\begin{aligned} {}^{(B)}\Omega_g^{(90\%)} &= 1.65\,\frac{5\pi}{4\rho_c}\left[2\hat{\tau}\int_0^\infty df\,\frac{\gamma(f)^2}{f^6 S_n^{(\mathrm{eff})}(f)^2}\right]^{-1/2} \\ &= 5.1\times 10^{-10}\,h_{75}^{-2}\,N\left(\frac{\hat{\tau}}{10^7\,\mathrm{s}}\right)^{-1/2}, \qquad (6.5) \end{aligned}$$

where h_{75} is the Hubble constant scaled to the value 75 $\mathrm{km\,sec^{-1}\,Mpc^{-1}}$, $\hat{\tau}$ is the observation time, the prefix (B) indicates broadband, and $N = S_m/10^{-48}\ \mathrm{Hz}^{-1}$. Note that this is the upper bound obtained from intersite cross correlations, since as argued in Sec. III B, it will not help much to include the information from intrasite correlations.

This bound (6.5) is a reasonably conservative estimate of what the LIGO sensitivity might be when the advanced detectors are operating. However the ultimate noise levels are largely unknown. This is because the experimental techniques and technology, which are presently somewhat far from the advanced level, will be steadily improving. To obtain an upper limit to ${}^{(B)}\Omega_g^{(90\%)}$ we take $S_n^{(\mathrm{eff})}(f) \approx S_{\mathrm{gas}}(f)$, which is a source of correlated noise that cannot be changed once the beam tube is built. Using $S_{\mathrm{gas}} \approx 2.2\times 10^{-50}\,\mathrm{Hz}^{-1}$ [28] gives as a rather firm upper limit for the LIGO sensitivity

$${}^{(B)}\Omega_g^{(90\%)} \geq 2\times 10^{-13} N h_{75}^{-2}\left(\frac{\hat{\tau}}{10^7\,\mathrm{sec}}\right)^{-1/2}. \qquad (6.6)$$

We now consider how ${}^{(B)}\Omega_g^{(90\%)}$ would change if we were to vary the angles β, δ, and Δ. The integral (5.6) is positive for $\beta = 27.2°$ with the noise power spectral density (6.1). Hence the optimal configuration of the detectors is configuration II, i.e., at $\delta = \Delta = 0°$ or $45°$. However LIGO already has $\delta \approx 45°$ (corresponding to parallel detectors because $\sigma_1 - \sigma_2 = 2\delta \approx 90°$). Also, as we have discussed in Sec. IV, the dependence of $\gamma(f)$ on Δ is very weak for $f \lesssim 100\,\mathrm{Hz}$ and at $\beta = 27°$, see Eq. (4.1) and Fig. 6. Because of these two facts the expected LIGO sensitivity at its planned configuration of $\delta = 44.9°$ and $\Delta = 28.2°$ is only $\sim 3\%$ less than the optimal sensitivity which would be attained at $\delta = \Delta = 45°$. This follows from Eq. (5.5) where we find that $B/A = 0.045$ and $C/A = 7.8\times 10^{-3}$.

The European VIRGO and GEO detectors, however, are currently planned to have $2\delta = \sigma_1 - \sigma_2 \approx 45°$. Such a configuration will allow VIRGO and GEO to extract information from both polarization components of burst and periodic gravitational waves, but will severely reduce the sensitivity to the SB. For these detectors $\beta = 9.75°$, and we find from Eq. (5.6) that $B/A = 4.0 \times 10^{-4}$ and $C/A = 1.7 \times 10^{-6}$. In Fig. 7 we plot the broadband signal-to-noise ratio (5.5) as a function of δ (assuming that Δ is optimally chosen); it can be seen that at $\delta = 45°$, the sensitivity to the SB is reduced by about 3 orders of magnitude. An intermediate choice of $\delta \sim 40°$ would give a broadband sensitivity to the SB comparable to that of LIGO, while still allowing the detectors to access two polarization components of the gravitational-wave field that are largely independent.

The total rotation angle Δ is unimportant for the broadband sensitivity to the SB in the case of detectors which are sufficiently close to each other [but *not* unimportant for the narrow-band sensitivity, cf. Eq. (5.4) above]. However, for detectors which are on separate continents, Δ will become important. For example in Fig. 8 we show the functions $\Theta_1(f)$ and $\Theta_2(f)$ for $\beta = 79.5°$, the value appropriate for one of the two LIGO-VIRGO cross correlations. Since Θ_2 is typically large compared to Θ_1 in this case, varying Δ will have a large effect on the sensitivity. This is confirmed by Eq. (5.5), where we find $B/A = -0.22$ and $C/A = 18.3$.

Next we consider how the broadband sensitivity to the SB varies with the angle β, assuming that the orientations are optimally chosen. Recall (cf. Sec. V) that this means either configuration I or II is chosen. In Fig. 1 we plot the SNR (5.5) as a function of β, for both configurations I and II, again assuming the noise curve (6.1). It can be seen that the sensitivity falls off rapidly when the detectors become more than a few thousand kilometers apart. The sensitivity for the LIGO separation is already

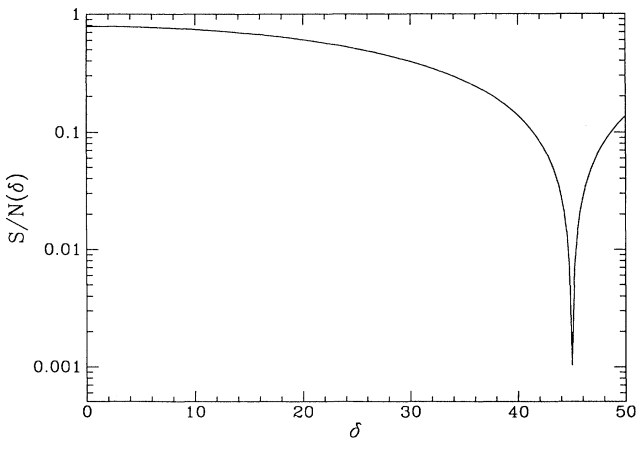

FIG. 7. The broadband signal-to-noise ratio for the VIRGO-GEO detector pair as a function of the relative rotation angle δ, assuming (i) that the total rotation angle Δ is optimally chosen, and (ii) the advanced detector noise curve of Ref. [1]. The normalization is to unity for coincident, aligned detectors.

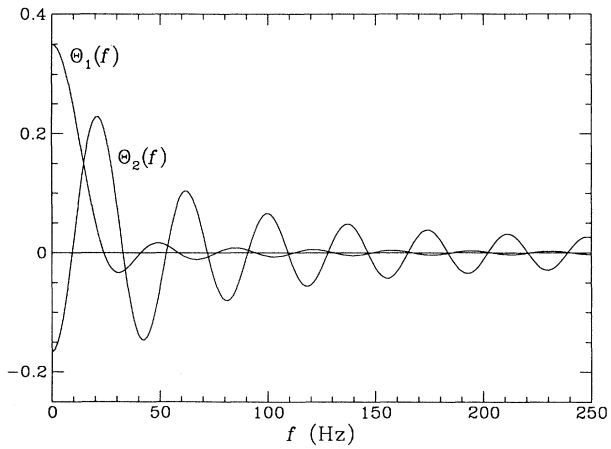

FIG. 8. The functions $\Theta_1(f)$ and $\Theta_2(f)$ appropriate for the LIGO detector in Hanford, Washington together with the VIRGO detector in Pisa, Italy. The fact that $\Theta_2(f)$ is large compared to $\Theta_1(f)$ for most frequencies implies that the sensitivity of the detector pair to the SB depends strongly in this case on the total rotation angle Δ of the two detectors with respect to the line joining them; see text.

a factor of ~ 4 worse than that for coincident detectors. Hence if there is ultimately a worldwide network of detectors in America, Europe, Japan, and Australia, only the cross correlations between proximate detector pairs like LIGO-LIGO or VIRGO-GEO will be important.

The various conclusions that we have drawn are relatively insensitive to the the detailed properties of the detector noise curve (6.1) that we have assumed (except for the value of the minimum frequency). This is because the

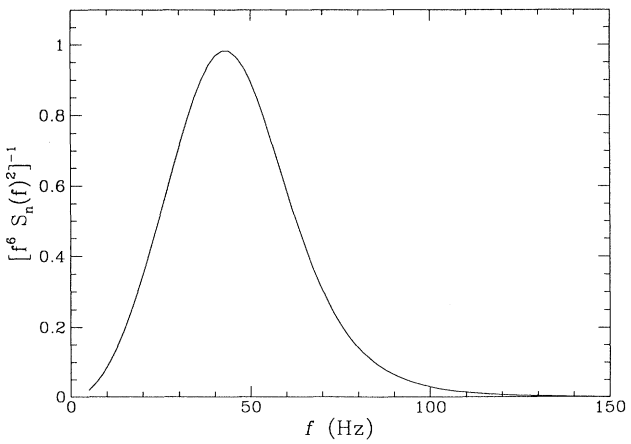

FIG. 9. A plot of the quantity $f^{-6} S_n(f)^{-2}$ with arbitrary normalization as a function of frequency. It is this quantity that must be integrated against the square of the overlap reduction function to determine the broadband response of a detector pair to the stochastic background, when one assumes a constant, Harrison-Zel'dovich spectrum; see Eq. (1.2). It is clear that the highest sensitivity is limited to a narrow band of frequency between ~ 20 Hz and ~ 70 Hz.

integral (1.2) is dominated by contributions in the narrow frequency band $20\,\mathrm{Hz} \lesssim f \lesssim 70\,\mathrm{Hz}$; see Fig. 9. At these frequencies thermal noise dominates over photon shot noise [1], so using sophisticated optical configurations of the interferometer such as dual recycling [29] to reduce the shot noise will not help much. It thus will be important for the detectors to have good low frequency sensitivity for the purposes of placing upper limits on the SB.

VII. CONCLUSION

To place upper bounds on the strength of the gravitational stochastic background, or perhaps to detect it, is but one of the aims of LIGO and other detector systems. Moreover, this background may well be very weak compared to waves from astrophysical sources; consider by analogy the weakness of the electromagnetic stochastic background in the visible region of the spectrum. However, these waves if detected would be amongst the most interesting that the detectors would see.

The sensitivity to these waves is therefore just one of various factors that need to be taken into account when choosing the fixed orientations of interferometric detectors. Nevertheless, in order to be able to make a wise choice, it is important to know the effect for the sensitivity of choosing this or that orientation. One of the main results of this paper is the simple expression (B5) describing this dependence. In particular, for frequencies such that the phase lag between the detectors is $\lesssim 1$, the sensitivity is determined solely by the overlap of the polarization tensors of the two detectors. At higher frequencies the direction of the vector which joins the detectors also becomes important.

We also determine the orientation of a pair of interferometers that will optimize the sensitivity of the pair to the SB, and show that the orientations which have been chosen for the LIGO detectors are close to optimal.

ACKNOWLEDGMENTS

The author is grateful to Kip Thorne for an introduction to the issue of detector orientations, for many helpful discussions and comments on the presentation, and also for the foresight of guessing, roughly, what the optimal filtering method and implications for LIGO and VIRGO might be, which to a large extent motivated and guided this paper's analysis. Thanks are also due to Eric Poisson for some helpful and detailed comments. This work was supported in part by NSF Grant No. PHY-9213508.

APPENDIX A: STATISTICAL FOUNDATIONS

1. Overview

In this appendix we justify the method of optimal filtering described in Sec. III and the resulting signal-to-noise ratio (3.12). We do this by deriving the probability distribution for the spectrum Ω_g given the detectors' output $\mathbf{h}(t)$, $p[\Omega_g|\mathbf{h}(t)]$.

First we transform from a continuous to a discrete description of the measured data. In the body of this paper we have treated the output of the detector network as a continuous vector random process $\mathbf{h}(t)$. However, a real, discrete measurement will be of a finite duration $\hat{\tau}$, and will contain frequencies only up to some maximum frequency determined by the time resolution of the sampling. In other words the output of the detectors will consist of the numbers $h_{aj} = h_a(t_j = t_{\mathrm{start}} + j\Delta)$ for $1 \le a \le n_d$ and $0 \le j \le N$, where n_d is the number of detectors, t_{start} is the starting time, and Δ is the time resolution, of the order of 10^{-4} sec. The number of samples per interferometer is $N = \hat{\tau}/\Delta$. We denote by $\mathbf{X}$ the vector of numbers h_{aj}, which we assume to have a multivariate Gaussian distribution with some variance-covariance matrix $\mathbf{\Sigma}'$. This matrix is essentially the correlation matrix (2.2):

$$\Sigma'_{ai\,bj} = C_h(t_i - t_j)_{ab}. \tag{A1}$$

However if we take a finite Fourier transform, which amounts to making a change of basis in the space of vectors $\mathbf{X}$, then $\mathbf{\Sigma}'$ corresponds instead to the spectral density matrix (2.3).

Now from Eq. (2.6), we have $\mathbf{\Sigma}' = \mathbf{\Sigma}'_n + \mathbf{\Sigma}'_s(\mathbf{\Omega})$, where the contribution $\mathbf{\Sigma}'_s$ from the SB depends on the spectrum $\Omega_g(f)$ as in Eq. (2.12). [Because of the finite frequency resolution of order $\sim 1/\hat{\tau}$, we represent the function $\Omega_g(f)$ as a finite vector $\mathbf{\Omega}$.] What we want to do is extract information about $\mathbf{\Omega}$ from a measurement of $\mathbf{X}$.

There are two general approaches to the task of quantifying the information obtained, in situations of this sort. First, one can in principle compute the probability distribution function (PDF) for $\mathbf{\Omega}$ given the measurement $\mathbf{X}$, $p(\mathbf{\Omega}|\mathbf{X})$. This PDF then contains complete information about the measurement. However, the calculation of the PDF $p(\mathbf{\Omega}|\mathbf{X})$ is frequently difficult in practice; and so one has to resort to instead calculating estimators (statistics) $\hat{\mathbf{\Omega}}(\mathbf{X})$ which are functions of $\mathbf{X}$, chosen so that the PDF for their values given some value of $\mathbf{\Omega}$, $p(\hat{\mathbf{\Omega}}|\mathbf{\Omega})$, is peaked near $\hat{\mathbf{\Omega}} = \mathbf{\Omega}$. There are standard criteria for choosing such estimators, see, for example, Refs. [30, 31].

Now suppose that, instead of one measurement of $\mathbf{X}$, we have n measurements $\mathbf{X}_1, \ldots, \mathbf{X}_n$. A standard statistical result is that in the Cramer-Frechet-Rao limit of $n \to \infty$ [31, 32], the two approaches discussed above yield the same unique result. More precisely, the PDF $p(\mathbf{\Omega}|\mathbf{X}_j)$ becomes a Gaussian centered at $\hat{\mathbf{\Omega}}_{\mathrm{ML}}(\mathbf{X}_j)$, where $\hat{\mathbf{\Omega}}_{\mathrm{ML}}$ is the so-called maximum likelihood estimator of $\mathbf{\Omega}$. The variance-covariance matrix $\mathbf{\Sigma}_\Omega$ of this probability distribution depends on the $\mathbf{X}_j$ only through $\hat{\mathbf{\Omega}}_{\mathrm{ML}}(\mathbf{X}_j)$, $\mathbf{\Sigma}_\Omega = \mathbf{\Sigma}_\Omega(\hat{\mathbf{\Omega}}_{\mathrm{ML}})$. Conversely, the PDF $p(\hat{\mathbf{\Omega}}_{\mathrm{ML}}|\mathbf{\Omega})$ of the estimator $\hat{\mathbf{\Omega}}_{\mathrm{ML}}$ given some value of $\mathbf{\Omega}$ becomes a Gaussian centered at $\mathbf{\Omega}$ with width $\mathbf{\Sigma}_\Omega(\mathbf{\Omega})$. Thus the two probability distributions, which are conceptually very different objects, become effectively the same, and all one really needs to calculate is the variance-covariance matrix $\mathbf{\Sigma}_\Omega(\mathbf{\Omega})$.

However, this simplifying Cramer-Frechet-Rao (CFR) limit does not apply in a straightforward manner to the

calculation of $p(\mathbf{\Omega}|\mathbf{X})$ for measurements of the SB; there are a number of subtleties and differences from the usual situations discussed in Refs. [31, 32]. First, we have only one measurement of $\mathbf{X}$ instead of a large number n of measurements. Nevertheless, something like a "large number of measurements limit" does apply, which we discuss further below. Second, one needs to address the issue of distinguishing between the contributions to $\mathbf{\Sigma}'$ from the detector noise and from the SB, since both are unknown *a priori*. Third, the CFR limit only applies to the extent that our *a priori* knowledge of the variables being measured is unimportant. However, our *a priori* knowledge about the correlated detector noise in intrasite correlations *is* important in the calculation of $p(\mathbf{\Omega}|\mathbf{X})$, and consequently this PDF is not approximately Gaussian [cf. Eqs. (A21) and (A33) below].

To resolve these complications, we now derive from first principles an approximate expression for $p(\mathbf{\Omega}|\mathbf{X})$. We also identify the conditions under which the approximations we make are valid, and show that they will all be satisfied in the LIGO and/or VIRGO context. We assume throughout that the effect of the SB on the detectors is small compared to the detectors intrinsic noise, so that from Eq. (3.1), $\Omega_g(f) \ll 10^{-6}$.

2. Calculation of $p(\mathbf{\Omega}|\mathbf{X})$

We start by considering the *frequency resolution* of the measurement of $\Omega_g(f)$. In a finite measurement of length $\hat{\tau}$ there are roughly $N = \hat{\tau}/\Delta$ independent frequencies f_j, and we get what amounts to one measurement of each $\Omega_j = \Omega_g(f_j)$. Now if $\hat{\tau}$ is doubled, we double the number of measurements, but we also double the number of variables measured. Clearly, the Cramer-Frechet-Rao limit of repeated measurements of the same variables cannot be attained in this context without further assumptions. Below we shall argue that in the LIGO and/or VIRGO context, to a good approximation,

$$\mathbf{C}_h(\tau) \approx 0 \quad \text{for } \tau > \tau_c, \tag{A2}$$

for some correlation time $\tau_c \ll \hat{\tau}$ [33]. In this case the unknown quantities to be measured are $\mathbf{C}_h(j\Delta)$ for $0 \le j \le n = \tau_c/\Delta$, or equivalently $\mathbf{S}_h(f)$ at n different frequencies; and the number of measurements of each of these variables is $\sim N/n = \hat{\tau}/\tau_c \gg 1$.

In order to explain Eq. (A2), we fix a value of τ_c, and decompose $\mathbf{C}_h(\tau)$ into

$$\mathbf{C}_h(\tau) = \mathbf{C}_h(\tau)\Theta(\tau_c - |\tau|) + \mathbf{C}_h(\tau)\Theta(|\tau| - \tau_c), \tag{A3}$$

where Θ is the step function. Let $\mathbf{S}_h = \mathbf{S}_h^< + \mathbf{S}_h^>$ be the corresponding decomposition in frequency space, so that $\mathbf{S}_h^<(f)$ is $\mathbf{S}_h(f)$ averaged over frequency scales of the order of $1/\tau_c$. Then from Eq. (2.6) there will be two contributions to $\mathbf{S}_h^>(f)$:

$$\mathbf{S}_h^>(f) = \mathbf{S}_s^>(f) + \mathbf{S}_n^>(f). \tag{A4}$$

The first term here will be small if $\Omega_g(f)$ is smooth over frequency scales $\sim 1/\tau_c$ [cf. Eq. (2.12)], which will be the case for currently conceived of SB sources when $\tau_c = 100\,\text{sec}$, for example. The second term in Eq. (A4) will also be small for this value of τ_c, except near isolated frequencies corresponding to high-Q resonances [34], the effect of which can be neglected.

We now calculate the PDF for $\mathbf{\Omega}$ using Baye's rule; by virtue of the condition (A2), something like the CFR limit will apply. Let $\hat{\mathbf{\Sigma}}' = \mathbf{X} \otimes \mathbf{X}$, which is the maximum likelihood estimator of $\mathbf{\Sigma}'$. This is an $Nn_d \times Nn_d$ matrix, where n_d is the number of detectors. Similarly the matrices $\mathbf{\Sigma}_n'$ and $\mathbf{\Sigma}_s'$ are of dimension Nn_d, and are given by formulas analogous to (A1). However because of condition (A2) the independent variables in $\mathbf{\Sigma}'$ can be formed into a smaller matrix $\mathbf{\Sigma}$, which is given by

$$\Sigma_{ai\,bj} = C_h[(i-j)\Delta]_{ab} \tag{A5}$$

for $0 \le i, j \le n = \tau_c/\Delta$. Note that since $\mathbf{C}_h(-\tau) = \mathbf{C}_h(\tau)^T$, the matrix $\mathbf{\Sigma}$ contains only $n\,n_d^2$ independent variables. We define the matrices $\mathbf{\Sigma}_n$, $\mathbf{\Sigma}_s$, and $\hat{\mathbf{\Sigma}}$ analogously in terms of $\mathbf{C}_n(\tau)$, $\mathbf{C}_s(\tau)$, and the estimator

$$\hat{C}_h(\tau)_{ab} \equiv \frac{1}{\hat{\tau}} \int_{t_{start}}^{t_{start}+\hat{\tau}} dt\, h_a(t+\tau) h_b(t), \tag{A6}$$

which is defined for $|\tau| < \tau_c$. The estimator $\hat{\mathbf{C}}_h$ is the quantity that will be measured; it is a sufficient statistic for $\mathbf{\Omega}$, and moreover its discrete counterpart $\hat{\mathbf{\Sigma}}(\mathbf{X})$ is the maximum likelihood estimator of $\mathbf{\Sigma}$.

The joint PDF for $\mathbf{\Sigma}_n$ and $\mathbf{\Sigma}_s$ given $\mathbf{X}$ is

$$p(\mathbf{\Sigma}_n, \mathbf{\Sigma}_s|\mathbf{X}) \propto p_n^{(0)}(\mathbf{\Sigma}_n)\, p_s^{(0)}(\mathbf{\Sigma}_s) \exp\left[-\frac{1}{2}\Lambda'(\mathbf{\Sigma}')\right], \tag{A7}$$

where

$$\Lambda'(\mathbf{\Sigma}') = \ln\det \mathbf{\Sigma}' + \hat{\mathbf{\Sigma}}' : \mathbf{\Sigma}'^{-1}, \tag{A8}$$

and $p_n^{(0)}$ and $p_s^{(0)}$ are the PDF's that represent our *a priori* knowledge. From the distribution (A7) we obtain the PDF $p(\mathbf{\Omega}|\mathbf{X})$ for $\mathbf{\Omega}$ given $\mathbf{X}$ by integrating over $\mathbf{\Sigma}_n$:

$$p(\mathbf{\Omega}|\mathbf{X}) = \mathcal{N} \int d\mathbf{\Sigma}_n\, p(\mathbf{\Sigma}_n, \mathbf{\Sigma}_s(\mathbf{\Omega})|\mathbf{X}), \tag{A9}$$

where $\mathcal{N}$ is a normalization constant.

The above formulas involve primed, $Nn_d \times Nn_d$ matrices. We now express them in terms of the corresponding, unprimed, $nn_d \times nn_d$ matrices, by using the fact that all the matrices will be approximately diagonal in frequency space. That is, they will be approximately diagonal in the indices i, j after a finite Fourier transform change of basis, but not in the indices a, b. For example, on a suitable basis,

$$(\Sigma_n)_{aI,bJ} \approx \delta_{IJ}\, S_n(f_I)_{ab}, \tag{A10}$$

where $f_I = I/\tau_c$ for $1 \le I \le n$. It is straightforward to verify using Eqs. (A2) and (A6) that [35]

$$\Lambda'(\mathbf{\Sigma}') \approx k\,\Lambda(\mathbf{\Sigma}), \tag{A11}$$

where $k = \hat{\tau}/\tau_c = N/n$ is the effective number of measurements, and

$$\Lambda(\mathbf{\Sigma}) = \ln\det\mathbf{\Sigma} + \hat{\mathbf{\Sigma}} : \mathbf{\Sigma}^{-1}. \tag{A12}$$

For example, the second term in $\Lambda'(\mathbf{\Sigma}')$ can be written as

$$\begin{aligned}\hat{\mathbf{\Sigma}}' : \mathbf{\Sigma}'^{-1} &= \left(\Sigma'^{-1}\right)^{\alpha\beta} X_\alpha X_\beta \\ &= 2\int_{-\infty}^{\infty} df\ \tilde{\mathbf{h}}(f)^\dagger \cdot \mathbf{S}_h(f)^{-1} \cdot \tilde{\mathbf{h}}(f) \\ &= \hat{\tau}\int_{-\infty}^{\infty} df\ \mathrm{tr}\left[\mathbf{S}_h(f)^{-1}\cdot\hat{\mathbf{S}}_h(f)\right]. \end{aligned}\tag{A13}$$

Here we have used Eq. (2.5), switched into the time domain, used Eq. (A6), and switched back [35]. The result is just $k\hat{\mathbf{\Sigma}} : \mathbf{\Sigma}^{-1}$, and the first term in Eq. (A8) transforms analogously.

Now Eqs. (A7), (A9), (A11), and (A12) together yield a formal expression for $p(\mathbf{\Omega}|\mathbf{X})$. To proceed further we need to invoke some approximations. The first of these is the quadratic approximation: we expand the likelihood function (A12) to give

$$\begin{aligned}\Lambda(\mathbf{\Sigma}) = \ln\det\hat{\mathbf{\Sigma}} + n\,n_d \\ +\frac{1}{2}\mathrm{tr}\left[\hat{\mathbf{\Sigma}}^{-1}\cdot\delta\mathbf{\Sigma}\cdot\hat{\mathbf{\Sigma}}^{-1}\cdot\delta\mathbf{\Sigma}\right] + O(\delta\mathbf{\Sigma}^3),\end{aligned}\tag{A14}$$

where $\delta\mathbf{\Sigma} = \mathbf{\Sigma} - \hat{\mathbf{\Sigma}} = \mathbf{\Sigma}_n + \mathbf{\Sigma}_s(\mathbf{\Omega}) - \hat{\mathbf{\Sigma}}$. Clearly this approximation will be good only for certain values of $\mathbf{\Sigma}_n$ and of $\mathbf{\Omega}$; we discuss below the implications of this restriction. First we derive the conditions under which most of the probability of the PDF,

$$p(\mathbf{\Sigma}|\mathbf{X}) \propto \exp\left[-\frac{k}{2}\Lambda(\mathbf{\Sigma})\right], \tag{A15}$$

will be concentrated in the region $\mathcal{Q}$ where the quadratic approximation (A14) is valid, so that the normalization of the PDF (A15) can be calculated correctly using Eq. (A14). If we let $\lambda_1,\ldots,\lambda_n$ be the eigenvalues of $\mathbf{\Sigma}\cdot\hat{\mathbf{\Sigma}}^{-1} - \mathbf{1}$, then, up to an additive constant,

$$\Lambda(\mathbf{\Sigma}) = -\sum_j\left[\ln(1+\lambda_j) + \frac{1}{1+\lambda_j}\right]. \tag{A16}$$

From this formula it can be seen that the quadratic regime $\mathcal{Q}$ is given by $\max_j|\lambda_j| < \varepsilon$ where ε is some small number. If we demand that the total probability in $\mathcal{Q}$ be $1-\delta$ for some $\delta \ll 1$, and assume that the *a priori* distribution $p_n^{(0)}$ is not sharply peaked, then we find the condition

$$N \gtrsim \frac{4}{\varepsilon^2}\left(\ln n + |\ln\delta|\right). \tag{A17}$$

This condition will be just barely satisfied for, e.g, $\varepsilon = \delta = 0.01$, $\hat{\tau} = 10^7\,\mathrm{sec}$, $\tau_c = 100\,\mathrm{sec}$, and $\Delta = 10^{-4}\,\mathrm{sec}$, which are values that are appropriate for LIGO and/or VIRGO. Now if $\mathbf{\Omega} \ll 10^{-6}$, then the intgeral over $\mathbf{\Sigma}_n$ in Eq. (A9) will be dominated by contributions from the region $\mathcal{Q}$, and so the approximation is valid. For $\mathbf{\Omega} \gtrsim 10^{-6}$, the exact probability $p(\mathbf{\Omega}\,|\,\mathbf{X})$ is very small, and so the quadratic approximation will still give a qualitatively correct result. In particular, the normalization of the PDF $p(\mathbf{\Omega}\,|\,\mathbf{X})$ resulting from Eq. (A14) will be approximately correct.

The next approximation involves consideration of the relative magnitudes of various components of the measured autocorrelation matrix $\hat{\mathbf{\Sigma}}$ [cf. Eq. (A6) above]. We introduce the following notation: for any matrix $\mathbf{A}$, $\mathbf{A}^{\|}$ denotes the matrix consisting of the diagonal subblocks (in the indices a,b) of $\mathbf{A}$ corresponding to intrasite correlations, and $\mathbf{A}^{\perp} = \mathbf{A} - \mathbf{A}^{\|}$ consists of the off-diagonal subblocks. We also define $\mathbf{A}^D$ to be the matrix of diagonal elements (in the indices a,b) of $\mathbf{A}$ and $\mathbf{A}^O = \mathbf{A}^{\|} - \mathbf{A}^D$. Thus, the estimator $\hat{\mathbf{\Sigma}}$ decomposes into $\hat{\mathbf{\Sigma}}^D + \hat{\mathbf{\Sigma}}^O + \hat{\mathbf{\Sigma}}^{\perp}$, where $\hat{\mathbf{\Sigma}}^D$ contains the detector noises $\hat{\mathbf{C}}_h(\tau)_{aa}$, $\hat{\mathbf{\Sigma}}^O$ contains the measured intrasite correlations, and $\hat{\mathbf{\Sigma}}^{\perp}$ the measured intersite correlations. Now we expect the contribution of correlated detector noise to $\hat{\mathbf{\Sigma}}^{\perp}$ to be very small, cf. the discussion in Sec. III C above. Hence from Eq. (3.1), $\hat{\mathbf{\Sigma}}^{\perp} \sim \varepsilon\,\hat{\mathbf{\Sigma}}^D$, where

$$\varepsilon \approx \frac{\Omega_g}{10^{-6}}. \tag{A18}$$

For example, if the SB is just barely detectable, then $\varepsilon \sim 10^{-4}$. We similarly define in order of magnitude the small parameter $\tilde{\varepsilon}$ by

$$\hat{\mathbf{\Sigma}}^O \sim \tilde{\varepsilon}\,\hat{\mathbf{\Sigma}}^D, \tag{A19}$$

which we expect to be of the order of $S_{\mathrm{gas}}(f)/S_n(f) \sim 10^{-2}$ or smaller if $\Omega_g \lesssim 10^{-8}$; see Eq. (6.3).

In the expression (A14) for the likelihood function, to leading order in ε we can replace the factors of $\hat{\mathbf{\Sigma}}^{-1}$ by $(\hat{\mathbf{\Sigma}}^{\|})^{-1}$, so that the cross terms between $\delta\mathbf{\Sigma}^{\perp}$ and $\delta\mathbf{\Sigma}^{\|}$ vanish. Note that to this order $(\hat{\mathbf{\Sigma}}^{-1})^{\|} = (\hat{\mathbf{\Sigma}}^{\|})^{-1}$. Hence, the PDF (A7) splits into a product of two factors which incorporate the measured intersite and the intrasite correlations. The *a priori* PDF for the detector noise in Eq. (A7) can be written as

$$p_n^{(0)}(\mathbf{\Sigma}_n) \approx \delta(\mathbf{\Sigma}_n^{\perp})\,p_{n,\|}^{(0)}(\mathbf{\Sigma}_n^{\|}), \tag{A20}$$

since we expect $\mathbf{\Sigma}_n^{\perp} \approx 0$. Using Eqs. (A7), (A9), (A11), (A14), and (A20) we obtain

$$p(\mathbf{\Omega}|\mathbf{X}) \approx \mathcal{N}_1\,p_s^{(0)}(\mathbf{\Omega})\,p_C(\mathbf{\Omega})\exp\mathrm{tr}\left\{-\frac{k}{4}\left[(\hat{\mathbf{\Sigma}}^{\|})^{-1}\cdot\left(\mathbf{\Sigma}_s(\mathbf{\Omega})^{\perp} - \hat{\mathbf{\Sigma}}^{\perp}\right)\right]^2\right\}, \tag{A21}$$

where $\mathcal{N}_1$ is a normalization constant. Here

$$\begin{aligned}p_C(\mathbf{\Omega}) \equiv \int d\mathbf{\Sigma}_n^{\|}\ p_{n,\|}^{(0)}(\mathbf{\Sigma}_n^{\|}) \\ \times\exp\mathrm{tr}\left\{-\frac{k}{4}\left[(\hat{\mathbf{\Sigma}}^{\|})^{-1}\cdot\delta\mathbf{\Sigma}^{\|}\right]^2\right\}\end{aligned}\tag{A22}$$

is a function representing the information from intrasite correlations, and $\delta\mathbf{\Sigma}^{\|} = \mathbf{\Sigma}_n^{\|} + \mathbf{\Sigma}_s(\mathbf{\Omega})^{\|} - \hat{\mathbf{\Sigma}}^{\|}$.

Now if $\tilde{\varepsilon} \gg \varepsilon$ as estimated above, then most of the information we obtain from the SB will come from the intersite correlation measurements. In this case the factor

$p_C(\mathbf{\Omega})$ in Eq. (A21) can be approximated to be constant [see Eq. (A33) below]. We discuss the implications of the resulting PDF for data processing in Sec. A 4. However it may turn out that $\tilde{\varepsilon} \lesssim \varepsilon$, which from Eqs. (A18) and (A19) will be the case if $\Omega_g \gtrsim 10^{-8}$. Even if $\Omega_g \ll 10^{-8}$, it may happen that the correlation coefficient between different detectors for residual gas noise in the beam tube will be small compared to unity, so that in the notation of Eq. (6.3), $S_n' \ll S_{\rm gas}$ and again $\tilde{\varepsilon} \lesssim \epsilon$. Alternatively, it may be possible to detect and take account of bursts of outgassing from the beam tube walls. More detailed discussions of the possible magnitude of intrasite correlated noise can be found in Ref. [17]. Because of the possibility that $\tilde{\varepsilon} \lesssim \varepsilon$, we now derive an approximate expression for $p_C(\mathbf{\Omega})$.

3. Information from intrasite correlations

To evaluate the expression (A22) we first make a change of variables. Let $\hat{\mathcal{C}}$ be the matrix of correlation coefficients

$$\hat{\mathcal{C}} = (\hat{\mathbf{\Sigma}}^D)^{-1/2} \cdot \hat{\mathbf{\Sigma}}^O \cdot (\hat{\mathbf{\Sigma}}^D)^{-1/2}, \tag{A23}$$

and similarly define matrices $\mathbf{N}$ and $\mathbf{G}$ by replacing $\hat{\mathbf{\Sigma}}^O$ in Eq. (A23) by $\mathbf{\Sigma}_n^D$ and $\mathbf{\Sigma}_s(\mathbf{\Omega})^{\|}$, respectively. Also let

$$\mathcal{C} = (\mathbf{\Sigma}_n^D)^{-1/2} \cdot \mathbf{\Sigma}_n^O \cdot (\mathbf{\Sigma}_n^D)^{-1/2}. \tag{A24}$$

We assume that the volume element in Eq. (A22) can be written as

$$p_{n,\|}^{(0)}(\mathbf{\Sigma}_n^{\|})\, d\mathbf{\Sigma}_n^{\|} \propto p_{\rm noise}(\mathbf{\Sigma}_n^D)\, p_{\rm corr}(\mathcal{C})\, d\mathbf{N}\, d\mathcal{C}. \tag{A25}$$

To leading order in $\tilde{\varepsilon}$ the cross term between $\delta\mathbf{\Sigma}^D$ and $\delta\mathbf{\Sigma}^O$ in the argument of the exponential in Eq. (A22) vanishes, and so it becomes proportional to

$$\mathrm{tr}\left[\left(\mathbf{N}+\mathbf{G}^D-\mathbf{1}\right)^2 + \left(\mathbf{N}^{1/2}\cdot\mathcal{C}\cdot\mathbf{N}^{1/2}+\mathbf{G}^O-\hat{\mathcal{C}}\right)^2\right]. \tag{A26}$$

Using the relations $k \gg 1$, $\hat{\mathcal{C}} \sim \tilde{\varepsilon} \ll 1$, $\mathbf{G} \sim \Omega_g/10^{-6} \ll 1$, and $|\mathcal{C}| \lesssim 1$, and assuming that the PDF $p_{\rm noise}(\mathbf{\Sigma}_n^D)$ is slowly varying, we can approximately carry out the integral over $\mathbf{N}$. From Eqs. (A22), (A25), and (A26), the result is

$$p_C(\mathbf{\Omega}) \approx \mathcal{N}_2 \int d\mathcal{C}\; p_{\rm corr}(\mathcal{C}) \times \exp\left(\mathrm{tr}\left\{-\frac{k}{4}\left[\mathcal{C}+\mathbf{G}^O-\hat{\mathcal{C}}\right]^2\right\}\right), \tag{A27}$$

where $\mathcal{N}_2$ is a constant. This expression is actually only a good approximation when $|\mathcal{C}| \ll 1$, but the integral (A27) is dominated by values of $\mathcal{C}$ close to $\hat{\mathcal{C}} - \mathbf{G}^O$ which is $\ll 1$.

Now the behavior of the function (A27) depends strongly on our *a priori* information about the correlation matrix $\mathcal{C}$. As in Eq. (A10), this matrix will be approximately diagonal on a frequency basis:

$$\mathcal{C}_{aI,bJ} \approx \delta_{IJ}\, \mathcal{C}_{abI}, \tag{A28}$$

where, from Eq. (A24),

$$\mathcal{C}_{abI} = \frac{S_n(f_I)^O_{ab}}{\sqrt{S_n(f_I)_{aa}\, S_n(f_I)_{bb}}}. \tag{A29}$$

Since $\mathbf{S}_h(f)$ is positive definite, we have $|\mathcal{C}_{abI}| < 1$. Moreover, the variables $\mathcal{C}_{abI}$ will be real whenever

$$\langle n_a(t+\tau)\, n_b(t)\rangle = \langle n_a(t)\, n_b(t+\tau)\rangle \tag{A30}$$

for all t and τ, and thus in particular they will be real if there are no sources of noise that preferentially excite one detector later than another one. If we assume that such sources of noise are negligible, then by inserting into Eq. (A27) the PDF,

$$p_{\rm corr}(\mathcal{C}) = \delta(\mathrm{Im}\,\mathcal{C})\, p'_{\rm corr}(\mathrm{Re}\,\mathcal{C}), \tag{A31}$$

we obtain an equation again of the form (A27), but where $\mathcal{C}$ and $\hat{\mathcal{C}}$ are now understood to be real. (A factor of $\exp[-k\,\mathrm{tr}\,(\mathrm{Im}\,\hat{\mathcal{C}})^2/4]$ is absorbed into $\mathcal{N}_2$.)

Next we assume that our *a priori* information about the correlation coefficients $\mathcal{C}_{abI}$ is of the form $\mathcal{C}_{\rm min} \le \mathcal{C}_{abI} \le \mathcal{C}_{\rm max}$, so that we can take [36]

$$p_{\rm corr}(\mathcal{C})d\mathcal{C} = \prod_{I=1}^{n}\prod_{a<b}\Theta(\mathcal{C}_{\rm max}-\mathcal{C}_{abI})\,\Theta(\mathcal{C}_{abI}-\mathcal{C}_{\rm min})\, d\mathcal{C}_{abI}, \tag{A32}$$

where Θ is the step function. Using Eqs. (A27) and (A32) and the replacement $\mathrm{tr} \to 2\sum_{I=1}^n$ [cf. Eq. (A13)], gives

$$p_C(\mathbf{\Omega}) \propto \prod_{I=1}^{n}\prod_{a<b}\left[\mathrm{erf}\left(\mathcal{C}_{\rm max}\sqrt{k}-\beta_{abI}\right) - \mathrm{erf}\left(\mathcal{C}_{\rm min}\sqrt{k}-\beta_{abI}\right)\right], \tag{A33}$$

where

$$\beta_{abI} = \sqrt{k}\left(\hat{\mathcal{C}}_{abI} - G^O_{abI}\right). \tag{A34}$$

If we define $\hat{\mathbf{S}}_h(f)$ to be twice the Fourier transform of the estimator (A6) as in Eq. (2.3), and put

$$\hat{\Omega}_{abI} = \frac{5\pi f^3}{4\rho_c}\,\hat{S}_h(f_I)^O_{ab}, \tag{A35}$$

then we find from Eq. (3.1) that

$$\beta_{abI} = \frac{4\rho_c}{5\pi f^3}\sqrt{\frac{k}{\hat{S}_h(f_I)_{aa}\hat{S}_h(f_I)_{bb}}}\left(\hat{\Omega}_{abI}-\Omega_I\right) \approx \frac{\sqrt{k}}{10^{-6}}\left(\hat{\Omega}_{abI}-\Omega_I\right), \tag{A36}$$

where $\Omega_I \equiv \Omega_g(f_I)$.

We now discuss the content of Eqs. (A33)–(A36). If we make no assumption about the amount of correlated noise present, and so take $\mathcal{C}_{\min} = -1$ and $\mathcal{C}_{\max} = 1$, then we see using $k \approx 10^5$ that $p_C(\mathbf{\Omega})$ is roughly constant for $0 \le \Omega_I \lesssim 10^{-6}$. Thus we obtain little information. More information about the correlated noise needs to be input in order to constrain the strength of the SB. One assumption which may be valid is that

$$S_n(f)_{ab} \ge 0, \tag{A37}$$

which can be enforced by setting $\mathcal{C}_{\min} = 0$ in Eq. (A32). Equation (A37) will hold if there are no sources of noise $\hat{n}(t)$ which contribute an amount $+\hat{n}(t)$ to the output of one detector, and an amount $-\hat{n}(t)$ to the output of another. As example of such a source of noise could be a mode of vibration of a suspension system that couples in a suitable way the vibrations of mirrors in two different interferometers. It may be a valid assumption that all such sources of noise will be negligible. In this case we can use in Eq. (A33) the values $\mathcal{C}_{\min} = 0$ and $\mathcal{C}_{\max} = 1$. (The value chosen for $\mathcal{C}_{\max}$ is unimportant as long as $\mathcal{C}_{\max} \gg 1/\sqrt{k}$.) Then from Eq. (A36) we see that $\beta_{abI} \ll \mathcal{C}_{\max}\sqrt{k}$ as $\hat{\Omega} \ll 10^{-6}$, and so, to a good approximation,

$$p_C(\mathbf{\Omega}) \propto \prod_{I=1}^{n} \prod_{a<b} \left[\frac{1}{2} - \mathrm{erf}(-\beta_{abI})\right]. \tag{A38}$$

Essentially this PDF gives an upper bound on each $\Omega_I = \Omega_g(f_I)$ of $\Omega_{I,\max} = \min_{a,b} \hat{\Omega}_{abI}$, which is of the order of $\sim 10^{-6}\tilde{\varepsilon}$ if $\Omega_g \lesssim 10^{-8}$. As already mentioned, this upper bound will be much worse than that obtained from intersite correlations, unless the dimensionless correlation coefficient $\tilde{\varepsilon}$ is $\lesssim 10^{-4}$. Moreover, the bound is only obtained by making the specific assumptions (A30) and (A37) about sources of correlated noise at one site. For these reasons, in the following subsection on data analysis we consider only intersite correlations and take $p_C(\mathbf{\Omega}) \approx$ const.

We now show in more detail that very little information about the SB is obtained if the assumptions (A30) and (A37) are dropped. As a simple model, we consider the opposite extreme of assuming equal probability for all relative phases in the contributions to the outputs of different detectors from correlated sources of noise. Thus, we take the *a priori* PDF to be of the form [36]

$$p_{\mathrm{corr}}(\mathcal{C})d\mathcal{C} = \prod_{I=1}^{n} \prod_{a<b} \times \exp\left[-\frac{|\mathcal{C}_{abI}|^2}{2\mathcal{C}_{\max}^2}\right] d(\mathrm{Re}\,\mathcal{C}_{abI})\, d(\mathrm{Im}\,\mathcal{C}_{abI}), \tag{A39}$$

where $\mathcal{C}_{\max}$ is the *a priori* maximum correlation. Then from Eq. (A27) we obtain

$$p_C(\mathbf{\Omega}) \propto \prod_{I=1}^{n} \prod_{a<b} \exp\left[-\frac{(\mathrm{Re}\,\beta_{abI})^2}{1 + k\mathcal{C}_{\max}^2}\right]. \tag{A40}$$

This function is of the form of Eq. (A42) below, where, from Eq. (A36),

$$\sigma_I \approx 10^{-6}\sqrt{\frac{1 + k\,\mathcal{C}_{\max}^2}{k}}. \tag{A41}$$

This is roughly the minimum detectable value of Ω_g that can be detected in a bandwidth of $\sim 1/\tau_c$. Thus, the function (A40) is qualitatively similar to the PDF (A42) obtained from intersite correlations, with the simple change that the minimum detectable value of Ω_g in any frequency band is increased by a factor of $\sqrt{1 + k\,\mathcal{C}_{\max}^2}$. If we take $\mathcal{C}_{\max}$ to be of order unity and so make no assumption about the amount of correlated noise present, then the upper bounds on Ω_g from intrasite correlations will be worse by a factor of $\sqrt{k} \sim 300$ than those obtained from intersite correlations. Only if $\mathcal{C}_{\max} \approx 1/\sqrt{k}$ will the two be comparable. However, because of the possibility of weak, unknown sources of noise, it is probably inappropriate to make such a strong assumption.

4. Implications for data processing

The distribution (A21) with $p_C = $ const is of the form

$$p(\Omega_I) = \mathcal{N}_3\, p_s^{(0)}(\Omega_I)\, \exp\left\{-\sum_{I=1}^{n} \frac{(\Omega_I - \hat{\Omega}_I)^2}{2\sigma_I^2}\right\}, \tag{A42}$$

where $\Omega_I = \Omega_g(f_I = I/\tau_c)$, because the matrices are all approximately diagonal in frequency. Using Eq. (2.12), we find that the argument of the exponential in Eq. (A42) becomes

$$-\frac{\hat{\tau}}{2}\int_0^\infty df\ \mathrm{tr}\left\{\left[\hat{\mathbf{S}}_h^{-1\parallel}\cdot\left(\Omega\bar{\gamma}^{\perp} - \hat{\mathbf{S}}_h^{\perp}\right)\right]^2\right\}, \tag{A43}$$

where $\bar{\gamma}_{ab} = 4\rho_c\gamma_{ab}/(5\pi f^3)$. Hence we find, using $\int_0^\infty df \to (1/\tau_c)\sum_I$, that

$$\frac{1}{\sigma_I^2} = k\,\mathrm{tr}\left[\left(\hat{\mathbf{S}}_h^{\parallel\,-1}(f_I)\cdot\bar{\gamma}\right)^2\right], \tag{A44}$$

and

$$\hat{\Omega}_I = k\,\sigma_I^2\,\mathrm{tr}\left[\hat{\mathbf{S}}_h^{\parallel\,-1}(f_I)\cdot\bar{\gamma}\cdot\hat{\mathbf{S}}_h^{\parallel\,-1}(f_I)\cdot\hat{\mathbf{S}}_h^{\perp}(f_I)\right]. \tag{A45}$$

One would like to summarize the information contained in Eq. (A42) by calculating some kind of signal-to-noise ratio. There are various, inequivalent ways of doing this. For example, one could calculate probability distributions for the quantities $\Omega_{\max} = \max_I \Omega_I$ or $\bar{\Omega} = \sum_I \Omega_I/n$; it is clear that the SNR for $\Omega_{\max}$ would be worse than that for $\bar{\Omega}$. What we shall in fact do is calculate the probability distribution $p(\Omega|\Omega_I = \Omega)$ of Ω assuming that $\Omega_1 = \cdots = \Omega_n = \Omega$, as this is the easiest to calculate. Now if we ignore the *a priori* information represented by $p_s^{(0)}$ in Eq. (A42) [so that $p(\Omega_I)$ is a multivariate Gaussian], then this is equivalent to calculating the PDF for the weighted average

$$\bar{\Omega}_1 = \sum_I \frac{\Omega_I}{\sigma_I^2} \Big/ \sum_I \frac{1}{\sigma_I^2}. \tag{A46}$$

However when we include the information contained in $p_s^{(0)}$, the main effect is to truncate and renormalize [18]

the PDF's for each Ω_I, since all the Ω's must be positive. Hence a simple interpretation of the PDF $p(\Omega|\Omega_I = \Omega)$ and the corresponding SNR in terms of the average (A46) not possible. Nevertheless we suspect that $p(\Omega|\Omega_I = \Omega)$ approximately represents the probability distribution of some type of average of $\Omega_g(f)$.

We now insert the assumption $\Omega_I = \Omega$ (const) into Eq. (A42), which gives a PDF of the form

$$p(\Omega \,|\, \Omega_I = \Omega) \propto \Theta(\Omega) \exp\left[\frac{(\Omega - \Omega_M)^2}{(2\sigma^2)}\right]. \tag{A47}$$

Here the quantity

$$\Omega_M = \sum_I \frac{\hat{\Omega}_I}{\sigma_I^2} \Big/ \sum_I \frac{1}{\sigma_I^2} \tag{A48}$$

is the statistic that should be calculated to estimate the value of Ω. We obtain

$$\Omega_M \propto \int_0^\infty df \,\mathrm{tr}\left[\hat{\mathbf{S}}_h^{\perp} \cdot \hat{\mathbf{S}}_h^{-1\,\parallel} \cdot \bar{\gamma}^{\perp} \cdot \hat{\mathbf{S}}_h^{-1\,\parallel}\right]. \tag{A49}$$

This can be written in the time domain as

$$\Omega_M \propto \int dt \int d\tau\, H_a(t+\tau) L_{ab}(\tau) H_b(t), \tag{A50}$$

where $\tilde{\mathbf{L}}(f) = \gamma(f)^{\perp}$, and

$$\tilde{\mathbf{H}}(f) = \frac{1}{f^{3/2}} \hat{\mathbf{S}}_h^{-1}(f)^{\parallel} \cdot \tilde{\mathbf{h}}(f). \tag{A51}$$

The functions $L_{ab}(\tau)$ are the sliding delay functions discussed in Sec. III. Note that Ω_M is constructed from the measured correlations in the following way: the intrasite correlations are used only to estimate the detector noise matrix $\mathbf{S}_n^{\parallel}$, and then this noise matrix is used together with the measured intersite correlations to estimate Ω_M. However, the matrix $\mathbf{S}_n^{\parallel}$ will typically be approximately diagonal, and so neglecting the intrasite correlations will give only a negligible error.

Finally the intersite SNR Ω_M/σ can be obtained from Eq. (A43) [18]. If we assume that $\hat{\mathbf{S}}_h(f)^{\perp} = \Omega_{\mathrm{real}}(f)\bar{\gamma}(f)^{\perp}$, so that all the intersite correlations we measure come from the SB, then the SNR takes the form

$$\frac{S^2}{N^2} = \hat{\tau}\left(\frac{4\rho_c}{5\pi}\right)^2 \int_0^\infty df\, \frac{\Omega_{\mathrm{real}}(f)^2}{f^6} \times \mathrm{tr}\left[\left(\gamma(f)^{\perp} \cdot \hat{\mathbf{S}}_h^{-1}(f)^{\parallel}\right)^2\right]. \tag{A52}$$

Using the fact that the matrix $\gamma(f)_{ab}$ will be constant on each subblock corresponding to two detector sites, one can derive Eqs. (3.12) and (3.13) from Eq. (A52).

APPENDIX B: CALCULATION OF THE OVERLAP REDUCTION FUNCTION

The ath detector is characterized by its position $\mathbf{x}_a$ and by the tensor $\mathbf{d} = (\mathbf{u}\otimes\mathbf{u} - \mathbf{v}\otimes\mathbf{v})/2$, where $\mathbf{u}$ and $\mathbf{v}$ are unit vectors in the direction of its arms. In terms of these quantities, the overlap reduction function is given, from Eq. (2.13), by

$$\gamma_{ab}(f) = \frac{5}{8\pi} \sum_A \int d^2\Omega\, (\mathbf{d}_a : \mathbf{e}^{A,\mathbf{n}})\, (\mathbf{d}_b : \mathbf{e}^{A,\mathbf{n}}) \times \exp\left[2\pi i f \mathbf{n}\cdot(\mathbf{x}_a - \mathbf{x}_b)\right]. \tag{B1}$$

If we write $\mathbf{x}_a - \mathbf{x}_b = d\,\mathbf{m}$ where $\mathbf{m}$ is a unit vector, and put $\alpha = 2\pi f d$ (in units in which $c = 1$), then we obtain

$$\gamma_{ab}(f) = d_{a\,ij}\ \Gamma_{ijkl}(\alpha, \mathbf{m})\ d_{b\,kl}, \tag{B2}$$

where

$$\Gamma_{ijkl}(\alpha, \mathbf{m}) = \frac{5}{8\pi} \sum_A \int d^2\Omega\, e_{ij}^{A,\mathbf{n}} e_{kl}^{A,\mathbf{n}}\, e^{i\alpha\mathbf{n}\cdot\mathbf{m}}. \tag{B3}$$

This integral can be evaluated by the standard method of writing down the most general possible answer:

$$\Gamma_{ijkl}(\alpha, \mathbf{m}) = A(\alpha)\delta_{ij}\delta_{kl} + B(\alpha)\left[\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right] + C(\alpha)\left[\delta_{ij}m_k m_l + \delta_{kl}m_i m_j\right] + D(\alpha)m_i m_j m_k m_l + E(\alpha)\left[\delta_{ik}m_j m_l + \cdots + \delta_{jl}m_i m_k\right]. \tag{B4}$$

One might expect to have to include a term proportional to $w_{il}w_{jk} + w_{ik}w_{jl}$ where $w_{ij} = \varepsilon_{ijk}m^k$, but in fact this tensor is a linear combination of the five tensors included above. Contracting Eq. (B4) with the five different tensorial expressions that appear on its right-hand side yields a system of linear equations for $A(\alpha)\cdots E(\alpha)$ which involves scalar integrals that are straightforward to evaluate. Solving these equations and substituting the results back into Eqs. (B2) and (B4) gives

$$\gamma_{ab}(f) = \rho_1(\alpha)\mathbf{d}_a : \mathbf{d}_b + \rho_2(\alpha)\mathbf{m}\cdot\mathbf{d}_a\cdot\mathbf{d}_b\cdot\mathbf{m} + \rho_3(\alpha)(\mathbf{m}\cdot\mathbf{d}_a\cdot\mathbf{m})(\mathbf{m}\cdot\mathbf{d}_b\cdot\mathbf{m}). \tag{B5}$$

The functions $\rho_j(\alpha)$ are linear combinations of the spherical Bessel functions:

$$\rho_1(\alpha) = 5j_0(\alpha) - 2j_1(\alpha)/\alpha + 5j_2(\alpha)/\alpha^2, \tag{B6}$$

$$\rho_2(\alpha) = -10j_0(\alpha) + 40j_1(\alpha)/\alpha - 50j_2(\alpha)/\alpha^2, \tag{B7}$$

and

$$\rho_3(\alpha) = 5j_0(\alpha)/2 - 25j_1(\alpha)/\alpha + 175j_2(\alpha)/(2\alpha^2). \tag{B8}$$

The result (B5) applies to any gravitational wave antennas, such as interferometers with nonperpendicular arms and arbitrary orientations, or resonant bar antennas. The special case of two resonant bar antennas, for which each $\mathbf{d}_a \propto \mathbf{1} - 3\mathbf{n}_a\otimes\mathbf{n}_a$ for some vector $\mathbf{n}_a$, has been previously derived by Michelson [16]. As can be seen from Fig. 10, the first term in Eq. (B5) dominates

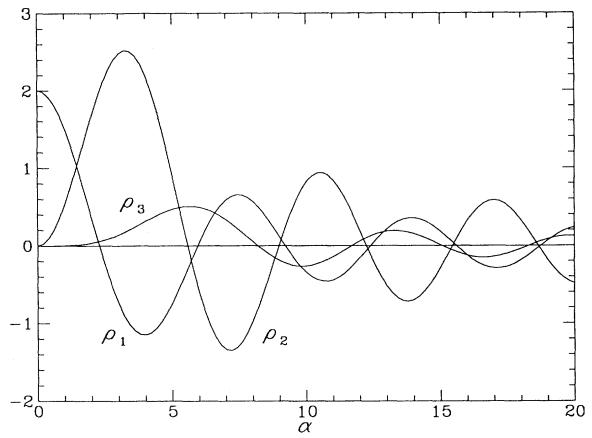

FIG. 10. A graph of the functions $\rho_1(\alpha)$, $\rho_2(\alpha)$, and $\rho_3(\alpha)$.

for $\alpha \lesssim 1$ (unless $\mathbf{d}_a : \mathbf{d}_b \approx 0$). A similar simplification applies for $\alpha \gg 1$: we have

$$\gamma_{ab}(f) = 5j_0(\alpha)\, \mathbf{d}_a^{\perp} : \mathbf{d}_b^{\perp} + \mathcal{O}(\alpha^{-2}), \tag{B9}$$

where $\mathbf{d}_a^{\perp}$ denotes the trace-free part of the projection $(\delta_{ik} - m_i m_k)(\delta_{jl} - m_j m_l) d_{a\,kl}$ of $\mathbf{d}_a$ perpendicular to $\mathbf{m}$. The fact that $\gamma(f)$ does not depend on the components of $\mathbf{d}_a$ and $\mathbf{d}_b$ parallel to $\mathbf{m}$ is due to the fact that the cross correlation at $\alpha \gg 1$ is dominated by modes whose wave vectors are nearly parallel to $\mathbf{m}$, as discussed in Sec. IV B.

To apply Eq. (B5) to terrestrial detectors, we need to choose a coordinate system and express the tensors $\mathbf{d}_1$, $\mathbf{d}_2$, and $\mathbf{m}$ in terms of the angles δ, Δ, and β defined in Sec. IV. A convenient choice is to take the detectors to be located at $\theta = \pi/2 \pm \beta/2$ and $\phi = 0$, where r, θ, ϕ are spherical polar coordinates with origin at the Earth's center, so that the unit vector in the direction joining the detectors is $\mathbf{m} = \mathbf{e}_z$. If we let $\mathbf{e}_{\hat{r}}$, $\mathbf{e}_{\hat{\theta}}$, and $\mathbf{e}_{\hat{\phi}}$ be the usual basis of orthonormal vectors, and define

$$\mathbf{d}(\sigma,\theta,\phi) = \sin(2\sigma)(\mathbf{e}_{\hat{\theta}} \otimes \mathbf{e}_{\hat{\theta}} - \mathbf{e}_{\hat{\phi}} \otimes \mathbf{e}_{\hat{\phi}})/2 - \cos(2\sigma)(\mathbf{e}_{\hat{\theta}} \otimes \mathbf{e}_{\hat{\phi}} + \mathbf{e}_{\hat{\phi}} \otimes \mathbf{e}_{\hat{\theta}})/2, \tag{B10}$$

then we can take $\mathbf{d}_1 = \mathbf{d}(\Delta + \delta, \pi/2 + \beta/2, 0)$ and $\mathbf{d}_2 = \mathbf{d}(\Delta - \delta, \pi/2 - \beta/2, 0)$. Inserting these expressions into the formula (B5) yields, after some manipulation, the result (4.1), where the functions $g_j(\alpha)$ are

$$g_1(\alpha) = \frac{5}{16}\mathbf{f}(\alpha) \cdot (-9, -6, 9, 3, 1), \tag{B11}$$

$$g_2(\alpha) = \frac{5}{16}\mathbf{f}(\alpha) \cdot (45, 6, -45, 9, 3), \tag{B12}$$

$$g_3(\alpha) = \frac{5}{4}\mathbf{f}(\alpha) \cdot (15, -4, -15, 9, -1), \tag{B13}$$

and

$$\mathbf{f}(\alpha) = (\alpha\cos\alpha, \alpha^3\cos\alpha, \sin\alpha, \alpha^2\sin\alpha, \alpha^4\sin\alpha)/\alpha^5. \tag{B14}$$

From Eq. (4.1) it is straightforward to evaluate the Fourier transform of the overlap reduction function, which gives the sliding delay function discussed in Sec. III,

$$L_{ab}(\tau) = \int_{-\infty}^{\infty} df\, e^{-2\pi i f\tau} \gamma_{ab}(f). \tag{B15}$$

The function $L_{ab}(\tau)$ vanishes for $|\tau| > d = 2r_E \sin(\beta/2)$. For $|\tau| < d$, $L_{ab}(\tau)$ is given by Eqs. (4.1)–(4.3), where the functions g_j are now

$$g_1(\tau) = \frac{5}{32d}(1 + 3v + \tfrac{3}{8}v^2), \tag{B16}$$

$$g_2(\tau) = \frac{5}{32d}(3 - 3v - \tfrac{15}{8}v^2), \tag{B17}$$

$$g_3(\tau) = \frac{5}{32d}(-4 + 8v - \tfrac{5}{2}v^2), \tag{B18}$$

and $v = 1 - \tau^2/d^2$.

[1] A. Abrmovici, W.E. Althouse, R.W.P. Drever, Y. Gürsel, S. Kawamura, F.J. Raab, D. Shoemaker, L. Sievers, R.E. Spero, K.S. Thorne, R.E. Vogt, R. Weiss, S.E. Whitcomb, and M.E. Zucker, Science **256**, 325 (1992).

[2] C. Bradaschia *et al.*, Nucl. Instrum. Methods A **289**, 518 (1990); also in *Gravitation 1990*, Proceedings of the Banff Summer Institute, Banff, Alberta, 1990, edited by R. Mann and P. Wesson (World Scientific, Singapore, 1991).

[3] *The Detection of Gravitational Waves*, edited by D.G. Blair (Cambridge University Press, Cambridge, England, 1991).

[4] L.A. Rosi and R.L. Zimmerman, Astrophys. Space. Sci. **45**, 447 (1976); D. Hils, P.L. Bender, and R.F. Webbink, Astrophys. J. **360**, 75 (1990).

[5] T. Vachaspati and A. Vilenkin, Phys. Rev. D **31**, 3502 (1985); F.R. Bouchet and D.B. Bennett, *ibid.* **41**, 720 (1990).

[6] C.J. Hogan, Mon. Not. R. Astron. Soc. **218**, 629 (1986); M.S. Turner and F. Wilczek, Phys. Rev. Lett. **65**, 3080 (1990); A. Kosowsky, M.S. Turner, and R. Watkins, Phys. Rev. D **45**, 4514 (1992); Phys. Rev. Lett. **69**, 2026 (1992); A. Kosowsky and M.S. Turner, Phys. Rev. D **47**, 4372 (1993); M. Kamionkowski, A. Kosowsky, and M.S. Turner (unpublished).

[7] L.P. Grishchuk, Zh. Eksp. Teor. Fiz. **67**, 825 (1974) [Sov. Phys. JETP **40**, 409 (1975)]; E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990), and references therein.

[8] L. Abbott and M. Wise, Nucl. Phys. **B224**, 541 (1984).

[9] L.M. Krauss and M. White, Phys. Rev. Lett. **69**, 869 (1992).

[10] K.S. Thorne, in *300 Years of Gravitation*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987), pp. 330–458.

[11] K.S. Thorne, in *Recent Advances in General Relativity*, edited by A. Janis and J. Porter (Birkhauser, Boston,

1991).
[12] N.L. Christensen, Ph.D. thesis, Massachusetts Institute of Technology, 1990.
[13] L.P. Grishchuk and M. Solokhin, Phys. Rev. D **43**, 2566 (1991).
[14] A.A. Starobinsky, Pis'ma Astron. Zh. **11**, 223 (1985) [Sov. Astron. Lett. **11**, 133 (1985)].
[15] R.L. Davies *et al.*, Phys. Rev. Lett. **69**, 1856 (1992).
[16] R.F. Michelson, Mon. Not. R. Astron. Soc. **227**, 933 (1987).
[17] N.L. Christensen, Phys. Rev. D **46**, 5250, (1992).
[18] The phrase *signal-to-noise ratio* usually means the following: From the measured data one can compute a number Y which has to a good approximation a Gaussian statistical distribution, and SNR is just the expected value of Y divided by the square root of its variance. In this paper we frequently encounter quantities Y whose distributions are *truncated* Gaussians:

$$p_Y(y) = \frac{N}{\sqrt{2\pi}\sigma} \exp\left[\frac{(y-\mu)^2}{2\sigma^2}\right] \Theta(y),$$

where Θ is the step function, and the normalization factor N depends on μ/σ. In this case we still use SNR to mean just μ/σ. The use of truncated Gaussian distributions affects the constants appearing in Eqs. (5.1) and (6.5) (where we assume $\mu \ll \sigma$).
[19] K.S. Thorne (private communication).
[20] The coefficients $s_{A,\mathbf{n}}(f)$ are related to the variables $\tilde{h}_A$ of Ref. [17] by $s_{A,\mathbf{n}}(f) = 2\pi^2 f \tilde{h}_A(\mathbf{n}, 2\pi f)$; Eqs. (2.8) and (2.9) correct Eq. (4.1) of that reference.
[21] For relic gravitational waves from the big bang, Grishchuk [13] has shown that Eq. (2.9) must be replaced by

$$\langle\, s_{A,\mathbf{n}}(f)\, s_{B,\mathbf{m}}(f')\,\rangle = \delta_{AB}\, \delta^2(\mathbf{n},\mathbf{m})\, \delta(f-f') D(f),$$

for $f, f' > 0$, where $D(f)$ is a complex function satisfying $|D(f)| \approx$ the corresponding term in Eq. (2.8). This is because the relic background is phase squeezed. However, Eq. (2.9) remains an adequate approximation for our purposes, since $D(f)$ is a rapidly oscillating function of f with a period of order H_0, the Hubble constant [13]. Observations of the SB by the methods discussed in this paper will have a frequency resolution $\gg H_0 \sim 10^{-18}$ Hz, and consequently will "average out" the effects of the phase squeezing.
[22] R.L. Forward, Phys. Rev. D **17**, 379 (1978).
[23] Equations (2.12) and (2.13) are essentially equivalent to Eqs. (4.5) and (4.6) of Ref. [17], and can be obtained from those by (i) taking a Fourier transform, (ii) dividing by the square of the transfer function $|\tilde{B}(f)|$ to convert from correlations of light intensities to correlations of strain amplitudes, (iii) replacing the cosine with a complex exponential, and (iv) correcting for the difference in normalization convention for $\gamma(f)$ between Ref. [17] and this paper.
[24] It is not quite true that the contributions $\mathbf{S}_n(f)$ and $\mathbf{S}_s(f)$ are indistinguishable. The structure of the matrix $\mathbf{S}_s(f)$ at a fixed frequency is tightly constrained by Eq. (2.12), so the whole (off-diagonal) matrix can be determined from just one element. Moreover all the off-diagonal elements must be real. If we measure the off-diagonal part of $\mathbf{S}_h(f)$ to be nonzero and it does not have the form (2.12), then we know that we are seeing noise that is not due to the SB. This fact will be helpful in identifying noise in intersite correlations (if there is any intersite correlated noise). It will not be as useful for identifying noise in intrasite correlations, because in this case the constraint is just that all the off-diagonal elements of $\mathbf{S}_A(f)$ be equal and positive, and this is likely to be satisfied by some sources of correlated noise.
[25] These designations must not be taken too literally, however, since for $\beta = 180°$, the angle Δ is what would normally be called the relative rotation; see Eq. (B10).
[26] Using Eq. (B5) and the fact that $\rho_1(0) = \rho_2(0) = 0$, we see that for *any* symmetric traceless tensors $\mathbf{e}_1$ and $\mathbf{e}_2$, the metric perturbation $\mathbf{h}(\mathbf{x}, t)$ due to the SB satisfies

$$\langle \mathbf{e}_1 : \mathbf{h}(\mathbf{0}, t)\ \mathbf{e}_2 : \mathbf{h}(\mathbf{0}, t')\rangle \propto \mathbf{e}_1 : \mathbf{e}_2.$$

This is a precise expression of the fact that orthogonal components of the strain tensor at a fixed point in space are statistically independent.
[27] LIGO team (private communication).
[28] S. Whitcomb (private communication).
[29] B.J. Meers, Phys. Rev. D **38**, 2317 (1988); K.A. Strain and B.J. Meers, Phys. Rev. Lett. **66**, 1391 (1991); A. Krolak, J.A. Lobo, and B.J. Meers, Phys. Rev. D **43**, 2470 (1991).
[30] L.A. Wainstein and V.D. Zubakov, *Extraction of Signals from Noise* (Dover, New York, 1962).
[31] C.W. Helstrom, *Statistical Theory of Signal Detection*, 2nd ed. (Pergamon, New York, 1968).
[32] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).
[33] If the condition (A2) were not valid, one could still in principle integrate over $\Omega_g^>(f)$ to derive the reduced PDF for $\Omega_g^<(f)$ [in the notation of Eq. (A4)]. However, in this case the CFR limit need not apply.
[34] A. Gillespie and F. Rabb, LIGO Report No. 93-2 (unpublished).
[35] There will be corrections due to aliasing which can be made small by increasing τ_c by a factor of a few, so that $\mathbf{C}_h(\tau) \approx 0$ for $|\tau| > \tau_c/$(a few).
[36] Strictly speaking $p_{\rm corr}(\mathcal{C})$ should have support only on positive definite matrices $\mathcal{C}$. However, this will be approximately true if $1/\mathcal{C}_{\max} + 1 \geq$ (number of detectors per site).

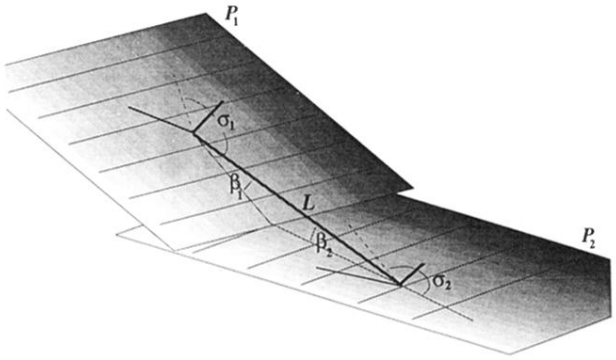

FIG. 3. The angles σ_1, σ_2, β_1, and β_2 formed by a pair of interferometric detectors and the line L which joins them.