Optimal strategies for gravitational wave stochastic background searches in pulsar timing data

Melissa Anholm,[1,*] Stefan Ballmer,[2,†] Jolien D. E. Creighton,[1,‡] Larry R. Price,[1,§] and Xavier Siemens[1,‖]

[1]*Center for Gravitation and Cosmology, Department of Physics, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201, USA*

[2]*LIGO Laboratory, California Institute of Technology, MS 18-34, Pasadena, California 91125, USA*

A low frequency stochastic background of gravitational waves may be detected by pulsar timing experiments in the next 5 to 10 yr. Using methods developed to analyze interferometric gravitational wave data, in this paper we lay out the optimal techniques to detect a background of gravitational waves using a pulsar timing array. We show that for pulsar distances and gravitational wave frequencies typical of pulsar timing experiments, neglecting the effect of the metric perturbation at the pulsar does not result in a significant deviation from optimality. We discuss methods for setting upper limits using the optimal statistic, show how to construct skymaps using the pulsar timing array, and consider several issues associated with realistic analysis of pulsar timing data.

PACS numbers: 04.30.Db, 04.80.Nn

I. INTRODUCTION

The search for gravitational waves is at the forefront of current fundamental physics research. The direct detection of gravitational waves will usher in a new era in astronomy and astrophysics. Gravitational waves will reveal information about black holes, supernovae, and neutron stars that cannot be gleaned from electromagnetic observations. Furthermore, the detection of a gravitational wave background will open an observational window onto a time in the early universe before recombination, prior to which the universe is opaque to electromagnetic waves. The scientific rewards for such a detection would be truly exceptional. Several international efforts are underway to detect gravitational waves and two of these efforts are expected to result in the detection of gravitational waves in the next 5 to 10 yr: Interferometric ground-based gravitational wave detectors and pulsar timing observations.

Neutron stars radiate powerful beams of radio waves from their magnetic poles. If a neutron star's magnetic poles are not aligned with its rotational axis, the beams sweep through space like the beacon on a lighthouse and the neutron star is said to be a pulsar. If the Earth lies within the sweep of a pulsar's beams, the star is observed as a point source in space emitting short, rapid bursts of radio waves [1]. Because of their enormous mass, neutron stars have a very large moment of inertia and the radio pulses we observe arrive at a very constant rate. Pulsar timing experiments exploit this regularity [2,3]. Fluctuations in the time of arrival of radio pulses, after all known effects have been subtracted out, could be due to the presence of gravitational waves. Since the 1970s, when these ideas were first conceived, pulsar timing precision has improved dramatically. Several known pulsars can now be timed with a precision of about 1 μs, and a handful can be timed with a precision around several hundred nanoseconds [4]. Recent work [5] has shown that the presence of nanohertz gravitational waves could be detected by observing 20 pulsars with timing precisions of 100 ns over a period of 5 to 10 yr. Nondetection would still improve current bounds on the low frequency stochastic gravitational wave background [6].

Gravitational waves from supermassive black hole binary systems could be detected via pulsar timing observations [7–11]. In addition, pulsar timing has the potential to measure the polarization properties of gravitational waves which could confirm (or even change) the current theory of gravity [12,13]. Gravitational wave observations in the nanohertz band could also yield information about the early universe [14]. Cosmic strings, linelike topological defects, could produce gravitational waves in the nanohertz band. Cosmic strings can form during phase transitions in the early universe due to the rapid cooling that takes place after the Big Bang [15–17]. Cosmic string production is generic in supersymmetric grand unified theories [18]. Additionally, in string theory motivated cosmological models cosmic strings may also form (dubbed cosmic superstrings to differentiate them from field theoretic cosmic strings) [19–24]. Cosmic strings and superstrings are expected to produce a background of stochastic gravitational waves and bursts of gravitational waves [25–30] that could be detected using pulsar timing observations. Pulsar timing observations are already producing some of the most interesting constraints on cosmic string models and a detection would have profound implications [29,30].

*anholm@gravity.phys.uwm.edu
†sballmer@caltech.edu
‡jolien@gravity.phys.uwm.edu
§larry@gravity.phys.uwm.edu
‖siemens@gravity.phys.uwm.edu

Recently Lommen [31] produced an upper limit on the stochastic gravitational wave background using observations of 3 ms pulsars spanning 17 yr. The methods used by Lommen were based on those developed by Kaspi and collaborators [32] and have been the subject of some criticism in the literature [6,28]. More recently Jenet and collaborators [5,6] developed a new technique for gravitational wave stochastic background searches in pulsar timing data and applied it to the Parkes pulsar timing array data [33,34].

In this paper we consider optimal strategies for extraction of a gravitational wave stochastic background signal using data from a pulsar timing array. Our methods are based on those developed for and used in ground-based interferometric gravitational wave detectors such as LIGO and Virgo [35–43], and improve on existing methods in several ways. In Sec. II we write the redshift of pulsar signals induced by passing gravitational waves, first derived by Detweiler [2], in a coordinate-independent way more suitable for our analysis, and discuss its form in the frequency domain including the long-wavelength limit. In Sec. III we construct the optimal cross correlation filter for a pulsar pair by maximizing the signal to noise. We find that the overlap reduction function is well approximated by a constant, or equivalently, that the metric perturbation at the pulsar can be neglected for values of pulsar distances and gravitational wave frequencies typical of pulsar timing experiments, without significant losses in sensitivity. In Sec. III we also show how to construct the optimal combination of cross correlations of pulsar pairs in a pulsar timing array and include a more sophisticated derivation of the optimal detection statistic based on the likelihood ratio. In Sec. IV we discuss upper limit and detection methods. In Sec. V we show how to construct skymaps using pulsar timing data—a pulsar timing radiometer. In Sec. VI we discuss several important issues relating to the realistic analysis of pulsar timing data, including the Lomb-Scargle periodogram for power spectrum estimation of unevenly sampled data, and optimal procedures for computing Fourier transforms. We conclude in Sec. A. Lommen, Romano, and Woan [44] will extend our work using a likelihood based approach developed in [45], and consider the case of stochastic backgrounds that are loud compared to the noise, closely examine time-domain implementations of the optimal statistic, and provide a detailed comparison of the optimal statistic described here with the methods of Jenet and collaborators [5,6].

II. THE SIGNAL

Gravitational waves affect pulsar timing measurements by creating perturbations in the null geodesics that the radio signals emitted from the pulsar travel on [2]. In this section we will describe the relationship between the metric perturbation and the signal measured in pulsar timing experiments.

A metric perturbation in a spatial, transverse, traceless gauge has a plane wave expansion given by [38]

$$h_{ij}(t, \vec{x}) = \sum_A \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} e^{i2\pi f(t - \hat{\Omega} \cdot \vec{x})} h_A(f, \hat{\Omega}) e_{ij}^A(\hat{\Omega}), \tag{1}$$

where f is the frequency of the gravitational waves, $\vec{k} = 2\pi f \hat{\Omega}$ is the wave vector, $\hat{\Omega}$ is a unit vector that points along the direction of travel of the waves, $i, j = x, y, z$ are spatial indices, and the index $A = +, \times$ labels polarizations. The polarization tensors $e_{ij}^A(\hat{\Omega})$ are

$$e_{ij}^+(\hat{\Omega}) = \hat{m}_i \hat{m}_j - \hat{n}_i \hat{n}_j, \tag{2a}$$

$$e_{ij}^\times(\hat{\Omega}) = \hat{m}_i \hat{n}_j + \hat{n}_i \hat{m}_j, \tag{2b}$$

where

$$\hat{\Omega} = (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta), \tag{3a}$$

$$\hat{m} = (\sin\phi, -\cos\phi, 0), \tag{3b}$$

$$\hat{n} = (\cos\theta \cos\phi, \cos\theta \sin\phi, -\sin\theta). \tag{3c}$$

Now consider the metric perturbation from a single gravitational wave traveling along the z-axis so that $\hat{\Omega} = \hat{z}$. The metric perturbation is given explicitly by

$$h_{ij}(t, \hat{\Omega} = \hat{z}) = \sum_A \int_{-\infty}^{\infty} df e^{i2\pi f(t-z)} h_A(f, \hat{z}) e_{ij}^A(\hat{z})$$

$$\equiv h_{ij}(t - z). \tag{4}$$

The physical metric due to the perturbation is given by

$$g_{ab} = \eta_{ab} + h_{ab}(t - z)$$

$$= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 + h_+ & h_\times & 0 \\ 0 & h_\times & 1 - h_+ & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{5}$$

where $\eta_{ab} = \text{diag}\{-1, 1, 1, 1\}$ is the Minkowski metric, a, b are spacetime indices, and

$$h_{+,\times} = h_{+,\times}(t - z) = \int_{-\infty}^{\infty} df e^{i2\pi f(t-z)} h_{+,\times}(f, \hat{z}). \tag{6}$$

In this background, a pulsar emitting pulses at frequency ν_0 and direction cosines α, β, and γ, with respect to the x-, y-, and z-axes, respectively, will be observed to change its frequency in the Solar System reference frame according to [2]

$$z(t, \hat{z}) \equiv \frac{\nu_0 - \nu(t)}{\nu_0}$$

$$= \frac{\alpha^2 - \beta^2}{2(1 + \gamma)} (h_+^p - h_+^e) + \frac{\alpha\beta}{1 + \gamma} (h_\times^p - h_\times^e), \tag{7}$$

where $h_{+,\times}^e$, $h_{+,\times}^p$ are the gravitational wave strains at the

Solar System barycenter and the pulsar, respectively. This central result was obtained by Detweiler [2], who generalized a result of Estabrook and Wahlquist [46] to include both gravitational wave polarizations and for pulsars at arbitrary locations. They, in turn, based their calculation on an earlier one by Kaufmann [47]. A detailed derivation of this result is provided in Appendix A.

Looking at Eq. (7) (and as shown in Appendix B) we can write the redshift $z(t, \hat{\Omega})$ of signals from a pulsar in the direction of the unit vector $\hat{p}$ produced by a gravitational wave coming from the direction $\hat{\Omega}$ as

$$z(t, \hat{\Omega}) = \frac{1}{2} \frac{\hat{p}^i \hat{p}^j}{1 + \hat{\Omega} \cdot \hat{p}} \Delta h_{ij}, \qquad (8)$$

where

$$\Delta h_{ij} \equiv h_{ij}(t_{\mathrm{p}}, \hat{\Omega}) - h_{ij}(t_{\mathrm{e}}, \hat{\Omega}), \qquad (9)$$

is the difference in the metric perturbation traveling along the direction $\hat{\Omega}$ at the pulsar and at the center of the Solar System. The vectors $(t_{\mathrm{e}}, \vec{x}_{\mathrm{e}})$ and $(t_{\mathrm{p}}, \vec{x}_{\mathrm{p}})$ give the spacetime coordinates of the Solar System barycenter and the pulsar, respectively. The metric perturbation at each location takes the form

$$h_{ij}(t, \hat{\Omega}) = \sum_A \int_{-\infty}^{\infty} df e^{i2\pi f(t - \hat{\Omega} \cdot \vec{x}_0)} h_A(f, \hat{\Omega}) e_{ij}^A(\hat{\Omega}), \quad (10)$$

for a fixed $\vec{x}_0$.

We choose a particular coordinate system by placing the Solar System barycenter at the origin and the pulsar some distance L away. With these conventions

$$t_{\mathrm{p}} = t_{\mathrm{e}} - L \equiv t - L, \qquad (11)$$

$$\vec{x}_{\mathrm{e}} = 0, \qquad (12)$$

$$\vec{x}_{\mathrm{p}} = L\hat{p}. \qquad (13)$$

If we assume that the amplitude of the metric perturbation is the same at the Solar System barycenter and the pulsar then we can use Eq. (10) to write out Δh_{ij} in our coordinate system as

$$\begin{aligned}
\Delta h_{ij} &= \int_{-\infty}^{\infty} df e^{i2\pi ft} (e^{-i2\pi fL(1 + \hat{\Omega} \cdot \hat{p})} - 1) \\
&\quad \times \sum_A h_A(f, \hat{\Omega}) e_{ij}^A(\hat{\Omega}) \\
&\equiv \Delta h_{ij}(t, \hat{\Omega}).
\end{aligned} \qquad (14)$$

Ultimately, we will be interested in the Fourier transform of this quantity which is simply

$$\Delta \tilde{h}_{ij}(f, \hat{\Omega}) = (e^{-i2\pi fL(1 + \hat{\Omega} \cdot \hat{p})} - 1) \sum_A h_A(f, \hat{\Omega}) e_{ij}^A(\hat{\Omega}). \qquad (15)$$

We can then write the Fourier transform of Eq. (8) as

$$\tilde{z}(f, \hat{\Omega}) = (e^{-i2\pi fL(1 + \hat{\Omega} \cdot \hat{p})} - 1) \sum_A h_A(f, \hat{\Omega}) F^A(\hat{\Omega}), \quad (16)$$

where we have defined

$$F^A(\hat{\Omega}) \equiv e_{ij}^A(\hat{\Omega}) \frac{1}{2} \frac{\hat{p}^i \hat{p}^j}{1 + \hat{\Omega} \cdot \hat{p}}. \qquad (17)$$

As shown in Appendix B the total redshift is given by summing over the contributions coming from gravitational waves in every direction

$$\tilde{z}(f) = \int_{S^2} d\hat{\Omega} \, \tilde{z}(f, \hat{\Omega}), \qquad (18)$$

and similarly for $z(t)$.

In fact, it is not the redshift, but a related quantity called the *residual* that gets reported in pulsar timing measurements. The residual, $R(t)$, is defined as the integral of the redshift

$$R(t) \equiv \int_0^t dt' z(t'). \qquad (19)$$

This simple relationship gives us the freedom to develop the data analysis for either variable and we henceforth limit our attention to the redshift, but the results here can be phrased in terms of the residual with minimal effort.

In the literature, searches for gravitational waves using pulsar timing data are typically performed in the time domain. The (unknown) metric perturbation at the pulsar in, say, Eqs. (7) or (8) is neglected because one can treat it as another noise term which averages to zero when performing correlations between measurements of different pulsars. In the frequency domain this is unnecessary. Equation (16) does not depend explicitly on the metric perturbation at the pulsar, rather the dependence is all in a distance and frequency dependent phase factor. It is then conceivable that if we could determine the distance to a pulsar L with sufficient accuracy we could use the metric perturbation at the pulsar to improve the sensitivity of our searches. Unfortunately, such measurements of pulsar distances are unavailable. We will show in Sec. III A, however, that for the case of a stochastic background search in pulsar timing data the phase factor can be neglected without any significant loss in sensitivity. It is unclear whether this is true for other types of gravitational wave searches.

From the ground-based interferometer perspective Eqs. (7) or (8) are somewhat counter-intuitive. This difficulty arises from the factor of $1 + \hat{\Omega} \cdot \hat{p}$ in the denominator; in Appendices A and B we show explicitly how this factor enters the expression. When $\hat{\Omega} \cdot \hat{p} = \pm 1$, i.e. the gravitational wave and the pulsar directions are parallel or antiparallel, Eqs. (16) and (17) lead to no redshifting of the pulsar signal for completely different reasons. When they are parallel the reason is the transverse nature of gravitational waves, and when they are antiparallel it is because

the pulsar signals "surf" the gravitational waves. Our surprise is a result of our long-wavelength limit intuition.

Equation (16) has an obvious long-wavelength limit. We can use this limit to compare the form of our results with those of ground-based interferometers such as LIGO. When $2\pi f L \ll 1$ we can Taylor expand the exponential and to first order Eq. (16) becomes

$$\tilde{z}(f, \hat{\Omega}) \approx -i\pi f L \hat{p}^i \hat{p}^j \sum_A h_A(f, \hat{\Omega}) e_{ij}^A(\hat{\Omega}). \tag{20}$$

Typical values of f are in the range $1/10$ yr^{-1} to 10 yr^{-1}. Typical values of the Earth pulsar distance L are in the range 100 ly to 10^4 ly. This means fL is in the range 10 to 10^5 and pulsar timing experiments are never in the long-wavelength limit. However, the Taylor expansion can also be done for large fL when the angle between $\hat{\Omega}$ and $\hat{p}$ is sufficiently close to π. In this case the pulsar signals can surf the gravitational waves and not undergo redshifting. Writing that angle as $\pi - \epsilon$ with $\epsilon \ll 1$, then the Taylor expansion is also valid when $\epsilon \ll (\pi f L)^{-1/2}$.

Taking the inverse Fourier transform of Eq. (20) yields

$$\tilde{z}(t) \approx -\frac{L}{2} \hat{p}^i \hat{p}^j \dot{h}_{ij}(t, \vec{x}_e), \tag{21}$$

which is the projection of the time derivative of the metric perturbation at the Solar System barycenter onto the unit vector that points to the pulsar. Note that unlike Eq. (8), this equation no longer depends on the direction of the gravitational wave and can be expressed in terms of the full metric perturbation (derivative). For the case of ground-based interferometers the signal, the so-called strain, is proportional to the difference in length of the two arms because the signal at the dark port of the interferometer depends on that difference. If the arms point in the directions of the unit vectors $\hat{X}$ and $\hat{Y}$ the strain is given by

$$h(t) \equiv h_{ij}(t, \vec{x}) \frac{1}{2} (\hat{X}^i \hat{X}^j - \hat{Y}^i \hat{Y}^j), \tag{22}$$

which is the metric perturbation $h_{ij}(t, \vec{x})$ projected onto the difference of the arms.

III. DETECTION STATISTIC

With an understanding of the signal in hand we now turn our attention to developing an optimal detection strategy. In this section we will first derive the optimal cross correlation statistic for a single pulsar pair using arguments based on maximizing signal to noise ratio (SNR). We will then determine the best way to combine measurements from multiple pulsar pairs to obtain the most constraining upper limit. This section will conclude with a more sophisticated derivation of the optimal detection statistic based on the likelihood ratio.

A. The optimal filter

In this section we will derive the optimal filter for detecting a stochastic background of gravitational waves from the cross correlation of redshift measurements of two different pulsars. This problem was addressed in detail by Allen and Romano [38], for the case of interferometric gravitational wave detectors and our analysis follows theirs closely.

Consider the signals from two pulsars

$$s_1(t) = z_1(t) + n_1(t), \tag{23}$$

$$s_2(t) = z_2(t) + n_2(t), \tag{24}$$

where $z_i(t)$ is the redshift and $n_i(t)$ is the noise intrinsic in the measurement. Throughout this work we will assume that each $n_i(t)$ is stationary and Gaussian, and is greater in magnitude than the redshift. Additionally we assume that

$$\langle n_i(t) \rangle = 0, \qquad \langle z_i(t) \rangle = 0,$$
$$\langle n_1(t) n_2(t) \rangle = 0, \qquad \langle n_i(t) z_j(t) \rangle = 0, \tag{25}$$

for all i and j, where the angle brackets denote an expectation value.

A stochastic background will show up in the data as correlated noise between measurements with different detectors. Our goal is to find a filter, $Q(t - t')$, that optimizes the cross correlation statistic

$$S \equiv \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt' s_1(t) s_2(t') Q(t - t'), \tag{26}$$

where T is the observation time. We will define the optimal filter to be the $Q(t - t')$ that maximizes the signal to noise ratio

$$\text{SNR} \equiv \frac{\mu}{\sigma}, \tag{27}$$

where μ and σ are the mean and square root of the variance, respectively, associated with the cross correlation signal defined in Eq. (26).

We start by assuming that the observation time is much greater than the separation of the two detectors and extend the limits of the integral over dt' to $\pm\infty$. Technically our assumption is not correct because pulsars are typically separated by distances far greater than the observation time. Later we will see that neglecting the phase terms that correspond to the metric perturbation at the pulsar location is an excellent approximation. In effect this makes our detectors colocated though not coaligned, and our assumption about the observation time is appropriate. We work in the frequency domain so that Eq. (26) becomes

$$S = \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T(f - f') \tilde{s}_1^*(f) \tilde{s}_2(f') \tilde{Q}(f'), \tag{28}$$

where $\delta_T(f - f')$ is the finite-time approximation to the delta function

$$\delta_T(f) \equiv \frac{\sin(\pi f T)}{\pi f}. \tag{29}$$

The mean of the cross-correlation is

$$\mu \equiv \langle S \rangle = \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T(f - f') \langle \tilde{z}_1^*(f) \tilde{z}_2(f') \rangle \tilde{Q}(f'). \tag{30}$$

Because of Eqs. (16), (18), and (25), taking the expectation value above requires us to evaluate $\langle h_A^*(f, \hat{\Omega}) h_{A'}(f', \hat{\Omega}') \rangle$. The assumptions that the stochastic background is stationary, unpolarized, and isotropic lead us to take

$$\langle h_A^*(f, \hat{\Omega}) h_{A'}(f', \hat{\Omega}') \rangle = \delta^2(\hat{\Omega}, \hat{\Omega}') \delta_{AA'} \delta(f - f') H(f), \tag{31}$$

where $H(f) = H(-f)$ is the gravitational wave spectrum. $H(f)$ is related to $\Omega_{gw}(f)$ through

$$\Omega_{gw}(f) \equiv \frac{1}{\rho_{crit}} \frac{d\rho_{gw}}{d \ln f}, \tag{32}$$

where $\rho_{crit} = 8\pi/3H_0^2$ and

$$\rho_{gw} = \frac{1}{32\pi} \langle \dot{h}_{ab}(t, \vec{x}) \dot{h}^{ab}(t, \vec{x}) \rangle, \tag{33}$$

is the energy density in gravitational waves. It follows from the plane wave expansion Eq. (1) along with Eqs. (31) and (32) in Eq. (33) that

$$H(f) = \frac{3H_0^2}{32\pi^3} |f|^{-3} \Omega_{gw}(|f|), \tag{34}$$

and therefore

$$\langle h_A^*(f, \hat{\Omega}) h_{A'}(f', \hat{\Omega}') \rangle = \frac{3H_0^2}{32\pi^3} \delta^2(\hat{\Omega}, \hat{\Omega}') \delta_{AA'} \delta(f - f')$$
$$\times |f|^{-3} \Omega_{gw}(|f|), \tag{35}$$

which is sometimes written in terms of the characteristic strain

$$h_c^2(f) = \frac{3H_0^2}{2\pi^2} \frac{1}{f^2} \Omega_{gw}(|f|). \tag{36}$$

The expectation value we set out to evaluate is then

$$\langle \tilde{z}_1^*(f) \tilde{z}_2(f') \rangle = \frac{3H_0^2}{32\pi^3} \frac{1}{\beta} \delta(f - f') |f|^{-3} \Omega_{gw}(|f|) \Gamma(|f|), \tag{37}$$

where we defined

$$\Gamma(|f|) = \beta \sum_A \int_{S^2} d\hat{\Omega} (e^{i2\pi f L_1(1 + \hat{\Omega} \cdot \hat{p}_1)} - 1)$$
$$\times (e^{-i2\pi f L_2(1 + \hat{\Omega} \cdot \hat{p}_2)} - 1) F_1^A(\hat{\Omega}) F_2^A(\hat{\Omega}), \tag{38}$$

the pulsar timing analogue of the overlap reduction function [38], which has a normalization factor β. The normal-

ization is chosen so that $\Gamma(|f|) = 1$ for coincident, coaligned detectors. As we show below, pulsar timing experiments are in a regime where the exponential factors in Eq. (38) can be neglected. In this situation, which we will assume henceforth, the normalization factor is easy to determine and we have that

$$\Gamma_0 \equiv \frac{3}{4\pi} \sum_A \int_{S^2} d\hat{\Omega} F_1^A(\hat{\Omega}) F_2^A(\hat{\Omega})$$
$$= 3\left\{ \frac{1}{3} + \frac{1 - \cos\xi}{2} \left[\ln\left(\frac{1 - \cos\xi}{2}\right) - \frac{1}{6} \right] \right\}, \tag{39}$$

where $\xi = \cos^{-1}(\hat{p}_1 \cdot \hat{p}_2)$ is the angle between the two pulsars. This quantity is proportional to the Hellings and Downs curve [48]. A detailed derivation of this result is provided for completeness in Appendix C 1.

The rationale given in the literature for throwing out the pulsar term in Eq. (7), or equivalently Eqs. (8) and (9), is that the unknown metric perturbation at the pulsars can be thought of as a kind of noise term which averages to zero when performing a correlation between different pulsars. The equivalent procedure in the frequency domain is to neglect the phase factors in Eq. (16), or in terms of our optimal filter, approximating Eq. (38) with Eq. (39). The regime where the approximate Eq. (39) is valid, is helpful in quantifying the accuracy of the rationale. Figure 1 shows the overlap reduction function for two pulsars a distance L from the Solar System barycenter. Since the distance to both pulsars is the same, the overlap reduction function can be written as just a function of fL. The top two (blue) curves show Eq. (38) with $\beta = 3/4\pi$ (solid line) and Eq. (39) (dashed line) for two pulsars at an angle $\xi = \pi/8$ as a function of fL. The middle (red) and bottom (green) curves show the same quantities for two pulsars at $\xi \approx 0.86$ and $\xi = \pi/2$ respectively. As discussed in the last section, the smallest value of the frequency $f_{min} \sim 0.1 \text{ yr}^{-1}$ and the closest pulsars used in timing experiments are at a distance of $L_{min} \sim 100$ ly so that $fL \gtrsim 10$. As shown in Fig. 1 this range of fL puts pulsar timing experiments in the regime where Eq. (39) is an excellent approximation to Eq. (38), and we can neglect the pulsar term while remaining close to optimal.

Returning to Eq. (30), we now have

$$\mu = \frac{3H_0^2}{32\pi^3} \frac{1}{\beta} T \int_{-\infty}^{\infty} df |f|^{-3} \Omega_{gw}(|f|) \Gamma(|f|)$$
$$\approx \frac{H_0^2}{8\pi^2} T \Gamma_0 \int_{-\infty}^{\infty} df |f|^{-3} \Omega_{gw}(|f|). \tag{40}$$

With the assumption that the noise is much greater than the signal, the variance, σ^2, depends only on the statistical properties of the noise in each detector. We have

$$\sigma^2 \equiv \langle S^2 \rangle - \langle S \rangle^2 \approx \langle S^2 \rangle$$
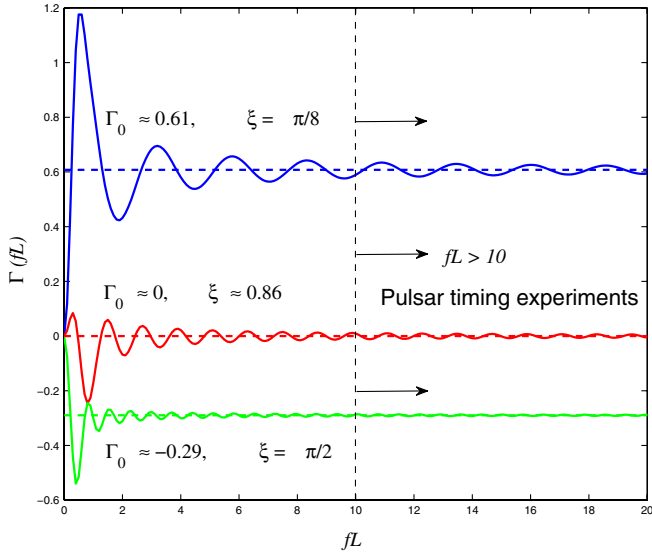$$\approx \frac{T}{4} \int_{-\infty}^{\infty} df P_1(|f|) P_2(|f|) |\tilde{Q}(f)|^2, \tag{41}$$

FIG. 1 (color online). Plot of the full overlap reduction, Eq. (38), along with the approximation Eq. (39) for two pulsars a distance L from the Solar System barycenter. The overlap reduction function is a function of fL. The top two (blue) curves show Eq. (38) with $\beta = 3/4\pi$ (solid line) and Eq. (39) (dashed line) for two pulsars at an angle $\xi = \pi/8$ as a function of fL. The middle (red) and bottom (green) curves show the same quantities for two pulsars at $\xi \approx 0.86$ and $\xi = \pi/2$ respectively. The smallest value of the frequency $f_{\min} \sim 0.1 \ \mathrm{yr}^{-1}$ and the closest pulsars used in timing experiments are at a distance of $L_{\min} \sim 100$ ly so that $fL \gtrsim 10$. This range of fL puts pulsar timing experiments in the regime where Eq. (39) is an excellent approximation to Eq. (38) and we are justified in throwing out the pulsar term while remaining close to optimal.

where

$$\langle \tilde{n}_i^*(f)\tilde{n}_i(f')\rangle = \tfrac{1}{2}\delta(f - f')P_i(|f|), \qquad (42)$$

is the (one-sided) noise power spectrum.

With μ and σ^2 in hand, we next define a positive-definite inner product using the noise power spectra of the two detectors

$$(A, B) \equiv \int_{-\infty}^{\infty} df A^*(f)B(f)P_1(|f|)P_2(|f|). \qquad (43)$$

With this definition it is easy to see that

$$\mu \approx \frac{H_0^2}{8\pi^2}T\left(\tilde{Q}, \frac{\Omega_{\mathrm{gw}}(|f|)\Gamma_0}{|f|^3P_1(|f|)P_2(|f|)}\right), \qquad (44)$$

$$\sigma^2 \approx \frac{T}{4}(\tilde{Q}, \tilde{Q}), \qquad (45)$$

from which it follows from the definition of SNR in Eq. (27) and Schwartz's inequality that the optimal filter is given by

$$\tilde{Q}(f) = \chi\frac{\Omega_{\mathrm{gw}}(|f|)\Gamma_0}{|f|^3P_1(|f|)P_2(|f|)}, \qquad (46)$$

for some normalization constant, χ. Our primary interest will be in stochastic backgrounds with power law spectra, $\Omega_{\mathrm{gw}}(f) = \Omega_\alpha f^\alpha$ (for constant Ω_α). In that case the normalization constant for the optimal filter, $\tilde{Q}_\alpha(f)$, is chosen so that

$$\mu = \Omega_\alpha T_0, \qquad (47)$$

where T_0 is some arbitrary constant with dimensions of time. From Eq. (44) it follows that

$$\chi = \Omega_\alpha\frac{T_0}{T}\frac{8\pi^2}{H_0^2}\left[\int_{-\infty}^{\infty}df\frac{\Omega_{\mathrm{gw}}^2(|f|)\Gamma_0^2}{f^6P_1(|f|)P_2(|f|)}\right]^{-1}. \qquad (48)$$

Finally, we can compute

$$\mathrm{SNR} \approx \frac{H_0^2}{4\pi^2}T^{1/2}\left[\int_{-\infty}^{\infty}df\frac{\Omega_{\mathrm{gw}}^2(|f|)\Gamma_0^2}{f^6P_1(|f|)P_2(|f|)}\right]^{1/2}. \qquad (49)$$

The differences between these results and those for interferometers can all be traced to the differing overlap reduction function $\Gamma(f) \approx \Gamma_0$. The normalization of Γ_0 means that the maximal SNR (for coincident, coaligned detectors) is only $5/6$ of that obtainable from interferometers, assuming the noise power spectra are the same in each case.

To construct the optimal filter, Eq. (46), the noise power spectra for the two pulsars $P_1(|f|)$ and $P_2(|f|)$ must be determined. These can either be modeled, or measured with the methods described in Sec. VI. Once constructed the optimal filter can be applied in the frequency domain. Section VI gives a prescription for taking Fourier transforms of unevenly sampled data. The optimal filter can also be inverse Fourier transformed and the correlation performed in the time domain. It is unclear which of these two methods is more robust and the authors of [44] will explore the time-domain approach in detail. In the derivation of the optimal filter we have assumed that the noise in pulsar timing data is uncorrelated between pulsars. While we expect the majority of noise sources (timing noise, scattering, dispersion, etc.) to be uncorrelated, it is conceivable that correlations in the measurements of pulsar timing residuals are introduced by the detector. If this correlated component turns out to be important the cross-spectral density would need to be estimated and included in the optimal filter.

B. The pulsar timing array

The question we would like to address in this section is: Given redshift measurements from N different pulsars (which each have a different noise profile), what is the best way to combine those measurements to produce the *most constraining* upper limit? One can consider the cross correlations between any even number of detectors, but it has been shown [38,49] that the optimal choice is the combination of pairwise cross correlations. As it turns out, the solution to this problem also solves the problem

of nonstationarity in the noise power spectra over periods longer than the typical observation time, T.

First let

$$^{(ij)}S_1, {}^{(ij)}S_2, \ldots, {}^{(ij)}S_{n_{ij}}, \tag{50}$$

be n_{ij} measurements of the cross correlation between the ith and jth pulsar. We will assume that each measurement is taken with an optimal filter normalized so that while searching for a background of the form $\Omega_{\rm gw}(f) = \Omega_\alpha f^\alpha$,

$$\langle {}^{(ij)}S_k \rangle = \Omega_\alpha T_0 \equiv \mu, \tag{51}$$

where T_0 is an arbitrary constant introduced for dimensional reasons. Each measurement therefore has the form

$$^{(ij)}S_k = \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T(f - f') \tilde{s}_{i,k}^*(f) \tilde{s}_{j,k}(f') {}^{(ij)}Q_k(f), \tag{52}$$

with

$$^{(ij)}Q_k(f) = {}^{(ij)}\chi_k \frac{\Omega_{\rm gw}(|f|)^{(ij)}\Gamma_0}{|f|^3 P_{i,k}(|f|)P_{j,k}(|f|)}, \tag{53}$$

where $^{(ij)}\Gamma_0$ is the overlap reduction function of the (ij) pulsar pair, $s_{i,k}$ is the kth measurement of the signal from the ith pulsar, and $P_{i,k}$ is the associated noise power spectrum. Additionally

$$^{(ij)}\chi_k = \Omega_\alpha \frac{T_0}{^{(ij)}T_k} \frac{8\pi^2}{H_0^2} \left[\int_{-\infty}^{\infty} df \frac{\Omega_{\rm gw}^2(|f|)^{(ij)}\Gamma_0^2}{f^6 P_{i,k}(|f|)P_{j,k}(|f|)} \right]^{-1}, \tag{54}$$

where $^{(ij)}T_k$ is the observation time of the kth measurement of the (ij) pulsar pair. Our task is to combine the $^{(ij)}S_k$ in a way that optimizes SNR. The first step is to form the sample mean for each set of measurements

$$^{(ij)}\hat{\mu} \equiv \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} {}^{(ij)}S_k, \tag{55}$$

which is both an unbiased estimator and random variable. It therefore has a mean

$$\mu_{ij} \equiv \langle {}^{(ij)}\hat{\mu} \rangle = \mu, \tag{56}$$

and a variance

$$\sigma_{ij}^2 \equiv \langle {}^{(ij)}\hat{\mu}^2 \rangle - \langle {}^{(ij)}\hat{\mu} \rangle^2 = \frac{^{(ij)}\sigma^2}{n_{ij}}, \tag{57}$$

where

$$^{(ij)}\sigma^2 = \sum_{k=1}^{n_{ij}} \frac{^{(ij)}T_k}{4} \int_{-\infty}^{\infty} df {}^{(ij)}\chi_k^2 \frac{\Omega_{\rm gw}^2(|f|)^{(ij)}\Gamma_0^2}{f^6 P_{i,k}(|f|)P_{j,k}(|f|)}. \tag{58}$$

The next step is combine the sample mean for each set of measurements into a single estimator we can use to deter-

mine an upper bound on Ω_α and hence $\Omega_{\rm gw} = \Omega_\alpha f^\alpha$. We do so by introducing an unbiased estimator consisting of a weighted average of the sample means

$$\hat{\mu} \equiv \frac{\sum_{i=1}^{l} \sum_{j<i}^{l} \lambda_{ij} {}^{(ij)}\hat{\mu}}{\sum_{i=1}^{l} \sum_{j<i}^{l} \lambda_{ij}}, \tag{59}$$

for some constants, λ_{ij}, which has mean

$$\mu_{\hat{\mu}} \equiv \langle \hat{\mu} \rangle = \mu, \tag{60}$$

and variance

$$\sigma_{\hat{\mu}}^2 \equiv \langle \hat{\mu}^2 \rangle - \langle \hat{\mu} \rangle^2 = \frac{\sum_{i=1}^{l} \sum_{j<i}^{l} \lambda_{ij}^2 \sigma_{ij}^2}{(\sum_{i=1}^{l} \sum_{j<i}^{l} \lambda_{ij})^2}. \tag{61}$$

The object is to now determine the λ_{ij} that maximizes the SNR of $\hat{\mu}$. The (squared) SNR of $\hat{\mu}$ is

$$\mathrm{SNR}_{\hat{\mu}}^2 \equiv \mu^2 \frac{(\sum_{i=1}^{l} \sum_{j<i}^{l} \lambda_{ij})^2}{\sum_{i=1}^{l} \sum_{j<i}^{l} \lambda_{ij}^2 \sigma_{ij}^2}. \tag{62}$$

To find the λ_{ij} that maximize the SNR, we exploit the same trick that led us to the optimal filter. Namely, we introduce an inner product

$$(A, B) \equiv \sum_{i=1}^{l} \sum_{j<i}^{l} A_{ij}^* B_{ij} \sigma_{ij}^2, \tag{63}$$

which allows us to write

$$\mathrm{SNR}_{\hat{\mu}}^2 \equiv \mu^2 \frac{(\lambda, \sigma^{-2})}{(\lambda, \lambda)}, \tag{64}$$

from which it follows that choosing

$$\lambda_{ij} \propto \sigma_{ij}^{-2}, \tag{65}$$

maximizes the SNR. The optimal statistic, choosing $\lambda_{ij} = \sigma_{ij}^{-2}$, is then given by

$$S_{\rm opt} = \frac{\sum_{i=1}^{l} \sum_{j<i}^{l} \sigma_{ij}^{-2(ij)}\hat{\mu}}{\sum_{i=1}^{l} \sum_{j<i}^{l} \sigma_{ij}^{-2}}$$
$$= \frac{\sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\sigma^{-2} \sum_{k=1}^{n_{ij}} {}^{(ij)}S_k}{\sum_{i=1}^{l} \sum_{j<i}^{l} n_{ij} {}^{(ij)}\sigma^{-2}}. \tag{66}$$

Because the estimator defined in Eq. (59) is unbiased and defined so that $\mu = \langle S_{\rm opt} \rangle = \Omega_\alpha T_0$, the estimate of $\hat{\Omega}_\alpha$ is found using

$$\hat{\Omega}_\alpha = \frac{\hat{S}_{\rm opt}}{T_0}, \tag{67}$$

where $\hat{S}_{\rm opt}$ is the measured value of the optimal statistic. The expected variance of $\hat{S}_{\rm opt}$ follows from Eq. (61)

$$\sigma_{\hat{\mu}}^2 = \left(\sum_{i=1}^{l} \sum_{j<i}^{l} \sigma_{ij}^{-2} \right)^{-1}. \tag{68}$$

Aside from maximizing the SNR, the linear combination of sample means that forms the optimal statistic in Eq. (66) serves two important and related purposes. First of all, as mentioned at the beginning of this section, weighing each $^{(ij)}\hat{\mu}$ by the inverse of the squared variance means that less noisy measurements (those with smaller variances) contribute more to the sum, which helps minimize the effect of long term nonstationarity. This is augmented by the normalization convention we chose in Eq. (51) for the mean of each measurement. Using Eqs. (45) and (46) with the λ that follows from Eq. (51), we see that $^{(ij)}\sigma^{-2} \propto {}^{(ij)}T$, and so longer observation times are also favored in the sum.

C. Computational procedure

In this subsection we describe how the quantities necessary for a stochastic background search are computed. The goal is to produce a measurement of the optimal statistic, $\hat{S}_{\text{opt}}$, using Eq. (66). The optimal statistic can then be used to make detection or upper limit statements (see Sec. IV).

First the power spectra for each pulsar (and each stretch), $P_{i,k}(|f|)$, must be determined. The spectra can either be modeled or measured with the methods described in Sec. VI. Then the overlap reduction functions, $^{(ij)}\Gamma_0$, need to be computed for each pulsar pair. To optimize the statistic for particular spectra the value of α (in $\Omega_{\text{gw}}(f) = \Omega_\alpha f^\alpha$) needs to be chosen. The normalizations, $^{(ij)}\chi_k$, can then be computed using Eq. (54). The normalizations allow us to compute the variances, $^{(ij)}\sigma^{-2}$, given by Eq. (58), in the numerator and denominator of Eq. (66), as well as the filters, $^{(ij)}Q_k(f)$, through Eq. (53). Note that the unknown factors of Ω_α cancel everywhere: From Eq. (54) it is easy to see the normalization $^{(ij)}\chi_k \propto \Omega_\alpha^{-1}$, so there is a cancellation factor of Ω_α in Eq. (53), and a factor of Ω_α^2 in Eq. (58). With these quantities in hand the cross correlations, $^{(ij)}S_k$, in Eq. (52) can be computed by taking Fourier transforms of the data (see Sec. VI). Alternatively, a set of time-domain filters, $^{(ij)}Q_k(t)$, can be created by taking inverse Fourier transforms of Eq. (53) and applied to the data in the time domain using Eq. (26).

Note that there is no dependence on the arbitrary constant T_0 introduced in Eq. (51) for dimensional reasons. The $^{(ij)}\chi_k$ are linear in T_0 and enter the variances quadratically [see Eq. (58)]. The dependence cancels in Eq. (66) because it is present in both numerator and denominator. T_0 also enters S_{opt} linearly through $^{(ij)}\chi_k$ in $^{(ij)}Q_k$ but cancels in Eq. (67) so that the point estimate of $\hat{\Omega}_\alpha$ is independent of T_0.

D. Likelihood approach

The detection statistic that has been derived is also an optimal statistic in the sense that it is the logarithm of the likelihood ratio, at least in the limit where the expected signal is smaller than the noise, and therefore it is the optimal statistic in both the Bayesian sense and by the Neyman-Pearson criterion. This section is based on the likelihood analysis of [45], generalized to consider multiple detector pairs.

As we did previously, we assume that the noise is stationary and Gaussian, as is the stochastic background. For any given pulsar i we assume that there are discrete samples of data which form a vector $\mathbf{s}_i$. Although the discussion below does not place requirements on the data sampling, we will assume that the observations of the pulsars all involve the same number of points N at the same evenly spaced sampling interval so that sample j of pulsar i is $\mathbf{s}_i[j] = s_i(j\Delta t)$ where Δt is the sampling interval. This signal vector is the sum of a noise vector $\mathbf{n}_i$ and the redshift vector $\mathbf{z}_i$, $\mathbf{s}_i = \mathbf{z}_i + \mathbf{n}_i$. The data is a combination of two random processes: the instrumental noise and the contribution from the stochastic background. The autocorrelation matrix $\mathbf{R}_i = \langle \mathbf{s}_i^\dagger \otimes \mathbf{s}_i \rangle$ is an $N \times N$ matrix which contains both of these contributions and, since we assume Gaussian noise and stochastic background, this matrix completely characterizes the distribution of the data. As we did previously, we assume that the measurement noise in a pulsar observation is independent of the noise in the observations of other pulsars; the stochastic background, however, is correlated amongst the pulsar signals. This correlation is characterized by the stochastic background correlation matrix $\epsilon^2 \mathbf{S}_{ij} = \langle \mathbf{z}_i^\dagger \otimes \mathbf{z}_j \rangle$. Here ϵ is an order parameter which we will use to expand the probability distribution in powers of the small stochastic background signal. It can also be interpreted as an overall amplitude parameter of the stochastic background. The probability distribution for the collection of all pulsar observations is given by a multidimensional Gaussian distribution

$$p(\mathbf{x}|\epsilon) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}\mathbf{x}^\dagger \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{x}\right), \quad (69)$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_l \end{bmatrix}, \quad (70)$$

is a column vector formed from all of the data vectors and

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{R}_1 & \epsilon^2\mathbf{S}_{12} & \cdots & \epsilon^2\mathbf{S}_{1l} \\ \epsilon^2\mathbf{S}_{21} & \mathbf{R}_2 & \cdots & \epsilon^2\mathbf{S}_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon^2\mathbf{S}_{l1} & \epsilon^2\mathbf{S}_{l2} & \cdots & \mathbf{R}_l \end{bmatrix}, \quad (71)$$

is the correlation matrix for the collective observation vector $\mathbf{x}$. In this weak signal limit we find

$$\Sigma^{-1} = \begin{bmatrix} \mathbf{R}_1^{-1} & 0 & \cdots & 0 \\ 0 & \mathbf{R}_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{R}_l^{-1} \end{bmatrix} - \epsilon^2 \begin{bmatrix} 0 & \mathbf{R}_1^{-1} \cdot \mathbf{S}_{12} \cdot \mathbf{R}_2^{-1} & \cdots & \mathbf{R}_1^{-1} \cdot \mathbf{S}_{1l} \cdot \mathbf{R}_l^{-1} \\ \mathbf{R}_2^{-1} \cdot \mathbf{S}_{21} \cdot \mathbf{R}_1^{-1} & 0 & \cdots & \mathbf{R}_2^{-1} \cdot \mathbf{S}_{2l} \cdot \mathbf{R}_l^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_l^{-1} \cdot \mathbf{S}_{l1} \cdot \mathbf{R}_1^{-1} & \mathbf{R}_l^{-1} \cdot \mathbf{S}_{l2} \cdot \mathbf{R}_2^{-1} & \cdots & 0 \end{bmatrix}$$

$$+ \epsilon^4 \begin{bmatrix} \sum_{\substack{m=1 \\ m\neq 1}}^{l} \mathbf{R}_1^{-1} \cdot \mathbf{S}_{1m} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{m1} \cdot \mathbf{R}_1^{-1} & \sum_{\substack{m=1 \\ m\neq 1,2}}^{l} \mathbf{R}_1^{-1} \cdot \mathbf{S}_{1m} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{m2} \cdot \mathbf{R}_2^{-1} & \cdots & \sum_{\substack{m=1 \\ m\neq 1,l}}^{l} \mathbf{R}_1^{-1} \cdot \mathbf{S}_{1m} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{ml} \cdot \mathbf{R}_l^{-1} \\ \sum_{\substack{m=1 \\ m\neq 2,1}}^{l} \mathbf{R}_2^{-1} \cdot \mathbf{S}_{2m} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{m1} \cdot \mathbf{R}_1^{-1} & \sum_{\substack{m=1 \\ m\neq 2}}^{l} \mathbf{R}_2^{-1} \cdot \mathbf{S}_{2m} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{m2} \cdot \mathbf{R}_2^{-1} & \cdots & \sum_{\substack{m=1 \\ m\neq 2,l}}^{l} \mathbf{R}_2^{-1} \cdot \mathbf{S}_{2m} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{ml} \cdot \mathbf{R}_l^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{\substack{m=l \\ m\neq l,1}}^{l} \mathbf{R}_l^{-1} \cdot \mathbf{S}_{lm} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{m1} \cdot \mathbf{R}_1^{-1} & \sum_{\substack{m=l \\ m\neq l,2}}^{l} \mathbf{R}_l^{-1} \cdot \mathbf{S}_{lm} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{m2} \cdot \mathbf{R}_2^{-1} & \cdots & \sum_{\substack{m=l \\ m\neq l}}^{l} \mathbf{R}_l^{-1} \cdot \mathbf{S}_{lm} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{ml} \cdot \mathbf{R}_l^{-1} \end{bmatrix}$$

$$+ O(\epsilon^6), \tag{72}$$

and

$$\ln \det \Sigma = \sum_{i=1}^{l} \ln \det \mathbf{R}_i$$

$$+ \epsilon^4 \sum_{i=1}^{l} \sum_{j<i}^{l} \mathrm{tr}(\mathbf{R}_i^{-1} \cdot \mathbf{S}_{ij} \cdot \mathbf{R}_j^{-1} \cdot \mathbf{S}_{ji}) + O(\epsilon^6). \tag{73}$$

The logarithm of the likelihood ratio is

$$\ln \Lambda = \ln p(\mathbf{x}|\epsilon) - \ln p(\mathbf{x}|0)$$

$$= \epsilon^2 \sum_{i=1}^{l} \sum_{j<i}^{l} \Re(\mathbf{s}_i^\dagger \cdot \mathbf{R}_i^{-1} \cdot \mathbf{S}_{ij} \cdot \mathbf{R}_j^{-1} \cdot \mathbf{s}_j)$$

$$+ \frac{1}{2}\epsilon^4 \sum_{i=1}^{l} \left\{ \sum_{j<i}^{l} \mathrm{tr}(\mathbf{R}_i^{-1} \cdot \mathbf{S}_{ij} \cdot \mathbf{R}_j^{-1} \cdot \mathbf{S}_{ji}) \right.$$

$$\left. - 2\sum_{j\leq i}^{l} \sum_{\substack{m=1 \\ m\neq i,j}}^{l} \Re(\mathbf{s}_i^\dagger \cdot \mathbf{R}_i^{-1} \cdot \mathbf{S}_{im} \cdot \mathbf{R}_m^{-1} \cdot \mathbf{S}_{mj} \cdot \mathbf{R}_j^{-1} \cdot \mathbf{s}_j) \right\}$$

$$+ O(\epsilon^6)$$

$$= \epsilon^2 \mathcal{S} - \frac{1}{2}\epsilon^4 \mathcal{N}^2 + O(\epsilon^6). \tag{74}$$

This is the optimal detection statistic for a weak stochastic background. We have identified $\mathcal{S}$ as the $O(\epsilon^2)$-term and $-2\mathcal{N}^2$ as the $O(\epsilon^4)$-term of the log-likelihood ratio.

The *locally-optimal* detection statistic is obtained in the $\epsilon \to 0$ limit; it is the leading $O(\epsilon^2)$ term,

$$\lim_{\epsilon\to 0} \frac{\ln \Lambda}{\epsilon^2} = \mathcal{S} = \sum_{i=1}^{l} \sum_{j<i}^{l} \Re(\mathbf{s}_i^\dagger \cdot \mathbf{R}_i^{-1} \cdot \mathbf{S}_{ij} \cdot \mathbf{R}_j^{-1} \cdot \mathbf{s}_j). \tag{75}$$

Although this presentation has been described in terms of observational vectors in the time domain, the derivation of the likelihood ratio has not explicitly required this choice of basis. It is convenient to perform a unitary transforma-

tion that diagonalizes the various correlation matrices. This transformation is called a Karhunen-Loeve transformation; for a stationary process with a correlation time much shorter than the time spanned by the l samples, the linear combinations of the time series that diagonalize the correlation matrices asymptotically approach the discrete Fourier transform. Therefore we can approximately express our result in the frequency domain where the $\mathbf{R}_i$ and $\mathbf{S}_{ij}$ matrices can be understood in terms of the power spectrum and the expectation value of the redshift cross correlation, respectively [cf. Eqs. (42) and (37)]. The locally-optimal detection statistic is therefore

$$\mathcal{S} = \frac{3H_0^2}{16\pi^3} \frac{1}{\beta} \sum_{i=1}^{l} \sum_{j<i}^{l} \int_{-\infty}^{\infty} \frac{\hat{\Omega}_{\mathrm{gw}}(|f|)^{(ij)}\Gamma(|f|)\tilde{s}_i^*(f)\tilde{s}_j(f)}{f^3 P_i(|f|)P_j(|f|)} df$$

$$= \frac{1}{2}\hat{\Omega}_\alpha T_0 \sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\sigma^{-2(ij)}S \tag{76}$$

where $\Omega_{\mathrm{gw}}(f) = \epsilon^2 \hat{\Omega}_{\mathrm{gw}}(f)$ and $\Omega_\alpha = \epsilon^2 \hat{\Omega}_\alpha$. This is the same optimal detection statistic S_{opt} of Eq. (66) (with the simplification of $n_{ij} = 1$) up to a normalization constant.

The locally-optimal statistic is optimal in the limit of weak signals. However, the likelihood ratio is only determined by this statistic up to a unknown factor which depends on the (unknown) strength of the signal. It is important now to distinguish between the assumed amplitude of the stochastic background, ϵ, and the true amplitude, ϵ_{true}. The true gravitational wave spectrum $\Omega_{\mathrm{gw}}(f)$ is now related to the template spectrum $\hat{\Omega}_{\mathrm{gw}}(f)$ via $\Omega_{\mathrm{gw}}(f) = \epsilon_{\mathrm{true}}^2 \hat{\Omega}_{\mathrm{gw}}$. To measure the strength of the stochastic background, given a set of pulsar observations, we can use the maximum likelihood estimator (MLE): the value of ϵ, ϵ_{MLE}, for which the likelihood ratio is a maximum. That is, we wish to find the value of ϵ_{MLE} for which $d\ln\Lambda/d\epsilon^2|_{\epsilon_{\mathrm{MLE}}} = 0$. From Eq. (74) we see that this estimate is

$$\epsilon_{\mathrm{MLE}}^2 = \mathcal{N}^{-2}\mathcal{S}, \tag{77}$$

where $\mathcal{N}^2$, from the $O(\epsilon^4)$-term of the log-likelihood ratio, is a normalizing factor which also includes the data. By substituting Eq. (77) into Eq. (74) we obtain the maximum likelihood detection statistic,

$$\max_\epsilon \ln\Lambda \simeq \frac{1}{2}\frac{\mathcal{S}^2}{\mathcal{N}^2}, \qquad (78)$$

where the terms of $O(\epsilon^6)$ have been discarded. Notice that this statistic is not simply the square of the cross correlation statistic. The data also appears in the factor $\mathcal{N}^{-2}$. This factor effectively suppresses elements of the pulsar network where the data measured greatly exceeds the normal noise level.

Some insight into the maximum likelihood detection statistic and the maximum likelihood amplitude estimate can be obtained by computing the expectation value of the log-likelihood ratio, Eq. (74). We find, to leading order in ϵ,

$$\langle \mathcal{N}^2 \rangle = \epsilon_{\text{true}}^{-2} \langle \mathcal{S} \rangle$$
$$= \left(\frac{3H_0}{32\pi^3}\right)^2 \frac{1}{\beta^2} T \sum_{i=1}^{l}\sum_{j<i}^{l} \int_{-\infty}^{\infty} df \frac{\hat{\Omega}_{\text{gw}}^2(|f|)^{(ij)}\Gamma^2(|f|)}{f^6 P_i(|f|)P_j(|f|)}. \qquad (79)$$

Therefore

$$\langle \ln\Lambda \rangle = \epsilon^2 \langle \mathcal{S} \rangle - \frac{1}{2}\epsilon^4 \langle \mathcal{N}^2 \rangle$$
$$= \epsilon^2\left(\epsilon_{\text{true}}^2 - \frac{1}{2}\epsilon^2\right)\left(\frac{3H_0}{32\pi^3}\right)^2$$
$$\times \frac{1}{\beta^2} T \sum_{i=1}^{l}\sum_{j<i}^{l} \int_{-\infty}^{\infty} df \frac{\hat{\Omega}_{\text{gw}}^2(|f|)^{(ij)}\Gamma^2(|f|)}{f^6 P_i(|f|)P_j(|f|)}$$
$$+ O(\epsilon^6). \qquad (80)$$

If we ignore the $O(\epsilon^6)$ terms, this is maximized when $\epsilon_{\text{MLE}} = \epsilon_{\text{true}}$, in which case

$$\max_\epsilon \langle \ln\Lambda \rangle \simeq \frac{1}{2}\left(\frac{3H_0}{32\pi^3}\right)^2$$
$$\times \frac{1}{\beta^2} T \sum_{i=1}^{l}\sum_{j<i}^{l} \int_{-\infty}^{\infty} df \frac{\Omega_{\text{gw}}^2(|f|)^{(ij)}\Gamma^2(|f|)}{f^6 P_i(|f|)P_j(|f|)}. \qquad (81)$$

This gives a scale of the value of the likelihood ratio we would expect to achieve.

IV. UPPER LIMITS AND DETECTION

Several methods exist in the LIGO literature that are appropriate for upper limit computation and detection using pulsar timing data [38–43]. These methods can be divided into two classes: Frequentist and Bayesian.

We expect that the optimal statistic Eq. (66) will be formed from a large number of pulsar pairs. For example,

the Parkes pulsar timing array [33,34] consists of 20 pulsars and the optimal statistic could be constructed from up to 190 cross correlation pairs. In this case we can make use of the central limit theorem: The distribution of $\hat{S}_{\text{opt}}$ should be well approximated by a Gaussian with a mean $\mu = \langle S_{\text{opt}} \rangle = \Omega_\alpha T_0$ and variance $\sigma_{\hat{\mu}}^2$ given by Eq. (68), namely,

$$p(\hat{S}_{\text{opt}}|\mu\sigma_{\hat{\mu}}) = \frac{1}{\sigma_{\hat{\mu}}\sqrt{2\pi}} \exp\left(\frac{-(\hat{S}_{\text{opt}} - \mu)^2}{2\sigma_{\hat{\mu}}^2}\right). \qquad (82)$$

A straightforward frequentist upper limit can then be set by finding the value of μ_{ul} such that in some predetermined fraction C (called the confidence) of hypothetical experiments, the value of the optimal statistic exceeds the actual value $\hat{S}_{\text{opt}}$ found in the search. In other words we would like to find the value μ_{ul} such that

$$\int_{\hat{S}_{\text{opt}}}^{\infty} dS_{\text{opt}} p(S_{\text{opt}}|\mu_{\text{ul}}\sigma_{\hat{\mu}}) = C. \qquad (83)$$

The solution to this is

$$\mu_{\text{ul}} = \hat{S}_{\text{opt}} + \sqrt{2}\sigma_{\hat{\mu}}\text{erfc}^{-1}(2(1 - C)). \qquad (84)$$

The assertion is that the real value of μ is less than μ_{ul} with confidence C, because if $\mu = \mu_{\text{ul}}$, a fraction C of the time we would have observed a value of S_{opt} greater than $\hat{S}_{\text{opt}}$. An equivalent, though potentially more robust, frequentist method to set upper limits involves performing simulated signal injections in the timing data set. Multiple injections are performed to determine the value of μ_{ul} such that a fraction C of the time the value of the optimal statistic measured in the data sets with injections exceeds the value found in the search. Frequentist detection methods such as Neyman-Pearson or maximum likelihood are well described in the literature (see, for example, [38] and references therein) and we will not discuss them here. Additionally Feldman and Cousins [50] provide a means to smoothly transition between upper limits and detection.

Bayesian upper limits can be computed by constructing a posterior distribution using the value of the optimal statistic found in the search, and variance along with priors. We begin by applying the product rule to the probability density of μ along with the measured value $\hat{S}_{\text{opt}}$ given $\sigma_{\hat{\mu}}$ to write

$$p(\mu\hat{S}_{\text{opt}}|\sigma_{\hat{\mu}}) = p(\mu|\hat{S}_{\text{opt}}\sigma_{\hat{\mu}})p(\hat{S}_{\text{opt}}|\sigma_{\hat{\mu}})$$
$$= p(\hat{S}_{\text{opt}}|\mu\sigma_{\hat{\mu}})p(\mu|\sigma_{\hat{\mu}}), \qquad (85)$$

then solve for $p(\mu|\hat{S}_{\text{opt}}\sigma_{\hat{\mu}})$ to obtain Bayes' theorem

$$p(\mu|\hat{S}_{\text{opt}}\sigma_{\hat{\mu}}) = p(\hat{S}_{\text{opt}}|\mu\sigma_{\hat{\mu}})\frac{p(\mu|\sigma_{\hat{\mu}})}{p(\hat{S}_{\text{opt}}|\sigma_{\hat{\mu}})}, \qquad (86)$$

the posterior probability density for μ, or equivalently Ω_α. One can then choose a prior $p(\mu|\sigma_{\hat{\mu}})$ (for example requiring $\mu > 0$) and normalize the probability distribution [the

probability $p(\hat{S}_{\text{opt}}|\sigma_{\hat{\mu}})$ does not depend on μ so it is a prior dependent normalization constant], and find the μ_{ul} such that

$$\mathcal{M} \int_{-\infty}^{\mu_{\text{ul}}} d\mu\, p(\mu|\hat{S}_{\text{opt}}\sigma_{\hat{\mu}}) = C, \tag{87}$$

where $\mathcal{M}$ is the normalization constant. For sufficiently simple choices of the prior distribution $p(\mu|\sigma_{\hat{\mu}})$ the integral equation (87) can be performed analytically to obtain the Bayesian analog of Eq. (84). As with frequentist methods [38,50], Bayesian detection methods involve selecting thresholds, in this case on the odds ratio, which is the ratio of the posteriors, suitably integrated over, say, different ranges of μ. For more details see Refs. [51,52].

V. A PULSAR TIMING RADIOMETER: CONSTRUCTING SKYMAPS OF THE STOCHASTIC BACKGROUND

A skymap may be created by computing $\Omega_{\text{gw}}(f)$ for a collection of pixels in the sky. We do this by assuming that the only signal present comes from a single location on the sky. We begin by relaxing the assumption that the stochastic background is isotropic. That is, we take

$$\langle h_A^*(f, \hat{\Omega})h_{A'}(f', \hat{\Omega}')\rangle = \frac{3H_0^2}{32\pi^3}\delta^2(\hat{\Omega}, \hat{\Omega}')\delta_{AA'}\delta(f - f')$$
$$\times P(\hat{\Omega})|f|^{-3}\Omega_{\text{gw}}(|f|), \tag{88}$$

where $P(\hat{\Omega})$ is the strength or brightness [53] of gravitational waves from the direction $\hat{\Omega}$.

In this case, the overlap reduction function takes the modified form

$$^{(ij)}\Gamma_P = \frac{3}{4\pi}\sum_A \int_{S^2} d\hat{\Omega}\, P(\hat{\Omega})F_i^A(\hat{\Omega})F_j^A(\hat{\Omega}). \tag{89}$$

where we have ignored the pulsar phase factors, and the optimal filter is given by

$$^{(ij)}\tilde{Q}_P(f) = {}^{(ij)}\chi\frac{^{(ij)}\Gamma_P\Omega_{\text{gw}}(|f|)}{|f|^3 P_i(|f|)P_j(|f|)}, \tag{90}$$

where we have suppressed the k index which specifies the particular measurement of the (ij) pulsar pair. We can further optimize for point sources by taking $P(\hat{\Omega}) = \delta^2(\hat{\Omega} - \hat{\Omega}')$. The optimal filter then becomes

$$^{(ij)}\tilde{Q}_{\hat{\Omega}}(f) = {}^{(ij)}\chi\frac{^{(ij)}\Gamma_{\hat{\Omega}}\Omega_{\text{gw}}(|f|)}{|f|^3 P_i(|f|)P_j(|f|)}, \tag{91}$$

with

$$^{(ij)}\Gamma_{\hat{\Omega}} = \frac{3}{4\pi}\sum_A F_i^A(\hat{\Omega})F_j^A(\hat{\Omega}). \tag{92}$$

Figure 2 shows two examples of the sky location dependent overlap reduction function. The top panel shows $|\Gamma_{\hat{\Omega}}|$ from

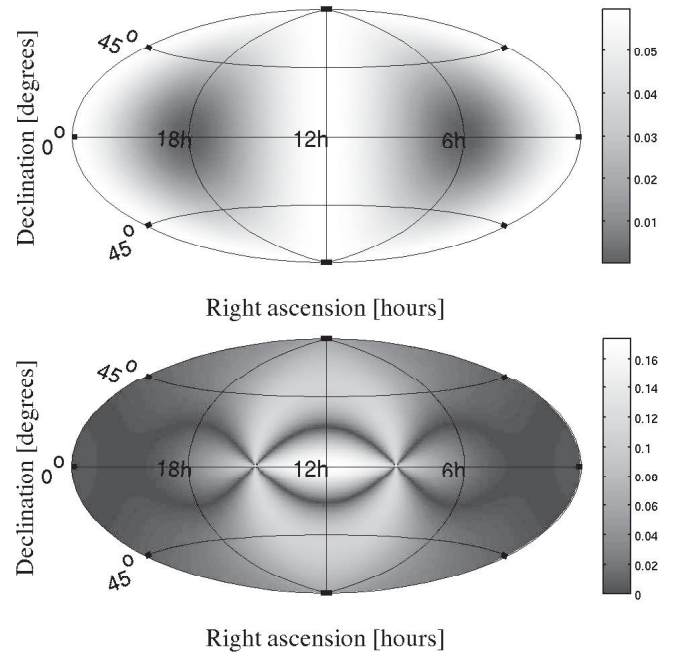

FIG. 2. Plots of $|^{(ij)}\Gamma_{\hat{\Omega}}|$ from Eq. (92) for two pulsars with $\xi = \pi/2$ (top panel) and $\xi = \pi$ degrees (bottom panel).

Eq. (92) for two pulsars with $\xi = \pi/2$ located at 0° declination (Dec) and 9 h and 15 h right ascension (RA), respectively. The bottom panel shows the same quantity for two pulsars with $\xi = \pi$ located at 0° Dec and 6 h and 18 h RA, respectively.

One could also imagine computing the overlap reduction function for each term in a multipole expansion of $P(\hat{\Omega})$. The overlap reduction function for the monopole term in the expansion (appropriate for an isotropic stochastic gravitational wave search) is the Hellings-Downs curve given by Eq. (39) in Sec. III. Surprisingly, the dipole overlap reduction function is given by a similarly simple equation. We find

$$\Gamma_{\text{dip}} = (\cos\alpha_1 + \cos\alpha_2)\left(2 - \frac{3}{2}\cos\xi + 6\tan^2\frac{\xi}{2}\ln\left(\sin\frac{\xi}{2}\right)\right), \tag{93}$$

where as before ξ is the angle between the two pulsars, and α_1 and α_2 are the angles each of two pulsars make to the direction of the dipole. A detailed derivation of this result is given in Appendix C 2. This result is relevant to searches for a dipole anisotropy in the gravitational wave sky using pulsar timing data.

The sky dependence of the sensitivity of a pulsar network can be estimated by computing the signal to noise for sources at the sky locations of interest. We start by taking the expectation value of the optimal statistic, Eq. (66), using the optimal filter for a sky location $\hat{\Omega}$ assuming the redshift data contain a stochastic signal from that location. We then divide by the square root of the variance given in
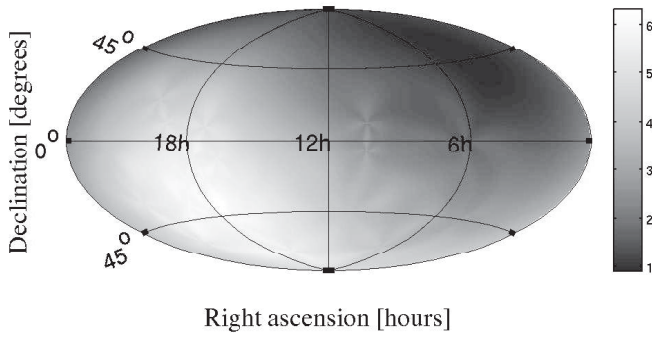
FIG. 3.   Skymap of the sensitivity, Eq. (94), for the Parkes pulsar timing array.

Eq. (68). The result is proportional to

$$G(\hat{\Omega}) = \left(\sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\Gamma_{\hat{\Omega}}^2 \right)^{1/2}, \qquad (94)$$

where we have assumed (for illustrative purposes) that the noise spectra of all pulsars is the same, the observation times for all pairs is the same, and $n_{ij} = 1$ for all pulsar pairs. Figure 3 shows the Eq. (94), for the 20 pulsars of the Parkes pulsar timing array [33,34]. Since most of the pulsars are in the Southern Hemisphere the Parkes pulsar timing array is most sensitive in that region.

Another quantity of interest is the point spread function, which measures the intrinsic spatial correlation of the sky-map, or equivalently, the ability of a pulsar network to locate a stochastic source of gravitational waves. We construct the point spread function by computing the signal to noise for a source at some sky location that we search for using the optimal filter for some other location. In particular, we take the expectation value of the optimal statistic, Eq. (66), using the optimal filter for a sky location $\hat{\Omega}$ assuming the redshift data contain a signal from another location $\hat{\Omega}'$, then we divide by the square root of the variance given in Eq. (68). The result is proportional to

$$A(\hat{\Omega}, \hat{\Omega}') = \frac{\sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\Gamma_{\hat{\Omega}}{}^{(ij)}\Gamma_{\hat{\Omega}'}}{\left(\sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\Gamma_{\hat{\Omega}}^2 \right)^{1/2}}, \qquad (95)$$

where we have again assumed that the noise spectra of all pulsars are the same, the observation times for all pairs are the same, and $n_{ij} = 1$ for all pulsar pairs. Figure 4 shows the point spread function, Eq. (95), for the 20 pulsars of the Parkes pulsar timing array [33,34] for a source at 6 h RA 45° Dec (top panel) and another at 18 h RA −45° Dec (bottom panel). Current earth-bound detectors operate in a different frequency band, around 100 Hz. This means that the time of arrival of gravitational waves at different terrestrial detectors can be used for additional angular resolution relative to pulsar timing experiments–the achievable angular resolution is given by the detector separation and observation frequency. For a pulsar timing array there is no
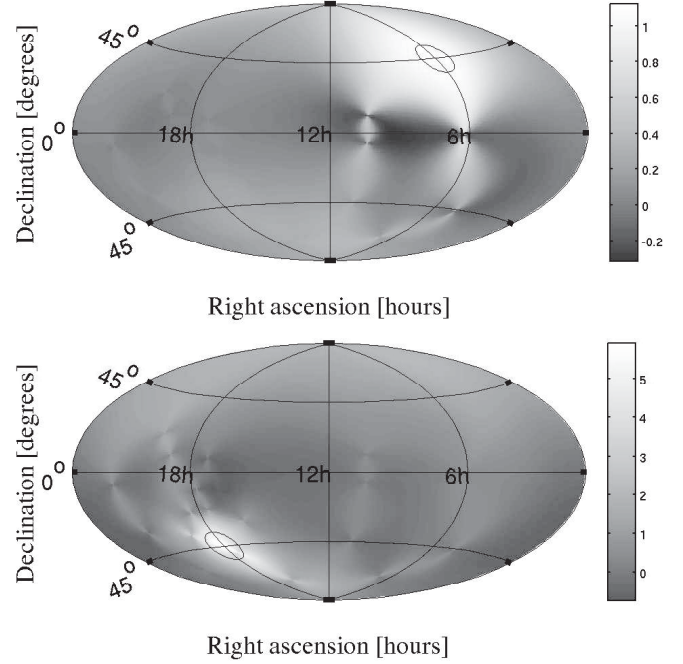

FIG. 4.   Plot of the point spread function Eq. (95) for the Parkes pulsar timing array for a source at 6 h RA 45° Dec (top panel) and 18 h RA −45° Dec (bottom panel).

observable time of arrival difference and the angular resolution is determined by the pulsar density in the direction of the signal (see Fig. 4).

The point spread function can be understood in terms of the likelihood ratio of Sec. III D: Suppose that the likelihood ratio is computed using the overlap reduction function $\Gamma_{\hat{\Omega}}$ appropriate for a stochastic signal coming from direction $\Gamma_{\hat{\Omega}}$ when the true signal is in fact coming from direction $\Gamma_{\hat{\Omega}}$. The expectation value of the log-likelihood ratio is [cf. Eq. (80)]

$$\langle \ln\Lambda \rangle = \epsilon^2 \epsilon_{\text{true}}^2 \sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\Gamma_{\hat{\Omega}}{}^{(ij)}\Gamma_{\hat{\Omega}'}{}^{(ij)}C$$

$$- \frac{1}{2}\epsilon^4 \sum_{i=1}^{l} \sum_{j<i}^{l} {}^{(ij)}\Gamma_{\hat{\Omega}}^2 {}^{(ij)}C + O(\epsilon^6), \qquad (96)$$

with

$$^{(ij)}C = \left(\frac{3H_0}{32\pi^3} \right)^2 T \int_{-\infty}^{\infty} df \frac{\hat{\Omega}_{\text{gw}}^2(|f|)}{f^6 P_i(|f|)P_j(|f|)}. \qquad (97)$$

If $^{(ij)}C$ is approximately the same for all pulsar pairs then

$$\max_{\epsilon} \langle \ln\Lambda \rangle \propto A^2(\hat{\Omega}, \hat{\Omega}'). \qquad (98)$$

In this sense the point spread function describes the degree to which the position of a point source of stochastic gravitational waves can be located in terms of the likelihood ratio.

VI. ISSUES WITH PULSAR TIMING DATA

We have derived the optimal statistic for detecting a stochastic background. In this section we would like to discuss some issues associated with departures from the idealizations made to arrive at the optimal statistic.

A. Colored noise and nonstationarity

In contrast to previous methods [5,6] the techniques presented here do not rely on the data being white. The power spectra $P_i(|f|)$ in the optimal statistic account for colored noise. However, the methods assume the data is stationary.

If the data is nonstationary over long time scales it can be divided into short stationary (or almost stationary) stretches and the power spectrum can be estimated using the Lomb-Scargle periodogram described below, or modeled for each stretch. The optimally filtered data stretches can then be combined along the lines discussed in Sec. III B. One concern associated with breaking the data up into small stretches is loss of low frequency information: Gravitational waves with periods larger than the length of the short stretches will be lost in this procedure. The problem can be avoided by first computing the quantities $s_i(f)/P_i(|f|)$ for each of the short stretches and then combining them using the Dirichlet kernel to construct full time baseline versions of these quantities.

If the spectrum is measured it can be smoothed by performing a running average over a small frequency window, which if the data are stationary in the stretch the spectrum is estimated, is equivalent to ensemble averaging.

B. Unevenly sampled data

The fact that pulsar timing measurements are not taken continuously leads to a data set that is unevenly sampled in time. This poses a problem for frequency-domain analyses not present in their time-domain counterparts. The authors of [44] will explore the time-domain approach in detail. It is unclear which of these two methods will turn out to be more robust. In what follows we address the specific issues of computing periodograms and Fourier transforms for unevenly sampled data sets which we think is useful in any case.

1. The Lomb-Scargle periodogram

The problem of constructing periodograms from unevenly sampled data comes up in the data analysis of variable stars. It was in precisely this context that Lomb [54] and Scargle [55] proposed a least-squares solution to the problem. The basic idea is as follows: Let $x(t_i)$ be a time series with zero mean sampled at $i = 0 \cdots N - 1$ unevenly spaced times. Now fit the time series by finding the coefficients $a_{\min}$ and $b_{\min}$ that minimize the square of the residual

$$r^2(f) \equiv \sum_{i=0}^{N-1} \{x(t_i) - a \cos[2\pi f(t_i - \tau(f))] - b \sin[2\pi f(t_i - \tau(f))]\}^2, \qquad (99)$$

where

$$\tan(4\pi f \tau(f)) = \frac{\sum_{i=0}^{N-1} \sin(4\pi f t_i)}{\sum_{i=0}^{N-1} \cos(4\pi f t_i)}. \qquad (100)$$

Then the periodogram is defined up to normalization by the difference

$$\Delta r^2(f) = \sum_{i=0}^{N-1} x^2(t_i) - r_{\min}^2(f)$$
$$= \frac{(\sum_{i=0}^{N-1} x(t_i) \cos[2\pi f(t_i - \tau(f))])^2}{\sum_{i=0}^{N-1} \cos^2[2\pi f(t_i - \tau(f))]}$$
$$+ \frac{(\sum_{i=0}^{N-1} x(t_i) \sin[2\pi f(t_i - \tau(f))])^2}{\sum_{i=0}^{N-1} \sin^2[2\pi f(t_i - \tau(f))]}, \qquad (101)$$

where $r_{\min}^2(f)$ is the quantity in Eq. (99) with $a = a_{\min}$ and $b = b_{\min}$. After normalization [56] and generalization to data with nonzero mean, we have the Lomb-Scargle periodogram

$$P_X^{\mathrm{LS}}(f) = \frac{1}{2\sigma_x^2} \left\{ \frac{(\sum_{i=0}^{N-1}[x(t_i) - \mu_x] \cos[2\pi f(t_i - \tau(f))])^2}{\sum_{i=0}^{N-1} \cos^2[2\pi f(t_i - \tau(f))]} \right.$$
$$\left. + \frac{(\sum_{i=0}^{N-1}[x(t_i) - \mu_x] \sin[2\pi f(t_i - \tau(f))])^2}{\sum_{i=0}^{N-1} \sin^2[2\pi f(t_i - \tau(f))]} \right\}, \qquad (102)$$

where μ_x and σ_x^2 are the mean and variance, respectively, of $x(t_i)$. Note that the definition of $\tau(f)$ in Eq. (100) ensures that the resulting periodogram is independent of where $t = 0$.

2. Fourier transforms

The idea of using a least-squares minimization is also useful for constructing Fourier transforms. To do so, we borrow an idea from the radar community [57]. Suppose we have a time series, $x(t_i)$, nonuniformly sampled at times $t_0 \cdots t_{N-1}$ and we wish to approximate its Fourier transform over M evenly spaced frequencies, $f_0 \cdots f_{M-1}$. The strategy we will employ is to use a least-squares procedure to find the best fit to the original time series after an inverse Fourier transform. That is, the (squared) residual to be minimized is given by

$$r^2 = \sum_{j=0}^{N-1} \left| x(t_j) - \sum_{k=0}^{M-1} \tilde{\xi}(f_k) e^{i2\pi f_k t_j} \right|^2, \qquad (103)$$

where the $\tilde{\xi}(f_k)$ are least-squares Fourier coefficients we set out to determine. Defining

$$A_{kj} = e^{i2\pi f_k t_j}, \qquad (104)$$

we can write

$$r^2 = \|\vec{x} - A\vec{\xi}^2\|. \tag{105}$$

The least-squares solution to this problem is given by

$$\vec{\xi} = (A^\dagger A)^{-1} A^\dagger \vec{x}. \tag{106}$$

The problem is then purely a computational one, which, because of the limited amount of pulsar timing data available, is completely tractable on a modern computer, regardless of the efficiency of the algorithm. On a final note, one can actually improve upon this procedure [57] by weighting the residual in Eq. (103) by the square root of the variance, $\sigma_{x(t_j)}$, associated with each data point,

$$r^2 = \sum_{j=0}^{N-1} \frac{1}{\sigma_{x(t_j)}} \left| x(t_j) - \sum_{k=0}^{M-1} \tilde{\xi}_k e^{i2\pi f_k t_j} \right|^2, \tag{107}$$

which has the advantage of automatically including the error bars associated with individual pulsar timing data points.

VII. SUMMARY AND CONCLUSIONS

A stochastic background of gravitational waves could be detected via pulsar timing observations in the next 5 to 10 yr. This background may be astrophysical, such as that produced by supermassive black holes, or cosmological, such as that produced by a network of cosmic (super) strings. In the latter case a detection would open a window onto a time in the early universe prior to recombination and could have profound consequences. Leveraging techniques developed for ground-based instruments such as LIGO and Virgo, in this paper we have shown how to optimally extract the signal produced by a stochastic background of gravitational waves using cross correlations of timing data from a pulsar timing array.

We started by considering the redshift induced by a gravitational wave on the frequency of arrival of radio pulses from a pulsar first derived by Detweiler [2]. The redshift is proportional to the difference in the metric perturbation at the pulsar (when a pulse is emitted) and at the Earth (when that pulse is received). Using a convenient coordinate-independent description of the signal we examined the form of the signal in the frequency domain. The term involving the metric perturbation at the pulsar is typically neglected because it can be treated as a sort of noise term which averages to zero in correlations of timing measurements of different pulsars. In the frequency domain the dependence on the metric perturbation at the pulsar is in a phase factor that depends on the distance to the pulsar. It is possible that if we could determine the distance to pulsars with sufficient accuracy we could use the metric perturbation at the pulsar to improve the sensitivity of our searches. Unfortunately, accurate measurements of pulsar distances are unavailable. By first finding

the optimal cross correlation filter, we have shown that for pulsar distances and gravitational wave frequencies typical of pulsar timing experiments, the metric perturbation at the pulsar can be neglected without a significant deviation from optimality. It is unclear whether this is true for other types of gravitational wave searches. We have also determined the optimal way to combine pulsar timing data from a pulsar timing array, which is constructed from pairs of optimally filtered cross correlations. We have assumed throughout that the gravitational wave is smaller than the noise in the sense that it does not affect our noise estimation. When the first detection of gravitational waves is made using pulsar timing data we expect to be in the regime where the signal is small, and the techniques presented here should be near optimal. As pulsar timing accuracies improve we may move into a regime where the gravitational wave signal is appreciable in the timing residuals of individual pulsars. At this point different techniques may become necessary [44].

We have discussed and illustrated frequentist and Bayesian methods for setting upper limits using the distribution of the optimal statistic. We have shown how to construct a pulsar timing radiometer: A map of the sky created by optimizing the cross correlation statistic for particular sky directions. We have also shown how to determine the intrinsic spatial correlation of such maps, which in turn determines the ability of a pulsar timing array to locate a source of stochastic gravitational waves.

We have ended with a discussion of some problems related to realistic analysis of pulsar timing data, particularly the issues of nonstationarity and uneven sampling. The optimal filter is constructed from power spectra of the pulsar timing data, which can be modeled or measured, and accounts for the effects of colored noise. We have described a technique, the Lomb-Scargle periodogram, for robust spectrum estimation that can be used to construct the optimal filter. The optimal filter can then be applied in the frequency domain and we have described a procedure for taking Fourier transforms of unevenly sampled data that accounts for error bars in the individual pulsar timing data points. The optimal filter can also be inverse Fourier transformed and applied in the time domain where uneven sampling is not an issue. Regardless of which method turns out to be more useful and robust for stochastic background searches, we believe the development of Fourier techniques for unevenly sampled data will be beneficial. Lommen, Romano, and Woan [44] will examine time-domain methods in detail.

APPENDIX A: DERIVATION OF DETWEILER'S FORMULA USING THE GEODESIC EQUATION

For completeness of presentation, we include a derivation of Detweiler's formula. We consider, as we did in Sec. II, the metric perturbation due to a single gravitational wave traveling in the $\hat{\Omega} = \hat{z}$ direction, so Eqs. (4)–(6) hold. Then, if a vector s^a is null in Minkowski space, the corresponding null vector, σ^a, in the perturbed spacetime $g_{ab} = \eta_{ab} + h_{ab}$ is given by [58]

$$\sigma^a = s^a - \tfrac{1}{2}\eta^{ab}h_{bc}s^c. \tag{A1}$$

The null vector in Minkowski space that points from the pulsar to the Solar Sytem is $s^a = \nu(1, -\alpha, -\beta, -\gamma)$, where, as before, α, β, and γ are the direction cosines with the x, y, and z directions, respectively. The corresponding perturbed vector is readily computed from Eq. (A1)

$$\sigma^a = \nu \begin{pmatrix} 1 \\ -\alpha(1 - \tfrac{1}{2}h_+) + \tfrac{1}{2}\beta h_\times \\ -\beta(1 + \tfrac{1}{2}h_+) + \tfrac{1}{2}\alpha h_\times \\ -\gamma \end{pmatrix}. \tag{A2}$$

The geodesic equation tells us that the t component of σ^a satisfies

$$\frac{d\sigma^t}{d\lambda} = -\Gamma^t_{ab}\sigma^a\sigma^b. \tag{A3}$$

It follows from the form of the metric perturbation in Eq. (5) that

$$\Gamma^t_{ab} = -\frac{1}{2}g^{tc}\left[\frac{\partial g_{bc}}{\partial x^a} + \frac{\partial g_{ac}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^c}\right] = \frac{1}{2}\dot{g}_{ab}$$

$$= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \dot{h}_+ & \dot{h}_\times & 0 \\ 0 & \dot{h}_\times & -\dot{h}_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{A4}$$

where the overdot denotes a derivative with respect to t. The geodesic equation then reads

$$\frac{d\sigma^t}{d\lambda} = -\frac{1}{2}\dot{g}_{ab}\sigma^a\sigma^b$$

$$= -\frac{1}{2}[\dot{g}_{xx}(\sigma^x)^2 - \dot{g}_{yy}(\sigma^y)^2] - \dot{g}_{xy}\sigma^x\sigma^y$$

$$= -\frac{1}{2}\dot{h}_+[(\sigma^x)^2 - (\sigma^y)^2] + \dot{h}_\times\sigma^x\sigma^y. \tag{A5}$$

After a little algebra Eq. (A2) leads to

$$(\sigma^x)^2 - (\sigma^y)^2 = \nu^2(\alpha^2 - \beta^2) + \mathcal{O}(h), \tag{A6}$$

as well as

$$\sigma^x\sigma^y = \nu^2\alpha\beta + \mathcal{O}(h), \tag{A7}$$

so that

$$-\frac{d\nu}{d\lambda} = \frac{1}{2}\dot{h}_+\nu^2(\alpha^2 - \beta^2) + \dot{h}_\times\nu^2\alpha\beta. \tag{A8}$$

We proceed by writing the time derivatives in Eq. (A8) as derivatives with respect to the affine parameter λ. In particular, since $h_{+,\times} = h_{+,\times}(t - z)$ we have

$$\frac{dh_{+,\times}}{d\lambda} = \frac{\partial h_{+,\times}}{\partial t}\frac{dt}{d\lambda} + \frac{\partial h_{+,\times}}{\partial z}\frac{dz}{d\lambda}. \tag{A9}$$

Moreover, we also have that the frequency $\nu = dt/d\lambda$, in addition to $\partial h_{+,\times}/\partial z = -\partial h_{+,\times}/\partial t$ and $dz/d\lambda = -\nu\gamma$. Therefore we can write Eq. (A9) as

$$\dot{h}_{+,\times} = \frac{1}{\nu(1 + \gamma)}\frac{dh_{+,\times}}{d\lambda}. \tag{A10}$$

Then Eq. (A8), the geodesic equation, becomes

$$-\frac{1}{\nu}\frac{d\nu}{d\lambda} = \frac{1}{2}\frac{\alpha^2 - \beta^2}{1 + \gamma}\frac{dh_+}{d\lambda} + \frac{\alpha\beta}{1 + \gamma}\frac{dh_\times}{d\lambda}, \tag{A11}$$

which we integrate to find

$$\frac{\nu(t)}{\nu_0} = \exp\left[-\frac{1}{2}\frac{\alpha^2 - \beta^2}{1 + \gamma}\Delta h_+ - \frac{\alpha\beta}{1 + \gamma}\Delta h_\times\right]. \tag{A12}$$

It is worth pointing out that the direction cosines are functions of the affine parameter. The dependence is in terms of $\mathcal{O}(h)$, and we have neglected this dependence in going from Eq. (A11) to Eq. (A12). The final result is obtained by expanding this expression to first order in $h_{+,\times}$,

$$\frac{\nu_0 - \nu(t)}{\nu_0} = \frac{1}{2}\frac{\alpha^2 - \beta^2}{1 + \gamma}\Delta h_+ + \frac{\alpha\beta}{1 + \gamma}\Delta h_\times, \tag{A13}$$

where $\Delta h_{+,\times} = h^p_{+,\times} - h^e_{+,\times}$ is the difference between the metric perturbation at the pulsar and the detector. This expression is precisely Eq. (7).

APPENDIX B: LINEARITY OF THE REDSHIFT

In this appendix we will explicitly demonstrate that the total redshift is the sum of the contributions from gravitational waves in every direction, as written in Eq. (18). The derivation is merely a generalization of the results derived in the previous appendix. We begin by considering a metric perturbation, h_{ab}, in a spatial transverse-traceless gauge comprised of the sum of metric perturbations, $h_{ab}^{(i)}$, from gravitational waves in N different directions, $\hat{\Omega}_{(i)}$. Namely,

$$h_{ab} = \sum_{(i)}^{N} h_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x}), \tag{B1}$$

where t and $\vec{x}$ form a four vector, x^a, in the background (Minkowski) geometry. Adjusting the notation from Appendix A

$$s^a = dx^a/d\lambda = \nu(1, -\alpha, -\beta, -\gamma) \equiv \nu(1, -\hat{p}), \tag{B2}$$

as a null vector in Minkowski space. The null geodesic is perturbed by Eq. (B1), resulting in a

$$\sigma^a = s^a + \delta s^a. \tag{B3}$$

As before, our interest is in the quantity

$$\frac{d\sigma^t}{d\lambda} = -\Gamma^t_{ab}\sigma^a\sigma^b. \tag{B4}$$

The spatial nature of the gauge we have chosen ensures that

$$\Gamma^t_{ab} = \tfrac{1}{2}\dot{g}_{ab}, \tag{B5}$$

$$= \tfrac{1}{2}\dot{h}_{ab}, \tag{B6}$$

which is evident from the first line of Eq. (A4). It follows from these definitions that

$$\begin{aligned}
\frac{d\sigma^t}{d\lambda} &= -\frac{1}{2}\dot{h}_{ab}\sigma^a\sigma^b \\
&= -\frac{1}{2}\dot{h}_{ab}(s^a + \delta s^a)(s^b + \delta s^b) \\
&= -\frac{1}{2}\dot{h}_{ab}s^a s^b \\
&= -\frac{1}{2}\dot{h}_{ij}(\nu^2 p^i p^j),
\end{aligned} \tag{B7}$$

where i and j are spatial indices. As before, we want to write the expression above in terms of the affine parameter along s^a. We begin by noting that for term in Eq. (B1)

$$\begin{aligned}
\frac{dh_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x})}{d\lambda} &= \frac{\partial h_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x})}{\partial t}\frac{dt}{d\lambda} \\
&\quad + \frac{\partial h_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x})}{\partial(\hat{\Omega}_{(i)} \cdot \vec{x})}\frac{d(\hat{\Omega}_{(i)} \cdot \vec{x})}{d\lambda} \\
&= \frac{\partial h_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x})}{\partial t}\nu \\
&\quad - \frac{\partial h_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x})}{\partial t}\hat{\Omega}_{(i)} \cdot \frac{d\vec{x}}{d\lambda} \\
&= \nu(1 + \hat{\Omega}_{(i)} \cdot \hat{p})\frac{\partial h_{ab}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x})}{\partial t},
\end{aligned} \tag{B8}$$

where we have used the $(t - \hat{\Omega}_{(i)} \cdot \vec{x})$ dependence of the metric perturbation to write the spatial derivatives as time derivatives along with Eq. (B2). Putting this together with Eq. (B7), the result is that

$$-\frac{1}{\nu}\frac{d\nu}{d\lambda} = \sum_{(i)}^{N}\frac{1}{2}\frac{\hat{p}^i \hat{p}^j}{1 + \hat{\Omega}_{(i)} \cdot \hat{p}}h_{ij}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x}), \tag{B9}$$

which can be integrated to give the redshift

$$z \equiv \frac{\nu_0 - \nu(t)}{\nu_0} = \sum_{(i)}^{N}\frac{1}{2}\frac{\hat{p}^i \hat{p}^j}{1 + \hat{\Omega}_{(i)} \cdot \hat{p}}\Delta h_{ij}^{(i)}(t - \hat{\Omega}_{(i)} \cdot \vec{x}), \tag{B10}$$

which is the discrete version of Eq. (18).

APPENDIX C: THE OVERLAP REDUCTION FUNCTION IN THE HIGH-FREQUENCY LIMIT

1. Derivation of the Hellings-Downs curve

In this section we derive the Hellings and Downs curve [48] given by Eq. (39). We begin with the definition of the overlap reduction function, Eq. (38), and we ignore the exponential factors. Thus we wish to evaluate

$$\Gamma_0 = \beta \sum_{A=+,\times}\int_{S^2} d\hat{\Omega}\, F_1^A(\hat{\Omega})F_2^A(\hat{\Omega}), \tag{C1}$$

and using the definition of $F^A(\hat{\Omega})$ given by Eq. (17) we find

$$\begin{aligned}
\Gamma_0 &= \frac{1}{4}\beta \sum_{A=+,\times}\int_{S^2} d\hat{\Omega}\, \frac{\hat{p}_1^i \hat{p}_1^j}{1 + \hat{\Omega} \cdot \hat{p}_1} \\
&\quad \times \frac{\hat{p}_2^k \hat{p}_2^l}{1 + \hat{\Omega} \cdot \hat{p}_2}e_{ij}^A(\hat{\Omega})e_{kl}^A(\hat{\Omega}).
\end{aligned} \tag{C2}$$

The two unit vectors $\hat{p}_1$ and $\hat{p}_2$ are those pointing from the Earth toward the first and second pulsar, respectively, and the polarization tensors $e_{ij}^+(\hat{\Omega})$ and $e_{ij}^\times(\hat{\Omega})$ for a gravitational wave traveling in direction $\hat{\Omega}$ are given by Eqs. (2a) and (2b) respectively. To evaluate the integral we choose a

coordinate system in which $\hat{p}_1$ is parallel to the z-axis and $\hat{p}_2$ is in the $x - z$ plane. Then

$$\hat{p}_1 = (0, 0, 1), \tag{C3a}$$

$$\hat{p}_2 = (\sin\xi, 0, \cos\xi), \tag{C3b}$$

$$\Gamma_0 = -\frac{1}{4}\beta \int_{S^2} d\hat{\Omega} \frac{\sin^2\theta(\sin^2\xi\sin^2\phi - \sin^2\xi\cos^2\theta\cos^2\phi - \cos^2\xi\sin^2\theta + 2\sin\xi\cos\xi\sin\theta\cos\theta\cos\phi)}{(1 + \cos\theta)(1 + \cos\xi\cos\theta + \sin\xi\sin\theta\cos\phi)}. \tag{C4}$$

Straightforward manipulation shows that this integral becomes

$$\Gamma_0 = \tfrac{1}{4}\beta(I + J), \tag{C5}$$

with

$$I = \int_{S^2} d\hat{\Omega}(1 - \cos\theta)(1 - \cos\xi\cos\theta - \sin\xi\sin\theta\cos\phi)$$

$$= 4\pi\left(1 + \frac{1}{3}\cos\xi\right), \tag{C6}$$

and

$$J = -2\sin^2\xi \int_0^\pi d\theta \sin\theta(1 - \cos\theta)K, \tag{C7}$$

where we have defined

$$K \equiv \int_0^{2\pi} d\phi \frac{\sin^2\phi}{1 + \cos\xi\cos\theta + \sin\xi\sin\theta\cos\phi}. \tag{C8}$$

K may be trivially evaluated by contour integration in the complex plane. The result is

$$K = 2\pi \frac{1 + \cos\xi\cos\theta + |\cos\xi + \cos\theta|}{\sin^2\xi\sin^2\theta}$$

$$= 2\pi\left(\frac{1 \mp \cos\xi}{\sin^2\xi}\right)\left(\frac{1 \mp \cos\theta}{\sin^2\theta}\right), \tag{C9}$$

where the negative sign applies when $0 < \theta < \pi - \xi$ and the positive sign applies when $\pi - \xi < \theta < \pi$. Hence we find that

$$J = -4\pi(1 - \cos\xi) \int_0^{\pi-\xi} d\theta \frac{(1 - \cos\theta)^2}{\sin\theta}$$

$$- 4\pi(1 + \cos\xi) \int_{\pi-\xi}^\pi d\theta \sin\theta$$

$$= 16\pi(1 - \cos\xi) \ln\left(\sin\frac{\xi}{2}\right). \tag{C10}$$

where ξ is the angular separation between the two pulsars. Because we have chosen coordinates in which $\hat{p} \cdot \hat{m} = 0$ [cf. Eq. (3b)], the $\times$-polarization terms vanish and our expression for Γ_0 becomes

Combining Eqs. (C5), (C6), and (C10), we obtain

$$\Gamma_0 = \frac{4\pi}{3}\beta\left\{1 + 3(1 - \cos\xi)\left[\ln\left(\sin\frac{\xi}{2}\right) - \frac{1}{12}\right]\right\}$$

$$= \frac{4\pi}{3}\beta\left\{1 + \frac{3}{2}(1 - \cos\xi)\left[\ln\left(\frac{1 - \cos\xi}{2}\right) - \frac{1}{6}\right]\right\}. \tag{C11}$$

The expression in braces achieves a maximum value of unity when $\xi = 0$, so the correct normalization constant is $\beta = 3/4\pi$. With this normalization we recover Eq. (39).

2. Generalization to a dipole stochastic background

We now generalize the Hellings-Downs curve to the case of a stochastic background with a dipole moment in the direction $\hat{D}$. We will start by defining the following quantities:

$$\hat{D} = (\sin\alpha_1\cos\eta, \sin\alpha_1\sin\eta, \cos\alpha_1), \tag{C12}$$

$$\hat{D} \cdot \hat{\Omega} \equiv \cos\chi = \cos\alpha_1\cos\theta + \sin\alpha_1\sin\theta\cos(\phi - \eta), \tag{C13}$$

$$\hat{D} \cdot \hat{p}_1 \equiv \cos\alpha_1, \tag{C14a}$$

$$\hat{D} \cdot \hat{p}_2 \equiv \cos\alpha_2$$

$$= \cos\alpha_1\cos\xi + \sin\alpha_1\sin\xi\cos\eta. \tag{C14b}$$

This derivation differs from the derivation of the Hellings-Downs curve only in that a factor $\hat{D} \cdot \hat{\Omega}$ must be included in the integral.

$$\Gamma_{\text{dip}} = \frac{1}{4}\beta \sum_{A=+,\times} \int_{S^2} d\hat{\Omega} \frac{\hat{p}_1^i \hat{p}_1^j \hat{p}_2^k \hat{p}_2^l e_{ij}^A(\hat{\Omega}) e_{kl}^A(\hat{\Omega})}{(1 + \hat{\Omega} \cdot \hat{p}_1)(1 + \hat{\Omega} \cdot \hat{p}_2)} \hat{D} \cdot \hat{\Omega}. \tag{C15}$$

This integral can be written as

$$\Gamma_{\text{dip}} = -\frac{1}{4}\beta \int_{S^2} d\hat{\Omega} \cos\chi \frac{\sin^2\theta(\sin^2\xi\sin^2\phi - \sin^2\xi\cos^2\theta\cos^2\phi - \cos^2\xi\sin^2\theta + 2\sin\xi\cos\xi\sin\theta\cos\theta\cos\phi)}{(1 + \cos\theta)(1 + \cos\xi\cos\theta + \sin\xi\sin\theta\cos\phi)}. \tag{C16}$$

As in the previous section, we write

$$\Gamma_{\text{dip}} = \tfrac{1}{4}\beta(I + J), \tag{C17}$$

where the first term is now given by

$$I = \int_{S^2} d\hat{\Omega}\cos\chi(1-\cos\theta)(1-\cos\xi\cos\theta-\sin\xi\sin\theta\cos\phi) = -\frac{4\pi}{3}(\cos\alpha_1+\cos\alpha_2), \tag{C18}$$

and J is as in Eq. (C7), but K is now given by

$$K = \int_0^{2\pi} d\phi \frac{\cos(\chi)\sin^2\phi}{1+\cos\xi\cos\theta+\sin\xi\sin\theta\cos\phi}, \tag{C19}$$

and may be evaluated by the same methods. The result is

$$K = \frac{2\pi}{\sin^2\xi}(a_\pm\cot\theta\csc\theta + b_\pm\cot^2\theta + c_\pm\csc^2\theta), \tag{C20}$$

where the following constant terms have been defined:

$$a_\pm = \cos\alpha_1(1\mp\cos\xi) \pm \frac{\cos\alpha_2-\cos\alpha_1\cos\xi}{\sin^2\xi}$$
$$\times (1\mp\cos\xi)^2, \tag{C21a}$$

$$b_\pm = \mp\cos\alpha_1(1\mp\cos\xi) - \frac{\cos\alpha_2-\cos\alpha_1\cos\xi}{2\sin^2\xi}$$
$$\times (1\mp\cos\xi)^2, \tag{C21b}$$

$$c_\pm = -\frac{\cos\alpha_2-\cos\alpha_1\cos\xi}{2\sin^2\xi}(1\mp\cos\xi)^2, \tag{C21c}$$

and a_+, b_+, and c_+ are to be used in the case where the inequality $0 < \theta < \pi - \xi$ holds, and a_-, b_-, and c_- are to be used otherwise. Thus, the integral J must again be split into two sections, and the result of the integration is

$$J = 4\pi(\cos\xi-1)(\cos\alpha_1+\cos\alpha_2)$$
$$- 16\pi(\cos\alpha_1+\cos\alpha_2)\tan^2\frac{\xi}{2}\ln\left(\sin\frac{\xi}{2}\right). \tag{C22}$$

Thus, we see that

$$\Gamma_{\text{dip}} = \pi\beta(\cos\alpha_1+\cos\alpha_2)\left(\cos\xi - \frac{4}{3} - 4\tan^2\frac{\xi}{2}\ln\left(\sin\frac{\xi}{2}\right)\right). \tag{C23}$$

Because we wish for Γ_{dip} to have maximal value of unity at $\xi = 0$ and $\xi = \pi$ (where it is clear that $\alpha_1 = \alpha_2 = 0$), we must select a normalization constant of $\beta = -3/2\pi$.

[1] R. A. Hulse and J. H. Taylor, Astrophys. J. **195**, L51 (1975).
[2] S. Detweiler, Astrophys. J. **234**, 1100 (1979).
[3] D. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy* (Cambridge University Press, Cambridge 2005).
[4] G. Hobbs, D. R. Lorimer, A. G. Lyne, and M. Kramer, Mon. Not. R. Astron. Soc. **360**, 974 (2005).
[5] F. A. Jenet, G. B. Hobbs, K. J. Lee, and R. N. Manchester, Astrophys. J. **625**, L123 (2005).
[6] F. A. Jenet *et al.*, Astrophys. J. **653**, 1571 (2006).
[7] J. S. B. Wyithe and A. Loeb, Astrophys. J. **590**, 691 (2003).
[8] A. H. Jaffe and D. C. Backer, Astrophys. J. **583**, 616 (2003).
[9] M. Enoki, K. T. Inoue, M. Nagashima, and N. Sugiyama, Astrophys. J. **615**, 19 (2004).
[10] M. Kramer, D. C. Backer, J. M. Cordes, T. J. W. Lazio, B. W. Stappers, and S. Johnston, New Astron. Rev. **48**, 993 (2004).
[11] A. Sesana, A. Vecchio, and C. N. Colacino, arXiv:0804.4476v2.
[12] D. M. Eardley, D. L. Lee, A. P. Lightman, R. V. Wagoner, and C. M. Will, Phys. Rev. Lett. **30**, 884 (1973).
[13] D. M. Eardley, D. L. Lee, and A. P. Lightman, Phys. Rev. D **8**, 3308 (1973).
[14] L. A. Boyle and A. Buonanno, Phys. Rev. D **78**, 043531 (2008).
[15] T. W. B. Kibble, J. Phys. A **9**, 1387 (1976).
[16] M. B. Hindmarsh and T. W. B. Kibble, Rep. Prog. Phys. **58**, 477 (1995).
[17] A. Vilenkin and E. Shellard, *Cosmic Strings and other Topological Defects* (Cambridge University Press, Cambridge 2000).
[18] R. Jeannerot, J. Rocher, and M. Sakellariadou, Phys. Rev. D **68**, 103514 (2003).
[19] N. T. Jones, H. Stoica, and S. H. H. Tye, J. High Energy Phys. 07 (2002) 051.
[20] S. Sarangi and S. H. H. Tye, Phys. Lett. B **536**, 185 (2002).
[21] G. Dvali and A. Vilenkin, J. Cosmol. Astropart. Phys. 03 (2004) 010.
[22] N. T. Jones, H. Stoica, and S. H. H. Tye, Phys. Lett. B **563**, 6 (2003).
[23] E. J. Copeland, R. C. Myers, and J. Polchinski, J. High Energy Phys. 06 (2004) 013.
[24] M. G. Jackson, N. T. Jones, and J. Polchinski, J. High Energy Phys. 10 (2005) 013.
[25] V. Berezinsky, B. Hnatyk, and A. Vilenkin, arXiv:astro-ph/0001213.
[26] T. Damour and A. Vilenkin, Phys. Rev. Lett. **85**, 3761 (2000).
[27] T. Damour and A. Vilenkin, Phys. Rev. D **64**, 064008 (2001).
[28] T. Damour and A. Vilenkin, Phys. Rev. D **71**, 063510 (2005).
[29] X. Siemens *et al.*, Phys. Rev. D **73**, 105001 (2006).

[30] X. Siemens, V. Mandic, and J. Creighton, Phys. Rev. Lett. **98**, 111101 (2007).

[31] A. N. Lommen, Ph.D. thesis, University of California, Berkeley, 2001.

[32] V. M. Kaspi, J. H. Taylor, and M. F. Ryba, Astrophys. J. **428**, 713 (1994).

[33] R. N. Manchester, in *40 Years of Pulsars: Millisecond Pulsars, Magnetars and More*, edited by C. Bassa, Z. Wang, A. Cumming, and V. M. Kaspi, AIP Conf. Proc. No. 983 (AIP, New York, 2008).

[34] Parkes Pulsar Timing Array Project, http://www.atnf.csiro.au/research/pulsar/ppta/.

[35] P. F. Michelson, Mon. Not. R. Astron. Soc. **227**, 933 (1987).

[36] N. Christensen, Phys. Rev. D **46**, 5250 (1992).

[37] E. E. Flanagan, Phys. Rev. D **48**, 2389 (1993).

[38] B. Allen and J. D. Romano, Phys. Rev. D **59**, 102001 (1999).

[39] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **69**, 122004 (2004).

[40] B. Abbott *et al.* (LIGO Collaboration), Phys. Rev. Lett. **95**, 221101 (2005).

[41] B. Abbott *et al.* (LIGO Collaboration), Astrophys. J. **659**, 918 (2007).

[42] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **76**, 082003 (2007).

[43] B. Abbott *et al.* (ALLEGRO Collaboration and LIGO Scientific Collaboration), Phys. Rev. D **76**, 022001 (2007).

[44] A. Lommen, J. Romano, and G. Woan (unpublished).

[45] B. Allen, J. D. E. Creighton, E. E. Flanagan, and J. D. Romano, Phys. Rev. D **67**, 122002 (2003).

[46] F. B. Estabrook and H. D. Wahlquist, Gen. Relativ. Gravit. **6**, 439 (1975).

[47] W. J. Kaufmann, Nature (London) **227**, 157 (1970).

[48] R. W. Hellings and G. S. Downs, Astrophys. J. **265**, L39 (1983).

[49] O. Malaspinas and R. Sturani, Classical Quantum Gravity **23**, 319 (2006).

[50] G. J. Feldman and R. D Cousins, Phys. Rev. D **57**, 3873 (1998).

[51] T. J. Loredo, in *Maximum Entropy and Bayesian Methods*, edited by P. F. Fougere (Kluwer, Dordrecht, 1990), p. 81.

[52] P. C. Gregory and T. J. Loredo, Astrophys. J. **398**, 146 (1992).

[53] S. Mitra, S. Dhurandhar, T. Souradeep, A. Lazzarini, V. Mandic, S. Bose, and S. Ballmer, Phys. Rev. D **77**, 042002 (2008).

[54] N. R. Lomb, Astrophys. Space Sci. **39**, 447 (1976).

[55] J. D. Scargle, Astrophys. J. **263**, 835 (1982).

[56] J. H. Horne and S. L. Baliunas, Astrophys. J. **302**, 757 (1986).

[57] M. G. House and P. D. Mountcastle, in *Proceedings of the IEEE Radar Conference, Long Beach, CA, 2002* (IEEE, New York, 2002), pp. 63–67.

[58] R. W. Hellings, Phys. Rev. D **23**, 832 (1981).