

Gravitational wave radiometry: Mapping a stochastic gravitational wave backgroundSanjit Mitra,^{1,2,*} Sanjeev Dhurandhar,^{1,†} Tarun Souradeep,^{1,‡} Albert Lazzarini,^{3,§} Vuk Mandic,^{3,||} Sukanta Bose,^{4,¶} and Stefan Ballmer^{3,**}¹*Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411007, India*²*Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France*³*LIGO Laboratory, California Institute of Technology, MS 18-34, Pasadena, California 91125, USA*⁴*Department of Physics, Washington State University, Pullman, Washington 99164-2814, USA*

(Received 21 August 2007; published 14 February 2008)

The problem of the detection and mapping of a stochastic gravitational wave background (SGWB), either cosmological or astrophysical, bears a strong semblance to the analysis of the cosmic microwave background (CMB) anisotropy and polarization, which too is a stochastic field, statistically described in terms of its correlation properties. An astrophysical gravitational wave background (AGWB) will likely arise from an incoherent superposition of unmodelled and/or unresolved sources and cosmological gravitational wave backgrounds (CGWB) are also predicted in certain scenarios. The basic statistic we use is the cross correlation between the data from a pair of detectors. In order to “point” the pair of detectors at different locations one must suitably delay the signal by the amount it takes for the gravitational waves (GW) to travel to both detectors corresponding to a source direction. Then the raw (observed) sky map of the SGWB is the signal convolved with a beam response function that varies with location in the sky. We first present a thorough analytic understanding of the structure of the beam response function using an analytic approach employing the stationary phase approximation. The true sky map is obtained by numerically deconvolving the beam function in the integral (convolution) equation. We adopt the maximum likelihood framework to estimate the true sky map using the conjugate gradient method that has been successfully used in the broadly similar, well-studied CMB map-making problem. We numerically implement and demonstrate the method on signal generated by simulated (unpolarized) SGWB for the GW radiometer consisting of the LIGO pair of detectors at Hanford and Livingston. We include “realistic” additive Gaussian noise in each data stream based on the LIGO-I noise power spectral density. The extension of the method to multiple baselines and polarized GWB is outlined. In the near future the network of GW detectors, including the Advanced LIGO and Virgo detectors that will be sensitive to sources within a thousand times larger spatial volume, could provide promising data sets for GW radiometry.

DOI: [10.1103/PhysRevD.77.042002](https://doi.org/10.1103/PhysRevD.77.042002)

PACS numbers: 04.80.Nn, 04.30.Db, 95.55.Ym, 98.70.Vc

I. INTRODUCTION

The existence of gravitational waves (GW) has long been verified “indirectly” through the observations of Hulse and Taylor [1]. However, direct observation of such waves with manmade gravitational wave detectors has been lacking. At present the laser interferometric detectors have achieved sensitivities close to that required for detecting such waves [2]. The space mission LISA [3] is also planned by the ESA and NASA to detect low frequency GW. The significance of the direct detection of GW lies, not only in the opening of an entirely new window into observational astronomy by probing phenomena in the regime of strong gravity; it further promises to test our present theories of gravitation.

Different types of GW sources have been predicted and may be directly observed by Earth-based detectors in the near future (see [4–10] and references therein for recent reviews): (i) Transient sources—such as binary systems of neutron stars (NS) and/or black holes (BH) in their inspiral phase, BH/BH and/or BH/NS mergers, and supernovae explosions, whose signals last for a time much shorter, typically between a few milliseconds and a few minutes, than the planned observational time; (ii) Continuous wave (CW) sources—e.g. rapidly rotating neutron stars, where a weak deterministic signal is continuously emitted, and (iii) Stochastic backgrounds of radiation, either of primordial or astrophysical origin.

In this paper we will address the problem of a spatially resolved search of the gravitational wave stochastic background. This approach was advocated in the LIGO technical note [11] and the basic analysis was recently implemented on the fourth science run data from the LIGO interferometers to prepare an upper limit map [12]. Our main focus will be on a stochastic astrophysical GW background (AGWB), which might arise from a superposition of a large number of independent and unresolved GW sources. The gravitational wave background can arise

* sanjit@oca.eu† sanjeev@iucaa.ernet.in‡ tarun@iucaa.ernet.in§ lazz@ligo.caltech.edu|| vmandic@ligo.caltech.edu¶ sukanta@mail.wsu.edu** sballmer@caltech.edu

from a variety of sources: supernovae with asymmetric core collapse, binary black hole (BBH) mergers, GWs from low-mass x-ray binaries (LMXBs) and hydrodynamical instabilities in neutron stars (r-modes), or even GWs from astrophysical objects that we never knew existed. When a collection of any subset of these sources is unresolvable, it can appear as a stochastic GW background (SGWB) of a variable duration in our detectors of interest. While an astrophysical background will provide information about our immediate neighborhood, cosmological GW backgrounds (CGWB) could probe the physics of the early universe. There exist cosmological scenarios (e.g., cosmic strings and superstring models) which predict CGWB that should be detectable by Advanced LIGO [13].

We propose and develop a data analysis method that measures and maps the power in the SGWB from a specific location in the sky—GW radiometry using a network of detectors. We find that the angular resolution essentially depends on the effective GW bandwidth and the linear size of the network. In this paper we will restrict ourselves to the network of the two 4 km LIGO detectors at Hanford (LHO) and Livingston (LLO). For the purposes of our analysis, we take their noise curves to be identical with the LIGO-I design power spectral density [14]. Our future plan is to include VIRGO and other detectors around the world in the numerical implementation of this analysis.

The basic statistic is the cross correlation between the data from a pair of detectors. In order to “point” the pair of detectors at different locations one must suitably delay the signal by the amount it takes for the GW to travel to both detectors corresponding to the source direction. This delay will be a function of the source position and will vary as the Earth rotates. Using the delay allows the detectors to sample the same wave front from the source. The cross spectrum formulation has been carried out in [15,16]. Methods for searching for isotropic backgrounds [17] using the cross correlation and for anisotropies using spherical harmonic decomposition [18] have been devised. Efforts have also been made to devise methods to measure the spherical harmonic moments of the SGWB anisotropy using a network of ground or space based detectors [19,20]. Here we focus on a spatially resolved search and the final goal is to make a map of the true SGWB sky. We achieve this goal by pixelizing the sky, that is, we use a pixel basis.

The advantage of a spatially resolved search is seen immediately if we examine the so-called overlap reduction factor, which partially determines the fractional power of source spectrum the search filter will receive at different frequency bands. The overlap reduction factor, normally denoted by $\gamma(f)$ in the literature [17] for the isotropic unpolarized background, becomes a time-dependent factor $\gamma(\hat{\Omega}, f, t)$ for the spatially resolved search. For the LIGO detectors, $\gamma(f)$ quickly reduces to zero beyond few tens of Hz, while $\gamma(\hat{\Omega}, f, t)$ has infinite bandwidth. So the band-

width of the spatially resolved search is essentially detector bandwidth limited. This is typically valid for a network of detectors and therefore important from the point of view of the sensitivity regime of GW detectors which lies in this region.

As in radio-interferometry, the correlation statistic so constructed produces a “dirty” map where a point source does not produce a point image, but one that is smeared by a beam response function (beam, for brevity). The “cleaned” GW sky map is obtained from the measured cross correlation statistic by deconvolving the beam. In other words, to obtain the GW power from each direction in the sky one needs to solve an integral equation where the measured power (data) is a convolution of the actual power with a kernel (beam). In order to understand the structure of the beam we carry out a numerical and an analytical study using the stationary phase approximation (SPA). We find that at low declinations (latitudes) of a point source, the kernel essentially has the shape of a “figure 8” with a bright spot at the intersection. The bright spot is at the location of the point source. The figure of 8 continuously changes and bifurcates into a “tear drop” as the point source moves to higher declinations. The declination at which this bifurcation occurs is determined by the half-angle of the cone traced out by the vector joining the two detectors. For the LIGO detectors this declination is about 26° . The size of the bright spot or effective sky patch, defined by a certain percentage of reduction in the beam response function, say 50%, is determined by the inverse of the bandwidth divided by the light (GW) travel time between the detectors. Considering a broadband source and LIGO detectors having kHz bandwidth with 10 ms light travel distance between them, the angular size is about 5° in radius. We find that these results agree very well with those obtained by applying singular value decomposition to the kernel matrix; the eigenvalues fall off steeply after a certain point which determines essentially the number of “degrees of freedom” of the kernel matrix and thereby the size of the sky patch.

We employ the maximum likelihood (ML) approach for deconvolving the sky map. The integral equation for a discrete pixelized sky leads to a set of linear algebraic equations. Several deconvolution algorithms exist in literature for solving such a problem. However, because of the broad similarity of our present problem with the cosmic microwave background (CMB) analysis, we have opted for techniques that have been successfully applied to deconvolve CMB sky maps. Moreover, the ML approach provides a framework to study the SGWB anisotropy in another basis of interest, e.g., the spherical harmonic basis. We find the ML estimate by employing the conjugate gradient method. To verify our method, we apply the analysis on simulated sky maps mapped by a GW radiometer consisting of the LIGO pair of detectors LHO and LLO. We generate “realistic” colored Gaussian noise

corresponding to the LIGO-I design sensitivity curve [14] detector noise and embed various simulated sky maps of the GW stochastic background into the noise. We demonstrate that the true sky maps can be recovered satisfactorily.

The paper is organized as follows: In Sec. II we briefly review the GW radiometer concepts and obtain an expression for the GW radiometric cross correlation signal, which is then optimized for the maximum signal-to-noise (SNR). In Sec. III, we set up the integral equation that must be solved in order to obtain the true sky map from the data. A time-frequency analysis is performed and the directed optimal filter is derived for anisotropic searches. Moreover, a stationary phase analysis is presented which provides us with the understanding of the kernel or the point spread function. In Sec. IV, we describe the maximum likelihood approach and the conjugate gradient method and apply it to simulated data to test the efficacy of this method. Later subsections outline the extension of the GW radiometer analysis to incorporate multiple baselines obtained with a network of detectors and the extension of the GW radiometer to search for polarized SGWB. The numerical implementation of the method is described in Sec. V. We conclude in Sec. VI.

II. GW RADIOMETER EMPLOYING EARTH ROTATION APERTURE SYNTHESIS

A. The principle of a radiometer

Radiometry or aperture synthesis is a well-known technique in radio astronomy and CMB experiments. The idea is to *point* a pair of detectors separated by a baseline to a desired direction in the sky by introducing an appropriate time delay between their data streams. This delay corresponds to the difference between the times of arrival of a GW signal if it were to arrive at those two detector sites from that direction. For a given source in the sky, this delay will change as the baseline orientation changes due to the rotation of the earth. The cross correlation of the data from the two detectors, appropriately time delayed, would cause potential GW signals arriving from the chosen direction to interfere constructively. Whereas signals from other directions will tend to cancel out because of destructive interference. This principle of Earth rotation aperture synthesis, which is well known in radio astronomy, could very well be used in GW astronomy using pairs of GW antennae. Figure 1 illustrates the principle on which the GW radiometer works.

We consider the Celestial Equatorial frame whose origin coincides with the center of the Earth. The axes are defined as follows: For a fixed but arbitrarily chosen time $t = 0$, the x -axis is directed towards the intersection of the equator and the longitude $\phi = 0$, which can be taken as the Greenwich meridian; the z -axis is directed towards the North Celestial Pole, and the y -axis is chosen orthogonal to the previous two axes forming a right-handed triad. The Earth rotates in this frame with the angular velocity $\omega_E =$

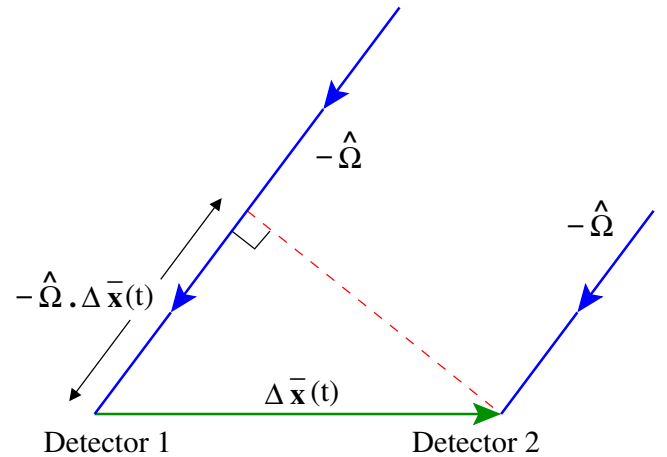


FIG. 1 (color online). Geometry of an elementary radiometer. Above, $\Delta \mathbf{x}(t)$ is the separation or baseline vector between the two detectors; as the Earth rotates, its direction changes, but its magnitude remains fixed. The direction to the source $\hat{\Omega}$ is also fixed in the barycentric frame. The phase difference between signals arriving at two detector sites from the same direction is also shown.

$2\pi/(1 \text{ sidereal day}) \sim 7.3 \times 10^{-5}$ radians/sec oriented along the z -axis. The two detectors are at locations \mathbf{x}_I (where $I = 1, 2$), and the baseline vector joining the two sites is $\Delta \mathbf{x} := \mathbf{x}_2 - \mathbf{x}_1$. $\hat{\Omega}$ is the unit vector in the direction to the source, which is fixed in the barycentric frame. The baseline $\Delta \mathbf{x}$ rotates with the Earth, but its magnitude $|\Delta \mathbf{x}|$, which is the distance between the detectors, remains constant. The map of the SGWB can be constructed by performing this synthesis for each location in the sky, patch by patch. An approximate size of the patch or resolution of the GW radiometer and number of patches required to cover the full sky can be estimated from the following simple argument. This simple model produces a fringe pattern with resolution given by the standard formula for the central width $\Delta\theta \sim \lambda_{\text{GW}}/(|\Delta \mathbf{x}| \sin\theta)$ where λ_{GW} is the GW wavelength and θ is the angle of incidence. This is, however, a naive estimate. A better estimate of the resolution would follow from considerations involving the pixel-to-pixel Fisher information matrix, in which case the solid angle resolution will scale inversely proportional to the square of the SNR.

In this paper we will consider a GW radiometer made using the two 4 km LIGO detectors LHO and LLO in the United States. However, this analysis is equally well applicable to any other baseline or a network of baselines involving other detectors such as VIRGO, TAMA, GEO, etc. The light travel time between the LHO and LLO is ~ 10 ms, which at a bandwidth of 1 kHz gives a resolution $\Delta\theta$ of few degrees. This resolution implies that a few thousand patches or pixels are required to completely cover the full sky. This estimate, in turn, will be useful for assessing the computational cost and numerical complica-

tions in handling large matrices for obtaining the full sky map. We make these estimates more robust below.

B. Basic framework and the statistical properties of the data

We first set up the notation and framework needed for investigating the problem.

In the transverse traceless (TT) gauge the metric perturbation $h_{ij}(t, \mathbf{x})$ in the equatorial frame can be expanded in terms of plane waves of the two polarizations $+$, \times in the following way:

$$h_{ij}(t, \mathbf{x}) = \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} \tilde{h}_A(f, \hat{\Omega}) e_{ij}^A(\hat{\Omega}) e^{2\pi i f(t + \hat{\Omega} \cdot \mathbf{x}/c)}, \quad (2.1)$$

where a tilde on a variable denotes its Fourier transform; the complex Fourier amplitudes satisfy the relation $\tilde{h}_A^*(f, \hat{\Omega}) = \tilde{h}_A(-f, \hat{\Omega})$ owing to the reality of $h_{ij}(t, \mathbf{x})$. Here $A = \{+, \times\}$ is the polarization index and summation over the repeated index A is implied. In terms of the spherical polar coordinates (θ, ϕ) , the source direction is given by

$$\hat{\Omega} = \cos\phi \sin\theta \hat{\mathbf{x}} + \sin\phi \sin\theta \hat{\mathbf{y}} + \cos\theta \hat{\mathbf{z}} \quad (2.2)$$

and, hence, the infinitesimal solid angle along the direction $\hat{\Omega}$ is $d\hat{\Omega} \equiv \sin\theta d\theta d\phi$. Note that the wave propagation direction is $-\hat{\Omega}$. The polarization tensors $e^A(\hat{\Omega})$ are defined by the following expressions:

$$e^+(\hat{\Omega}) = \hat{\mathbf{e}}_\theta \otimes \hat{\mathbf{e}}_\theta - \hat{\mathbf{e}}_\phi \otimes \hat{\mathbf{e}}_\phi, \quad (2.3a)$$

$$e^\times(\hat{\Omega}) = \hat{\mathbf{e}}_\theta \otimes \hat{\mathbf{e}}_\phi + \hat{\mathbf{e}}_\phi \otimes \hat{\mathbf{e}}_\theta, \quad (2.3b)$$

where

$$\begin{aligned} \hat{\mathbf{e}}_\theta &= \cos\phi \cos\theta \hat{\mathbf{x}} + \sin\phi \cos\theta \hat{\mathbf{y}} - \sin\theta \hat{\mathbf{z}}, \\ \hat{\mathbf{e}}_\phi &= -\sin\phi \hat{\mathbf{x}} + \cos\phi \hat{\mathbf{y}}, \end{aligned}$$

are the two orthonormal basis vectors on a two-sphere.

The statistics of the GW signal can be best described in the Fourier domain: If we assume the signal to be stochastic and uncorrelated in the two polarizations [21], different frequencies, and different directions, then the Fourier components of the GW strain obey,

$$\begin{aligned} \langle \tilde{h}_A^*(f, \hat{\Omega}) \tilde{h}_{A'}(f', \hat{\Omega}') \rangle &= \delta_{AA'} \delta(f - f') \delta^2(\hat{\Omega} - \hat{\Omega}') \\ &\times \mathcal{P}_A(\hat{\Omega}) H(f), \end{aligned} \quad (2.5)$$

where $\mathcal{P}_A(\hat{\Omega})$ is proportional to the strength of the SGWB in the direction $\hat{\Omega}$ and $H(f)$ is the *two-sided* GW source power spectral density (PSD). The interpretation of the quantity $\mathcal{P}_A(\hat{\Omega})$ can be made apparent by relating it to the *specific intensity* or *brightness* [22] of GW $I_{\text{GW}}(f, \hat{\Omega})$. Specific intensity is defined as (c times) the incident energy density per unit frequency interval per unit solid angle, or,

equivalently, specific intensity is the (normally incident) flux per unit solid angle. Following the convention commonly used in SGWB analysis, if the incident GW energy density is expressed in the units of critical energy density for a flat universe $\rho_c := 3c^2 H_0^2 / (8\pi G)$, where H_0 is the Hubble constant at the current epoch and G is the Newton's gravitational constant, it can be shown that [23]:

$$I_{\text{GW}}(f, \hat{\Omega}) = \frac{4\pi^2 c}{3H_0^2} f^2 H(f) [\mathcal{P}_+(\hat{\Omega}) + \mathcal{P}_\times(\hat{\Omega})]. \quad (2.6)$$

Therefore, $\mathcal{P}_A(\hat{\Omega})$ can be interpreted as the specific intensity of SGWB for the corresponding polarization (up to a certain proportionality constant).

In general, we cannot separate $H(f)$ from $\mathcal{P}_A(\hat{\Omega})$, because the frequency power spectrum $H(f)$ could also depend on the direction $\hat{\Omega}$. A more general treatment would use a quantity like $\mathcal{P}_A(\hat{\Omega}, f)$ which describes both frequency and angular distribution of SGWB power together. But for a small enough bandwidth the assumption may be justified, as the signal is expected to have a smooth profile of the power spectra. Further, it should be noted that \mathcal{P}_A is actually a second rank tensor and should be represented by two indices as $\mathcal{P}_{AA'}$, but because we assume that the two polarizations are uncorrelated the quantity $\mathcal{P}_{+\times}$ is zero and thus absent. To avoid unnecessary indices and with a slight abuse of notation, we therefore write \mathcal{P}_A with a single index instead of $\mathcal{P}_{AA'}$.

We consider two GW detectors located at $\mathbf{x}_I(t)$, $I = 1, 2$. The detector frames are denoted by the coordinates (X_I, Y_I, Z_I) where the arms of the respective detectors lie along the (X_I, Y_I) axes. Then the detector tensors d_I are given by

$$d_I = \frac{1}{2} [\hat{\mathbf{X}}_I \otimes \hat{\mathbf{X}}_I - \hat{\mathbf{Y}}_I \otimes \hat{\mathbf{Y}}_I]. \quad (2.7)$$

The factor $\frac{1}{2}$ is inherited from the geodesic deviation equation. Owing to the Earth's rotation, the detector coordinates and d_I are functions of t . Thus in matrix form:

$$d_I(t) = \mathcal{R}^T(t) \cdot d_I(t=0) \cdot \mathcal{R}(t), \quad (2.8)$$

where $\mathcal{R}(t)$ is the rotation matrix describing a rotation of angle $\omega_E t$ around the Earth's rotation axis, namely, the z -axis. Here ω_E is the Earth's sidereal angular velocity $\approx 7.3 \times 10^{-5}$ radians/sec.

The strain $h_I(t)$ in the I^{th} detector is given by

$$h_I(t) = h_{ij}(t, \mathbf{x}_I(t)) d_I^{ij}(t). \quad (2.9)$$

We define the antenna pattern functions as

$$F_I^A(t; \hat{\Omega}) = e_{ij}^A(\hat{\Omega}) d_I^{ij}(t) \quad (2.10)$$

for a wave incident from the direction $\hat{\Omega}$.

Contracting Eq. (2.1) with the detector tensors d_I , the signal amplitudes can be expressed in terms of the antenna pattern functions as

$$h_I(t) = \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} \tilde{h}_A(f, \hat{\Omega}) F_I^A(t; \hat{\Omega}) e^{2\pi i f(t + \hat{\Omega} \cdot \mathbf{x}/c)}. \quad (2.11)$$

The baseline $\Delta \mathbf{x}$ at any time t (in matrix form) is given by

$$\Delta \mathbf{x}(t) = \mathcal{R}^T \cdot \Delta \mathbf{x}(t=0). \quad (2.12)$$

The detector output $s_I(t)$ is a sum of the GW signal $h_I(t)$ and noise $n_I(t)$

$$s_I(t) = h_I(t) + n_I(t). \quad (2.13)$$

Statistically, the gravitational wave strains $h_I(t)$ are uncorrelated with the detector noise $n_I(t)$; that is, the following four correlations, in the time domain are zero:

$$\langle h_I(t) n_J(t') \rangle = 0, \quad I, J = 1, 2, \quad (2.14)$$

where t, t' are any two time instants. We also assume that the noise in different detectors is uncorrelated; that is, $\langle n_1(t) n_2(t') \rangle = 0$. This assumption is not unreasonable when the detectors are widely separated. Thus the only possible correlation is between $h_1(t)$ and $h_2(t')$. The statistic that we construct in the next subsection is based on this fact.

C. The cross correlation statistic for the directed search

Stochastic GW signals inherently can arrive from any direction with any amplitude. Moreover, they are characterized by the statistical expectation values of energy density. The noise in the two detectors are assumed to be essentially independent. In this situation the cross correlation of the data from the two detectors is an appropriate statistic for detecting and observing stochastic GW. In order to optimize the SNR, the cross correlation statistic for the directed search involves a direction-dependent filter function $Q(t, \hat{\Omega}; t', t'')$, which connects sidereal time t' of one detector's data to t'' of the other detector's data to match the phases of the GW strains in the detectors. The filter does an inverse noise weighting using the PSDs of the two detectors in order to suppress noisy frequency bands and enhance the SNR by assigning relatively large weight to the sensitive regimes of the detectors and the bands where the source is expected to emit more power. (In general, Q will depend on $\mathcal{P}_A(\hat{\Omega})$. However, for the directed search that we envisage here, as we will see later, the $\mathcal{P}_A(\hat{\Omega})$ are delta functions and, therefore, the filter Q only depends on $\hat{\Omega}$.)

The sampling interval with which the data are sampled is determined by the Nyquist frequency of the stochastic signals of interest, and can be well below a millisecond, corresponding to several kilohertz. While it is possible to compute the filter function on a time segment of this sampling interval, it is erroneous to do so for the physical problem at hand. This is because the signals at the two detectors will be incoherent on time scales smaller than the

light travel time, t_d , of that baseline, which is at most a few tens of milliseconds for ground-based detectors. Thus t_d sets the lower limit on the duration of the time segment on which the filter function should be computed. The upper limit, τ , is set by the smaller of the time scales, on which the data are stationary and the time scale on which the detector orientation changes appreciably. We thus divide the data into time segments, Δt , such that $t_d \ll \Delta t \ll \tau$. The time segments used currently in LSC data analysis vary from 32 to 192 seconds, and are consistent with these limits.

The final statistic for the full observation time T is obtained by linearly combining the cross correlations over the smaller time intervals as a weighted sum. The filter is optimized for each time segment, $I_k = [t_k - \Delta t/2, t_k + \Delta t/2]$, at sidereal time t_k and the statistic for the k th segment is given by

$$\begin{aligned} \Delta S(t = t_k, \hat{\Omega}) &\equiv \Delta S_k(\hat{\Omega}) \\ &= \int_{I_k} dt' \int_{I_k} dt'' s_1(t') s_2(t'') Q(t_k, \hat{\Omega}; t', t''). \end{aligned} \quad (2.15)$$

The final cross correlation statistic S for all the $n = T/\Delta t$ sidereal time bins can then be obtained by combining the ΔS_k as a weighted sum as follows:

$$S(\hat{\Omega}) = \sum_{k=1}^n w_k(\hat{\Omega}) \Delta S_k, \quad (2.16)$$

where, the n quantities $w_k(\hat{\Omega})$ are to be chosen so that the SNR for the statistic $S(\hat{\Omega})$ is maximized. We denote the SNR of S by $\rho_S = \mu_S/\sigma_S$, where μ_S and σ_S are the mean and standard deviation of S , respectively. We use normalized weights $\sum_{k=1}^n w_k = 1$.

The μ_S , σ_S , and the SNR ρ_S are given in terms of the means $\mu_k = \langle \Delta S_k \rangle$ and variances $\sigma_k^2 = \langle \Delta S_k^2 \rangle - \langle \Delta S_k \rangle^2$ of the individual *mutually uncorrelated* time segments as follows:

$$\begin{aligned} \mu_S &= \sum_{k=1}^n w_k \mu_k, & \sigma_S^2 &= \sum_{k=1}^n w_k^2 \sigma_k^2, \\ \rho_S &= \left[\sum_{k=1}^n w_k \mu_k \right] / \left[\sum_{k=1}^n w_k^2 \sigma_k^2 \right]^{1/2}. \end{aligned} \quad (2.17)$$

The maximization is achieved elegantly by invoking the Schwarz inequality. We consider an n -dimensional Euclidean space equipped with the usual scalar product in which \vec{w} is a vector with components w_k . We define $\vec{\kappa}$ and $\vec{\rho}$ having the components $\kappa_k = w_k \sigma_k$ and $\rho_k = \mu_k/\sigma_k$, respectively. Then, we may write $\rho_S \equiv \hat{\kappa} \cdot \vec{\rho}$ where the vector $\hat{\kappa}$ is a unit vector in the direction of $\vec{\kappa}$. ρ_S is maximized when $\vec{\kappa}$ points in the direction of $\vec{\rho}$. Thus, $w_k \propto \mu_k/\sigma_k^2$, the proportionality factor being $[\sum_{k=1}^n \mu_k \sigma_k^{-2}]^{-1}$. We have the results:

$$S = \frac{\sum_{k=1}^n \mu_k \sigma_k^{-2} \Delta S_k}{\sum_{k=1}^n \mu_k \sigma_k^{-2}}, \quad (2.18)$$

$$\rho_S = |\vec{\rho}| = \left\{ \sum_{k=1}^n \mu_k^2 / \sigma_k^2 \right\}^{1/2}. \quad (2.19)$$

Further, we are free to choose the normalization of the filter Q . We choose the normalization such that all μ_k are made identical, equal to μ , say. This simplifies the statistic S and its SNR ρ_S :

$$S = \frac{\sum_{k=1}^n \Delta S_k \sigma_k^{-2}}{\sum_{k=1}^n \sigma_k^{-2}}, \quad (2.20)$$

$$\rho_S = \mu \left\{ \sum_{k=1}^n \sigma_k^{-2} \right\}^{1/2}, \quad (2.21)$$

which are their final forms. Normally the GW signal is expected to be weak so that even after its integration over Δt it is still smaller than the noise, that is, $|\mu_k| \ll \sigma_k$. Therefore the variance in each segment obeys $\sigma_k^2 \approx \langle \Delta S_k^2 \rangle$. The signal, however, builds up when we integrate over all the time segments during the observation time. Second, from Eq. (2.20), it is evident that noisy time intervals contribute less to the statistic S leading to its optimal character.

III. THE CONVOLUTION EQUATION FOR THE SKY MAP

Since the GW power spectra and the detector noise are modeled in the frequency domain, it is convenient to formulate the whole analysis in that domain. However, since the detector output is a time series, it is pertinent to ask over what time duration must one compute their Fourier transforms. As discussed in the last section, typical acceptable segment sizes are a few tens to a few hundreds of seconds. Fourier transforms computed over such “small” (*vis à vis* the total observation) time scales are termed as *short-term Fourier transforms (SFTs)*. Each SFT becomes a function of time t as well because t is essentially the identifier of the segment. Thus we have a time-frequency representation of the data. As we shall see, this representation is most suitable for further analysis and has also been used in previous literature [17–19].

A. Time-frequency analysis of the signal and noise

The (approximate) SFT of a segment of detector output can be defined as [18],

$$\tilde{s}_I(t; f) := \int_{t-\Delta t/2}^{t+\Delta t/2} dt' s_I(t') e^{-2\pi i f t'}. \quad (3.1)$$

We retain here the convention of using “tilde” over a symbol to denote Fourier transform—the distinction should be evident from the context. Most importantly, by taking the inverse Fourier transform,

$$\begin{aligned} & \int_{-\infty}^{\infty} df \tilde{s}_I(t; f) e^{2\pi i f t} \\ &= \int_{-\infty}^{\infty} df e^{2\pi i f t} \int_{t-\Delta t/2}^{t+\Delta t/2} dt' s_I(t') e^{-2\pi i f t'} = s_I(t), \end{aligned} \quad (3.2)$$

the exact time series segment $s_I(t)$ can be recovered. The same notation will be used for several other quantities as well in this analysis.

While constructing the statistic, it was assumed that the true GW strains $h_I(t)$ in the detectors are correlated, but that the noise streams are uncorrelated. The expression for the correlation between the GW strains in two detectors will be derived here, which is necessary for the derivation of the optimal filter in the next subsection. The detector parameters can be approximated to be stationary over the period of the time segment. Hence we may consider the quantities \mathbf{d}_I and \mathbf{x}_I to be nearly constant over a time segment and regard them as functions of the time t labeling the time segment. The SFT of the GW strain in detector I over a time segment is given by

$$\begin{aligned} \tilde{h}_I(t; f) &= \int_{t-(\Delta t/2)}^{t+(\Delta t/2)} dt' \int_{S^2} d\hat{\Omega} \tilde{h}_A(f', \hat{\Omega}) F_I^A(\hat{\Omega}, t) \\ &\quad \times e^{2\pi i [f' t' - f t + f'(\hat{\Omega} \cdot \mathbf{x}_I(t)/c)]} \\ &= \int_{S^2} d\hat{\Omega} F_I^A(\hat{\Omega}, t) \int_{-\infty}^{\infty} df' \tilde{h}_A(f', \hat{\Omega}) \\ &\quad \times e^{2\pi i [(f' - f)t + f'(\hat{\Omega} \cdot \mathbf{x}_I(t)/c)]} \delta_{\Delta t}(f - f'), \end{aligned} \quad (3.3)$$

where summation over A is implied. The $\delta_{\Delta t}(f)$ is the finite time delta function (sinc function) defined by

$$\delta_{\Delta t}(f) := \int_{-\Delta t/2}^{\Delta t/2} dt e^{-2\pi i f t} = \frac{\sin \pi f \Delta t}{\pi f}. \quad (3.4)$$

The finite time delta function $\delta_{\Delta t}(f)$ behaves as the Dirac delta function $\delta(f)$ in the limit $\Delta t \rightarrow \infty$, but has the property $\delta_{\Delta t}(0) = \Delta t$. Hence for a large time segment Δt the SFT from detector I takes the simple form:

$$\tilde{h}_I(t, f) = \int_{S^2} d\hat{\Omega} \tilde{h}_A(f, \hat{\Omega}) F_I^A(\hat{\Omega}, t) e^{2\pi i f \hat{\Omega} \cdot \mathbf{x}_I(t)/c}. \quad (3.5)$$

The important result of this subsection is the expectation of cross correlation between the SFTs of time segments of detector outputs at time t from the two detectors 1 and 2, which can be obtained from Eqs. (2.5) and (3.3) as

$$\langle \tilde{h}_1^*(t, f) \tilde{h}_2(t, f') \rangle = e^{2\pi i t(f - f')} \int_{-\infty}^{\infty} df'' H(f'') \gamma(t, f''; \Delta \Omega, d\mathcal{P}_A) \delta_{\Delta t}(f'' - f) \delta_{\Delta t}(f'' - f'), \quad (3.6)$$

$$\gamma(t, f; \Delta\Omega, d\mathcal{P}_A) = \int_{\Delta\Omega=S^2} d\hat{\Omega} [F_1^+(\hat{\Omega}, t) F_2^+(\hat{\Omega}, t) \mathcal{P}_+(\hat{\Omega}) + F_1^\times(\hat{\Omega}, t) F_2^\times(\hat{\Omega}, t) \mathcal{P}_\times(\hat{\Omega})] e^{2\pi i f \hat{\Omega} \cdot \Delta \mathbf{x}(t)/c}. \quad (3.7)$$

The general overlap reduction function defined by Eq. (3.7) is a generalization of the usual overlap reduction function for the isotropic SGWB case, first constructed by Nelson Christensen [15] and formally written in a closed form by Flanagan [16]. In this case it is a more complex object. Besides the frequency, it is also a function of the segment time t . It also depends on $\mathcal{P}_A(\hat{\Omega})$ which is integrated over the full sky S^2 . Thus it is a functional of $\mathcal{P}_A(\hat{\Omega})$. In general, when we construct our directed filters, the integral will be restricted to a small patch of the sky, $\Delta\Omega \subset S^2$. (In principle, $\Delta\Omega$ can be any (measurable) subset of S^2 .) Accordingly, $\mathcal{P}_A(\hat{\Omega})$ plays the part of a weight function over the sky and we may define the measures

$$d\mathcal{P}_A = \mathcal{P}_A(\hat{\Omega}) d\hat{\Omega}. \quad (3.8)$$

Thus, the γ in general becomes a functional of the SGWB power in both polarizations coming from the patch $\Delta\Omega$. We therefore separate the function arguments, t and f , from the nonfunction arguments, $\Delta\Omega$ and \mathcal{P}_A , by a semicolon. Finally, the exponential term in Eq. (3.6) before the integral is just the time shift term in the Fourier transform of the segment at time t .

In the limit of a large time segment, Eq. (3.6) takes the simple form:

$$\langle \tilde{h}_1^*(t, f) \tilde{h}_2(t, f') \rangle = \delta(f - f') H(f) \gamma(t, f; S^2, d\mathcal{P}_A). \quad (3.9)$$

The advantage of expressing $\langle \tilde{h}_1^*(t, f) \tilde{h}_2(t, f) \rangle$ by Eq. (3.6) can be readily realized if we put $f = f'$. In this case, the correlation given by Eq. (3.9) diverges in the limit $\Delta t \rightarrow \infty$. But, in practice, Δt is finite, and hence, we expect a finite value for the correlation. Equation (3.6) lets us compute that finite value of $\langle \tilde{h}_1^*(t, f) \tilde{h}_2(t, f) \rangle$ at $f = f'$. We use the large Δt limit and replace one of the finite time delta functions $\delta_{\Delta t}(f'' - f)$ in the integrand of Eq. (3.6) by the Dirac delta function $\delta(f'' - f)$, while treating the other $\delta_{\Delta t}(f'' - f)$ as a normal function and put $\delta_{\Delta t}(0) = \Delta t$. We get

$$\langle \tilde{h}_1^*(t, f) \tilde{h}_2(t, f) \rangle = \Delta t H(f) \gamma(t, f; S^2, d\mathcal{P}_A). \quad (3.10)$$

This formula was derived by following the same procedure as prescribed in [17], so, not surprisingly, for isotropic backgrounds our result matches the formula obtained in [17]. This result is important for injecting test signals in the detector output [24].

Next we describe the properties of detector noise in a finite time segment.

The noise in the segment labeled by t is a time series $n_I(t)$ in a detector I . Its SFT is given by

$$\tilde{n}_I(t; f) := \int_{t-\Delta t/2}^{t+\Delta t/2} dt n_I(t) e^{-2\pi i f t}. \quad (3.11)$$

We take the mean to be zero: $\langle n_I(t) \rangle = \langle \tilde{n}_I(t; f) \rangle = 0$. Since $n_I(t)$ is real, its SFT obeys the reality condition, $\tilde{n}_I^*(t; f) = \tilde{n}_I(t; -f)$. The noise in a detector is uncorrelated with the noise in another detector and with the GW signal, i.e., $\langle n_1(t) n_2(t') \rangle = \langle h_1(t) n_2(t') \rangle = \langle n_1(t) h_2(t') \rangle = 0$ [see Eq. (2.14)]. These relations also hold for their corresponding SFTs. The length of the time segment is usually kept a few tens of seconds long, over which the detector noise can be regarded as stationary. Thus $\langle n_I(t') n_I(t'') \rangle$ is a function of $t'' - t'$, provided both t' , t'' are in the same segment centered at time t . Then, using the fact that $n_I(t)$ is real, we have

$$\langle n_I(t') n_I(t'') \rangle = \frac{1}{2} \int_{-\infty}^{\infty} df P_I(t; |f|) e^{2\pi i f(t'' - t')}, \quad (3.12)$$

where $P_I(t; f)$ is the *one-sided* noise PSD. This noise PSD is also a function of time t as detector noise is nonstationary. The correlation between the corresponding SFTs can be easily obtained from the above relations:

$$\langle \tilde{n}_I^*(t; f) \tilde{n}_I(t; f') \rangle = \frac{1}{2} \int_{-\infty}^{\infty} df'' P_I(t; |f''|) \delta_{\Delta t}(f'' - f) \times \delta_{\Delta t}(f'' - f'). \quad (3.13)$$

In the limit of large length of the time segment, we arrive at the usual formula

$$\langle \tilde{n}_I^*(t; f) \tilde{n}_I(t; f') \rangle = \frac{1}{2} \delta(f - f') P_I(t; |f|). \quad (3.14)$$

Again, the advantage of using Eq. (3.13) in expressing $\langle \tilde{n}_I(t; f) \tilde{n}_I(t; f') \rangle$ becomes evident when we set $f = f'$. The usual formula, Eq. (3.14), involving Dirac delta function diverges, whereas, in practice, that expression is actually finite because of a finite observation time. However, in Eq. (3.13), if we replace one finite time delta function by the Dirac delta function and treat the other as a normal function we obtain a finite result,

$$\langle |\tilde{n}_I(t; f)|^2 \rangle = \frac{1}{2} \Delta t P_I(t; |f|). \quad (3.15)$$

This is used in our work when generating simulated noise for our test analyses.

B. Optimal filters for anisotropic searches and the directed search

The aim of this subsection is to construct an optimal filter to maximize the SNR of the cross correlation statistic over the small time segments. We essentially generalize the analysis presented in [17] for isotropic background search to the case of anisotropic background search.

The optimal filter depends on the theoretical model of the SGWB, that is, on $P_A(\hat{\Omega})$ and $H(f)$. First we will derive the filter for the general case of the anisotropic search and then specialize to the directed search. Our first goal is to compute the SNR $\rho(t)$ of the statistic $\Delta S(t)$ over the time segments of length Δt at time t .

The equation for the general case is the generalization of Eq. (2.15):

$$\Delta S(t) = \int_{I(t)} dt' \int_{I(t)} dt'' s_1(t') s_2(t'') Q(t; t', t''), \quad (3.16)$$

where $I(t)$ is the interval $[t - \Delta t/2, t + \Delta t/2]$. Here we have suppressed the model dependence of Q . Assuming that the noise in both detectors and the Earth are stationary within the duration of each time segment, we may write $Q(t; t', t'') = Q(t; t' - t'')$, which allows us to expand the filter in terms of its SFT $\tilde{Q}(t, f)$ as

$$Q(t; t' - t'') = \int_{-\infty}^{\infty} df e^{2\pi i f(t' - t'')} \tilde{Q}(t, f). \quad (3.17)$$

Thus in terms of SFTs the statistic can be expressed as

$$\Delta S(t) = \int_{-\infty}^{\infty} df \tilde{s}_1^*(t; f) \tilde{s}_2(t; f) \tilde{Q}(t, f). \quad (3.18)$$

The mean of the statistic $\Delta S(t)$ is

$$\mu(t) := \langle \Delta S(t) \rangle = \int_{-\infty}^{\infty} df \langle \tilde{s}_1^*(t; f) \tilde{s}_2(t; f) \rangle \tilde{Q}(t, f). \quad (3.19)$$

Replacing s_I by h_I within the ensemble average, since the relevant correlations of signal and noise are zero except for $\langle \tilde{h}_1^*(t; f) \tilde{h}_2(t; f) \rangle$, we obtain from Eq. (3.10),

$$\mu(t) = \Delta t \int_{-\infty}^{\infty} df H(f) \gamma(t, f; S^2, d\mathcal{P}_A) \tilde{Q}(t, f). \quad (3.20)$$

Here the H and \mathcal{P}_A are the actual quantities pertaining to the SGWB source. The corresponding quantities of the theoretical model are hidden inside the filter Q .

The variance after a routine but fairly involved calculation is obtained as

$$\begin{aligned} \sigma^2(t) &= \langle [\Delta S(t) - \langle \Delta S(t) \rangle]^2 \rangle \\ &= \frac{\Delta t}{4} \int_{-\infty}^{\infty} df P_1(t; |f|) P_2(t; |f|) |\tilde{Q}(t, f)|^2. \end{aligned} \quad (3.21)$$

From these results the SNR $\rho(t)$ for the segment at time t can be computed. However, we need to maximize the SNR over each time segment. Here again we follow the prescription presented in [17]—we invoke the Schwarz inequality. To this end it is convenient to define a scalar product of two functions A and B on each time segment, labeled by t , as

$$(A, B)(t) := \int_{-\infty}^{\infty} df P_1(t; |f|) P_2(t; |f|) \tilde{A}^*(t; f) \tilde{B}(t; f). \quad (3.22)$$

The norm of a function $A(t)$ is defined as $\|A\|^2 = (A, A)$. Then the mean and variance are given in terms of the scalar product as follows:

$$\mu(t) = \Delta t \left(\frac{H(f) \gamma(t, f; S^2, d\mathcal{P}_A)}{P_1(t; |f|) P_2(t; |f|)}, Q(t) \right), \quad (3.23)$$

$$\sigma^2(t) = \frac{1}{4} \Delta t \|Q\|^2(t). \quad (3.24)$$

The SNR $\rho(t)$ is just the ratio $\mu(t)/\sigma(t)$. The SNR is maximized when the “signal” and the “filter” vectors are parallel, which happens when

$$\tilde{Q}(t, f) = \lambda(t) \frac{H(f) \gamma^*(t, f; S^2, d\mathcal{P}_A)}{P_1(t; |f|) P_2(t; |f|)}, \quad (3.25)$$

where $\lambda(t)$ is a (real) proportionality constant for the time segment. It is a function of the segment time t . This function will be chosen so that the SNR of S for the full observation is maximized.

For the optimal filter, the expression for the mean given by Eq. (3.20) simplifies to

$$\mu(t) = \frac{\Delta t \|Q\|^2(t)}{\lambda(t)}. \quad (3.26)$$

We exploit the freedom of choosing $\lambda(t)$ by setting $\mu(t) = 1$ for each time segment. This immediately gives

$$\lambda(t) = \Delta t \|Q\|^2(t), \quad (3.27)$$

$$\sigma^2(t) = \frac{\lambda(t)}{4}. \quad (3.28)$$

We further require to find $\lambda(t)$ explicitly. This is done by computing $\|Q\|^2$. We find $\|Q\|^2 = \lambda^2 P_{NW}^2$ where

$$P_{NW}^2(t) = \int_{-\infty}^{\infty} df \frac{H^2(f) |\gamma(t, f; S^2, d\mathcal{P}_A)|^2}{P_1(t; |f|) P_2(t; |f|)}. \quad (3.29)$$

The integral on the right-hand side (rhs) of the above equation is positive definite: Its integrand contains the fourth power of the GW amplitude. Therefore, it may be denoted by the square of a real quantity P_{NW} , which is determined by the GW power accessible to the network of the two detectors. The above equation gives the normalizations

$$\begin{aligned} \|Q(t)\|^2 &= [\Delta t P_{NW}(t)]^{-2}, \\ \lambda(t) &= [\Delta t P_{NW}^2(t)]^{-1} \equiv 4\sigma^2(t). \end{aligned} \quad (3.30)$$

The optimal statistic is then easily obtained as in Eq. (2.20). We replace the sum in that equation by an integral over the segment time t :

$$S = \frac{\int \Delta S(t) \sigma^{-2}(t) dt}{\int \sigma^{-2}(t) dt}, \quad (3.31)$$

where the integration is over the observation time (which could consist of disconnected time intervals).

The filter given by Eq. (3.25) is *the general optimal filter* to search for any anisotropic SGWB, which, unfortunately, requires the knowledge of $H(f)$ and $\mathcal{P}_A(\hat{\Omega})$. In practice, we do not have an exact *a priori* model for $H(f)$ and $\mathcal{P}_A(\hat{\Omega})$, which anyway we are trying to measure. We then must use models for those quantities, $H'(f)$ and $\mathcal{P}'_A(\hat{\Omega})$, to search for different anisotropic backgrounds. These models will be used to construct the model-dependent overlap reduction function $\gamma(t, f; \Delta\Omega, d\mathcal{P}'_A)$ and the filter Q .

The angular power distribution for only one unit point source on the sky in the direction $\hat{\Omega}$ with equal power in both the polarizations can be expressed as

$$\mathcal{P}_A(\hat{\Omega}') = \delta(\hat{\Omega}' - \hat{\Omega}). \quad (3.32)$$

This immediately simplifies the expression for the overlap reduction function because now the integral in Eq. (3.7) simplifies owing to the delta function, and γ becomes a function of $\hat{\Omega}$ as well. Therefore, we have

$$\gamma(\hat{\Omega}, t, f) = \Gamma(\hat{\Omega}, t) e^{2\pi i f \hat{\Omega} \cdot \Delta \mathbf{x}(t)/c}, \quad (3.33)$$

$$\Gamma(\hat{\Omega}, t) = \sum_A F_1^A(\hat{\Omega}, t) F_2^A(\hat{\Omega}, t). \quad (3.34)$$

Unlike the time-independent overlap reduction function of the isotropic SGWB case, the *direction-dependent overlap reduction function*, $\gamma(\hat{\Omega}, t, f)$, accepts power from *all* the frequencies and in fact has infinite bandwidth in the limit of vanishing pixel area. So the bandwidth of the filter Q in this case would only be limited by the bandwidth of the detectors through the coefficients $P_i(t, f)$.

If, instead, the source has a finite spatial extent, the bandwidth would be limited, because the integral in Eq. (3.7) would have to be performed over the solid angle $\Delta\Omega$ subtended by the source. If one takes a small patch of the sky of size $(\Delta\theta, \Delta\phi)$ around some fixed (source) direction $\hat{\Omega}_0 := (\theta_0, \phi_0)$, it is easy to show that γ is essentially a product of sinc functions in $(\Delta\theta, \Delta\phi)$. In fact,

$$\begin{aligned} & \int_{\Delta\Omega} d\hat{\Omega} e^{2\pi i f (\hat{\Omega} - \hat{\Omega}_0) \cdot \Delta \mathbf{x}/c} \\ &= |\Delta\Omega| \text{sinc} \frac{\pi f \Delta\theta \Delta \mathbf{x} \cdot \hat{\mathbf{e}}_\theta}{c} \text{sinc} \frac{\pi f \Delta\phi \Delta \mathbf{x} \cdot \hat{\mathbf{e}}_\phi}{c}, \end{aligned} \quad (3.35)$$

where $|\Delta\Omega| = \sin\theta \Delta\theta \Delta\phi$ is the solid angle subtended by the patch $\Delta\Omega$. In the integral, the factor $\Gamma(\hat{\Omega}, t)$ remains nearly constant. The sinc functions go to zero when their arguments reach π radians. Taking this definition as the bandwidth and taking $|\Delta \mathbf{x}|/c \sim 10$ ms for the two LIGO

detectors, the bandwidth is about 750 Hz for a square source of side 10° on the sky.

If there were known models for the anisotropic SGWB, the optimal filter for the general anisotropic case would have included $\mathcal{P}'_A(\hat{\Omega})$ and we would perform a *full sky search for an anisotropic background*. However, no reasonable model for the anisotropic SGWB sky exists in literature and, so, blind estimations are currently the only possible alternatives.

Directed search is one blind estimation approach, where the strength of each point (pixel) of the sky is “observed” using a direction-dependent filter, assuming that the other points on the sky do not contribute towards the observed value. So, in the directed case, Q becomes a function of $\hat{\Omega}$:

$$\tilde{Q}(\hat{\Omega}, t, f; H) = \lambda(\hat{\Omega}, t) \frac{H(f) \gamma^*(\hat{\Omega}, t, f)}{P_1(t; |f|) P_2(t; |f|)}, \quad (3.36)$$

where $\lambda(\hat{\Omega}, t)$ is the normalization constant, which now varies from pixel to pixel. It is given by Eq. (3.30), but with the γ in the expression for P_{NW}^2 in Eq. (3.29) replaced by the simpler Γ of Eq. (3.34). The directed filter given in Eq. (3.36) is an optimal filter if there is a single point source in the direction $\hat{\Omega}$ and *no sources elsewhere in the sky*. If there are other sources in the sky, as in a general anisotropic background, the filter becomes *suboptimal* as it stands. However, we use the above filter to make a “dirty” map of the sky, which is a convolution of the actual anisotropic background with the beam function and contains additive noise. We intend to extract information about the true background by the process of deconvolution. The convolution equation will be derived in the next subsection.

The working principle of the above filter is evidently similar to the earth rotation aperture synthesis often used in CMB and radio astronomy to make a map of a certain portion or the whole sky. The phase lag between two detectors, separated by a distance $\Delta \mathbf{x}(t)$, in receiving a plane wave front from a certain direction $\hat{\Omega}$, as shown in Fig. 1, is compensated in the filter via the phase factor $\exp[2\pi i f \hat{\Omega} \cdot \Delta \mathbf{x}(t)/c]$. As the earth rotates this factor adjusts, so that, waves from the given direction are coherently added, while the waves from other directions tend to cancel out. Note that, we did not introduce the phase factor by hand, it appeared automatically through the process of the maximization of the SNR. Though the whole radiometer analysis is based on this principle, the idea is clearly realized in the directed search analysis.

C. The integral equation for the directed search

In this subsection we set up the convolution equation, which is an integral equation for the statistic $S(\hat{\Omega})$ —the dirty map of directed search. The kernel of the integral equation consists of beam functions that we define below. The goal is to obtain the power in both polarizations

$\mathcal{P}_A(\hat{\Omega})$, given the statistic $S(\hat{\Omega})$. To this end we take the expectation value of the statistic $S(\hat{\Omega})$ in Eq. (3.31) and use Eq. (3.20) for the expectation value of ΔS inside the integral sign. We must also use expressions for the source overlap reduction function from Eq. (3.7) and the directed filter from Eq. (3.36). The final form of the (noiseless) convolution equation is given by

$$s(\hat{\Omega}) \equiv \langle S(\hat{\Omega}) \rangle = \int_{S^2} d\hat{\Omega}' [B^+(\hat{\Omega}, \hat{\Omega}') \mathcal{P}_+(\hat{\Omega}') + B^\times(\hat{\Omega}, \hat{\Omega}') \mathcal{P}_\times(\hat{\Omega}')], \quad (3.37)$$

where the beam functions $B^A(\hat{\Omega}, \hat{\Omega}')$ are defined as

$$B^A(\hat{\Omega}, \hat{\Omega}') = \Lambda(\hat{\Omega}) \int dt \int_{-\infty}^{\infty} df \frac{H(f)H'(f)}{P_1(t; |f|)P_2(t; |f|)} \times F_1^A(\hat{\Omega}', t) F_2^A(\hat{\Omega}', t) \Gamma(\hat{\Omega}, t) \times e^{-2\pi i f(\Delta\Omega \cdot \Delta\mathbf{x}(t)/c)} \quad (3.38)$$

$$\Lambda^{-1}(\hat{\Omega}) := \frac{1}{\Delta t} \int \lambda^{-1}(t, \hat{\Omega}) dt. \quad (3.39)$$

The function $H'(f)$ is the model power spectrum of the source we insert in the kernel and $\Delta\Omega = \hat{\Omega} - \hat{\Omega}'$. We measure $\langle S(\hat{\Omega}) \rangle$ and from the kernels $B^A(\hat{\Omega}, \hat{\Omega}')$, we propose to solve the integral equation for $\mathcal{P}_A(\hat{\Omega})$. Physically, we may expect the power in both polarizations to be the same, that is, $\mathcal{P}_+(\hat{\Omega}) = \mathcal{P}_\times(\hat{\Omega}) = \mathcal{P}(\hat{\Omega})$, say, the kernel then is just the sum of the two individual kernels of each polarization,

$$B(\hat{\Omega}, \hat{\Omega}') = B^+(\hat{\Omega}, \hat{\Omega}') + B^\times(\hat{\Omega}, \hat{\Omega}') = \Lambda(\hat{\Omega}) \int dt \int_{-\infty}^{\infty} df \frac{H(f)H'(f)}{P_1(t; |f|)P_2(t; |f|)} \times \Gamma(\hat{\Omega}', t) \Gamma(\hat{\Omega}, t) e^{-2\pi i f \Delta\Omega \cdot \Delta\mathbf{x}(t)/c}. \quad (3.40)$$

Our numerical deconvolution strategy is described in the next section. But before we do that we examine the kernel in Eq. (3.40) and try to understand it from a physical point of view. This will afford us some insight into the beam patterns associated with directed filters.

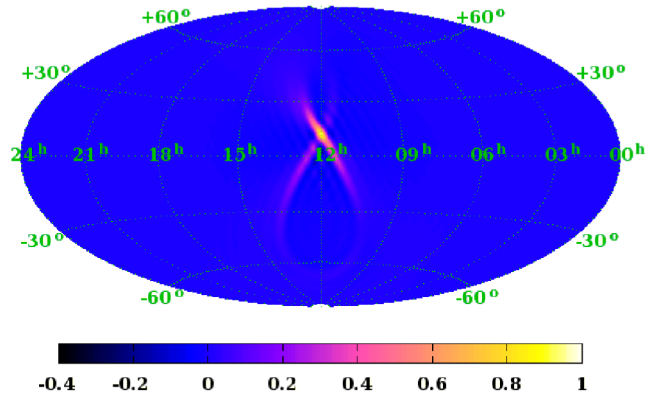
It is worth noting that the beam function $B(\hat{\Omega}, \hat{\Omega}')$ is not symmetric only due to the leading normalization factor $\Lambda(\hat{\Omega})$, which comes from the normalization of the statistic $S(\hat{\Omega})$. We make use of this observation to introduce a symmetric kernel in Sec. V, which is advantageous for numerical deconvolution.

D. The stationary phase approximation of the kernel and its singular value decomposition

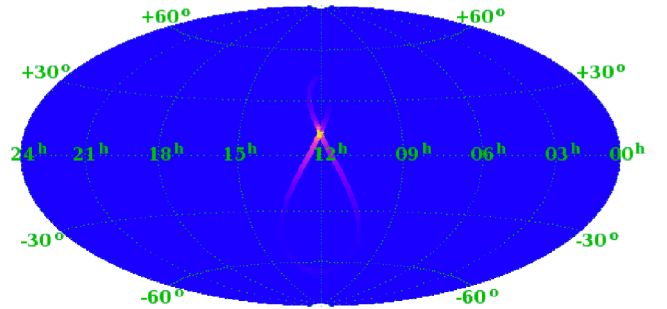
The GW radiometer beams are not pointed but have a spread out profile, which varies with sky position. Thus in order to make progress towards deconvolving the GW sky map we try to understand the beam pattern. We find that the

stationary phase approximation (SPA) of $B(\hat{\Omega}, \hat{\Omega}')$, given in Eq. (3.40), yields useful results. It is essentially the exponential term $\exp[-2\pi i f \Delta\Omega \cdot \Delta\mathbf{x}(t)/c]$ containing the phase that determines the integral—the integrand constructively contributes when the phase in the integral in Eq. (3.40) is stationary. We also use the fact that the rest of the functions in the integrand vary slowly with time, so that they effectively behave as constants as far as the integral is concerned.

We obtain the beam function for a unit point source at $\hat{\Omega}_0 \equiv (\theta_0, \phi_0)$. We write $\Delta\Omega := \hat{\Omega} - \hat{\Omega}_0$. Note that $\Delta\Omega$ may not be necessarily small; the points $\hat{\Omega}, \hat{\Omega}_0$ can lie anywhere on the unit sphere. By performing a numerical computation for an observation time of one sidereal day, we find that for low declinations, the beam is shaped like the figure of 8, as shown in Fig. 2(a), while as one goes higher in declination, the 8 smoothly turns into a tear drop.



(a) Numerical beam pattern.



(b) Theoretical beam pattern obtained by stationary phase approximation (SPA)

FIG. 2 (color online). Illustration of the agreement between numerical and theoretical GW radiometer beam patterns at declination $+12^\circ$ for the LIGO detectors at Livingston and Hanford (with white noise, upper cutoff frequency of 1024 Hz, $H(f) = \text{constant}$ and observation time of one sidereal day). Same color map has been used for both panels. Note that, the *shape* of the beam does not depend on the right ascension of the pointing direction for observation time of full sidereal day(s).

With the application of the SPA we can explain the shape of the beam. The integrand in the kernel usually oscillates rapidly, because of the exponential phase term, except when the phase is stationary. The kernel is a double integral over f and t and therefore the SPA must be carried out in two dimensions. Setting the first derivative of the phase with respect to both variables f and t equal to zero, we obtain

$$\Delta \boldsymbol{\Omega} \cdot \Delta \mathbf{x}(t) = 0, \quad (3.41)$$

$$\Delta \boldsymbol{\Omega} \cdot \Delta \dot{\mathbf{x}}(t) = 0, \quad (3.42)$$

where $\Delta \dot{\mathbf{x}}(t) := d\Delta \mathbf{x}(t)/dt$. The detector separation vector $\Delta \mathbf{x}(t)$ rotates about the Earth's rotation axis (z -axis in our coordinate system) with the angular velocity ω_E . Geometrically $\Delta \mathbf{x}(t)$ traces out a right circular cone with z -axis as its symmetry axis [see Fig. 3]. It is explicitly given by

$$\Delta \mathbf{x}(t) = \Delta R(\sin\Theta \cos\omega_E t, \sin\Theta \sin\omega_E t, \cos\Theta), \quad (3.43)$$

where $\Delta R = |\Delta \mathbf{x}(t)|$ is the constant distance between the detectors. As $\omega_E t$ ranges from 0 to 2π , $\Delta \mathbf{x}(t)$ traces out a cone with half-angle Θ . The half-angle Θ of the cone, $0 \leq \Theta \leq \pi$, is given by

$$\cos\Theta = \frac{\cos\theta_1 - \cos\theta_2}{\sqrt{2[1 - \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2)]}}, \quad (3.44)$$

where (θ_i, ϕ_i) are the detector coordinates. For the LIGO pair of detectors, $\Theta \sim 27^\circ$.

From Eqs. (3.41) and (3.42) it is clear that the phase is stationary when $\Delta \mathbf{x}(t)$, $\Delta \dot{\mathbf{x}}(t)$, and $\Delta \boldsymbol{\Omega}$ form an orthogonal triad. Since the unit vector normal to the baseline at any given time as well as to the cone

$$\hat{\mathbf{n}}_{\text{cone}}(t) = \frac{\Delta \mathbf{x}(t) \times \Delta \dot{\mathbf{x}}(t)}{|\Delta \mathbf{x}(t) \times \Delta \dot{\mathbf{x}}(t)|}, \quad (3.45)$$

the SPA condition is satisfied when

$$\Delta \boldsymbol{\Omega} = \Delta \Omega \hat{\mathbf{n}}_{\text{cone}}(t), \quad (3.46)$$

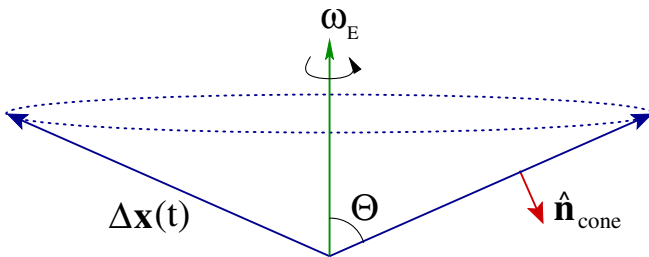


FIG. 3 (color online). The baseline formed by two detectors, $\Delta \mathbf{x}(t)$, traces out a cone as the earth rotates. A schematic diagram is shown here. The vector $\hat{\mathbf{n}}_{\text{cone}}(t)$ is normal to the baseline as well as the cone.

where $\Delta \Omega = |\Delta \boldsymbol{\Omega}|$ can take *both positive or negative* values. Since both $\hat{\boldsymbol{\Omega}}_0$ and $\hat{\boldsymbol{\Omega}}(t) = \hat{\boldsymbol{\Omega}}_0 + \Delta \boldsymbol{\Omega}(t)$ are constrained to lie on the unit sphere and thus both have unit norm, it follows from Eq. (3.46) that the SPA solution $\hat{\boldsymbol{\Omega}}(t)$ is a curve on the unit sphere given by [23],

$$\hat{\boldsymbol{\Omega}}(t) = \hat{\boldsymbol{\Omega}}_0 - 2[\hat{\boldsymbol{\Omega}}_0 \cdot \hat{\mathbf{n}}_{\text{cone}}(t)]\hat{\mathbf{n}}_{\text{cone}}(t). \quad (3.47)$$

The trajectory has been parametrized in terms of the sidereal time t . One can even obtain an approximate analytical expression for the beam function along the SPA trajectory using standard SPA techniques as [23]

$$B(\hat{\boldsymbol{\Omega}}(t), \hat{\boldsymbol{\Omega}}_0) = \Lambda(\hat{\boldsymbol{\Omega}}(t))\Gamma(\hat{\boldsymbol{\Omega}}(t); t)\Gamma(\hat{\boldsymbol{\Omega}}_0; t) \frac{\sqrt{f_u} - \sqrt{f_l}}{\omega_E} \times \sqrt{\frac{8c}{|[\hat{\mathbf{z}} \cdot \Delta \mathbf{x}(t)][\hat{\mathbf{z}} \cdot (\hat{\boldsymbol{\Omega}}(t) - \hat{\boldsymbol{\Omega}}_0)]|}}. \quad (3.48)$$

As t is varied over a full sidereal day, the shaded figure of 8 is generated through Eqs. (3.47) and (3.48) as shown in Fig. 2(b). Clearly, SPA results match very well with the numerical beam pattern shown in Fig. 2(a).

The case where Eq. (3.48) does not apply (though the analysis still remains valid) is when the detectors are at the same latitude, as the normal to the baseline cone, $\hat{\mathbf{n}}_{\text{cone}}(t)$, always remains parallel to the $\hat{\mathbf{z}}$ axis, causing the denominator of Eq. (3.48) to vanish. In this case the whole SPA trajectory shrinks to a point, which is the image of the pointing direction about the equatorial plane, $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Omega}}_0 - 2[\hat{\boldsymbol{\Omega}}_0 \cdot \hat{\mathbf{z}}]\hat{\mathbf{z}}$. The value of the beam function at the image point,

$$\int_0^T dt \Gamma(\hat{\boldsymbol{\Omega}}_0, t)\Gamma(\hat{\boldsymbol{\Omega}}_0 - 2[\hat{\boldsymbol{\Omega}}_0 \cdot \hat{\mathbf{z}}]\hat{\mathbf{z}}, t) / \int_0^T dt [\Gamma(\hat{\boldsymbol{\Omega}}_0, t)]^2, \quad (3.49)$$

is also quite easy to compute. Therefore, a sky map produced by such a baseline will be a superposition of the (blurred) true sky and its (differently blurred) reflection about the equatorial plane. In practice, a situation like this arises for the LHO-Virgo pair, as their latitudes are quite close, $46^\circ 27''\text{N}$ and $43^\circ 37''\text{N}$, respectively.

The SPA solution given in Eq. (3.48) does not remain finite very close to the pointing direction, as the denominator vanishes. However, close to the pointing direction a better approximation is obtained by expanding the phase term $\exp[-2\pi i f \Delta \boldsymbol{\Omega} \cdot \Delta \mathbf{x}(t)/c]$ up to the *second* order, which can also be used for more accurate modeling of the core of the beam.

The stationary phase analysis also indicates an approximate resolving power of the radiometer. Since near the maximum of the beam function, the phase must not vary too much over the bandwidth, say no more than a radian, this implies that the resolving power is determined by $|\Delta \boldsymbol{\Omega}| \sim \lambda/|\Delta \mathbf{x}|$, where the λ corresponds to the bandwidth $f = \Delta f$ of the detectors. For the LIGO detectors,

$|\Delta \mathbf{x}| \sim 3000$ km. If we take the bandwidth to be ~ 1 kHz, then $\lambda \sim 300$ km, so the radiometer resolution is ~ 0.1 radians, that is, ~ 6 degrees, which we find is consistent with the numerically obtained beam profiles. A contour plot of the core of the beam of the radiometer formed by the LIGO pairs (with white noise and upper cutoff frequency of 1024 Hz) is shown in Fig. 4. The plot confirms that the beam size in this case is $6^\circ \sim 0.1$ rad.

The width of the beam estimated above is seen to be consistent with the “number of degrees of freedom” present in the kernel (beam). A widely used method that identifies the linearly independent modes in a linear transformation is the singular value decomposition (SVD) [25]. The decomposition identifies linear combinations of modes that have almost zero eigenvalues—the null subspace. It then projects out the solution orthogonal to the null subspace which spans the true degrees of freedom. The singular values of the kernel for the LIGO detectors at Hanford and Livingston for white noise with upper cutoff frequency $f_u = 1024$ Hz are plotted in Fig. 5 (solid line). The figure shows that the eigenvalues become essentially negligible after ~ 1000 implying that this is the number of degrees of freedom in the kernel. The numerical and the SPA analysis shows that the size of the central spot or the resolution is ~ 0.1 radian, which means that there are

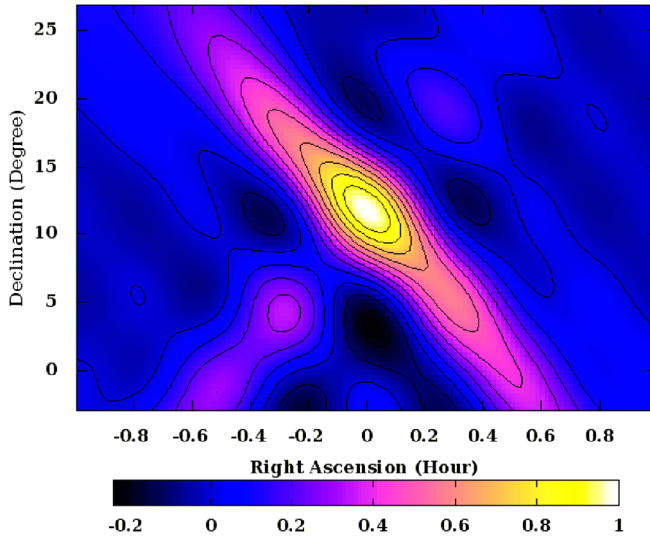


FIG. 4 (color online). Contour plot of GW radiometer beam pattern at declination $+12^\circ$ revealing the approximate size and the (negative) side lobes of the beam for the LIGO detectors with white noise, upper cutoff frequency $f_u = 1024$ Hz, $H(f) = \text{constant}$, and observation time of one sidereal day. The contours are drawn starting from the highest level of 0.9 with a difference in levels of 0.1. The beam falls by $1/e$ in between the 5th and the 6th contour (from the highest value), which is a good indicator for the beam size. Clearly, the narrowest beam size (along the “minor axis”) is in reasonable agreement with the theoretical prediction of $6^\circ \sim 0.1$ rad. The beam becomes broader (not shown in the figure) for real detector noise and negative source spectral index, e.g., $H(f) \propto f^{-3}$.

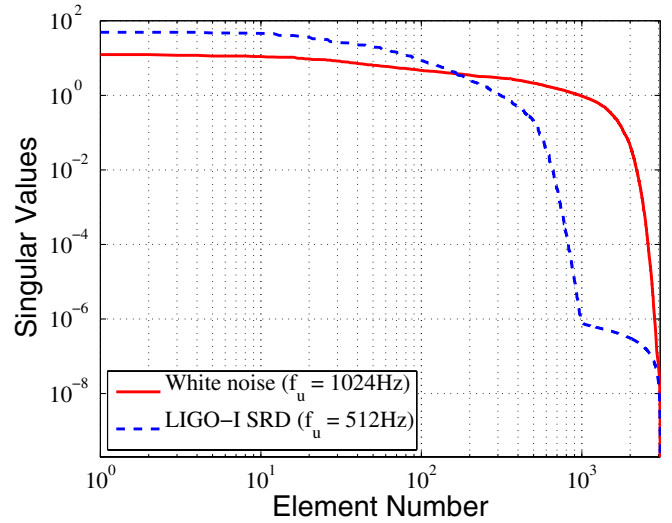


FIG. 5 (color online). The solid line is a plot of the singular values of the kernel for the LIGO pair of detectors with white noise, upper cutoff frequency of $f_u = 1024$ Hz, and observation time of one sidereal day. From the plot it is evident that the singular values are negligible after ~ 1000 . These results agree with the numerical and SPA results which give the size of the central spot ~ 0.1 radian. The dotted line shows the singular values for the same detector pair, but with LIGO-I noise PSD and upper cutoff frequency of $f_u = 512$ Hz. Clearly, the latter curve indicates a loss of angular resolution due to relatively poorer higher frequency response of the detectors.

$4\pi/(0.1)^2 \sim 1000$ independent patches in the sky. So the SVD results are consistent with the size of the beam obtained by numerical and theoretical methods.

In practice, however, the higher frequency response of the detectors is not as good, hence the achievable angular sensitivity of the radiometer becomes relatively poor. The plot of singular values for the same detector pair both having LIGO-I goal noise PSD with an upper cutoff frequency of $f_u = 512$ Hz is overlaid (dashed line) on Fig. 5. Clearly, the number of degrees of freedom, which represents the amount of information content in a map, is less in this case.

IV. THE MAXIMUM LIKELIHOOD DECONVOLUTION

A. Unpolarized background and single baseline

We first consider the simpler case of detecting and deconvolving the signal from an unpolarized SGWB using one pair of detectors. We later indicate in the subsections that follow how this method can be extended to the more general cases of SGWBs and detector baselines.

The observed data construct, $S(\hat{\Omega})$, consists of a signal and additive noise, namely,

$$S(\hat{\Omega}) = s(\hat{\Omega}) + n(\hat{\Omega}). \quad (4.1)$$

The first term on the rhs, the expectation of the dirty map

given in Eq. (3.37), is a convolution of the true power in SGWB $\mathcal{P}_A(\hat{\Omega})$ in the two polarizations arriving from a direction $\hat{\Omega}$ in the sky with corresponding beam response functions $B^A(\hat{\Omega}, \hat{\Omega}')$. Note that, though the definitions of the above quantities involve complex Fourier transforms, these quantities are all *real* owing to the fact that they were originally derived from real time series, so that, $\tilde{s}_I^*(t; -f) = \tilde{s}_I(t; f)$, $\gamma^*(\hat{\Omega}, t, -f) = \gamma(\hat{\Omega}, t, f)$, and so on.

In this section, we construct the maximum likelihood (ML) estimator for the angular power distribution $\mathcal{P}_A(\hat{\Omega})$ given the measured data $S(\hat{\Omega})$. For simplicity and clarity of presentation, we first limit the analysis to the simple case where both the polarizations follow the same angular power distribution $\mathcal{P}(\hat{\Omega})$. This simplifies the form of the construct to

$$S(\hat{\Omega}) = \int_{S^2} d\hat{\Omega}' [B_+(\hat{\Omega}, \hat{\Omega}') + B_\times(\hat{\Omega}, \hat{\Omega}')] \mathcal{P}(\hat{\Omega}') + n(\hat{\Omega}). \quad (4.2)$$

In practice, the sky is divided into a finite number of pixels. Then, the observed data vector is denoted by \mathbf{S} , whose component $S_i := S(\hat{\Omega}_i)$ is the signal measured at the i th pixel. We similarly define the vectors \mathcal{P} and \mathbf{n} , with components $\mathcal{P}(\hat{\Omega}_i)$ and $n(\hat{\Omega}_i)$, respectively. In this notation the convolution leads to a set of linear algebraic equations,

$$\mathbf{S} = \mathbf{B} \cdot \mathcal{P} + \mathbf{n}, \quad (4.3)$$

where \mathbf{B} is the *known* [26] beam matrix, expressed as $B_{ij} := B_+(\hat{\Omega}_i, \hat{\Omega}_j) + B_\times(\hat{\Omega}_i, \hat{\Omega}_j)$. In the weak-signal approximation, the variances of the signal-noise cross terms $\tilde{h}_{1,2}^*(t; f)\tilde{n}_{2,1}(t; f)$ are much smaller than the variance of the noise-noise cross term $\tilde{n}_1^*(t; f)\tilde{n}_2(t; f)$. So the observed pixel noise is strongly dominated by the noise-noise term and can be written as

$$\begin{aligned} n_i &:= n(\hat{\Omega}_i) \\ &= \left[\int dt \lambda^{-1}(t, \hat{\Omega}_i) \right]^{-1} \int dt \int_{-\infty}^{\infty} df \tilde{n}_1^*(t; f) \tilde{n}_2(t; f) \\ &\quad \times \frac{\tilde{Q}(\hat{\Omega}_i, t, f; H)}{\lambda(t, \hat{\Omega}_i)}; \end{aligned} \quad (4.4)$$

the cross terms $\tilde{h}_{1,2}^*(t; f)\tilde{n}_{2,1}(t; f)$ have been dropped. The pixel noise is a sum of a large number of zero mean random numbers, where none of the addend strongly dominate (statistically) over the others. Hence, following the generalized central limit theorem [27], one can argue that the pixel noise tends to be zero mean Gaussian. If this argument is used to calculate the variance of $n(\hat{\Omega})$, we consistently get the same result as expected from Eq. (2.16). After a routine but fairly involved algebra, the pixel-to-pixel noise covariance matrix, $\mathbf{N} \equiv N_{ij}$, of the dirty map

turns out to be

$$N_{ij} := \langle n_i n_j \rangle = \frac{\Delta t}{4} \left[\int dt \lambda^{-1}(t, \hat{\Omega}_j) \right]^{-1} B_{ij}. \quad (4.5)$$

The deconvolution problem is, of course, a very standard problem in many areas of science. In particular, Eq. (4.3) is identical in structure to the set of equations that arise in the map-making stage of CMB experiments. The temperature anisotropy ΔT_i in a direction is inferred from time stream data, d_t using a linear model $d_t = \sum_i A_{ti} \Delta T_i + n_t$. The convolution kernel A_{ti} that relates the time domain to the pixel domain is determined by the pointing or scan strategy as well as the beam response function of the antenna. The noise n_t is (assumed to be) Gaussian and described by the noise covariance matrix $N_{tt'}$. As described below, a ML solution for the sky map ΔT_i is readily obtained in this linear model under the assumption that the noise is Gaussian. We adapt this technique to solve our problem since it has been applied with great success in the CMB field and there exists an extensive literature [28] and also public domain package for implementing it numerically [29,30]. However, it should be noted that the problem differs in two important aspects. First, in our case there is the simplicity that the kernel connects two vectors which are both in sky pixel space. This implies the kernel is a square matrix for the case of single baseline and unpolarized background. Second, in our case the statistics of the noise is potentially nontrivial; the noise in this situation is a complex object built by integrating the product of two random variables corresponding to the noise streams in each detector. The Gaussianity of the noise has to arise from the generalized central limit theorem [27].

We proceed assuming that the joint probability distribution of the elements of \mathbf{n} is a multivariate Gaussian distribution [31] given by the probability distribution function:

$$\begin{aligned} \mathfrak{P}(\mathbf{n}) &= \frac{1}{(2\pi)^{N_{\text{pix}}/2} \|\mathbf{N}\|^{1/2}} \exp\left[-\frac{1}{2} \mathbf{n}^T \cdot \mathbf{N}^{-1} \cdot \mathbf{n}\right] \\ &= \frac{1}{(2\pi)^{N_{\text{pix}}/2}} \exp\left[-\frac{1}{2} (\mathbf{n}^T \cdot \mathbf{N}^{-1} \cdot \mathbf{n} + \text{Tr}[\ln \mathbf{N}])\right], \end{aligned} \quad (4.6)$$

where N_{pix} is the total number of pixels and $\mathbf{N} := \langle \mathbf{n} \mathbf{n}^T \rangle$ is the *known* noise covariance matrix [see Eq. (4.5)]. Thus, given a signal vector \mathcal{P} , the probability of observing radiometer output \mathbf{S} is [29]

$$\begin{aligned} \mathfrak{P}(\mathbf{S}|\hat{\mathcal{P}}) &= (2\pi)^{-N_{\text{pix}}/2} \\ &\quad \times \exp\left[-\frac{1}{2} (\mathbf{S} - \mathbf{B} \cdot \hat{\mathcal{P}})^T \cdot \mathbf{N}^{-1} \cdot (\mathbf{S} - \mathbf{B} \cdot \hat{\mathcal{P}}) \right. \\ &\quad \left. + \text{Tr}[\ln \mathbf{N}]\right]. \end{aligned} \quad (4.7)$$

Solving for $\partial \mathfrak{P}(\mathbf{S}|\hat{\mathcal{P}}) / \partial \hat{\mathcal{P}} = \mathbf{0}$, using the fact that \mathbf{N} is symmetric and positive definite, it is straightforward to show that the above probability is maximum when

$$\hat{\mathcal{P}} = (\mathbf{B}^T \mathbf{N}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{N}^{-1} \cdot \mathbf{S}, \quad (4.8)$$

which is, therefore, the well-known result for the desired ML estimator of the true sky map of SGWB anisotropy. The deconvolved map will also have pixel noise, given by

$$\begin{aligned} \mathbf{n} &:= \hat{\mathcal{P}} - \mathcal{P} = (\mathbf{B}^T \mathbf{N}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{N}^{-1} \cdot (\mathbf{B} \cdot \mathcal{P} + \mathbf{n}) - \mathcal{P} \\ &= (\mathbf{B}^T \mathbf{N}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{N}^{-1} \cdot \mathbf{n} \end{aligned} \quad (4.9)$$

and the pixel-to-pixel noise covariance matrix of the ML map is obtained as

$$\boldsymbol{\Sigma} := \langle \mathbf{n} \mathbf{n}^T \rangle = (\mathbf{B}^T \mathbf{N}^{-1} \mathbf{B})^{-1}. \quad (4.10)$$

Therefore, to obtain the ML map estimate one has to first compute the inverse of the pixel-to-pixel noise covariance matrix $\boldsymbol{\Sigma}^{-1} = \mathbf{B}^T \mathbf{N}^{-1} \mathbf{B}$. The ML map is then obtained as a solution to the linear algebraic equations,

$$\boldsymbol{\Sigma}^{-1} \hat{\mathcal{P}} = \mathbf{B}^T \mathbf{N}^{-1} \cdot \mathbf{S}. \quad (4.11)$$

In solving this equation, we have a choice of either using one of a number of direct methods or one of the iterative methods. In the low resolution regime, the matrix inversion looks feasible with reasonable accuracy. However, in general, direct methods are computationally more expensive and iterative methods are preferred, provided their convergence to the solution is rapid. For sparse matrices the iterative conjugate gradient (CG) method is well suited for this inversion [25,32] problem. The CG method solves linear systems with symmetric positive definite matrix and have been found to be more advantageous compared to other iterative methods such as the Jacobi method [28,33]. Starting with a guess solution, the convergence of the method can be often greatly improved by ‘‘preconditioning’’ the system of equations. i.e., multiplying both sides with a suitable matrix (say, the inverse of diagonal elements of the matrix). Our choice is also motivated by the fact that the CG method has been successfully imple-

mented for map making in CMB experiments [28,29,34–36].

The above method can be extended to include multiple baselines and also to estimate power in each polarization component. We discuss these extensions in the following subsections.

B. Multiple baselines

The above analysis can be extended to a set of N_b GW radiometer baselines. Let $\mathbf{S}^{(i)}$ be the observed map by the i th baseline with beam matrix $\mathbf{B}^{(i)}$ and observed noise $\mathbf{n}^{(i)}$. Then Eq. (4.3) can still be written as

$$\mathbf{S} = \mathbf{B} \cdot \mathcal{P} + \mathbf{n},$$

where

$$\mathbf{X} := \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \\ \vdots \\ \mathbf{X}^{(N_b)} \end{pmatrix}, \quad (4.12)$$

with \mathbf{X} representing the matrices \mathbf{S} , \mathbf{B} , and \mathbf{n} . Note that \mathbf{S} , \mathbf{n} are now $1 \times N_{\text{pix}} N_b$ vectors and \mathbf{B} is a $N_{\text{pix}} \times N_{\text{pix}} N_b$ matrix, while \mathcal{P} (the true SGWB sky) remains unchanged. This is similar to CMB experiments where each pixel is visited by the detector several times. In the multibaseline GW radiometer case each pixel is visited by different baselines, and, unlike a CMB experiment, each pixel is visited an equal number of times. In this case too the maximum likelihood estimate [37] and the pixel-to-pixel noise covariance matrix of the ML map are given by

$$\hat{\mathcal{P}} = \boldsymbol{\Sigma} \mathbf{B}^T \mathbf{N}^{-1} \cdot \mathbf{S}; \quad \boldsymbol{\Sigma}^{-1} := \mathbf{B}^T \mathbf{N}^{-1} \mathbf{B}, \quad (4.13)$$

but the noise covariance matrix $\mathbf{N} := \langle \mathbf{n} \mathbf{n}^T \rangle$ of the raw sky map has to be modified. Let $n_i(\hat{\boldsymbol{\Omega}})$ be the pixel noise from the radiometer baseline i with detectors I and I' and we follow the same convention for λ_i and Q_i . Then,

$$\begin{aligned} \langle n_i(\hat{\boldsymbol{\Omega}}_1) n_j(\hat{\boldsymbol{\Omega}}_2) \rangle &= \left[\int dt \lambda_i^{-1}(t, \hat{\boldsymbol{\Omega}}_1) \right]^{-1} \left[\int dt \lambda_j^{-1}(t, \hat{\boldsymbol{\Omega}}_2) \right]^{-1} \int dt_1 \int dt_2 \int_{-\infty}^{\infty} df_1 \int_{-\infty}^{\infty} df_2 \int_{-\infty}^{\infty} df'_1 \\ &\times \int_{-\infty}^{\infty} df'_2 \delta_{\Delta I}(f_1 - f'_1) \delta_{\Delta I'}(f_2 - f'_2) \langle \tilde{n}_I^*(t_1, f_1) \tilde{n}_J^*(t_2, f_2) \tilde{n}_{I'}(t_1, f'_1) \tilde{n}_{J'}(t_2, f'_2) \rangle \frac{\tilde{Q}_i(\hat{\boldsymbol{\Omega}}_1, t_1, f'_1; H)}{\lambda_i(t_1, \hat{\boldsymbol{\Omega}}_1)} \\ &\times \frac{\tilde{Q}_j(\hat{\boldsymbol{\Omega}}_2, t_2, f'_2; H)}{\lambda_j(t_2, \hat{\boldsymbol{\Omega}}_2)}. \end{aligned} \quad (4.14)$$

If i and j denote the same baseline we get back the previous result [Eq. (4.5)]. However, if i and j denote different baselines, at least one of the two detector pairs will be different (i.e. either $I \neq J$ or $I' \neq J'$), so in that case $\langle n_i(\hat{\boldsymbol{\Omega}}_1) n_j(\hat{\boldsymbol{\Omega}}_2) \rangle = 0$. Hence, one can write

$$\langle n_i(\hat{\boldsymbol{\Omega}}_1) n_j(\hat{\boldsymbol{\Omega}}_2) \rangle = \delta_{ij} \langle n_i(\hat{\boldsymbol{\Omega}}_1) n_j(\hat{\boldsymbol{\Omega}}_2) \rangle. \quad (4.15)$$

Thus, the matrix \mathbf{N} for a network of baselines will be a

block diagonal matrix, where each diagonal block $\mathbf{N}^{(i)}$ is the noise covariance matrix for the corresponding baseline, $N_{kk'}^{(i)} := \langle n_i(\hat{\boldsymbol{\Omega}}_k) n_i(\hat{\boldsymbol{\Omega}}_{k'}) \rangle$, provided in Eq. (4.5).

The above algebra suggests that it is fairly straightforward to combine the observations made by multiple baselines for the estimation of the true SGWB angular power distribution. In [17,38] it was shown that the optimal way to search for an isotropic SGWB using a network of

detectors (with uncorrelated noise) is to linearly combine correlations from pairs of detectors, instead of computing higher order correlations using data from more than two detectors. We can extrapolate the same logic to the directed search and argue that the procedure described above to combine data from a network of detectors is also optimal.

The search for a SGWB using a network of detectors is becoming progressively relevant as other kilometer scale detectors, namely, Virgo, GEO, and LCGT, are expected to reach their initial goal sensitivity in the next few years. A network of detectors can enhance the directed search in many ways. The resolution of a radiometer is proportional to the length of the baseline. Inclusion of a detector at a distance like Virgo, which is further away of the LIGO detectors than the mutual separation of the LIGO sites, will clearly increase the highest baseline separation and hence the resolution of the baseline. However, more important enhancement would be realized due to better coverage of the sky. An analogy with radio astronomy using an array of antennas may be appropriate to mention in this context [39]. As the Earth rotates, the projections of the radio antenna baselines on the plane perpendicular to the source direction sample different points on the two-dimensional Fourier plane (commonly known as the u - v plane). The sampled Fourier plane is then inverse transformed to generate the image. While the *highest resolution* of the network is limited by the projection of maximum antenna separation, the addition of more antennas to the network ensures better sampling of the u - v plane reducing the side lobes, thereby producing a more faithful image of the sky. A detailed introduction to the basic principles of Earth rotation aperture synthesis can be found in most of the standard texts on radio astronomy, e.g., [40]. In GW radiometry with a network of detectors we expect that a similar scenario will arise—better coverage of the sky should be possible due to different orientations of the baselines with respect to the source. Moreover, since the true power distribution will be estimated from an overconstrained set of equations, the error in the estimated quantities will be reduced. In addition, a radiometer search can benefit from certain technical advantages that a network of detectors can provide: A detector at a third site joining the LIGO detectors will boost the “single-baseline integration time,” i.e., the single-baseline duty cycle in a three-site network will be at least as good as, but will likely be better than, that in a two-site network. Also, owing to common instrumental noise sources in the LIGO detectors, certain frequency bands are currently notch-filtered in computing the cross correlation statistic. Some of these noise sources are known not to affect Virgo and, therefore, the LIGO-Virgo cross correlation statistics. For example, the powerline noise affects the LIGO detectors at the harmonics of 60 Hz, whereas it affects the Virgo detector at the harmonics of 50 Hz. Therefore, a radiometer search that benefits from the LIGO-Virgo baseline’s contribution will probe

the presence of astrophysical signals over a larger set of frequencies than one limited to the baseline consisting of the LIGO pair of detectors.

C. Polarization map

We may also choose to extract power from different polarizations separately. The discrete convolution equation,

$$\mathbf{S} = \mathbf{B}_+ \cdot \mathcal{P}_+ + \mathbf{B}_\times \cdot \mathcal{P}_\times + \mathbf{n}, \quad (4.16)$$

can also be expressed by Eq. (4.3)

$$\mathbf{S} = \mathbf{B} \cdot \mathcal{P} + \mathbf{n},$$

where

$$\mathbf{B} := \begin{bmatrix} \mathbf{B}_+ & \mathbf{B}_\times \end{bmatrix}; \quad \mathcal{P} := \begin{pmatrix} \mathcal{P}_+ \\ \mathcal{P}_\times \end{pmatrix}. \quad (4.17)$$

Again in this case the ML estimator of sky map can be expressed by Eq. (4.13):

$$\hat{\mathcal{P}} = \Sigma \mathbf{B}^T \mathbf{N}^{-1} \cdot \mathbf{S}; \quad \Sigma^{-1} := \mathbf{B}^T \mathbf{N}^{-1} \mathbf{B}.$$

This case can also be generalized for a network of detectors by retaining the same definition of \mathcal{P} [Eq. (4.17)], but redefining the beam matrix as

$$\mathbf{B} := \begin{bmatrix} \mathbf{B}_+^{(1)} & \mathbf{B}_\times^{(1)} \\ \mathbf{B}_+^{(2)} & \mathbf{B}_\times^{(2)} \\ \vdots & \vdots \\ \mathbf{B}_+^{(N_b)} & \mathbf{B}_\times^{(N_b)} \end{bmatrix} \quad (4.18)$$

and, again, using the same ML estimation formula given in Eq. (4.13).

V. IMPLEMENTATION AND NUMERICAL RESULTS

In this work we have numerically implemented the maximum likelihood estimation algorithm on simulated data using the MATLAB® software package [41] to estimate the “true” (unpolarized) SGWB sky observed with a single-baseline ground-based GW radiometer. The details of the computation scheme, the numerical deconvolution algorithm, the simulated data, and the deconvolved maps are presented in this section.

A. Preparation of simulated dirty maps

The data are simulated in the frequency domain for each time segment. Since the noise of ground-based interferometric detectors is very high at frequencies greater than a few 100 Hz and the computation cost increases with the number of frequency bins, we use an upper cutoff frequency of $f_u = 512$ Hz and bin width of $\Delta f = 2$ Hz for testing of the algorithm. The length of each time segment is chosen as $\Delta t = 192$ sec and the total integration time

is $T = 86\,400$ sec [42]. The sky is pixelized using the Hierarchical Equal Area isoLatitude Pixelization (HEALPix) [43] scheme, which divides the 2-sphere (S^2) in $12n_{\text{side}}^2$ pixels, where n_{side} is an integer power of 2. Since the radiometer beam width is greater than $\sim 6^\circ$, we chose $n_{\text{side}} = 16$, which corresponds to a pixel width of $\sim 3^\circ$ and a total of 3072 pixels. The HEALPix scheme also allows fast spherical harmonic transform on a sphere, which may become useful for more advanced analysis in future. Note that, the algorithm is *independent* of the pixelization scheme; other equal area pixelization schemes can also be used in the analysis.

We generate the detector noise $\tilde{n}_I(t; f)$ using a Gaussian pseudo random number generator for each time segment. The noise is colored using the (one-sided) noise PSD $P_I(t; f)$ of the corresponding detector according to Eq. (3.15). MATLAB® software’s pseudo random number generator `randn` can generate very long sequences of random numbers, so we relied on that routine for simulating detector noise. For each of $T/\Delta t = 86\,400/192 = 450$ time segments, we generated a complex random sequence (that is, two real random sequences) of $f_u/\Delta f = 512/2 = 256$ real numbers. The total number of random numbers, $2(T/\Delta t)(f_u/\Delta f) = 225\,000$, is much less than the period of `randn`, which is $2^{1492} \geq 10^{449}$ [44].

Signal is also generated directly in the frequency domain. However, the GW strain in each detector, $\tilde{h}_I(t; f)$, is not generated independently; rather the product of the strains in the detectors, $\tilde{h}_1^*(t; f)\tilde{h}_2(t; f)$, is generated directly using the statistical properties of the strain correlation described in subsection III A, in particular, Eq. (3.10). We may write the product of the strains as a sum of its expectation value and statistical fluctuations:

$$\tilde{h}_1^*(t, f)\tilde{h}_2(t, f) = \langle \tilde{h}_1^*(t, f)\tilde{h}_2(t, f) \rangle + \text{fluctuations}. \quad (5.1)$$

Since our main aim is to generate

$$\tilde{s}_1^*(t, f)\tilde{s}_2(t, f) = [\tilde{h}_1^*(t, f) + \tilde{n}_1^*(t, f)][\tilde{h}_2(t, f) + \tilde{n}_2(t, f)] \quad (5.2)$$

and since statistically the variation in the signal terms are much weaker than the *zero mean uncorrelated* detector noise terms, we may simply drop the signal “fluctuations” term from Eq. (5.1)—that is, we may approximate the product of the detector outputs using the formula [45]:

$$\tilde{s}_1^*(t, f)\tilde{s}_2(t, f) = \langle \tilde{h}_1^*(t, f)\tilde{h}_2(t, f) \rangle + \tilde{n}_1^*(t, f)\tilde{n}_2(t, f). \quad (5.3)$$

For all the cases considered in this paper we have used flat source PSDs, i.e., $H(f) = \text{constant}$.

In this analysis we assume the sky to be a collection of *uncorrelated point sources* of different strengths placed at every pixel. Moreover, the numerical analysis has been restricted to the case of equal power in each polarization. So the (injected) true sky is constructed by putting

$$\mathcal{P}_{\text{true}}(\hat{\Omega}) = \sum_k \mathcal{P}_k \delta(\hat{\Omega} - \hat{\Omega}_k), \quad (5.4)$$

where \mathcal{P}_k is the strength of the point source placed at pixel k , located in the direction $\hat{\Omega}_k$ (in order to inject only one point source at pixel k_0 , we make all the $\mathcal{P}_k = 0$ except for $k = k_0$). In this setup, the expression for the overlap reduction function [Eq. (3.7)] for the true SGWB strain becomes

$$\gamma(t, f; \Delta\Omega, d\mathcal{P}_A) = \sum_k \Gamma(\hat{\Omega}_k, t) \mathcal{P}_k e^{2\pi i f \hat{\Omega}_k \cdot \Delta\mathbf{x}(t)/c}, \quad (5.5)$$

where we use our usual notation for $\Gamma(\hat{\Omega}, t)$ defined by Eq. (3.34).

We substitute the above in Eq. (3.10) and inject that simulated signal in noise using Eq. (5.3) to generate products of outputs from two detectors. In order to preserve the reality of time series data, the products of signals are generated only for positive frequencies and setting the negative frequencies equal to the complex conjugates of their positive frequency counterparts, that is, $\tilde{s}_I(t, -f) = \tilde{s}_I^*(t, f)$.

The radiometer analysis is then run on the simulated data by placing filters $\tilde{Q}(\hat{\Omega}_k, t, f; H)$ at each pixel k to generate the dirty maps.

B. Deconvolution: Clean maps

Any deconvolution routine requires the beam function at all the points of interest on the sky. For a GW radiometer, the filters (and hence the beam functions) are dependent on the data set itself. So, if the beam function is calculated for each sky pixel, apparently the computational cost should go up by a factor of the number of pixels ($N_{\text{pix}} \sim 3000$) times the cost to make one sky map of beam for a given pointing direction. However, we can use algebraic tricks to make this method computationally viable. A possible way to implement this method for the simple case of one baseline and equal power in each polarization is demonstrated below.

The beam and noise covariance matrices are given by

$$B_{ij} = 2\Delta f \left[\sum_{t=0}^T \frac{1}{\lambda(t, \hat{\Omega}_i)} \right]^{-1} \sum_{t=0}^T \Gamma(\hat{\Omega}_j, t) \Gamma(\hat{\Omega}_i, t) \times \Re \left[\sum_{f=0}^{f_u} e^{2\pi i f (\hat{\Omega}_j - \hat{\Omega}_i) \cdot \Delta\mathbf{x}(t)/c} G(t, f) \right], \quad (5.6)$$

$$N_{ij} = \frac{\Delta t}{4} \left[\sum_{t=0}^T \frac{1}{\lambda(t, \hat{\Omega}_j)} \right]^{-1} B_{ij}, \quad (5.7)$$

where

$$\Gamma(\hat{\Omega}, t) := F_1^+(\hat{\Omega}, t)F_2^+(\hat{\Omega}, t) + F_1^\times(\hat{\Omega}, t)F_2^\times(\hat{\Omega}, t), \quad (5.8)$$

$$G(t, f) := H^2(f)/[P_1(t, f)P_2(t, f)].$$

Now one can see that the beam matrix is a summation of parts which depend on either $\hat{\Omega}_i$ or $\hat{\Omega}_j$. So, it is possible to precompute the arrays $\lambda(\hat{\Omega}, t)$, $\Gamma(\hat{\Omega}, t)$, $G(t, f)$, and $\Delta \mathbf{x}(t)$, only once, and then use them efficiently to evaluate the whole \mathbf{B} matrix. Since each of these arrays are functions of *any two* of the three variables $\hat{\Omega}$, t , f , the memory size required for each two-dimensional array will not be too large. Once the arrays are computed, the next step is to obtain each element of the beam matrix separately. Each element requires the evaluation of two sum loops (like matrix multiplication) involving one exponential. The symmetry of the beam matrix (without the normalization constant) can be utilized here to reduce the computation cost by a factor close to 2. To summarize, a significant amount of CPU time required to make a sky map of the beam for one pointing direction could be utilized to make the beam maps for all other pointing directions. This is true for the noise covariance matrix as well, which, in this case, is proportional to the (unnormalized) beam matrix. A fast Fourier transform (FFT) with interpolation trick [46] and assumption of stationarity of noise for deconvolution (*not* for making the dirty map, where nonstationarity will be accounted for) can also be incorporated in the future to reduce the CPU time.

In this simple case, since the beam matrix is a square matrix, so that $\Sigma = (\mathbf{B}^T \mathbf{N}^{-1} \mathbf{B})^{-1} = \mathbf{B}^{-1} \mathbf{N} (\mathbf{B}^{-1})^T$, the estimated map given by Eq. (4.13) is just the least square solution:

$$\hat{\mathcal{P}} = \mathbf{B}^{-1} \cdot \mathbf{S}. \quad (5.9)$$

Even for the general cases of multiple baselines and polarized background, the estimation equation takes the above simple form [see Sec. IV].

The first task was to compute the beam matrix, which happens to be the computationally most intensive task. A typical beam matrix for the LIGO baseline using 192 HEALPix pixels is shown in Fig. 6. By construction, $B_{kk} = 1$ and $|B_{kk'}| < 1$ for $k \neq k'$, hence the matrix is diagonal dominated. The “stripes” in the matrix are related to the pixelization scheme. The beam is stronger if the pixels are closer to the pointing direction and it weakens as the distance between the pixels and pointing direction increases. In other words, the pixels closer to a point source will have stronger contamination from the point source. However, since we have used an isoLatitude pixelization scheme, the indices of two neighboring pixels at different latitudes differ by the total number of pixels on that latitude. This fact is reflected in the plot of the beam matrix—the matrix is sparse with certain “periodic” behavior which produces the stripes in the plot. The matrix becomes even more sparse for finer resolutions as a greater number of isoLatitude pixel rings pass through the core of the beam. Making a legible plot of the beam matrix for higher resolution is difficult, so the plot presented here is restricted to lower resolution—192 pixels instead of 3072.

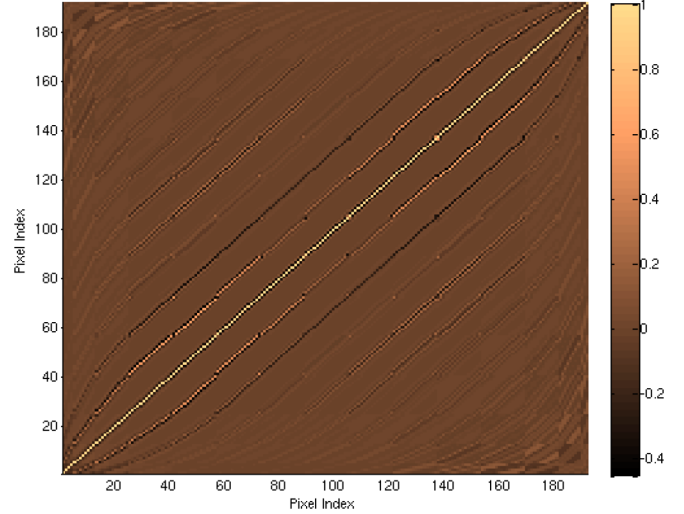


FIG. 6 (color online). A typical beam matrix for the LIGO Hanford-Livingston baseline at a low resolution (192 pixels) is shown in this figure. Each row of the matrix is the beam response function for the pointing direction that corresponds to the row index. The matrix is diagonal dominated as the radiometer receives maximum contribution from the pointing direction. The stripes are related to the isoLatitude pixelization scheme—the indices of the neighboring pixels at different latitudes differ by the total number of pixels on that latitude. The possibility of making the beam matrix more diagonal dominated using a nested pixelization scheme, where the indices of the neighboring pixels are close, is being explored.

Since the sparseness of the beam matrix depends on the pixelization scheme, it may be possible to make the beam matrix significantly diagonal by using a nested pixelization scheme, where the indices of the neighboring pixels are close. This possibility is being explored.

Sparse matrices are computationally easier to invert; however, the stability of inversion of such matrices is a numerical challenge. Therefore, as mentioned in Sec. IV, instead of evaluating $\hat{\mathcal{P}} = \mathbf{B}^{-1} \cdot \mathbf{S}$ [Eq. (5.9)], we choose to algebraically solve for $\hat{\mathcal{P}}$ from the set of linear equations

$$\mathbf{B} \cdot \hat{\mathcal{P}} = \mathbf{S}, \quad (5.10)$$

with the same number of unknowns as the number of equations (i.e., the system is *not* under or overconstrained). The above equation can be cast into a linear system with symmetric kernel,

$$\mathfrak{b} \cdot \hat{\mathcal{P}} = \mathfrak{s}, \quad (5.11)$$

by introducing two new quantities,

$$\mathfrak{b}_{ij} := \left[\sum_{t=0}^T \frac{1}{\lambda(t, \hat{\Omega}_i)} \right] B_{ij}; \quad \mathfrak{s}_i := \left[\sum_{t=0}^T \frac{1}{\lambda(t, \hat{\Omega}_i)} \right] S_i. \quad (5.12)$$

Comparing with Eq. (5.6) one can clearly see that \mathfrak{b} is a

symmetric matrix. A symmetric kernel is always preferred by most of the algorithms for solving numerical linear equations. Moreover, it is possible to show from the algebra presented in subsection IV B that, to incorporate observations from multiple baselines, one can simply replace \mathfrak{z} and \mathfrak{b} by the sum of those respective quantities over all the baselines and solve Eq. (5.11) for $\hat{\mathcal{P}}$ (which is, of course, independent of any baseline). Thus the extension of this analysis to incorporate a network of detectors becomes straightforward with these new quantities.

Following the discussion presented in Sec. IV, we use a conjugate gradient (CG) iterative technique to solve the above set of linear equations. CG algorithm solves a set of linear equations $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$, where \mathbf{A} is a square matrix and \mathbf{x} , \mathbf{b} are vectors, by minimizing the quadratic form $\frac{1}{2} \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} - \mathbf{b} \cdot \mathbf{x}$. We use the *minimum residual* method, which efficiently utilizes the fact that \mathfrak{b} is symmetric and does not require \mathfrak{b} to be positive definite. The minimum residual method aims to minimize the residual $|\mathbf{A} \cdot \mathbf{x} - \mathbf{b}|^2$ itself, instead of the quadratic form $\frac{1}{2} \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} - \mathbf{b} \cdot \mathbf{x}$. Further details on CG methods can be found in standard literature, e.g., [47].

The clean maps also contain pixel noise—partly due to the random noise present in the data and partly due to the numerical errors introduced at each stage of the pipeline, mainly during the process of deconvolution. There are pixels in the deconvolved map, which have negative values, even though the injected map is positive. To reduce the noise in the clean maps, we introduce an additional step: We compute the root-mean-square (RMS) noise “ σ ” in a map when there is no injected source. Then in the clean map (with source) we *mask*, that is, set to zero, all the pixels that have values less than a threshold of few σ . The number of iterations for deconvolution and the threshold for masking can be adjusted according to the tolerable levels of false alarm and false dismissal probabilities.

The above can be easily extended to handle real data where we have no control on the injections. One can calculate the equivalent of RMS noise for no injection by shifting the data streams from different detectors by a large time lag (much smaller than the segment duration), say, 1 sec, that corresponds to distances much greater than the earthly distances, so that, true GW signals are not added coherently. To be more careful, one can perform this exercise for a few large time shifts and confirm that the noise levels are not significantly different for different shifts.

It is, however, not so straightforward to measure the quality of deconvolution. The SNR of individual pixels do not carry enough significance, as the neighboring pixels are highly correlated. It is also difficult to define a quantity that can take into account the pixel-to-pixel noise covariance due to the difficulty in inverting the beam matrix. In this paper, we use a rather simplistic measure to quantify the quality of deconvolution, which is often used in image processing to measure the reconstruction error. We use a

quantity known as the “normalized mean square error” (NMSE) [48], expressed in terms of the injected \mathcal{P} and the estimated $\hat{\mathcal{P}}$ maps as

$$\text{NMSE} := \frac{|\hat{\mathcal{P}} - \mathcal{P}|^2}{|\mathcal{P}|^2}. \quad (5.13)$$

Obviously, the lower the NMSE, the better the reconstruction.

The whole analysis was tested for different kinds of injected maps consisting of localized sources and diffuse sources. In all these cases each pixel k of the injected map was assigned a value \mathcal{P}_k between 0 and 1 with a source PSD $H(f) = 5 \times 10^{-47}/\text{Hz}$. This means that, if a pixel of a test map has strength 1, the standard deviation of the Fourier transform of stochastic GW coming from that pixel is $\sqrt{H(f)} \sim 7 \times 10^{-24}/\sqrt{\text{Hz}}$. This standard deviation is about one-third of the standard deviation of Fourier transform of noise at the most sensitive frequency band of the LIGO-I detectors which is about $2 - 3 \times 10^{-23}/\sqrt{\text{Hz}}$. To our knowledge, the strength of anisotropic astrophysical GW background has not so far been predicted theoretically. However, if we try to extend the results from the all-sky averaged (isotropic) astrophysical background [49] to have a crude estimate of the strength of the anisotropic background, it turns out that, the PSD of the anisotropic astrophysical background in the Universe is weaker than the $H(f)$ we have injected by roughly a few orders of magnitude. In the present work we have used an observation time of *one day* to demonstrate the method. With longer observation times and employing several baselines comprising of the upcoming advanced (more sensitive) detectors, the difference between the expected background and the detectable background would diminish or altogether disappear.

We first inject a 4-pixel wide localized source near the Virgo cluster, a potential point source of SGWB, as illustrated in Fig. 7. Figure 7(a) shows the injected map. Figure 7(b) shows the dirty map made from simulated data. Figure 7(c) shows the clean map, obtained by deconvolving the above dirty map with the beam using 15 conjugate gradient iterations and Fig. 7(d) shows the same clean map masked using a 5σ threshold. It is evident that deconvolution has successfully localized the source in a relatively smaller area as compared to the dirty map. Still, one should note that, the deconvolution routines do not perform well when the injected source is like a delta function. This causes high NMSE, 1.22 for the unmasked and 0.64 for the masked clean maps, and significant loss of the peak strength of the reconstructed point source, as indicated in Fig. 7. Moreover, increasing the number of iterations beyond a certain level actually deteriorates the quality of deconvolution due to noise amplification, and this level is dependent on the kind of source one is searching for. In this basic analysis we have used 15 iterations to search for localized sources and 40 iterations to search for

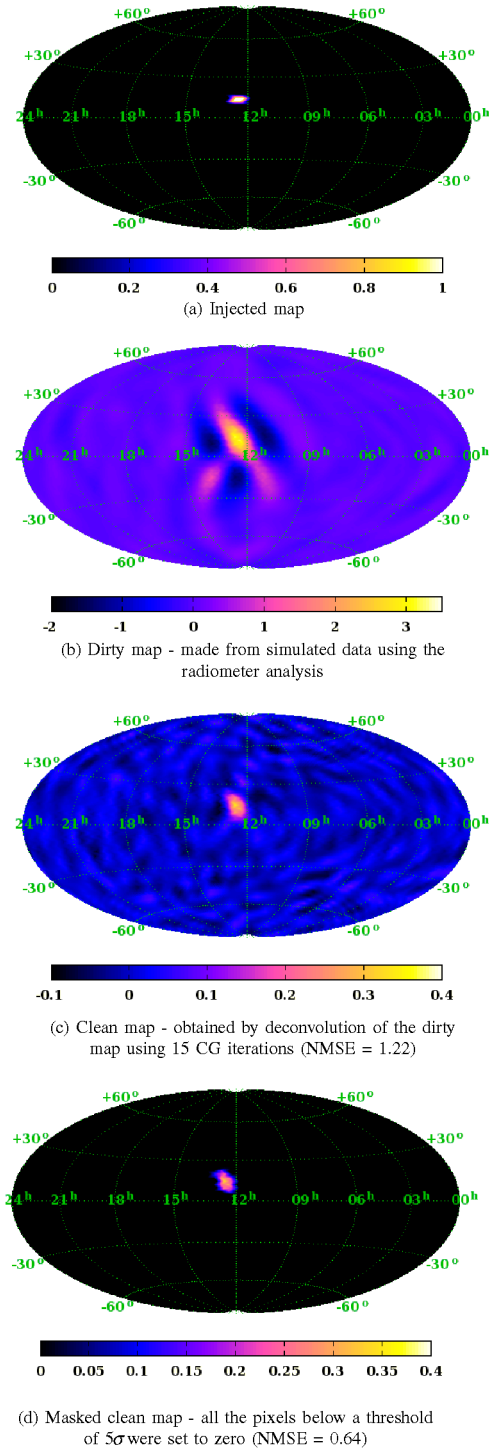


FIG. 7 (color online). Illustration of deconvolution for localized sources: A 4-pixel wide localized source was injected near the location of the Virgo cluster—a potential source of SGWB.

broad sources, which offer reasonably clean deconvolution and comparatively low NMSE. The introduction of a minimum error criterion [50] to terminate the iteration process is being considered. Several other deconvolution algorithms are being explored in order to identify the one which is *best* suited for the GW radiometer analysis.

Next, we inject two kinds of diffuse sources, viz., one that is nearly equatorial and another that is distributed across multiple declinations, as illustrated in Fig. 8. We injected modified (using FTOOLS [51]) galactic foreground seen in CMB temperature anisotropy measurements as our test patterns for the diffuse SGWB sources. The left panels of Fig. 8 correspond to a modified form of the temperature anisotropy map measured by the WMAP satellite [52]. We emphasize that the sky looks different in barycentric coordinates, similar to what is shown in the top-right panel of Fig. 8. We omit the coordinate transformation step intentionally in order to get a diffuse equatorial source. The right panels of Fig. 8 correspond to a modified form of the temperature anisotropy map in the barycentric coordinates generated by the Planck simulator [53]. One of the main modifications applied to both of these maps was to mask the brightest part of the Galaxy. This step reveals more structures in the maps, which is useful for testing a deconvolution algorithm. Figure 8(a) shows the injected toy maps. Figure 8(b) shows the dirty maps obtained by the radiometer analysis. One can see that the dirty maps have lost all the fine structures present in the injected maps. Furthermore, they show certain features that were not even present in the injected map. Also, the pixel values in the dirty maps are spread over a range consisting of large positive and almost equally negative values. Figure 8(c) shows the clean maps recovered by 40 CG iterations. Clearly, many of the features of the injected maps have been recovered in the clean maps, which is also evident from the lower values of NMSE, 0.33 and 0.22, respectively. Also, the positivity of the estimated map has been vastly improved—the pixel values of the clean maps lie mostly on the positive side, as one should expect. Finally, Fig. 8(d) shows the masked clean maps obtained by using a 4σ threshold. Though the masked maps give better visual impression, masking can actually discard several pixels which have weak sources, thereby increasing the NMSE. In Fig. 8(d), for example, masking increases NMSE to 0.36 and 0.33, respectively, though the masked maps look more similar to the injected maps shown in Fig. 8(a) than the unmasked clean maps in Fig. 8(c).

VI. CONCLUSION

The stochastic astrophysical GW background is likely to be dominated by sources in the nearby anisotropic universe, so the detection of localized sources is more favorable than the all-sky-averaged search. Making a skymap of the SGWB sky has been a long standing ambition of stochastic GW research. Different analysis methods have been proposed to create sky maps by measuring the first few spherical harmonic multipoles of the sky. Here we have presented a direct approach of directed GW radiometer analysis. In this approach, the whole sky is decomposed in a discrete set of pixels and the contribution from each pixel is measured separately by correlating phase shifted

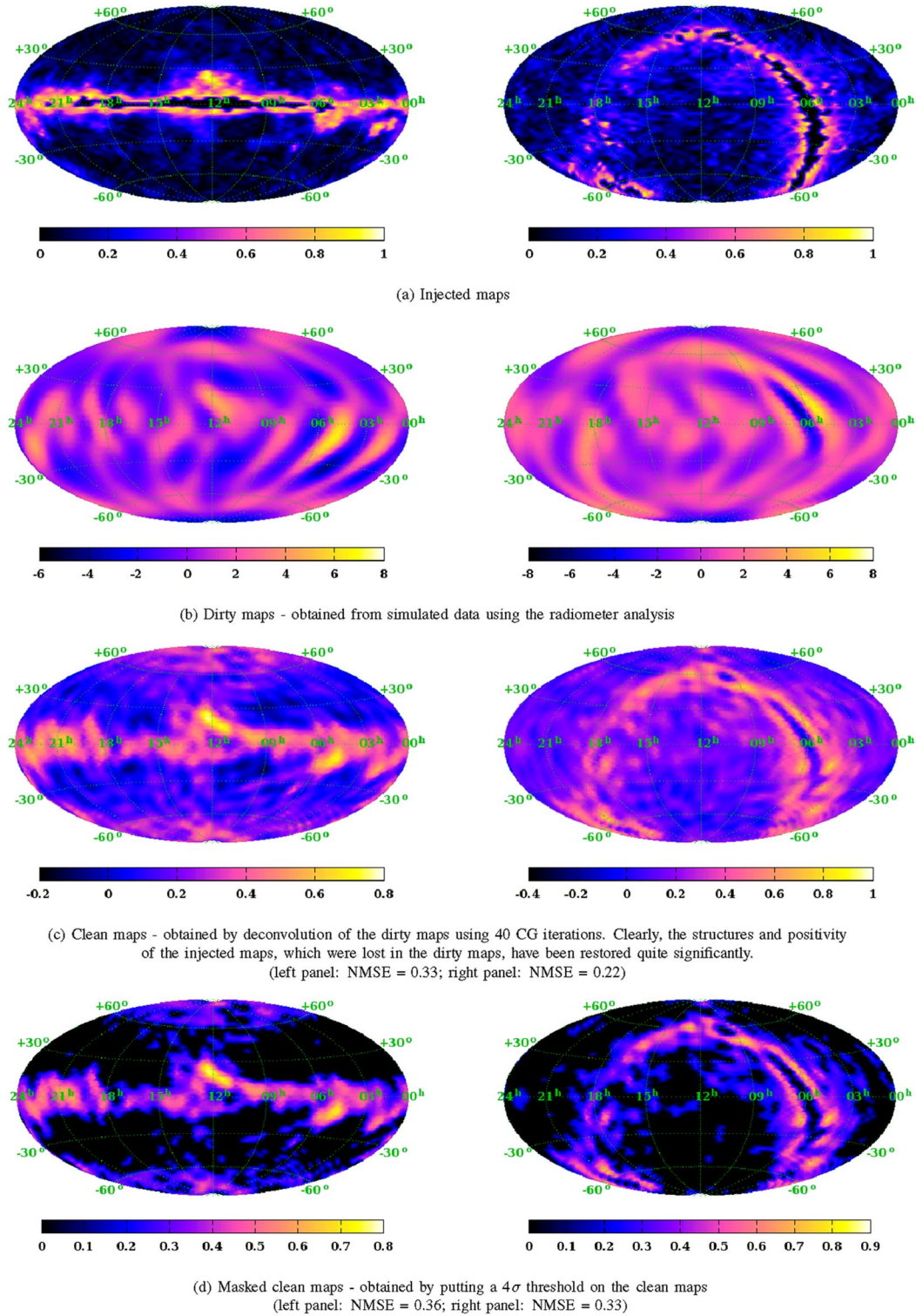


FIG. 8 (color online). Illustration of deconvolution for broad sources: Maps similar to CMB temperature anisotropy sky with the galactic foreground were injected as toy maps. The left panels correspond to a map measured by the WMAP satellite [52], as seen in the galactic coordinates, as a toy model for an equatorial source and the right panels correspond to a map generated by the Planck simulator [53], as a toy model for a multideclination extended source.

detector outputs to generate the whole sky map, which is a clear application of earth rotation aperture synthesis. Specifically, for the AGWB detection statistic, we have defined a correlation statistic with a directed optimal filter that targets a fixed point in the sky by adjusting the time delay across a baseline to track its rotation with the Earth. This statistic, however, provides us with a dirty map of the sky which we numerically deconvolve to obtain the true sky map. For this purpose, we have employed the conjugate gradient method. We numerically implement the deconvolution on simulated unpolarized GW sky maps obtained with the LIGO detector baseline. The success of this method is demonstrated by the recovery of simulated source distributions, namely, (i) of a point source, (ii) of a diffuse source in the equatorial plane, and (iii) of a diffuse source off the equatorial plane.

This work needs to be implemented on other baselines of the upcoming/future network of detectors such as LIGO-VIRGO, LIGO-LCGT, LISA, etc. The outline of the analysis has been presented here. However, a further detailed analysis and the implementation is a future goal. Even for the single baseline, SPA analysis shows that, perhaps a more efficient method of deconvolution, yielding better accuracy and convergence, lower computational costs, and convenience of application may be possible using a more sophisticated analysis, for example, one involving basis functions. The maximum likelihood framework presented here is, in fact, independent of any particular choice of basis. So once a suitable basis is chosen, the rest of the analysis can be applied without requiring any major change. It may also be possible to deconvolve only a patch of the sky using a similar method [54].

The work presented in this paper should also benefit two other searches. First, since the long-duration integration of the data will essentially comprise a sum over short stretches, a large signal in a short stretch will constitute a candidate for a transient or burst (short-duration) event. Unlike the coincidence search being currently conducted for such events, our work will combine coherently the outputs of several detectors and, thus, improve their detectability. Second, the long-duration integration of the data should be able to find gravitational wave signals

from modeled sources, such as pulsars. Although our method is optimal for searching unmodelled sources, it is not so for pulsars, the signals from which can be matched filtered. The problem with the latter method is that owing to the very large parameter space volume, an all-sky, all-frequency search for pulsars with matched filtering is not computationally viable. Our proposed method is not handicapped by this problem since it does not use the intrinsic source parameters for the search; rather it uses the data from one detector to “filter” that from others in the network, after appropriately time shifting them and weighting them with the respective antenna patterns. Thus, our method can be used as the first step in a two-step hierarchical search for pulsars, where triggers from our method are followed up with matched filtering.

ACKNOWLEDGMENTS

S. Mitra would like to acknowledge the Council of Scientific and Industrial Research (India) and Centre National d’Etudes Spatiales (France) for supporting his research and Caltech for supporting his visit to LIGO Laboratory, Caltech in 2006, where part of this work was done. He further thanks Stuart Anderson for helpful suggestions and Kent Blackburn and Patrick Sutton for providing useful help with the software computing facilities. Some of the results in this paper have been derived using the HEALPix [55] package, the Planck Simulator [53], and FTOOLS [56] and some of the plots were made using a colormap, specifically designed for compatibility with grayscale printing, included in the GNUPLOT package [57]. The LIGO Data Analysis System (LDAS) at Caltech and the High Performance Computing (HPC) facility at IUCAA were used for the numerical implementation. This work was supported in part by the Department of Science and Technology Grant No. DST/INT/(US-RP077)/2001 and the National Science Foundation Grants No. INT-01-38459, No. PHY-0630121, and No. PHY-0239735, and the NSF LIGO Laboratory Cooperative Agreement No. PHY-0107417. This paper has LIGO Document No. LIGO-P070033-Z.

-
- [1] R. A. Hulse and J. H. Taylor, *Astrophys. J.* **195**, L51 (1975); J. H. Taylor, L. A. Fowler, and J. M. Weisberg, *Nature (London)* **277**, 437 (1979); J. H. Taylor, *Rev. Mod. Phys.* **66**, 711 (1994).
- [2] A. Abramovici *et al.*, *Science* **256**, 325 (1992); C. Bradaschia *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **289**, 518 (1990); K. Danzmann, in *Gravitational Wave Experiments*, edited by E. Coccia, G. Pizzella, and F. Ronga (World Scientific, Singapore, 1995), p. 100;

- K. Tsubono, in *Gravitational Wave Experiments*, edited by E. Coccia, G. Pizzella, and F. Ronga (World Scientific, Singapore, 1995), p. 112; R. J. Sandeman *et al.*, A.I.G.O. Prospectus (unpublished).
- [3] P. L. Bender *et al.*, Report No. MPQ 233, 1995).
- [4] K. S. Thorne, in *300 Years of Gravitational*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987), p. 330.
- [5] K. S. Thorne, in *Proceedings of the Snowmass 95 Summer*

- Study on Particle and Nuclear Astrophysics and Cosmology*, edited by E. W. Kolb and R. Peccei (World Scientific, Singapore, 1995), p. 398.
- [6] E. Flanagan, in *Gravitation and Relativity: At the Turn of the Millennium*, edited by N. Dadhich and J. Narlikar (IUCAA, Pune, 1998), p. 177.
- [7] B.F. Schutz, *Classical Quantum Gravity* **16**, A131 (1999).
- [8] K. S. Thorne, arXiv:gr-qc/9704042.
- [9] V. Kalogera, R. Narayan, D. N. Spergel, and J. H. Taylor, *Astrophys. J.* **556**, 340 (2001).
- [10] B.F. Schutz, *Classical Quantum Gravity* **6**, 1761 (1989).
- [11] A. Lazzarini, R. Weiss, LIGO Technical Document No. LIGO-T0401-40-00-Z, 2004.
- [12] B. Abbott *et al.* (LIGO Scientific Collaboration), *Phys. Rev. D* **76**, 082003 (2007).
- [13] M. Maggiore, *Phys. Rep.* **331**, 283 (2000). Also see the references therein.
- [14] Plots and data files of the initial LIGO noise curves can be found at http://www.ligo.caltech.edu/~jzweizig/distribution/LSC_Data/.
- [15] P.F. Michelson, *Mon. Not. R. Astron. Soc.* **227**, 933 (1987); N. Christensen, Ph.D. thesis, MIT, 1990; *Phys. Rev. D* **46**, 5250 (1992).
- [16] E. Flanagan, *Phys. Rev. D* **48**, 2389 (1993).
- [17] B. Allen and J. Romano, *Phys. Rev. D* **59**, 102001 (1999).
- [18] B. Allen and A. Ottewill, *Phys. Rev. D* **56**, 545 (1997).
- [19] N.J. Cornish, *Classical Quantum Gravity* **18**, 4277 (2001).
- [20] H. Kudoh and A. Taruya, *Phys. Rev. D* **71**, 024025 (2005); A. Taruya and H. Kudoh, *Phys. Rev. D* **72**, 104015 (2005); A. Taruya, *Phys. Rev. D* **74**, 104022 (2006).
- [21] It is important to note that, even if the polarizations are independent in a certain direction, there can be mixing between polarizations in other directions due to the rotation of basis vectors. This complication has not been considered here, although it may not be too difficult to incorporate this in our analysis.
- [22] G.B. Rybicki and A.P. Lightman, *Radiative Processes in Astrophysics* (John Wiley & Sons, New York, 1979).
- [23] S. Mitra, Ph.D. thesis, IUCAA, Pune, India [LIGO technical document No. LIGO-P070034-00-Z, 2006].
- [24] S. Bose *et al.*, *Classical Quantum Gravity* **20**, S677 (2003).
- [25] G.H. Golub and C.F. Van Loan, *Matrix Computations* (John Hopkins, Baltimore, 1996), 3rd ed.
- [26] As mentioned in the text, we assume here that we know the correct source PSD $H(f)$, which is of course *not* known prior to a detection. However, as long as it is known that $H(f)$ follows a power law with a finite set of possible spectral indices, one can estimate/constrain SGWB anisotropy for each possible spectral index.
- [27] W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1965), Vol. II.
- [28] T. Poutanen *et al.*, *Astron. Astrophys.* **449**, 1311 (2006).
- [29] J. Borrill, in Proceedings of the 5th European SGI/Cray MPP Workshop, Bologna, Italy, arXiv:astro-ph/9911389 (1999)
- [30] Microwave Anisotropy Dataset Computational Analysis Package (MADCAP), <http://crd.lbl.gov/borrill/cmb/madcap/>.
- [31] D.F. Morrison, *Multivariate Statistical Methods* (McGraw-Hill, New York, 1990), 3rd ed.
- [32] J. Reid, *Large Sparse Sets of Linear Equations*, edited by J. Reid (Academic, New York, 1971).
- [33] C. Armitage and B. Wandelt, *Phys. Rev. D* **70**, 123007 (2004).
- [34] E.L. Wright, in Proceedings of the IAS CMB Data Analysis Workshop, Princeton, 1996 (unpublished).
- [35] A. Challinor *et al.*, *Mon. Not. R. Astron. Soc.* **331**, 994 (2002).
- [36] N. Jarosik *et al.*, *Astrophys. J. Suppl. Ser.* **170**, 263 (2007).
- [37] Of course, one could combine the maps from different baselines with suitable pixel dependent weight factors \mathbf{w}_i ,
- $$\mathbf{S} = \sum_i \mathbf{w}_i \cdot \mathbf{S}^{(i)},$$
- which may also reduce the noise by a factor of $\sim\sqrt{N_b}$. But, it would have issues of combining data from baselines with different beam functions. Further analysis might be required to assess which method would be more advantageous.
- [38] O. Malaspina and R. Sturani, *Classical Quantum Gravity* **23**, 319 (2006).
- [39] Notably, aperture synthesis technique using earth rotation was first introduced by Martin Ryle for radio observation of cosmic sources.
- [40] A.R. Thompson, J.M. Moran, and G.W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy* (Wiley, New York, 2001), 2nd ed.
- [41] MATLAB® Matrix Laboratory, <http://www.mathworks.com/>.
- [42] For convenience we have used 86 400 sec for one sidereal day, instead of 86 164 sec. It hardly affects the accuracy of the results presented in this paper, as the same duration has been used for both injecting signals and analyzing data. Even for analyzing real data *the same segments can be used*, however, Earth's rotation frequency should be accurately supplied in order to establish correct correspondence between time and celestial coordinates.
- [43] <http://healpix.jpl.nasa.gov/>.
- [44] MATLAB® function reference: rand, <http://www.mathworks.com/access/helpdesk/help/techdoc/ref/rand.html>.
- [45] Note that, it is also possible to generate the correlated detector strains $h_I(t)$ independently and construct $\tilde{s}_I(t; f) = \tilde{h}_I(t; f) + \tilde{n}_I(t; f)$ for each detector I separately using simulated colored noise $\tilde{n}_I(t; f)$ to test the analysis. However, we chose the above method in order to reduce complications in the primary testing of the analysis presented in this paper.
- [46] S.W. Ballmer, *Classical Quantum Gravity* **23**, S179 (2006).
- [47] W. Press *et al.*, *Numerical Recipes: the Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1992).
- [48] J.R. Fienup, *Appl. Opt.* **36**, 8352 (1997).
- [49] D. Coward and T. Regimbau, *New Astron. Rev.* **50**, 461 (2006).
- [50] H. Tang, L.W. Cahill, in TENCON '92. Technology Enabling Tomorrow: Computers, Communications and

- Automation towards the 21st Century. 1992 IEEE Region 10 International Conference, 1992, Vol. 2, p. 578.
- [51] <http://heasarc.gsfc.nasa.gov/ftools/>.
- [52] G. Hinshaw *et al.*, *Astrophys. J. Suppl. Ser.* **170**, 288 (2007).
- [53] Planck Simulator, <http://www.g-vo.org/planck/>.
- [54] C. Burigana and D. Saez, *Astron. Astrophys.* **409**, 423 (2003).
- [55] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, *Astrophys. J.* **622**, 759 (2005).
- [56] K. J. Blackburn, in *Astronomical Data Analysis Software and Systems IV*, edited by R. A. Shaw, H. E. Payne, and J. J. E. Hayes, ASP Conference Series, Vol. 77 (ASP, San Francisco, 1995), p. 367.
- [57] T. Williams *et al.*, GNUPLOT, <http://gnuplot.info>.