PHYSICAL REVIEW D **81,** 103009 (2010)

Possible use of self-calibration to reduce systematic uncertainties in determining distance-redshift relation via gravitational radiation from merging binaries

Tarun Deep Saini*

Indian Institute of Science, Bangalore, 560 012, India

Shiv K. Sethi†

McWilliams Center for Cosmology Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh, Pennsylvania 15213, USA‡

Varun Sahni§

Inter-University Centre for Astronomy and Astrophysics (IUCAA), Ganeshkhind, Pune 411 007, India

(Received 2 November 2009; published 28 May 2010)

By observing mergers of compact objects, future gravity wave experiments would measure the luminosity distance to a large number of sources to a high precision but not their redshifts. Given the directional sensitivity of an experiment, a fraction of such sources (gold plated) can be identified optically as single objects in the direction of the source. We show that if an approximate distance-redshift relation is known then it is possible to statistically resolve those sources that have multiple galaxies in the beam. We study the feasibility of using gold plated sources to iteratively resolve the unresolved sources, obtain the self-calibrated best possible distance-redshift relation and provide an analytical expression for the accuracy achievable. We derive the lower limit on the total number of sources that is needed to achieve this accuracy through self-calibration. We show that this limit depends exponentially on the beam width and give estimates for various experimental parameters representative of future gravitational wave experiments DECIGO and BBO.

DOI: 10.1103/PhysRevD.81.103009 PACS numbers: 98.62.Py, 04.30.Tv, 95.85.Sz, 97.80.−d

Establishing the nature of dark energy is a paramount objective of modern cosmology. A precise knowledge of cosmic distance to sources at moderate redshifts ($z \lesssim$ few) is essential for success in this endeavor [1]. It has been suggested that gravitational radiation from merging binaries [neutron star (NS)-NS, NS-black hole (BH), and BH-BH] could be a "standard siren" and a complementary means (to standard candles and rulers) for probing cosmic expansion [2–6]. Indeed, a knowledge of the underlying physics of gravitational radiation from binaries could help establish the luminosity distance to a NS-NS binary to a precision of 2% (ignoring, for the time being, the redshift-dependent error from gravitational lensing). In order to serve as a cosmological probe however, the luminosity distance should be known *as a function of redshift*. Therefore, unlike other probes of distance, the main systematic uncertainty in this case is the identification (and redshift determination) of galaxies hosting the binaries (see e.g. [4–6]).

The space-borne gravitational wave observatory LISA [7] is expected to achieve an angular resolution of about $1'$ (for a detailed discussion see [4]). The volume bounded by this angle (inclusive of the redshift uncertainty from luminosity distance errors) is expected to contain roughly 30 objects at $z \simeq 1$. The directional sensitivity of next generation gravitational wave (GW) observatories such as DECIGO [3], the Big Bang Observer (BBO) [8], and ASTROD [9] is likely to be even better ($\sim$ few arc seconds), in which case, in only a small fraction of cases would more than a single galaxy lie within the observational beam [10].

The main source of uncertainty in using standard sirens to probe cosmic expansion comes from misidentifying galaxies hosting the standard sirens. Clearly, the larger the number of galaxies within the observational beam, the greater the chances for this to happen. Given the enormous potential of using gravity wave standard sirens to determine the nature of dark energy, it would clearly be desirable to minimize this source of systematics. One possible approach rests in determining an association between merging binaries and the gas in the surrounding medium. Unique signatures of an "afterglow" from such an event in the electromagnetic spectrum could help in identifying the source galaxy (e.g. [11]). Statistical determination of redshift using clustering properties of galaxies constitute another possible approach [12].

In this paper, we present a statistical iterative method to isolate the source of the GW signal. This formulation assumes no prior knowledge of the relation between the luminosity distance and redshift (DZ relation). The method

*tarun@physics.iisc.ernet.in
†sethi@rri.res.in
‡On leave from Raman Research Institute, C. V. Raman Avenue, Sadashivanagar, Bangalore 560 080, India.
§varun@iucaa.ernet.in

we propose can iteratively improve upon the errors on DZ relation. Our method is briefly summarized as follows: Beams containing a single galaxy with redshift consistent with the expected distance-redshift relation would accurately and reliably portray the DZ relationship. We shall call such sources "gold plated" (GP) following [10]. The expected number of such sources depends crucially on directional sensitivity. If more than one galaxy falls within the beam we propose to rule out nonsources through an iteratively improved DZ relation. In our proposal we do not use other means of constraining the DZ relation such as supernova type Ia, since supernova type Ia systematics is likely to be far more complex than previously thought [13]. Instead, we shall use GP sources to provide an independent approximate starting DZ relation, which we iterate upon. To what precision this can be achieved depends upon the details of cosmology as well the experimental parameters. In this paper we investigate the efficacy as well as limitations of this iterative self-calibration.

Redshift Error-box—The directional accuracy of a GW signal is determined by the experimental beam width $\Delta\Omega$, giving rise to the possibility that several sources might lie within the beam. To single out an object as the *unique* source of the GW signal we either require a smoking gun signal or a criterion by which other objects in the beam can be ruled out as possible sources. In the presence of statistical uncertainty in DZ it is impossible to establish precisely the redshift z for a source of gravitational radiation. One must settle instead for an uncertainty $z \pm \Delta z_a$ where

$$\Delta z_a \simeq \frac{\sigma_{\rm m}(z)\Phi(z)}{D_L} = \eta_{\rm m}\Phi(z), \qquad (1)$$

$\Phi(z) = (d\log D_L/dz)^{-1}$, and $\sigma_{\rm m}(z)$ is the redshift-dependent standard deviation in the luminosity distance to a single source. The dimensionless standard deviation $\eta_{\rm m}(z) = \sigma_{\rm m}(z)/D_L$ is partly due to instrumental noise and partly due to weak lensing. The dominant uncertainty is due to lensing and although lensing produces an asymmetric distribution of magnifications, for our purposes we shall assume that the distribution is described by a Gaussian with a dimensionless standard deviation $\eta_{\rm wl}(z) = 0.042z$, derived from the results of [14]. In this paper, we also add in quadrature a fixed value of instrumental noise $\eta_{\rm inst} = 0.02$ to the lensing standard deviation to obtain $\eta_{\rm m}^2 = \eta_{\rm wl}^2 + \eta_{\rm inst}^2$. Δz_a also contains an important *cosmology-dependent* contribution, so that the redshift error box is finally given by (see Appendix)

$$\Delta z_a \simeq \frac{\sqrt{\sigma_{\rm m}^2(z) + \sigma_{\rm c}^2(z)}}{D_L}\Phi(z) = \eta_{\rm m}(z)\Phi(z)\left[1 + \frac{\eta_{\rm c}^2(z)}{\eta_{\rm m}^2(z)}\right]^{1/2}, \qquad (2)$$

where $\sigma_{\rm c}(z)$ is the standard deviation in DZ reflecting uncertainty in our knowledge of the expansion history and $\eta_{\rm c} = \sigma_{\rm c}/D_L$.

Though multiple objects might lie within the beam, it may still be possible, with a small enough value of Δz_a, to single out a source purely on statistical grounds. However, since measurement errors on a single source are fixed, the only way to lower the redshift error box given by Eq. (2), is by reducing the second term in that equation.

Occupation Number—The redshift range Δz_a together with the beam width, $\Delta\Omega$, determine the expected number of galaxies lying within the beam that are statistically consistent both with the approximate DZ relation as well as the measurement uncertainty. In order to calculate the occupation number $\bar{n}$, defined as the mean number of galaxies that satisfy this criterion, we have adopted the number density of sources from [10,15,16]. The mean number of galaxies in the redshift range $2\Delta z_a$ turns out to be

$$\bar{n}(z) \simeq \frac{8N_\Omega}{h(z)\sqrt{\pi}} r(z)\exp[-r^4(z)]\Delta\Omega\Delta z_a, \qquad (3)$$

where we have assumed a small Δz_a so the linear approximation suffices. Here $r(z) = \int_0^z dz/h(z)$ is the c/H_0 normalized coordinate distance, $h(z) = H(z)/H_0$ and $N_\Omega = 1000\ {\rm arcmin}^{-2}$ is the projected number density of galaxies consistent with the Hubble ultra deep field [17]. Substituting Δz_a from Eq. (2) we obtain

$$\bar{n} = \nu(z)\left[1 + \frac{\eta_{\rm c}^2}{\eta_{\rm m}^2}\right]^{1/2}, \qquad (4)$$

where we have defined the *minimum occupation number* $\nu(z)$ as the value of $\bar{n}$ when $\eta_{\rm c} = 0$, namely,

$$\nu(z) = \frac{8\Phi(z)N_\Omega}{h(z)\sqrt{\pi}} r(z)\exp[-r^4(z)]\eta_{\rm m}\Delta\Omega. \qquad (5)$$

We shall assume that galaxies falling within the beam are distributed uniformly randomly, however, at the end of this paper we briefly discuss how the clustering of galaxies affects our analysis. Knowing the occupation number $\bar{n}$, the probability that there are k galaxies, apart from the source galaxy, within the beam is given by ${\rm Pr}(k) = \bar{n}^k\exp(-\bar{n})/k!$. If there is only a single object in the redshift error box then we shall assume that it is the source of the signal. The probability for such instances is given by ${\rm Pr}(0) = \exp(-\bar{n})$. Clearly the limiting fraction of sources that cannot be resolved statistically is $1 - \exp(-\nu(z))$, which for $\nu(z) \ll 1$ is approximately given by $\nu(z)$.

Method—A GP source would measure the DZ relation with an accuracy $\eta_{\rm m}$ at a redshift z. Let us consider a redshift bin, $\Delta z_{\rm bin}$, centered at the redshift z. The total number of sources in this bin is given by $\Delta N(z) = N_0 f(z)\Delta z_{\rm bin}$, where N_0 is the total number of GW sources at all redshifts and $f(z)\Delta z_{\rm bin}$ is the fraction occurring in the bin $\Delta z_{\rm bin}$. The value of N_0 (NS-NS binaries) for GW space missions is expected to range from $N_0 \sim 10^3$ (LISA) to

$N_0 \sim 10^6$ (BBO). Let us assume that there are $\Delta N_{\mathrm{GP}}(z)$ gold plated sources in the bin. These sources furnish a first estimate of the DZ relation. Since each source has a measurement error given by η_{m} then clearly the zeroth error on cosmology is given by $\eta_{\mathrm{c}0} = \eta_{\mathrm{m}}/\sqrt{\Delta N_{\mathrm{GP}}(z)}$.

We note that if there happen to be no GP sources in the redshift bin Δz_{bin} then $\eta_{\mathrm{c}0}$ can be obtained by fitting a dark energy model to the GP sources at other redshifts. However, with no *a priori* reason to assume a given behavior for dark energy, we advocate this self-consistently obtained DZ relation where each redshift is dealt with independently.

Since we now have the zeroth order information about cosmology we can use $\eta_{\mathrm{c}0}$ to calculate the occupation number $\bar{n}_0 = \nu(z)\sqrt{1 + \eta_{\mathrm{c}0}^2/\eta_{\mathrm{m}}^2}$. The zeroth knowledge of the DZ relationship resolves some sources to give the new value of resolved sources (GP sample and statistically resolved sources) as $\Delta N^{(1)}_{\mathrm{resolved}} = \Pr(0)\Delta N(z) = \Delta N(z) \times \exp(-\bar{n}_0)$, and thus provides us with a first improved estimate $\eta_{\mathrm{c}1} = \eta_{\mathrm{m}}/\sqrt{\Delta N(z)\exp(-\bar{n}_0)}$. With this refinement in η_{c} we can recalculate the occupation number at the first iteration as

$$\bar{n}_1 = \nu(z)\left[1 + \frac{1}{\Delta N(z)\exp(-\bar{n}_0)}\right]^{1/2}. \tag{6}$$

It is clear that iterating further we shall obtain the recurrence relation

$$\bar{n}_{i+1} = \nu(z)\left[1 + \frac{1}{\Delta N(z)\exp(-\bar{n}_i)}\right]^{1/2}. \tag{7}$$

The iteration terminates when $\bar{n}_{i+1} = \bar{n}_i$, and therefore the saturation cosmological uncertainty, which is the second term inside parenthesis in the previous expression, is given by

$$\left(\frac{\eta_{\mathrm{c}}}{\eta_{\mathrm{m}}}\right)_s = \frac{1}{\sqrt{N_0 f(z)\Delta z_{\mathrm{bin}}\exp(-\bar{n}_s)}}, \tag{8}$$

where we substituted for $\Delta N(z)$ to explicitly show the dependence of the saturation occupation number on the bin size and the subscript s denotes saturation value.

The uncertainty decreases as Δz_{bin} increases. However, since the bin size cannot be arbitrarily large, this ratio has a lower bound, which we parametrize as $\eta_{\mathrm{c}}^{\min}/\eta_{\mathrm{m}} \equiv \min(\eta_{\mathrm{c}}/\eta_{\mathrm{m}})_s = \epsilon(z)$, where the minimum value is obtained by choosing the largest allowed bin size. The occupation number in this case is given by $\bar{n}_s = \nu(z) \times \sqrt{1 + \epsilon^2(z)}$, and using Eq. (8) it follows that the bin size required to achieve this accuracy is given by

$$\Delta z_{\mathrm{bin}} = \frac{\exp[\nu(z)\sqrt{1 + \epsilon^2(z)}]}{N_0 f(z)\epsilon^2(z)}. \tag{9}$$

Averaging sources in the bin Δz_{bin} introduces a systematic bias η_c^{sys} in the DZ relation. If the bin size is small we can assume that the sources are distributed uniformly in the bin. By Taylor expanding the luminosity distance D_L, and taking its average over the bin we can easily obtain

$$\eta_c^{\mathrm{sys}} = \frac{\langle D_L(z_0)\rangle - D_L(z_0)}{D_L} \simeq \frac{1}{24}\frac{D_L''}{D_L}\Delta z_{\mathrm{bin}}^2, \tag{10}$$

which is correct up to third order in Δz_{bin}, and $\langle\rangle$ denotes averaging over the bin. If we demand $\eta_c^{\min} = \epsilon(z)\eta_{\mathrm{m}} > \eta_c^{\mathrm{sys}}$ then we obtain

$$\epsilon(z)\exp\left[-\frac{2}{5}\nu(z)\sqrt{1 + \epsilon^2(z)}\right] \geq \frac{1}{24^{1/5}N_0^{2/5}f(z)^{2/5}\eta_{\mathrm{m}}^{1/5}}\left(\frac{D_L''}{D_L}\right)^{1/5}, \tag{11}$$

where we have substituted Δz_{bin} from Eq. (9). This formula encapsulates our main result and determines the limit of self-calibration.

In Fig. 1 we plot $\delta D_L/D_L = \epsilon(z)\eta_{\mathrm{m}}$ for BBO, assuming an equality in the above expression, i.e. assuming that the systematic term is equal to the random error. We have taken a flat ΛCDM model with $\Omega_{\mathrm{m}} = 0.3$ for this figure. For this plot we have taken N_0, $f(z)$, and $\Delta\Omega_{\mathrm{BBO}}(z)$ from [10]. The same figure shows the accuracy obtainable for a degraded beam by applying a constant multiplying factor to the BBO beam value.

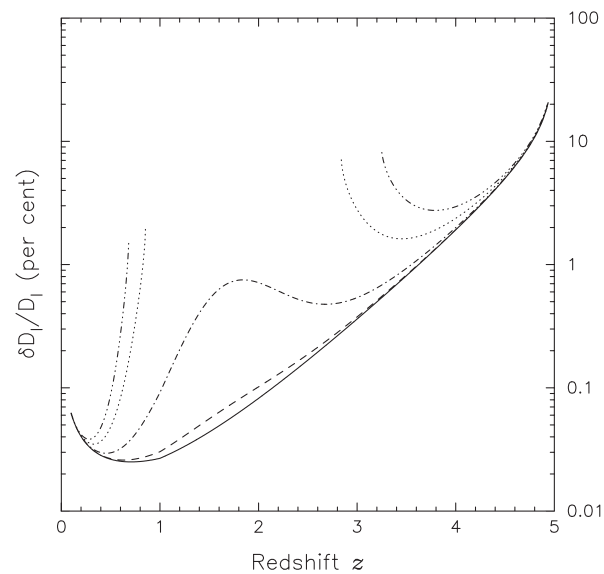

FIG. 1. The maximum achievable accuracy on the DZ relation as a function of redshift using self-calibration for the BBO case. The solid line corresponds to perfect pointing and the dashed curve is for the BBO pointing accuracy where $\Delta\Omega_{\mathrm{BBO}}$ is taken from [10]. The other curves correspond to degrading the pointing accuracy to $10\Delta\Omega_{\mathrm{BBO}}$, $50\Delta\Omega_{\mathrm{BBO}}$, $100\Delta\Omega_{\mathrm{BBO}}$, corresponding to curves with increasing values of $\delta D_L/D_L$. The last two curves have a region in the middle where self-calibration does not work since the required number of sources from Eq. (12) is larger than 3×10^5. The BBO accuracy is almost as good as the case for perfect pointing, with small departures at intermediate redshifts.

Necessary condition for self-calibration—The left-hand side of Eq. (11) has a redshift-dependent upper bound that we denote as $b(z)$. Self-calibration would work only if the left-hand side is larger than the right -hand side, leading to

$$N_0 > \frac{1}{24^{1/2} b^{5/2}(z) f(z) \eta_{\rm m}^{1/2}} \left(\frac{D_L''}{D_L}\right)^{1/2}. \qquad (12)$$

If this condition is not satisfied then the bin size required is too large and the systematic term would dominate the random error. Therefore, in an experiment the total number of sources N_0 should satisfy this inequality to self-calibrate at a given redshift. In Fig. 2 we plot this for a few redshifts as a function of $\Delta\Omega$.

Gain factor—We now give a rough estimate of the number of sources resolved through this method. At a given redshift z a single source measures the luminosity distance at a fractional accuracy $\eta_{\rm m}$, which we can take as the cosmological accuracy $\eta_{\rm c}$, leading to the occupation number at the beginning to be $\bar{n}_0 = \nu(z)\sqrt{1 + \eta_{\rm c}^2/\eta_{\rm m}^2} = \sqrt{2}\nu(z)$. To obtain the resolved fraction at the end of iteration we shall consider the extreme case when $N_0 \gg 1$, which along with Eq. (8) gives $\epsilon_s \simeq 0$, therefore the saturation occupation number is $\bar{n}_s \simeq 0.41\nu(z)$. Since for a given occupation number $\bar{n}$, the fraction of total resolved sources is given by $\exp(-\bar{n})$, it is clear that the ratio of resolved sources at the end to that at the beginning (of iteration) is given by $\exp[0.414\nu(z)]$. Since $\nu(z)$ is proportional to the pointing accuracy, the gain is an exponential function of the beam size. As an example, at $z \simeq 1.75$, for BBO the minimum occupation number $\nu = 0.63$ [18], giving a maximum gain factor of $\sim$1.3, while for DECIGO, assuming a beam linear size about a factor three worse, the gain factor is about $\sim$11, showing the extreme sensitivity of gain to the directional sensitivity.

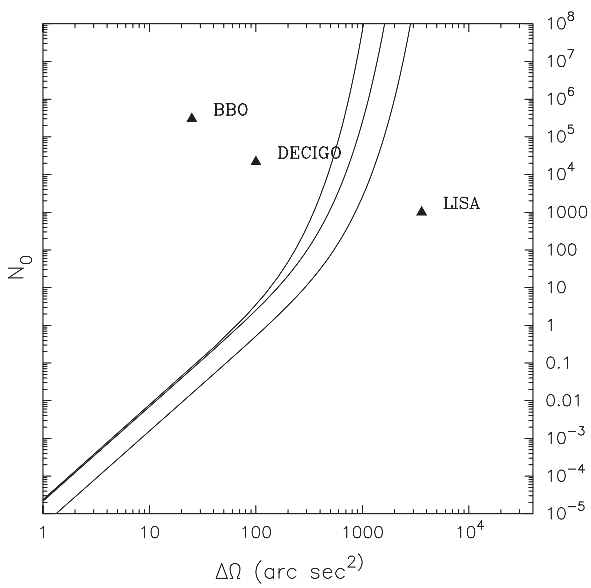

FIG. 2. The minimum number of sources N_0 required to self-calibrate the DZ relation at a given redshift [see Eq. (12)] is shown as a function of the pointing accuracy $\Delta\Omega$. The curves (right to left) correspond to $z = 0.5$, 1, 2. The value of N_0 rises steeply with the beam width since an increasing bin size, $\Delta z_{\rm bin}$, necessary to suppress cosmology errors, conflicts with the requirement of unbiased calibration. (The latter cannot be satisfied if most GW sources are associated with multiple optical counterparts in the pointing beam.)

Effect of clustering—In the discussion so far we have neglected the impact of galaxy clustering. The effect of the galaxy clustering can be taken into account by replacing $\bar{n}$ with

$$\bar{n}\left(1 + \frac{1}{\Delta V}\int \xi dV\right). \qquad (13)$$

Here ΔV is the volume bounding the redshift and angular error box (Δz and $\Delta\Omega$) in the determination of the source. ξ is the two-point correlation function of galaxy clustering and the integral extends over ΔV.

Here we give estimates of the impact of clustering [second term of Eq. (13)] for BBO and DECIGO configurations at $z \simeq 2$. In the approximation, valid for these cases, in which the (comoving) linear size corresponding to angular resolution $l_\perp$ is much smaller than the distance corresponding to the radial distance $l_\parallel$ of the (minimum, i.e. when $\eta_{\rm c} \to 0$) redshift error, one can readily show that

$$\frac{\Delta\bar{n}}{\bar{n}} = \frac{1}{\Delta V}\int \xi dV \simeq \left(\frac{l_\perp}{0.2\ {\rm Mpc}}\right)^{-0.8}\left(\frac{l_\parallel}{300\ {\rm Mpc}}\right)^{-1}. \qquad (14)$$

$\Delta\bar{n}/\bar{n} \simeq \{1, 0.45\}$ for the BBO and the DECIGO at $z \simeq 2$. Equation (14) shows that this term scales inversely with $l_\parallel$, and therefore the effect of clustering would be less important in the beginning of the iteration process when $\eta_{\rm c}$ could be appreciable but would be increasingly important as the maximum achievable precision is approached.

We have shown that by iterating over a self-consistently obtained DZ relation from resolved gravity wave sources it is possible to improve the DZ relation and therefore isolate those sources that initially are unidentifiable (owing to multiple objects in the pointing beam).

However, due to the fact that in this process only the cosmological errors are reduced, the limiting resolved set crucially depends on the pointing accuracy at a given redshift. We have derived analytical expressions for the final accuracy reached on the DZ relationship as well as the condition for successful self-calibration [Eq. (12) and Fig. 2]. Our formulation will help future GW probes grapple with the issue of redshift measurement uncertainty due to the presence of multiple objects within their beam (Fig. 1). A comprehensive analysis using simulated data to estimate cosmological constraints arising from future GW experiments will be presented in a companion paper.

APPENDIX

We now derive the probability distribution function (PDF) for the source redshift given a cosmological model and a distance measurement to a GW source. We first derive the general formula and then specialize to the local approximation used in this paper.

Let the measured distance be given by d_m. To quantify cosmology errors we employ a linear model for the DZ relation,

$$D_L(z, \mathbf{h}) = \sum_i^N h_i f_i(z) = \mathbf{h}^T \mathbf{f}, \tag{A1}$$

where $\mathbf{h}$ are the N parameters of the model, f_i are N arbitrary functions of redshift, and we have defined $\mathbf{f} = [f_1(z), f_2(z), \ldots, f_N(z)]$. The unknown redshift is to be treated as a parameter of the model. The simplest choice is $f_i = z^{i-1}$, leading to a polynomial form for $D_L(z)$. The parameters $\mathbf{h}$ are not known precisely and are described by the Gaussian distribution

$$P(\mathbf{h}) = \frac{1}{(2\pi)^{N/2}\sqrt{\det \mathbf{C}}} \exp\left[-\frac{1}{2}(\mathbf{h}^T - \mathbf{h}_0^T)\mathbf{C}^{-1}(\mathbf{h} - \mathbf{h}_0)\right], \tag{A2}$$

where $\mathbf{C}$ is the covariance matrix, obtained by fitting the model to the resolved sources (or to other data sets), and $\mathbf{h}_0$ are the best fit parameters. We employ bold lower case letters to denote column matrices and bold capital letters to denote second rank matrices. Employing the Bayes theorem we can write down the posterior probability for the parameters of the model as

$$P(z, \mathbf{h}|d_m) \propto P(d_m|z, \mathbf{h})P(\mathbf{h})P(z), \tag{A3}$$

where $P(\mathbf{h})$ is the prior PDF for the parameters $\mathbf{h}$ given by Eq. (A2), the prior $P(z)$ is assumed to be flat, and

$$P(d_m|z, \mathbf{h}) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left[-\frac{(d_m - \mathbf{h}^T\mathbf{f})^2}{2\sigma_m^2}\right]. \tag{A4}$$

The PDF in Eq. (A3) is a function of redshift z and $\mathbf{h}$, therefore the posterior PDF for the source redshift can be obtained by integrating over $\mathbf{h}$

$$P(z|d_m) \propto \int P(d_m|z, \mathbf{h})P(\mathbf{h})d^N h, \tag{A5}$$

which can be expressed through variables $\mathbf{g} = \mathbf{h} - \mathbf{h}_0$ and $\chi = d_m - \mathbf{h}_0^T\mathbf{f}$ as

$$P(z|d_m) \propto \int \exp\left[-\frac{\chi^2 - 2\chi\mathbf{g}^T\mathbf{f} + \mathbf{g}^T\mathbf{f}\mathbf{g}}{2\sigma_m^2} - \frac{1}{2}\mathbf{g}^T\mathbf{C}^{-1}\mathbf{g}\right]d^N g, \tag{A6}$$

where the coefficients that do not contain z and $\mathbf{g}$ have been dropped, and we have defined a second rank matrix $\mathbf{f} = \mathbf{f} \otimes \mathbf{f}$. If we define a matrix $\mathbf{S} = \mathbf{C}^{-1} + \mathbf{f}/\sigma_m^2$ then this equation takes the form

$$P(z|d_m) \propto \int \exp\left[-\frac{\chi^2 - 2\chi\mathbf{g}^T\mathbf{f}}{2\sigma_m^2} - \frac{1}{2}\mathbf{g}^T\mathbf{S}\mathbf{g}\right]d^N g. \tag{A7}$$

Translating the coordinate system in the parameter space $\mathbf{g} = \mathbf{u} + \mathbf{u}_0$, where $\mathbf{u}_0$ is such that $\mathbf{S}\mathbf{u}_0 = \chi\mathbf{f}/\sigma_m^2$, we finally obtain

$$P(z|d_m) \propto \int \exp\left[\frac{1}{2}\left(-\frac{\chi^2}{\sigma_m^2} + \frac{\chi}{\sigma_m^2}\mathbf{u}_0^T\mathbf{f} - \mathbf{u}^T\mathbf{S}\mathbf{u}\right)\right]d^N u. \tag{A8}$$

Carrying out the integration and dropping all terms that do not depend on the redshift we obtain

$$P(z|d_m) \propto \det(\mathbf{S}) \exp\left[-\frac{\chi^2}{2\sigma_m^2}\left(1 - \frac{\mathbf{f}^T\mathbf{S}^{-1}\mathbf{f}}{\sigma_m^2}\right)\right]. \tag{A9}$$

Recalling that $\chi = d_m - \mathbf{h}_0^T\mathbf{f}$, we find that the redshift probability distribution is centered at the redshift predicted for the distance d_m by the best fit model $d(z; \mathbf{h}_0)$. Since the functions $\mathbf{f}$ are redshift dependent, the precise behavior of this function is complicated. $P(z|d_m)$ can be normalized in the range $z = 0$ to $z = z_{\max}$, and would, in general, produce an asymmetric distribution, due to the manner in which the cosmological errors scale with redshift.

Local approximation—Since the PDF peaks at $\chi = 0$, we can define a redshift z_0 through $d_m = \mathbf{h}_0^T\mathbf{f}(z_0)$. If the cosmology is determined precisely then we can assume the redshift PDF to decline rapidly away from z_0, and therefore we can replace $\mathbf{h}_0^T\mathbf{f}(z) = \mathbf{h}_0^T\mathbf{f}(z_0) + \mathbf{h}_0^T\mathbf{f}'(z_0)(z - z_0)$, implying $\chi = -\mathbf{h}_0^T\mathbf{f}'(z_0)(z - z_0)$. Then, at the same level of accuracy we can replace $\mathbf{f} \equiv \mathbf{f}_0 = \mathbf{f}(z_0)$ in $\mathbf{f}^T\mathbf{S}^{-1}\mathbf{f}$. To evaluate $\mathbf{f}_0^T\mathbf{S}_0{}^{-1}\mathbf{f}_0$, we need an expression for $\mathbf{S}_0{}^{-1}$. Noting that $\mathbf{F}_0$ is a rank one matrix, we have [19]

$$\mathbf{S}_0{}^{-1} = \mathbf{C} - \frac{1}{1+g}\frac{\mathbf{C}\mathbf{F}_0\mathbf{C}}{\sigma_m^2}, \tag{A10}$$

where $g = \mathrm{tr}\mathbf{F}_0\mathbf{C}/\sigma_m^2$. Noting that $\sigma_c^2 = \mathrm{tr}\mathbf{F}_0\mathbf{C}$, it can be readily shown that

$$\mathbf{f}_0\mathbf{S}_0{}^{-1}\mathbf{f}_0 = \frac{\sigma_m^2\sigma_c^2}{\sigma_m^2 + \sigma_c^2}. \tag{A11}$$

The redshift probability distribution function can be now written explicitly as

$$P(z|d_m) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left[-\frac{(z - z_0)^2}{2\sigma_z^2}\right], \tag{A12}$$

where $\sigma_z = \sqrt{(\sigma_m^2 + \sigma_c^2)}/D_L'(z_0)$.

[1] V. Sahni and A. A. Starobinsky, Int. J. Mod. Phys. D **9**, 373 (2000); P. J. E. Peebles and B. Ratra, Rev. Mod. Phys. **75**, 559 (2003); T. Padmanabhan, Phys. Rep. **380**, 235 (2003); V. Sahni, Lect. Notes Phys. **653**, 141 (2004); E. J. Copeland, M. Sami, and S. Tsujikawa, Int. J. Mod. Phys. D **15**, 1753 (2006); V. Sahni and A. A. Starobinsky, Int. J. Mod. Phys. D **15**, 2105 (2006); J. A. Frieman, M. S. Turner, and D. Huterer, Annu. Rev. Astron. Astrophys. **46**, 385 (2008).
[2] B. F. Schutz, Nature (London) **323**, 310 (1986).
[3] N. Seto, S. Kawamura, and T. Nakamura, Phys. Rev. Lett. **87**, 221103 (2001); S Kawamura *et al.*, J. Phys. Conf. Ser. **122**, 012006 (2008).
[4] D. E. Holz and S. A. Hughes, Astrophys. J. **629**, 15 (2005).
[5] K. Arun, B. Iyer, B. Sathyaprakash, S. Sinha, and C. Van Den Broeck, Phys. Rev. D **76**, 104016 (2007); **76**, 129903 (2007).
[6] B. S. Sathyaprakash, B. F. Schutz, and C. Van Den Broeck, arXiv:0906.4151.
[7] T. Apostolatos, C. Cutler, G. Sussman, and K. Thorne, Phys. Rev. D **49**, 6274 (1994).
[8] E. S. Phinney *et al.*, Big Bang Oberver Mission Concept Study (NASA) (2003).
[9] Wei-Tou Ni, Mod. Phys. Lett. A **25**, 922 (2010).
[10] C. Cutler and D. E. Holz Phys. Rev. D**80**, 104009 (2009).
[11] B. Kocsis, Z. Haiman, and K. Menou, Astrophys. J. **684**, 870 (2008).
[12] C. L. MacLeod and C. L. Hogan, Phys. Rev. D **77**, 043512 (2008).
[13] R. P. Kirshner, arXiv:0910.0257; J. D. Neill *et al.*, Astrophys. J. **707**, 1449 (2009).
[14] D. E. Holz and E. V. Linder, Astrophys. J. **631**, 678 (2005).
[15] N. Kaiser, Astrophys. J. **388**, 272 (1992).
[16] W. Hu, Astrophys. J. **522**, L21 (1999).
[17] S. V. W. Beckwith *et al.*, Astron. J. **132**, 1729 (2006).
[18] This number is twice that in Fig 6 of [10], which we suspect is due to a difference in interpreting Δz_a as the allowed range. In our analysis the allowed range for a standard deviation Δz_a is twice the standard deviation.
[19] K. S. Miller, Mathematics Magazine **54**, 67 (1981).