LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Document Type	**LIGO-T020115-00-Z**	12 August 2001

Stochastic Sources Upper Limit Group E7 Report

Bruce Allen, Warren Anderson, Sukanta Bose, Nelson Christensen, Ed Daw, Mario Diaz, Ronald Drever, Sam Finn, Peter Fritschel, Joe Giaime, Bill Hamilton, Siong Heng, Warren Johnson, Robert Johnston, Erik Katsavounidis, Sergei Klimenko, Michael Landry, Albert Lazzarini, Martin McHugh, Tom Nash, Adrian Ottewill, Patricia Pérez, Tania Regimbau, Jamie Rollins, Joseph Romano, Bernard Schutz, Antony Searle, Peter Shawhan, Charlie Torres, Erick Vallarino, Alberto Vecchio, Rai Weiss, John Whelan, Bernard Whiting

Distribution of this draft:

Stochastic Sources Upper Limit Group

California Institute of Technology
LIGO Project - MS 51-33
Pasadena CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project - MS 20B-145
Cambridge, MA 01239
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

WWW: http://www.ligo.caltech.edu/

RCS Id: E7report.tex,v 1.42 2002/08/19 16:14:58 joe Exp — Processed with LaTeX on 2002/08/19

1 Introduction

From 28 Decemeber 2001 to 14 January 2002, the LIGO Hanford and LIGO Livingston Observatories (LHO [1] and LLO [2]) took coincident engineering data (E7) [3]. Since that time, the inspiral [4], burst [5], periodic [6], and stochastic [7] sources working groups of the LIGO Scientific Collaboration [8] (in collaboration with the LIGO Lab [9]) have been developing code and analyzing that data to search for gravitational wave signals. This report describes the E7 analyses performed by the stochastic sources working group, which have ultimately led to an upper limit of $\Omega_{\rm gw}(f)\,h_{100}^2 \leq 7.7 \times 10^4$ on the strength of stochastic gravitational wave signals in the frequency band $40\ {\rm Hz} < f < 215\ {\rm Hz}$.

2 Preliminaries

Here we briefly describe the standard optimally-filtered cross-correlation technique that was used to search for a stochastic background of gravitational radiation. Readers interested in more details should consult the original papers [10, 11, 12] or longer review articles (e.g., [13, 14, 15]) for a more indepth discussion.

2.1 Stochastic background spectrum

A stochastic background of gravitational radiation is a *random* gravitational wave signal produced by a large number of weak, independent, unresolved gravitational wave sources. Its spectral properties are described by the dimensionless quantity

$$\Omega_{\rm gw}(f) := \frac{1}{\rho_{\rm critical}}\,\frac{d\rho_{\rm gw}}{d\,\ln\,f}\ , \tag{1}$$

which is the ratio of the energy density in gravitational waves contained in a bandwidth $\Delta f = f$ to the total energy density required (today) to close the universe:

$$\rho_{\rm critical} = \frac{3c^2H_0^2}{8\pi G}\ . \tag{2}$$

H_0 is the Hubble expansion rate (today):

$$H_0 = h_{100}\cdot 100\,\frac{\rm km}{\rm sec\cdot Mpc} \approx 3.24\times 10^{-18}\,h_{100}\,\frac{1}{\rm sec}\ , \tag{3}$$

and h_{100} is a dimensionless factor, included to account for the different values of H_0 that are quoted in the literature.[1] Note that $\Omega_{\rm gw}(f)\ h_{100}^2$ is *independent* of the actual Hubble expansion rate, and for this reason we will often focus attention on this quantity, rather than $\Omega_{\rm gw}(f)$ alone. In addition, $\Omega_{\rm gw}(f)$ is related to the one-sided power spectral density $S_{\rm gw}(f)$ via[2]

$$S_{\rm gw}(f) = \frac{3H_0^2}{10\pi^2}\,f^{-3}\Omega_{\rm gw}(f)\ . \tag{4}$$

Thus, for a stochastic gravitational wave background with $\Omega_{\rm gw}(f) = {\rm const}$, the power in gravitational waves falls off like $1/f^3$.

[1] h_{100} almost certainly lies within the range $1/2 < h_{100} < 1$.

[2] $S_{\rm gw}(f)$ is defined by $\frac{1}{T}\int_0^T |h(t)|^2\,dt = \int_0^\infty S_{\rm gw}(f)\,df$, where $h(t)$ is the gravitational wave strain in a single detector due to the stochastic background signal.

2.2 Statistical assumptions

The spectrum $\Omega_{\rm gw}(f)$ completely specifies the statistical properties of a stochastic background of gravitational radiation provided we make enough additional assumptions. Here, we assume that the stochastic background is: (i) isotropic, (ii) unpolarized, (iii) stationary, and (iv) Gaussian. Anisotropic or non-Gaussian backgrounds (e.g., due to an incoherent superposition of gravitational waves from a large number of unresolved white dwarf binary star systems in our own galaxy, or a "pop-corn" stochastic signal produced by gravitational waves from supernova explosions [16, 17, 18]) will require different data analysis techniques than the one we present here. (See, e.g. [19, 20] for a detailed discussion of these different techniques.)

In addition, we will assume that the intrinsic detector noise is: (i) stationary, (ii) Gaussian, (iii) uncorrelated between different detectors and with the stochastic gravitational wave signal, and (iv) much greater in power than the stochastic gravitational wave background.

2.3 Cross-correlation statistic

The standard method of detecting a stochastic gravitational wave signal is to *cross-correlate* the output of two gravitational wave detectors [10, 11, 12, 13, 14, 15]:

$$Y_Q = \int_0^T dt_1 \int_0^T dt_2\, h_1(t_1)\, Q(t_1 - t_2)\, h_2(t_2) \tag{5}$$

$$= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df'\, \delta_T(f - f')\, \tilde{h}_1^*(f) \tilde{Q}(f') \tilde{h}_2(f')\,, \tag{6}$$

where T is the observation time and $\delta_T(f - f')$ is a finite-time approximation to the Dirac delta function $\delta(f - f')$.[3] Assuming that the detector noise is uncorrelated between the detectors, it follows that the expected value of Y_Q depends only on the cross-correlated stochastic signal:

$$\mu = \frac{T}{2} \int_{-\infty}^{\infty} df\, \gamma(|f|) S_{\rm gw}(|f|)\, \tilde{Q}(f)\,, \tag{7}$$

while the variance of Y_Q is dominated by the noise in the individual detectors:

$$\sigma^2 \approx \frac{T}{4} \int_{-\infty}^{\infty} df\, P_1(|f|)\, |\tilde{Q}(f)|^2\, P_2(|f|)\,. \tag{8}$$

($P_1(|f|)$ and $P_2(|f|)$ are again one-sided power spectral densities.) The integrand of Eq. (7) contains a factor $\gamma(f)$, called the *overlap reduction function* [12], which characterizes the reduction in sensitivity to detecting a stochastic background due to: (i) the separation time delay, and (ii) the relative orientation of the two detectors. (For coincident and coaligned detectors, $\gamma(f) = 1$ for all frequencies.) Plots of the overlap reduction function for correlations between LIGO Livingston and the other major interferometers and ALLEGRO are shown in Fig. 1.

2.4 Optimal filter

Given Eqs. (7) and (8), it is relatively straightforward to show that the SNR ($= \mu/\sigma$) is maximized when

$$\tilde{Q}(f) = \lambda \frac{\gamma(|f|) S_{\rm gw}(|f|)}{P_1(|f|) P_2(|f|)} \propto \frac{\gamma(|f|)\Omega_{\rm gw}(|f|)}{|f|^3 P_1(|f|) P_2(|f|)}\,, \tag{9}$$

[3] $\delta_T(f) := \int_{-T/2}^{T/2} dt\, e^{-i2\pi ft} = \sin(\pi f T)/\pi f$.

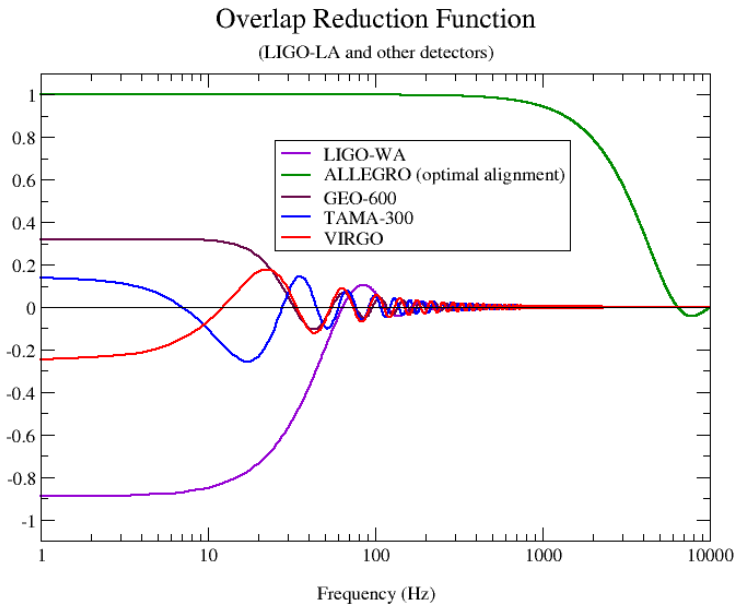

Figure 1: Overlap reduction function between LIGO Livingston and the other major interferometers plus ALLEGRO (in an optimal alignment of 72° East of North).

where λ is a (real) overall normalization constant. Such a $\tilde{Q}(f)$ is called the *optimal filter* for the cross-correlation statistic. For such a $\tilde{Q}(f)$, the expected SNR is

$$\mathrm{SNR} \approx \frac{3H_0^2}{10\pi^2}\,\sqrt{T}\,\left[\int_{-\infty}^{\infty} df\;\frac{\gamma^2(|f|)\Omega_{\mathrm{gw}}^2(|f|)}{f^6 P_1(|f|)P_2(|f|)}\right]^{1/2}\,, \tag{10}$$

which grows like the square-root of the observation time T.

2.5 Observational constraints

(i) The strongest observational constraint on $\Omega_{\mathrm{gw}}(f)$ comes from the high degree of isotropy observed in the CMBR. The one-year[21, 22], two-year[23], and four-year[24] data sets from the Cosmic Background Explorer (COBE) satellite place very strong restrictions on $\Omega_{\mathrm{gw}}(f)$ at very low frequencies:

$$\Omega_{\mathrm{gw}}(f)\,h_{100}^2 \leq 7\times 10^{-11}\left(\frac{H_0}{f}\right)^2 \quad \text{for} \quad H_0 < f < 30H_0\,. \tag{11}$$

Since $H_0 \approx 3.24\times 10^{-18}\,h_{100}$ Hz, this limit applies only over a narrow band of frequencies (10^{-18} Hz $< f < 10^{-16}$ Hz), which is far below any frequency band accessible to investigation by either earth-based (10 Hz $\lesssim f \lesssim 10^3$ Hz) or space-based (10^{-4} Hz $\lesssim f \lesssim 10^{-1}$ Hz) detectors.

(ii) Another observational constraint comes from roughly a decade of monitoring the radio pulses arriving from a number of stable millisecond pulsars[25]. These pulsars are remarkably stable clocks, and the regularity of their pulses places tight constraints on $\Omega_{\mathrm{gw}}(f)$ at frequencies on the order of the inverse of the

observation time of the pulsars ($\sim 10^{-8}$ Hz):

$$\Omega_{\rm gw}(f = 10^{-8}\ {\rm Hz})\ h_{100}^2 \leq 10^{-8}\ . \tag{12}$$

Like the constraint on the stochastic gravitational wave background from the isotropy of the CMBR, the millisecond pulsar timing constraint is irrelevant for earth-based and space-based detectors.

(iii) The third and final observational constraint on $\Omega_{\rm gw}(f)$ comes from the standard model of big-bang nucleosynthesis[26]. This model provides remarkably accurate fits to the observed abundances of the light elements in the universe, tightly constraining a number of key cosmological parameters. One of the parameters constrained in this way is the expansion rate of the universe at the time of nucleosynthesis. This places a constraint on the energy density of the universe at that time, which in turn constrains the energy density in a cosmological background of gravitational radiation:

$$\int_{f>10^{-8}\ {\rm Hz}} d\ln f\ \Omega_{\rm gw}(f)\ h_{100}^2 \leq 10^{-5}\ . \tag{13}$$

This constraint corresponds to a 95% confidence upper bound on $\Omega_{\rm gw}(f)$ of roughly 10^{-7} in the frequency band of earth-based interferometers.

2.6 Upper limits

In addition to the above observational constraints, there are a couple of (much weaker) upper limits on $\Omega_{\rm gw}(f)$ that have been set directly using gravitational wave data: (i) An upper limit from a correlation measurement between the Garching and Glasgow prototype interferometers[27]:

$$\Omega_{\rm gw}(f)\, h_{100}^2 \leq 3 \times 10^5 \quad \text{for} \quad 100 < f < 1000\ {\rm Hz}\ , \tag{14}$$

(ii) An upper limit from data taken by a single resonant bar detector[28]:

$$\Omega_{\rm gw}(f = 907\ {\rm Hz})\, h_{100}^2 \leq 100\ . \tag{15}$$

(iii) An upper limit from a correlation measurement between the EXPLORER and NAUTILUS resonant bar detectors[29, 30]:

$$\Omega_{\rm gw}(f = 907\ {\rm Hz})\, h_{100}^2 \leq 60\ . \tag{16}$$

Note that these last two upper limits are for $\Omega_{\rm gw}(f)$ evaluated at a *single* frequency ($f = 907$ Hz), which is near the resonant frequency of the bar detectors.

2.7 Data analysis pipeline

In an actual search for a stochastic background signal, we work with discretely-sampled data broken up into segments $T = 90\ {\rm sec}$ in length. Integrations in the previous formulas are replaced by summations and continuous Fourier transforms by discrete Fourier transforms. Within the LIGO data analysis system (LDAS), we request the gravitational wave data (LSC-AS_Q from LHO and LLO) in 15-minute chunks, originally sampled at 16384 Hz. We then down-sample the data to 1024 Hz, and estimate power spectra for each detector, which are used in the calculation of the optimal filter for a stochastic background signal with $\Omega_{\rm gw}(f) = {\rm const} := \Omega_0$. (The normalization factor λ [cf. Eq. (9)] is chosen so that the expected mean value is $\mu = \Omega_0\, T$.) The data are then split into 10 (90-second) segments, each of which is windowed in

the time domain (e.g., with a Hann window; see Sec. 3.3), zero-padded to twice its length, and discrete Fourier-transformed. A value of the optimally-filtered cross-correlation statistic is then calculated for each $T = 90$-sec data stretch ($Y_{QI1}, Y_{QI2}, \cdots Y_{QI10}$), along with the theoretical variance σ_I^2 [cf. Eq. (8)] of the individual cross-correlation statistics, valid for the whole 15-minute chunk.[4] (Here, I labels the different 15-minute chunks $I = 1, 2, \cdots, M$, where $M \sim 200$ for the 70 hours of clean, coincident H2-L1 E7 data.) For each 15 minute chunk, we calculate the sample mean

$$\bar{Y}_{QI} := \frac{1}{10} \sum_{J=1}^{10} Y_{QIJ} \tag{17}$$

and sample standard deviation

$$s_I := \left(\frac{1}{9} \sum_{J=1}^{10} (Y_{QIJ} - \bar{Y}_{QI})^2 \right)^{1/2} \tag{18}$$

of the 10 cross-correlation statistic values. We then form a weighted average[5]

$$\bar{Y}_Q := \frac{\sum_{I=1}^{M} \lambda_I \bar{Y}_{QI}}{\sum_{J=1}^{M} \lambda_J} \,, \quad \text{where} \quad \lambda_I = \sigma_I^{-2} \tag{19}$$

to obtain a point estimate of the stochastic background signal strength. For $\Omega_{\rm gw}(f) = \Omega_0 = \text{const}$, the point estimate of $\Omega_0\, h_{100}^2$ is given by

$$\hat{\Omega}_0\, h_{100}^2 := \bar{Y}_Q / T \tag{20}$$

with standard error

$$\hat{\sigma} := \frac{1}{\sqrt{10}} \frac{\left(\sum_{I=1}^{M} \lambda_I^2 s_I^2 \right)^{1/2}}{\sum_{J=1}^{M} \lambda_J} \,. \tag{21}$$

A schematic diagram of the data analysis pipeline is shown in Fig. 2.

2.8 Statistical approach

We use frequentist methods to convert our composite cross-correlation measurement into an upper limit on $\Omega_0\, h_{100}^2$. Ordinarily, the fact that $\Omega_0\, h_{100}^2$ must physically be positive, while the observed cross-correlation could be positive or negative, could lead to unphysical formal bounds unless we applied a technique like that described in [31]. However, as was illustrated by the analysis of the hardware-injected signals [cf. Sec. 3.6], the instrument response function of each detector is only known up to an overall sign. This means that $\bar{Y}_Q/T$ is a point estimate of the product of $\Omega_0\, h_{100}^2$ and an unknown sign ζ which represents our ignorance of the overall sign of the product of the two instrument response functions. (I.e., if both signs are correct, or both are incorrect, $\zeta = +1$, while if one is correct and the other incorrect, $\zeta = -1$.) This means that the quantity related to the results of our measurement, $\zeta \Omega_0\, h_{100}^2$, can be either positive or negative.

We could thus set a two-sided confidence limit on $\zeta \Omega_0\, h_{100}^2$, as illustrated in Fig. 3. Given a physical value $y = \zeta \Omega_0\, h_{100}^2 T/\hat{\sigma}$ of the unknown stochastic background strength in units of the standard error $\hat{\sigma}$, our

[4] We also calculate frequency series representing the integrand of the optimally-filtered cross-correlation statistic for each $T = 90$-sec data stretch. These spectra provide detailed information about the frequency bins that contribute most to the cross-correlation statistic values.

[5] The weighting factor $\lambda_I = \sigma_I^{-2}$ is chosen to maximize the signal-to-noise ratio for the overall average in the presence of nonstationary noise, as described in [14].

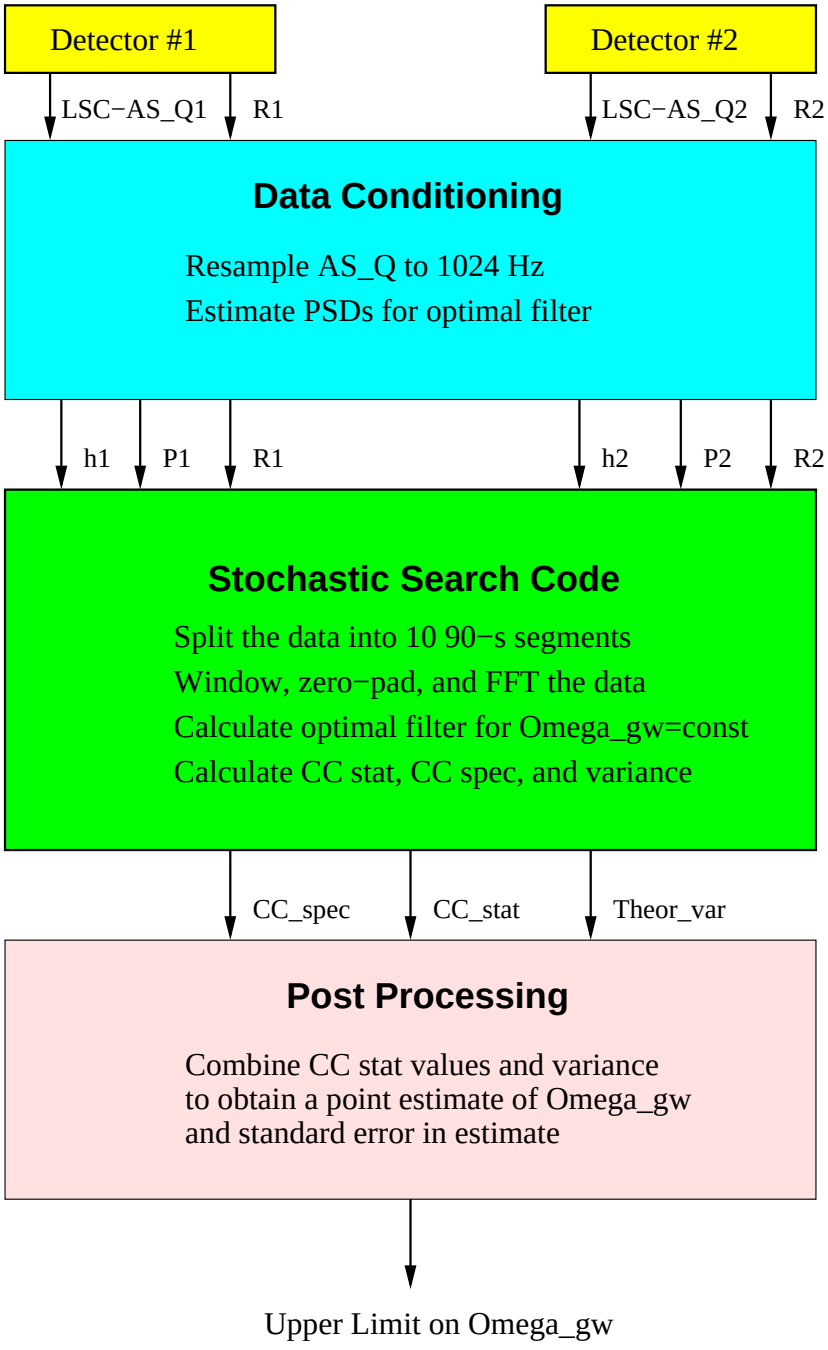

Figure 2: Data analysis pipeline for the stochastic search. LSC-AS_Q1,2 and R1,2 denote the raw gravitational wave output and response functions (containing calibration information) for detectors 1 and 2, respectively. h1,2 and P1,2 denote the resampled gravitational wave data and estimated power spectral densities for the two detectors.

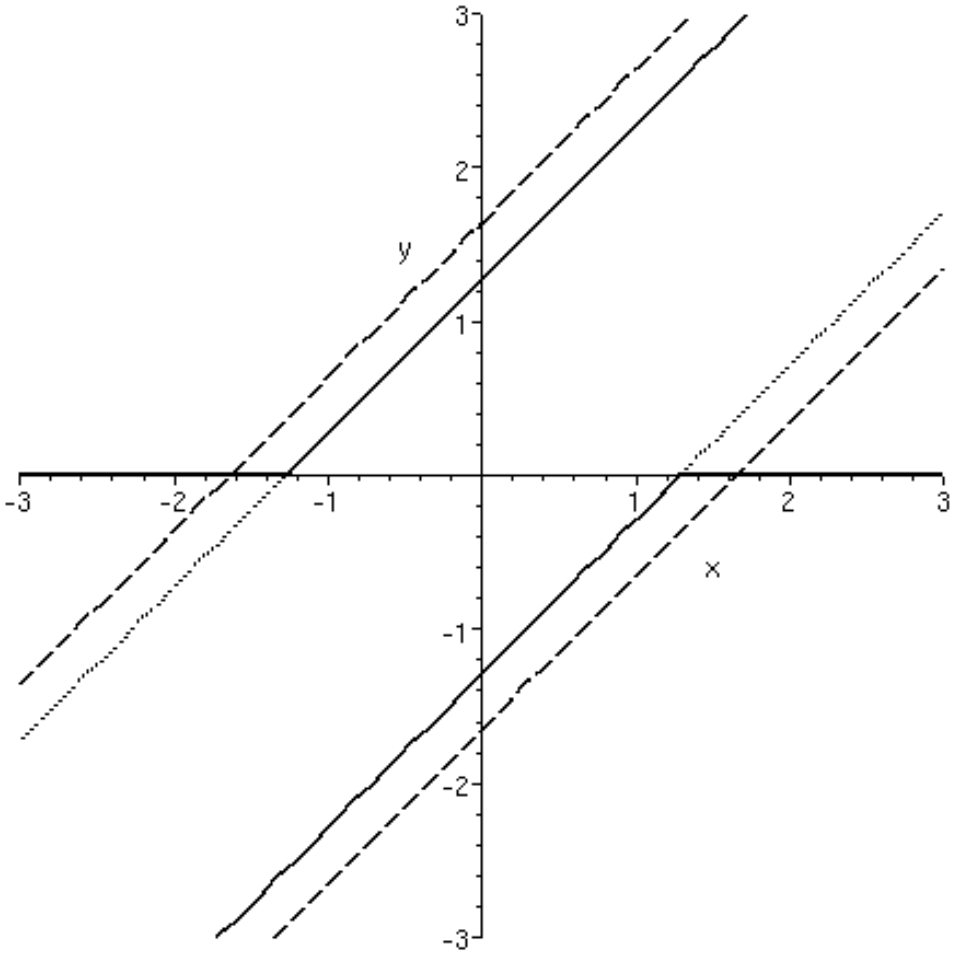

Figure 3: 90% confidence belts for cross-correlation measurement, in terms of the actual value of $y = \zeta\Omega_0\, h_{100}^2 T/\hat{\sigma}$ and the measured value of $x = \bar{Y}_Q/\hat{\sigma}$. The solid lines indicate our chosen confidence intervals. For reference, the dashed lines indicate 90%CL central confidence belts, the leftmost dotted line indicates the confidence interval for a 90%CL upper limit, and the rightmost dotted line corresponds to a 90%CL lower limit. (Both dotted lines are inifinite straight lines and include sections which are obscured by part of the solid lines.)

cross-correlation measurement $x = \bar{Y}_Q/\hat{\sigma}$ is a random Gaussian variable of mean y and unit variance, so the middle 90% of the distribution of x values for a given y values lies between the dashed lines. The classical confidence interval for the actual value of y based on a measurement of x is also between the dashed lines (the range of ys between the intersections with the dashed lines of a vertical line at the observed value of x).

Similarly, if we were seeking to set an upper limit on y, we would use confidence intervals defined by the leftmost of the two dotted lines; for a given y, 90% of the probability distribution of possible x measurements lies to the right of the leftmost dotted line. A lower limit on y would likewise be associated with the area to the left of the rightmost dotted line.

However, since the aim of this work is to set an upper limit on the stochastic background strength $\Omega_0\, h_{100}^2 = |y|\hat{\sigma}/T$, we choose a different set of confidence belts. These correspond to an upper limit on y for $y > 0$ and a lower limit on y for $y < 0$, and are illustrated by the solid lines in Fig. 3. For each $y \neq 0$, these confidence belts cover 90% of the probability distribution of x values. (The situation at $y = 0$ can be handled carefully by making a continuous transition between the two regimes for $|y| \leq \epsilon$; we are basically looking at the limit $\epsilon \to 0$.) This means that our limits on y are set as follows

$$\begin{aligned} \text{If } x < -1.28\,, \quad & \text{then} \quad x - 1.28 < y < 0\,; \\ \text{if } -1.28 < x < 1.28\,, \quad & \text{then} \quad x - 1.28 < y < x + 1.28\,; \\ \text{if } x > 1.28\,, \quad & \text{then} \quad 0 < y < x + 1.28\,. \end{aligned} \tag{22}$$

3 E7 investigations

3.1 Playground data

Roughly two hours of E7 data were set aside for stochastic upper limit playground analyses. This data corresponded to GPS times 693961597 to 693968475, and contained 5 900-seconds stretches of clean coincident data between the Hanford-2km and 4-km and Livingston-4km interferometers. All investigations that could have biased our upper limit on the stochastic background signal strength were initially performed on the playground data. These included the effect of different choices of windows (Sec. 3.3), high pass filtering (Sec. 3.4), and time-shifting (e.g., Sec. 3.6) the data streams. In addition, data analysis pipeline scripts and ldas user commands were originally tested on the playground data.

3.2 Expected upper limit

The *expected* 90% confidence level upper limit using 70 hours of clean, coincident data from the Hanford 2-km and Livingston 4-km interferometers is

$$\Omega_{\rm gw}(f)\, h_{100}^2 \leq 1.4 \times 10^5 \quad \text{for} \quad 40\ \text{Hz} < f < 215\ \text{Hz}\,. \tag{23}$$

This upper limit was calculated using Eq. 10, solving for $\Omega_{\rm gw}(f) = \text{const}$ with SNR set to 1.28 (for 90% confidence), $40\ \text{Hz} < f < 215\ \text{Hz}$, and typical E7 power spectra substituted for $P_1(f)$ and $P_2(f)$ (see Fig. 4). It corresponds to the upper limit we would expect to set if the measured cross-correlation was zero. Note that, due to variations in the noise floor over the course of the engineering run, the value quoted in Eq. (23) could differ from the actual measured value by an order of magnitude in either direction. Thus, 1.4×10^5 was simply an order of magnitude *estimate* of the upper limit we had expected to set using the E7 coincident data, *before* doing the actual analysis (Secs. 3.7 and 3.8).

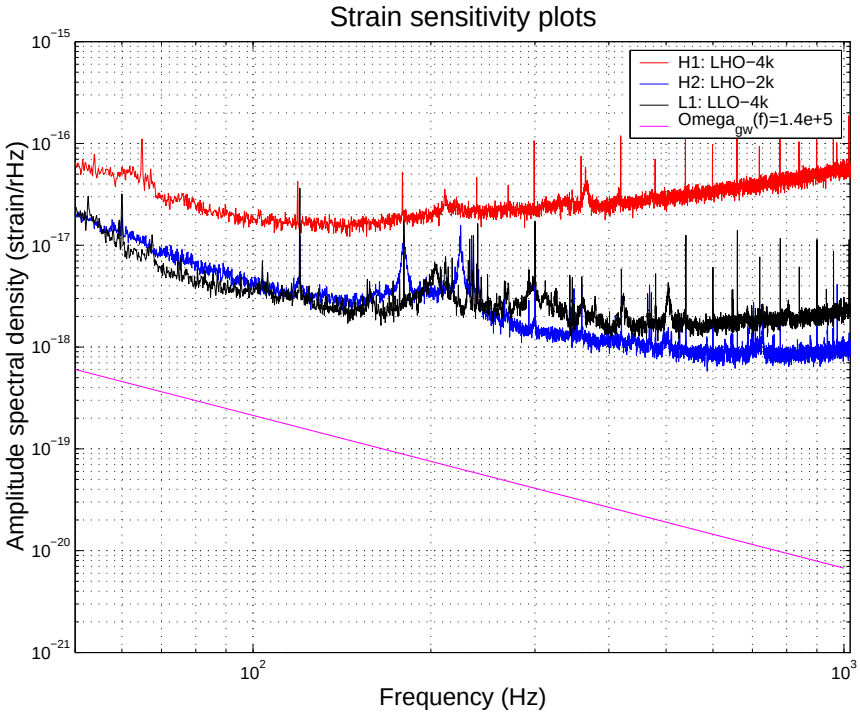

Figure 4: Typical strain sensitivity plots for E7. The straight line indicates the signal which would be generated in a single detector by a stochastic background with $\Omega_{gw}(f) = 1.4 \times 10^5$. Note that this is more than an order of magnitude (in amplitude) below the noise level of the quietest detector, illustrating the additional sensitivity arising from the cross-correlation measurement in around 70 hours of operation.

3.3 Windowing

Initial investigations of the amplitude and phase of the discrete Fourier-transformed data revealed spurious correlations at high frequencies. Ultimately, this was due to the large dynamic range of the gravitational wave data and the leakage of power introduced by rectangular windows, demonstrating the need to window the individual $T = 90$ sec data segments in the time-domain before applying the optimal filter.

Careful consideration of the effects of discrete sampling and windowing on the cross-correlation statistic [32] leads to modifications of the continuous-approximation expressions (Eqs. 7 and 8) for the expected mean and variance of the statistic. Provided that the windowing is sufficient to prevent significant leakage of power across the frequency range, the effect is to multiply the expected value of μ by $\overline{w_1 w_2}$ and the expected value of σ by $\overline{(w_1 w_2)^2}$, where w_1 and w_2 are the windowing functions and the bar indicates an average over the window. The effect is to multiply the signal-to-noise by a factor of

$$\frac{\overline{w_1 w_2}}{\sqrt{\overline{(w_1 w_2)^2}}} \,, \tag{24}$$

which when w_1 and w_2 are both Hann windows is approximately equal to $\sqrt{18/35} = 0.717$.

Non-overlapping Hann windows were subsequently agreed upon for the E7 upper limit analysis, despite the reduction in signal-to-noise ratio. *Spliced Hann windows*, consisting of a Hann window split in half with a constant section of all 1's in the middle[6] , were proposed as a way to combat this reduction in signal-to-noise ratio—e.g., one could use a spliced Hann window that would allow one to analyze more than 99% of the data—but it was not clear at the time of the E7 analysis if the leakage in power introduced by such a spliced window was sufficiently small to justify its use in the presence of spectral line noise. (See Fig. 5.) However, analysis of playground data with different spliced Hann windows (Fig. 6) seemed to indicate that as long as the overall Hann section was longer than around 0.1 seconds, the leakage was small enough to avoid the sorts of pathologies seen with rectangular windows.

There was also some discussion within our group regarding the use of *overlapping* Hann windows to avoid the reduction in signal-to-noise ratio, but the effect of overlapping windows on the statistics of the optimally-filtered cross-correlation statistic was not worked out in time for E7. Both of these suggestions (i.e., spliced Hann windows and overlapping Hann windows) are items of study for S1.

3.4 High-pass filtering

In order to reduce the effect of low frequency seismic noise on the variance of the optimally-filtered cross-correlation statistic, we had originally thought that we would need to high-pass filter the data, with low-frequency cutoff of $\sim$ 25 Hz. However, subsequent investigations using a 7th-order Butterworth filter showed that high-pass filtering was unnecessary if we limited the range of frequencies in the calculation of the cross-correlation statistic to 40 to 215 Hz. (The high frequency cutoff was chosen so that $\sim 90\%$ of the total signal-to-noise ratio come from this band.) The fractional difference in cross-correlation statistic values with and without high-pass filtering was roughly 1 part in 10^3. It may still be necessary to high-pass filter the data in the future, however, when the high-frequency sensitivity improves.

[6] Such a window has the appealing feature that the transition between the Hann and constant pieces at the top is just as smooth as the transition between the implicit 0 values outside the window and the Hann beginning and end. In effect, using a spliced Hann window is like using a Hann window which is shorter than the full length of the segment being windowed out, sacrificing some of the "windowing power" of the Hann window in order to use more of the data with non-overlapping windows.

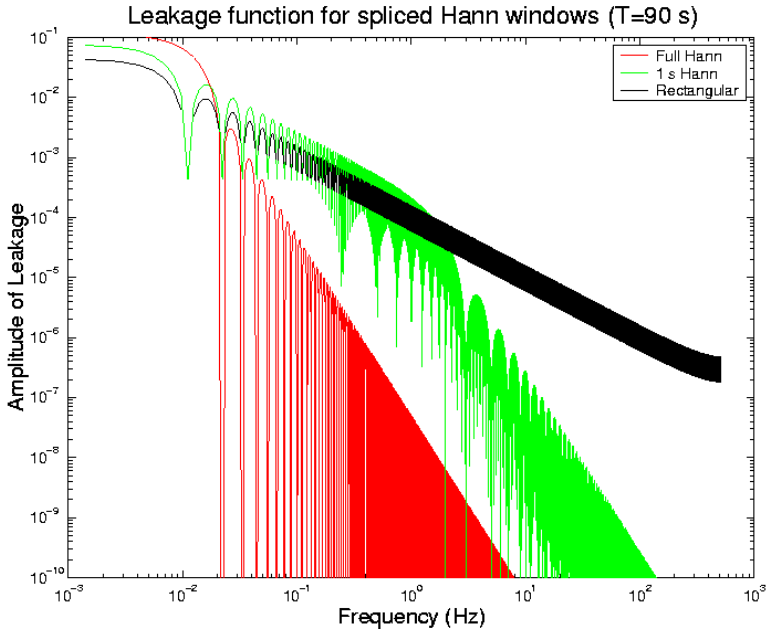

Figure 5: Leakage function for various windows. The black curve represents a rectangular window, the red curve a standard Hann window of width 90 seconds, and the green curve a "spliced Hann window" consisting of a 1-second Hann window with a 89-second flat section spliced into the middle.

3.5 60 Hz mains correlations

Initial studies by Klimenko et al. revealed long-term correlations in site-site power lines. Lazzarini, Schofield, and Viceré [33] subsequently explored whether these long-term correlations might arise from how the U.S. power grids are frequency-stabilized at 60 Hz. They learned that there may be long-term coherent effects because frequency and phase errors are constrained and corrected by the GPS time stamp. In particular, they obtained δ-frequency data from the U.S. East and West power grids (time series data with sample period of 1/2 hr) and noted 4 hr, 12 hr, and 24 hr periodicity in the coherence. (See Figs. 7 and 8.)

After subsequently analyzing all of the E7 power monitor data (PEM-LVEA_V1), Klimenko and Castiglione subsequently found that the 60 Hz line is *not* coherent over a time scale of 180 hours [34]. The coherence fall-off like $1/\sqrt{T}$, where T is the total observation time, consistent with uncorrelated data. (See Fig. 9.)

The same behavior was also observed for the 60 Hz line extracted from the gravitational wave data (LSC-AS_Q) using the DMT LineMonitor [35] . They thus concluded that such a correlation should not shift the mean value of the cross-correlation statistic. However, the power line noise may significantly increase the cross-correlation variance, increasing the value of the upper limit. In order to reduce the variance and hence improve the upper limit, a suppression of the lines by a factor of ~ 100 is required. Using the optimal filtering method described above, the lines are effectively suppressed by the optimal filter, which works as a notch filter at the power line frequencies. This suppression should be adequate for the E7 data. However, an accurate power line removal method is desirable for future runs (like S1), where better detector sensitivities and/or larger data sets are expected.

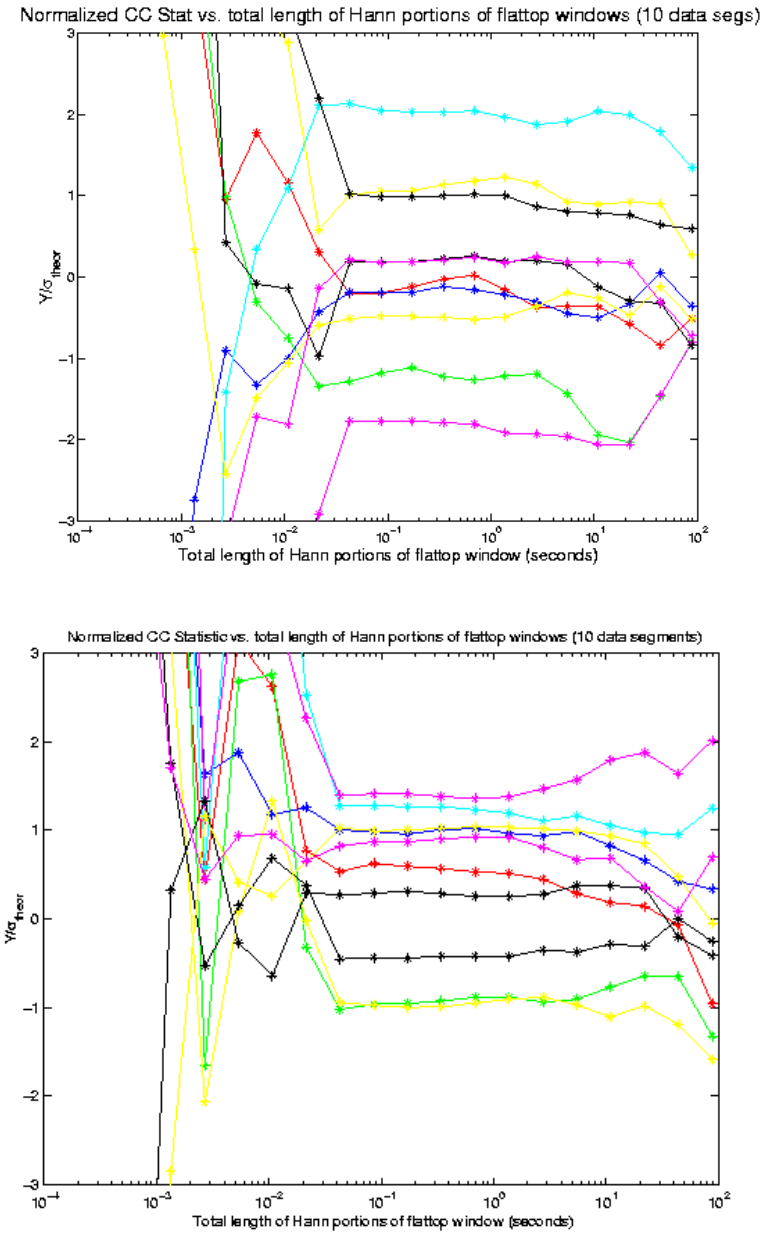

Figure 6: Effects of various spliced Hann windows on the cross-correlation statistic for two 15-minute chunks of playground data. In each plot, the cross-correlation statistics (calculated over the range 40-1000Hz for data downsampled to 2048Hz) for the ten segments, in units of the theoretical standard deviation for a segment, are shown as a function of the length of the Hann portion of the window. Note that for windows with a Hann portion shorter than around 0.1 seconds, the statistic values fluctuate wildly due to the spurious high-frequency correlations arising from a nearly rectangular window. For windows with a Hann portion longer than around 1 second (1.1% of the total segment), the statistics are well-behaved, but have different values, because a different (smaller) set of data is being used in the calculation. For spliced windows in between these two extremes, the cross-correlation statistic depends very little on the details of the window, because the Hann portion is (apparently) long enough to prevent substantial leakage but short enough that almost all of the data are contributing equally in each case.

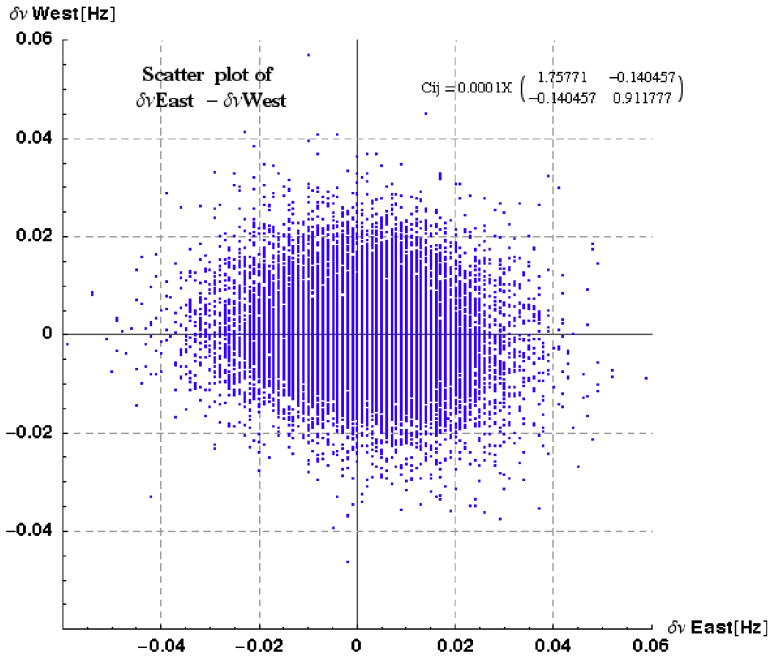

Figure 7: Scatter plot of δ-frequency between the East and West power grids over a period of one year measured at 1/2 hr intervals. The measured cross-correlation is roughly -1.4×10^{-5}.

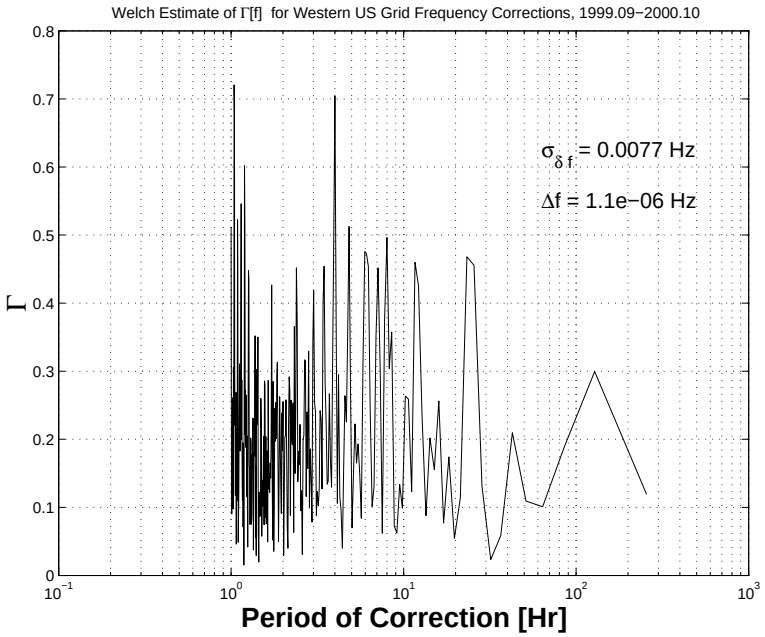

Figure 8: Coherence of cross-correlation as a function of period (i.e., inverse frequency).

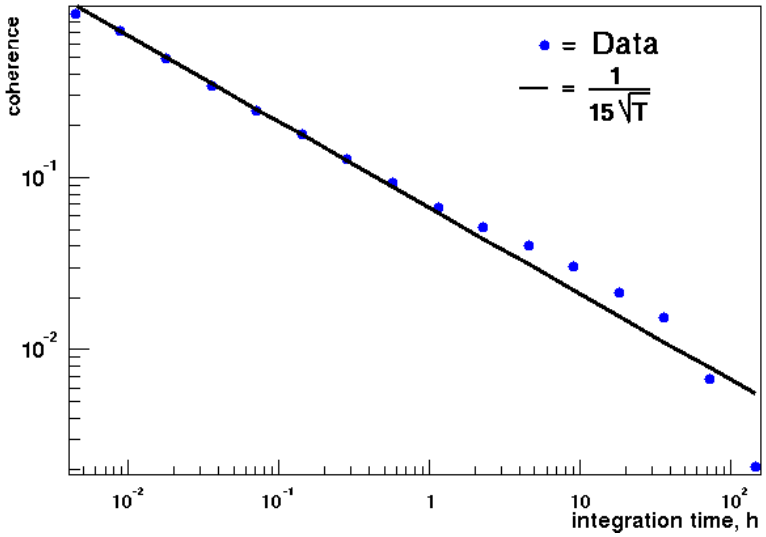

Figure 9: Coherence between power line monitors L0:PEM-LVEA_V1 and H0:PEM-LVEA_V1.

3.6 Analysis of hardware injected signals

Simulated stochastic background signals with $\Omega_{\rm gw}(f) = {\rm const}$ were injected into two 512-second stretches of E7 data (GPS start times 694311500 and 694318750) in the Hanford-4km and Livingston-4km interferometers.[7] Values of the optimally-filtered cross-correlation statistic calculated for these two injections are shown in Fig. 10. For these analyses, we split the data into 10 (45-second) segments, starting at GPS 694311515 and 694318765, respectively, which corresponds to waiting 15 seconds after the injection started before analyzing the data. The measured signal-to-noise ratios for the two injections were -8.93 and -19.3, corresponding to point estimates of $\zeta\Omega_0\, h_{100}^2$ given by -4.23×10^7 and -5.94×10^9, respectively. Note that both point estimates are *negative* indicating that the hardware injections and/or calibrations were such that $\zeta = -1$ for the injections.

Results of time shifting the two data stretches relative to one another are shown in Fig. 11. As expected, the point estimates of $\zeta\Omega_0\, h_{100}^2$ are much smaller for all time-shifts greater than the light travel time (~ 10 msec) between the Hanford and Livingston detectors.

We also plan to hardware-inject simulated stochastic background signals during S1. The injections will have durations of 1024 second with logarithmically varying amplitudes—e.g., corresponding to signal-to-noise ratios of 0.5, 1, 2, 4, 8, and 16.

3.7 Upper limit for E7 data (Hann windowing)

Based on analysis of 205 15-minute chunks of E7 data, each broken into 10 (90-second) segments, we find, as described in Sec. 2.7, a value of $\bar{Y}_Q = -1.6685 \times 10^6\,{\rm sec}$ corresponding to a point estimate $\zeta\hat{\Omega}_0\, h_{100}^2 = -1.8539 \times 10^4$. The standard error is $\hat{\sigma} = 4.0853 \times 10^6\,{\rm sec} = 4.5392 \times 10^4\, T$. Applying the methods of

[7] Note that the stochastic signal was injected into the 4km interferometer at Hanford (H1), while our overall upper limit measurement was done using the Hanford-2km (H2), which turned out to be more sensitive.

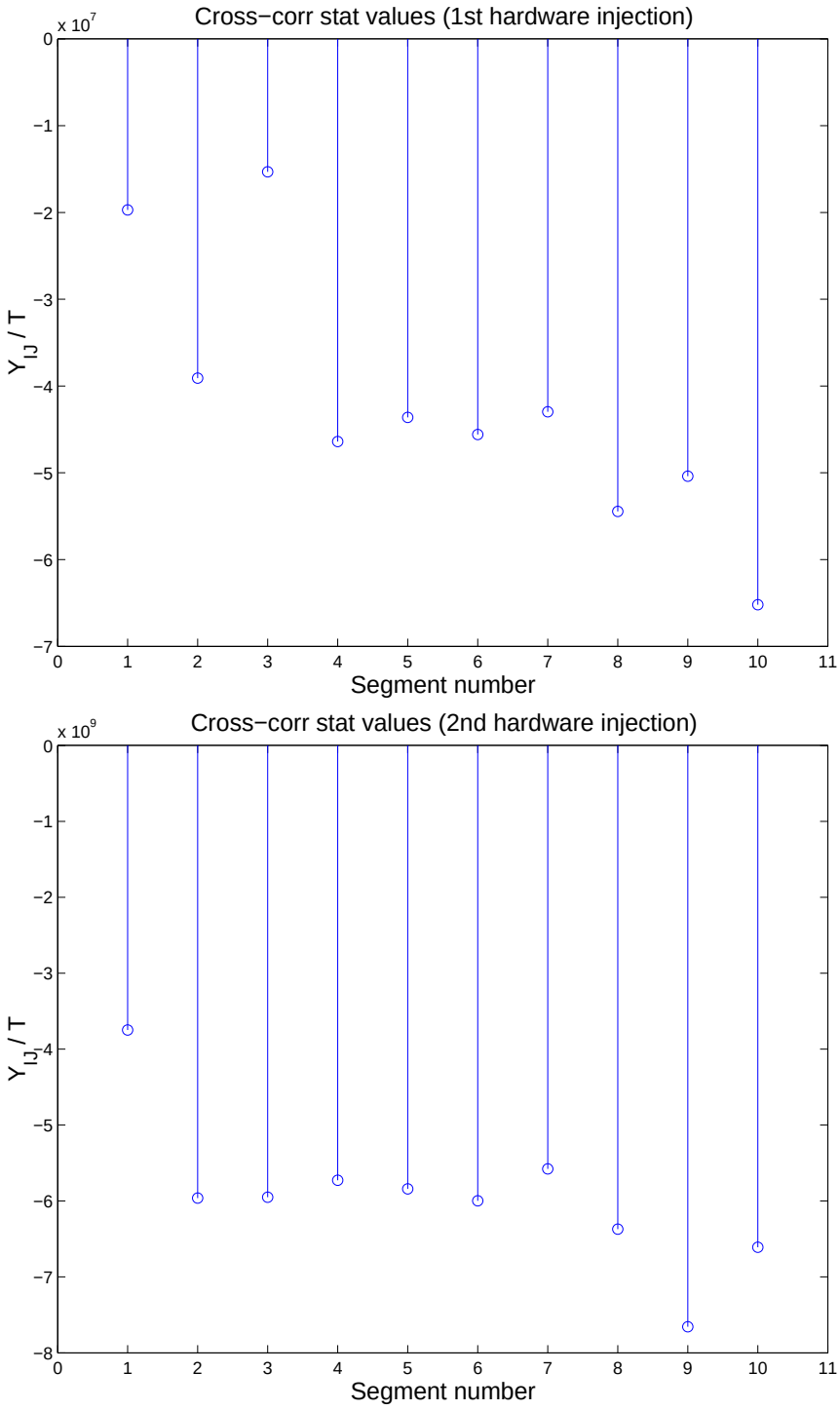

Figure 10: Optimally-filtered cross-correlation statistic values $\{Y_{QIJ}\}$ for the 1st and 2nd hardware injections (GPS start times 694311515 and 694318765). The normalization has been chosen so that the vertical axis represents a point estimate of $\zeta\Omega_0\, h_{100}^2$ for each segment.

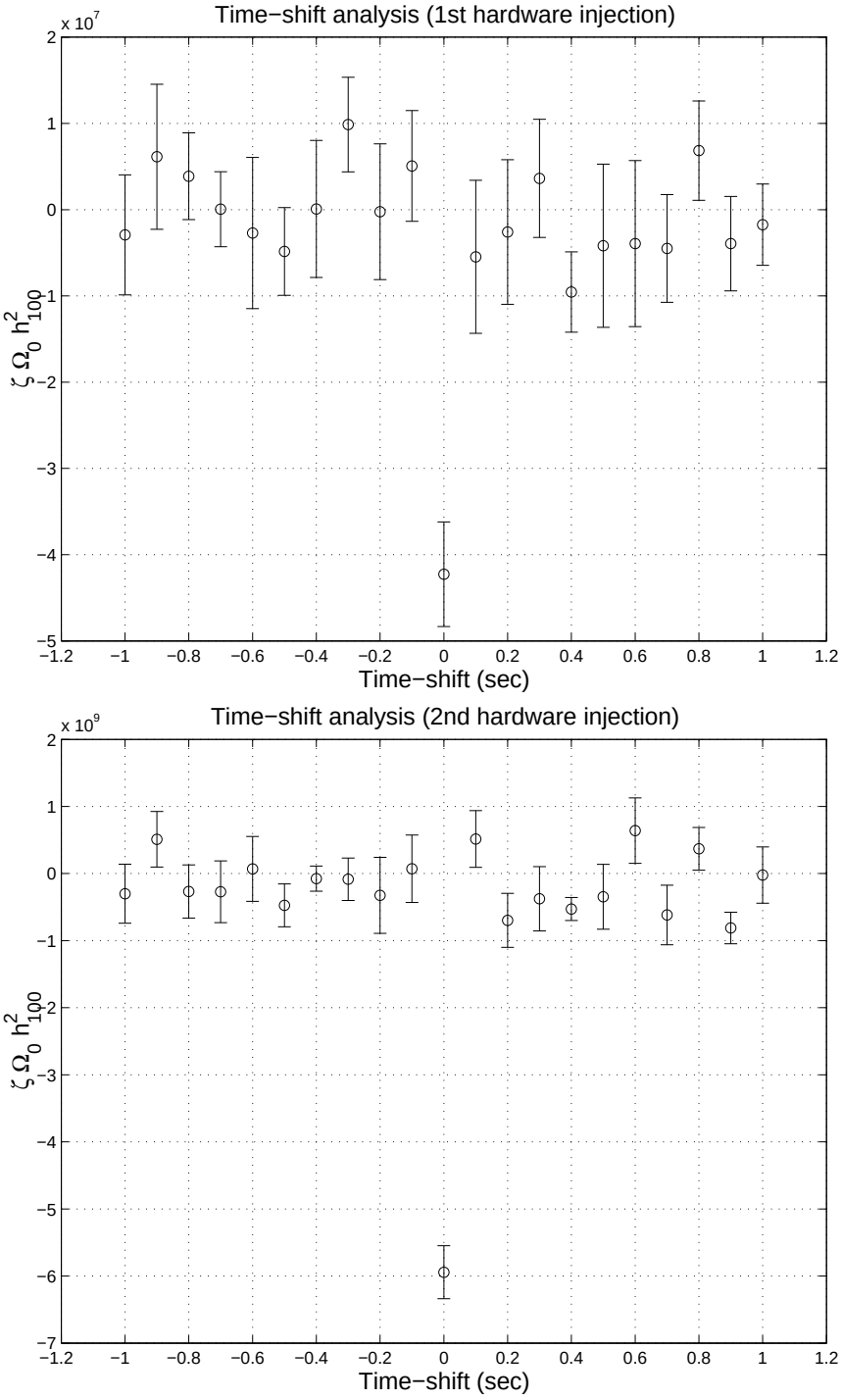

Figure 11: Time-shift analysis for the 1st and 2nd hardware injections (GPS start times 694311515 and 694318765). Plotted are point estimates of $\zeta\Omega_0\, h_{100}^2$ (circles) with 90% confidence level error bars for time-shifts between -1 and 1 second in steps of 0.1 second.

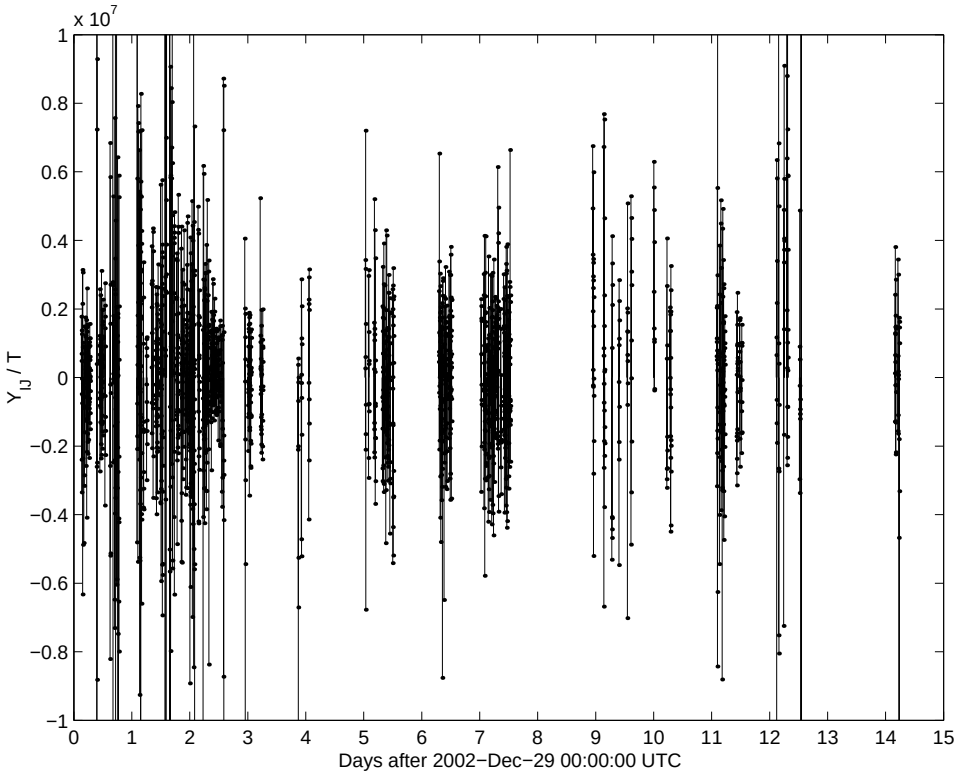

Figure 12: The cross-correlation statistic values $\{Y_{QIJ}\}$ for the 2050 (90-second) Hann-windowed data segments, as a function of segment time. The normalization has been chosen so that the vertical axis represents a point estimate of $\zeta\Omega_0\, h_{100}^2$ for each segment.

Sec. 2.8, we obtain the following 90% confidence level limit:

$$-\,7.6712 \times 10^4 < \zeta\Omega_0\, h_{100}^2 < 3.9633 \times 10^4 \tag{25}$$

this corresponds, given the unknown sign of ζ, to an upper limit of $\Omega_0\, h_{100}^2 < 7.7 \times 10^4$.

The 2050 statistic values $\{Y_{QIJ}\}$ are shown in Fig. 12. Figure 13 shows the 205 15-minutes averages $\{\bar{Y}_{QI}\}$, with 1-sigma errorbars given by the measured standard deviation s_I for the 10 jobs in each chunk. Figure 14 provides a measure of the accuracy of the theoretical σ_I used in weighting the contributions from the different data chunks, by comparing it to the actual measured standard deviation s_I. Note the lack of apparent systematic bias: the ratio is above unity about as often as it is below, which among other things forms an confirmation of the theoretical adjustments to the expected standard deviation which we have made to compensate for the windowing of the data. Figure 15 shows the weighting factors $\{\lambda_I = \sigma_I^{-2}\}$, which also act as a measure of the sensitivity of the experiment over the course of the run. Finally, Fig. 16 provides a measure of the Gaussianity of the underlying data. The quantities

$$x_{IJ} = \frac{Y_{QIJ} - \bar{Y}_Q}{\sigma_I} \tag{26}$$

should, according to our model, be realizations of a Gaussian random variable with zero mean and unit variance. Using all 2050 data segments, we calculate the mean of x_{IJ} as -0.005 and the standard deviation as 1.22. However, the standard deviation in particular is heavily influenced by outlying values coming from

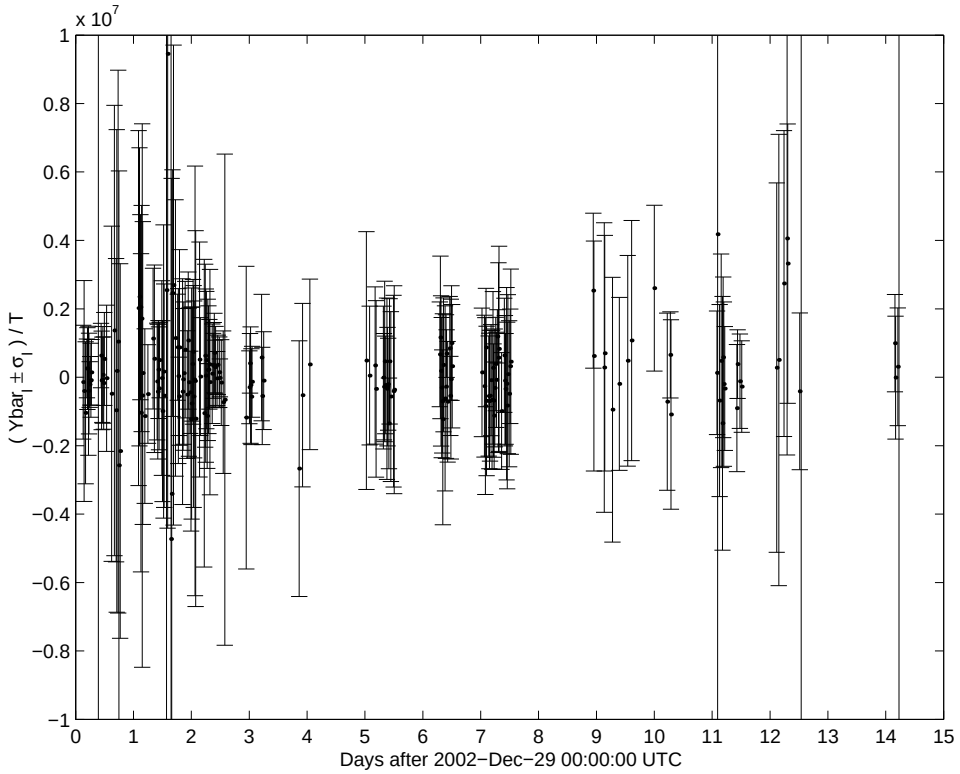

Figure 13: The averages $\{\bar{Y}_{QI}\}$ of the statistic values for the 10 Hann-windowed segments in each of the 205 900-second data chunks, as a function of chunk time. The normalization has been chosen so that the vertical axis represents a point estimate of $\zeta\Omega_0\, h_{100}^2$ for each chunk. The error bars represent the measured standard deviation s_I ($\pm 1\sigma$) of the ten Y_{QIJ} values for each I.

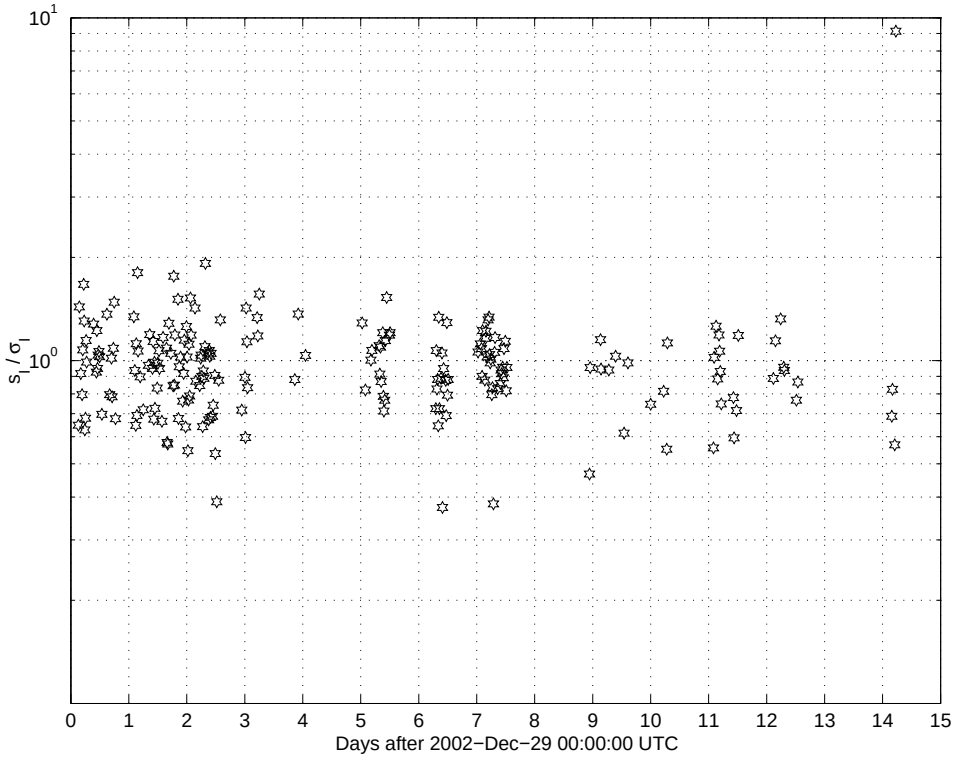

Figure 14: The ratio of the measured standard deviation s_I to its theoretical value σ_I for each of the 205 900-second data chunks, as a function of chunk time. Note that with the exception of the last chunk, for which the actual standard deviation is around 9 times its theoretical value, the ratio is between 0.5 and 2 and shows no systematic trend above or below unity.

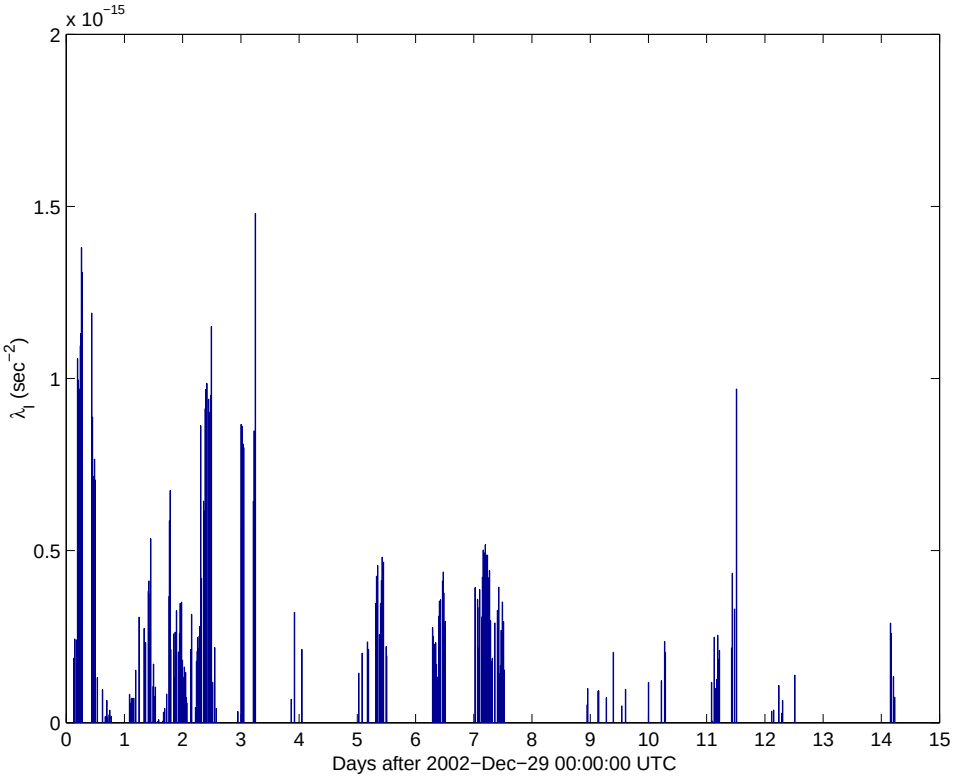

Figure 15: The weighting factor $\lambda_I = \sigma_I^{-2}$ corresponding to each of the 205 900-second data chunks, as a function of chunk time. This is also a measure of the sensitivity of the measurement, as a function of time.

the last data chunk, for which the actual standard deviation was much larger than the theoretical one. In general, we might expect the noisier stretches of data to be less well-behaved. Those stretches contribute little to the weighted averages we used to obtain the upper upper limit, while every data segment contributes equally in the statistics of x_{IJ}. It is therefore instructive to restrict attention to the 1290 data segments corresponding to the 129 most heavily-weighted chunks, which make up 90% of the weighted average. The x_{IJ} values for those have a mean of -0.003 and a standard deviation of 1.02. The higher moments of each distribution are shown graphically in the normal probability plots in Fig. 16.

3.8 Spliced Hann window analysis

We repeated the analysis on the same 205 15-minute chunks of data, this time dividing each into 10 segments using spliced Hann windows as described in Sec. 3.3 (one-second Hann windows with an 89-second-long constant stretch "spliced" into the middle of each). We found a value of $\bar{Y}_Q = -7.3633 \times 10^5\,\mathrm{sec}$ corresponding to a point estimate $\zeta\hat{\Omega}_0\,h_{100}^2 = -8.7285 \times 10^2$. The standard error is $\hat{\sigma} = 2.8371 \times 10^6\,\mathrm{sec} = 3.1524 \times 10^4\,T$. Applying the methods of Sec. 2.8, this could be converted into the following 90% confidence level upper limit:

$$-4.8580 \times 10^4 < \zeta\Omega_0\,h_{100}^2 < 3.2218 \times 10^4 \tag{27}$$

this would correspond, given the unknown sign of ζ, to an upper limit of $\Omega_0\,h_{100}^2 < 4.9 \times 10^4$.

So, using spliced Hann windows would have improved our upper limit on $\Omega_0\,h_{100}^2$ by a factor of 1.57, which is comparable to the theoretical factor of $\sqrt{35/17} \approx 1.43$ arising from the longer effective observing time. In fact, most of the discrepancy arises from the difference in the point estimates; the ratio of the standard errors in the two analyses is 1.44, which is very close to the theoretical value. (It's worth noting that the theoretical standard deviations $\{\sigma_I\}$ were used only to obtain the relative weightings of the different chunks in the overall averages; the standard error is based on the standard deviations $\{s_I\}$ calculated from the individual segment measurements within each chunk.)

4 Future plans

4.1 Line removal

Although there is general agreement that line removal (in particular, for power lines) will reduce the variance of the optimally-filtered cross-correlation statistic, and thus improve the upper limit, we were not able to sufficiently characterize the line removal method currently implemented in the data conditioning API to include it as part of our data analysis pipeline for the E7 upper limit analysis. This method, which *regresses* the gravitational wave channel against selected physical and environmental monitors (PEMs), has excellent suppression for high signal-to-noise ratio lines[36], but more systematic investigations, quantifying the effect of line removal on the values of the cross-correlation statistic, have yet to be performed. In particular, we need to know the level of spurious broad-band cross-correlations (if any) introduced by this method. These investigations are scheduled for S1.

4.2 Monte Carlo simulations

Although code exists within LAL and LALWrapper for simulating stochastic background signals, this functionality was not sufficiently tested within LDAS to be used for the E7 analysis. It is planned for S1.

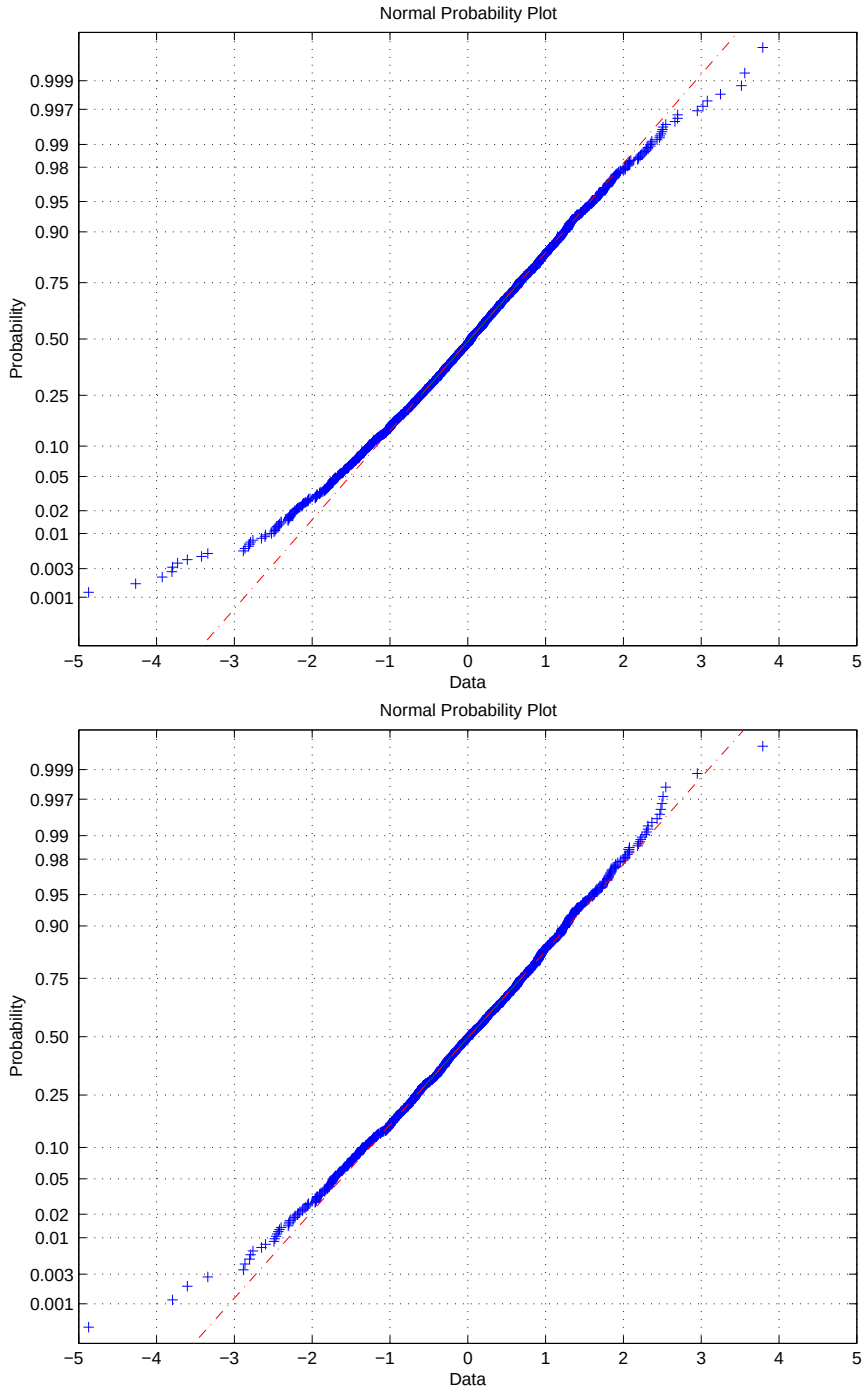

Figure 16: Normal probability plots for $(Y_{QIJ} - \bar{Y}_Q)/\sigma_I$, which should on theoretical grounds be drawn from a Gaussian distribution of zero mean and unit variance. The red line indicates the Gaussian which best fits the data, and curvature away from that is a sign of non-Gaussianity. The upper plot was made using the full set of 2050 data segments, for which the mean is -0.005 and the standard deviation 1.22. (Two points with values of -25 and -19 do not appear on the plot, but contribute to the statistics.) The lower plot was made using only the 129 "quietest" data chunks out of 205, which make up 90% of the contribution to the weighted average, and has a mean of -0.003 and a standard deviation of 1.02.

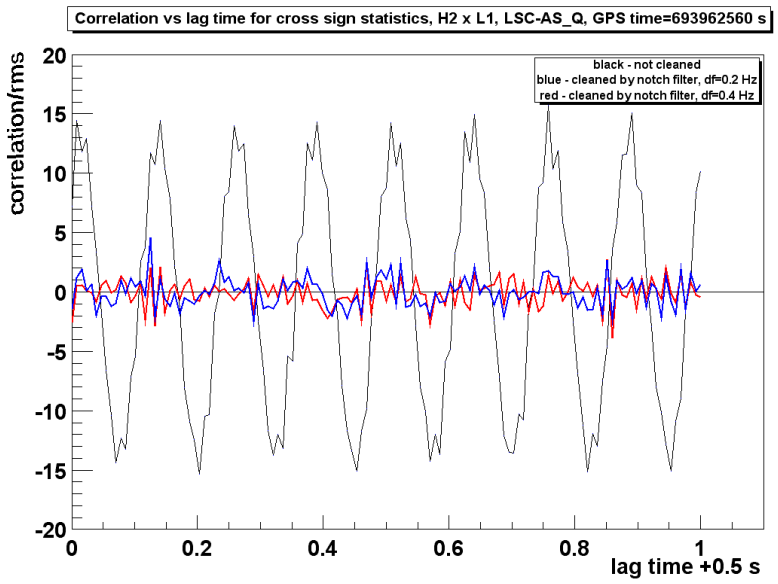

Figure 17: Normalized sign cross-correlation coefficient for gravitational wave data as a function of lag time. Red and blue curves show the cross-correlation coefficient when the power lines are removed.

4.3 Sign correlation test

Klimenko et al. have implemented a complementary method to search for stochastic gravitational wave signals [37]. It is a *robust* cross-correlation technique in the *wavelet domain*, which uses only the sign of the wavelet coefficients. The probability distribution of the correlation coefficients is known for the sign statistic, which allows for a robust calculation of the cross-correlation significance for non-Gaussian data. If the detector noise is Gaussian, the sign correlation method is roughly a factor of two *less efficient* than the optimally-filtered cross-correlation method with the spliced Hannwindow. However, it allows for a useful and independent cross-check when the detector noise is non-Gaussian, which we observed in the E7 data.

Wavelets and the sign transform are implemented in the data conditioning API, and have already been used to analyze the E7 playground data. Figure 17 is a plot of the sign correlation coefficient (normalized by the measured standard deviation) for gravitational wave data with and without power lines removed. Figure 18 is another plot of the normalized sign correlation coefficient, this time for the 2nd hardware injection. In both cases, the normalized cross-correlation coefficients are plotted versus lag time, which is an important signature of the correlation process.

Note that Fig. 17 also shows that the sign correlation method cannot be used without accurate line removal. Although for long observation times the line noise should not shift the mean value of the cross-correlation, it can significantly increase (~ 100 times) the variance of the cross-correlation (see Sec. 4.4). We plan to use the sign correlation method on E7 and S1 data as soon as the line removal algorithm is fully implemented and characterized.

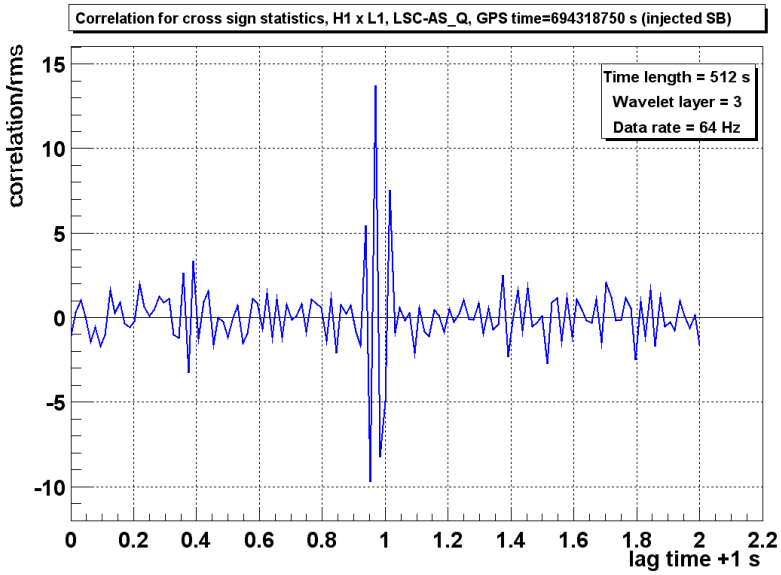

Figure 18: Normalized sign cross-correlation coefficient for hardware injected stochastic signal as a function of lag time.

4.4 Cross-correlated noise study

In addition to the power mains, there could be other sources of correlated noise coming from seismic events, lightning strikes, etc. Unlike power lines, these sources may introduce a *broad-band* correlated noise, which may not be removed or suppressed by the optimal filter. A more general study and estimation of correlated environmental noise is thus needed, and should be implemented for S1.

One way to do it is to use the sign correlation test described in Sec. 4.3. If correlated noise is present in the data, the autocorrelation function of the product of the sign statistics (for each wavelet layer) will differ from a delta function, and have a non-trivial (i.e., $\neq 1$) variance ratio $\nu(T)$ as a function of the integration time T [37]. For large integration times, $\nu(T)$ is proportional to the correlation time of the process, $T_c = \nu(T)\Delta t/2$, where Δt is the sampling period.

Figure 19 shows how this method works on data contaminated with correlated noise from power lines. The figure shows $\nu(T)$ for three different E7 data segments. Lower curves show $\nu(T)$ when the lines are removed. They are close to unity, as expected for uncorrelated data.

4.5 ALLEGRO-LLO correlation

As described in [38], the ALLEGRO [39] resonant bar detector took data in three different orientations during E7: (i) It operated for approximately two days in an orientation of 48° West of North, which was close to the default orientation used to align ALLEGRO with the other bars in the IGEC collaboration. (ii) It operated for approximately two days in an orientation parallel to the Y arm of LIGO Livingston, 18° West of North. (iii) It operated for approximately two days in an orientation halfway between the X and Y arms of LLO (63° West of North). The overlap reduction functions for these three alignments, along with a fourth alignment parallel to the LLO X arm, are shown in figure 20. As described in [40], combining data from

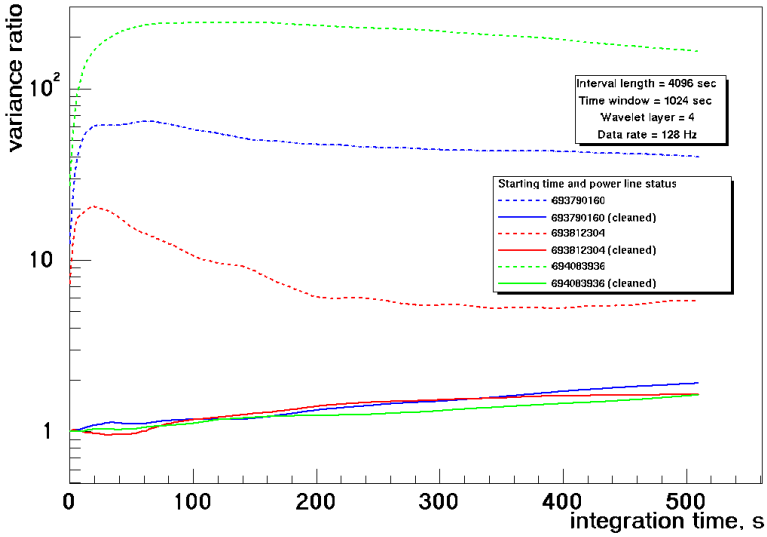

Figure 19: Variance ratio $\nu(T)$ as a function of integration time T for three data segments with correlated noise from power lines (3 upper curves) and with power lines removed (3 lower curves).

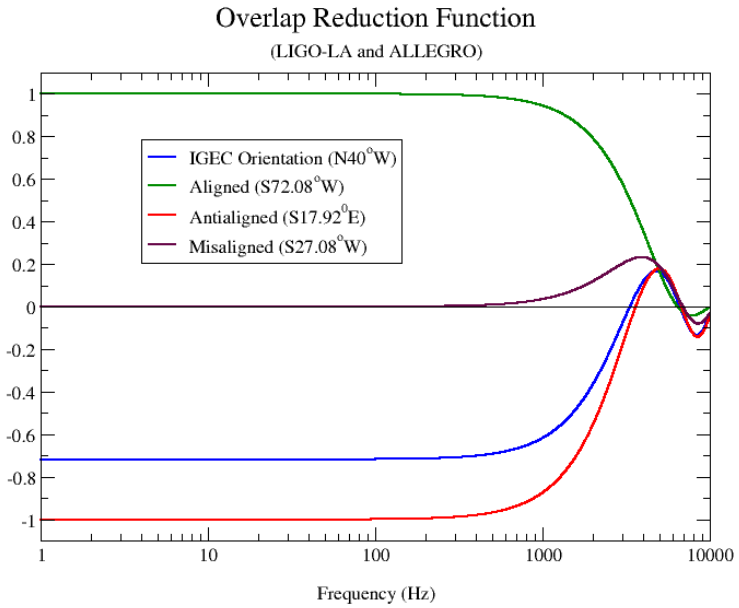

Figure 20: Overlap reduction function between LIGO Livingston and ALLEGRO in four alignments.

the “aligned” and “anti-aligned” configurations, which respond to gravitational waves in the opposite way, would allow one to cancel out the effects of orientation-independent correlated noise in obtaining an estimate of the stochastic background strength. Additionally, operating in the null or “misaligned” configuration allows one to effectively point the ALLEGRO-LLO detector pair “off-source”, obtaining an estimate of cross-correlated environmental noise component in a configuration insensitive to stochastic gravitational waves.

Since ALLEGRO is not scheduled to take data during S1, we will be using LDAS to analyze the coincident ALLEGRO-LLO data from E7. This will produce an upper limit on $\Omega_{\mathrm{gw}}(f)$ in a different frequency range from the LLO-LHO analysis (around 900Hz as opposed to 40-215 Hz) and will test the mechanics of the analysis in anticipation of more sophisticated investigations during S2 and beyond.

4.6 GEO-LIGO correlation

The GEO-600 [41] interferometer operated in coincidence with LIGO during E7, but no correlation measurements were performed. Richard Ingley (at Ph.D. student at Birmingham) has been building a data analysis pipeline within GEO for stochastic searches, which should to be ready in time to analyze S1 data. Although a GEO-LIGO correlation will not improve the upper limit by much due to the small overlap between the GEO and LIGO interferometers, it should provide insight about inter-continental cross-correlated environmental noise.

References

[1] http://www.ligo-wa.caltech.edu/

[2] http://www.ligo-la.caltech.edu/

[3] http://blue.ligo-wa.caltech.edu/engrun/E7/

[4] http://www.lsc-group.phys.uwm.edu/iulgroup/

[5] http://www.ligo.caltech.edu/∼ajw/bursts/bursts.html

[6] http://www.lsc-group.phys.uwm.edu/pulgroup/

[7] http://www.phys.utb.edu/stochastic/

[8] http://www.ligo.org/

[9] http://www.ligo.caltech.edu/

[10] P.F. Michelson, *Mon. Not. Roy. Astron. Soc.* **227**, 933, 1987.

[11] N. Christensen, "Measuring the stochastic gravitational radiation background with laser interferometric antennas," *Phys. Rev.* **D46**, pp. 5250-5266, 1992.

[12] É.É. Flanagan, "The Sensitivity of the laser interferometer gravitational wave observatory (LIGO) to a stochastic background, and its dependence on the detector orientations," *Phys. Rev.* **D48**, pp. 2389-2407, 1993.

[13] B. Allen, "The stochastic gravity-wave background: sources and detection," in *Proceedings of the Les Houches School on Astrophysical Sources of Gravitational Waves*, eds. J.A. Marck and J.P. Lasota, Cambridge, pp. 373-417, 1995.

[14] B. Allen and J.D. Romano, "Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities," *Phys. Rev.* **D59**, pp. 102001-102041, 1999.

[15] M. Maggiore, "Gravitational wave experiments and early universe cosmology," *Phys. Rept.* 331, pp. 283-367, 2000.

[16] D.G. Blair and L. Ju, "A cosmological background of gravitational waves produced by supernovae in the early universe," *Mon. Not. R. Astron. Soc.*, 1996.

[17] D.G. Blair, R. Burman, L. Ju, S. Woodings, M. Mulder, M.G. Zadnik, "The supernova cosmological background of gravitational waves," Preprint, University of Western Australia, 1997.

[18] V. Ferrari, "Astrophysical sources of gravitational waves," in *Proceedings of the XII Italian conference on GR and Gravitational Physics*, World Scientific, 1997.

[19] B. Allen and A.C. Ottewill, "Detection of anisotropies in the gravitational wave stochastic background," *Phys. Rev.* **D56**, pp. 545-563, 1997.

[20] S. Drasco and É.É. Flanagan, "Detecting a non-Gaussian Stochastic Background of Gravitational Radiation," in *Proceedings of the Ninth Marcel Grossmann Meeting on General Relativity*, edited by V. G. Gurzadyan, R. T. Jantzen and R. Ruffini, World Scientific, 2002.

[21] G.F. Smoot et al., "Structure in the COBE DMR first year maps," *Astrophys. J.* **396**, pp. L1-L5, 1992.

[22] E.L. Wright et al., "Interpretation of the CMB anisotropy detected by the COBE DMR," *Astrophys. J.* **396**, pp. L13-L18, 1992.

[23] C.L. Bennett et al., "Cosmic temperature fluctuations from two years of COBE DMR observations," *Astrophys. J.* **436**, pp. 423-442, 1994.

[24] C.L. Bennett et al., "Four year COBE DMR cosmic microwave background observations: Maps and basic results," *Astrophys. J.* **464**, pp. L1-L4, 1996.

[25] V. Kaspi, J. Taylor, and M. Ryba, *Astrophys. J.* **428**, 713, 1994.

[26] E.W. Kolb and M. Turner, *The Early Universe*, Frontiers in Physics, Addison Wesley, 1990.

[27] K. Compton, D. Nicholson, and B.F.S̃chutz, in *Proceedings of the Seventh Marcel Grossman Meeting on General Relativity,* World Scientific, 1078, 1994.

[28] P. Astone et al., "Upper limit for a gravitational wave stochastic background with the Explorer and Nautilus resonant detectors," *Phys. Lett.* **B385**, pp. 421-424, 1996.

[29] P. Astone et al., *Astronomy and Astrophysics,* **351**, pp. 811-814, 1999.

[30] P. Astone, V. Ferrari, M. Maggiore, and J.D. Romano, "Stochastic background of gravitational waves," *Int. J. Mod. Phys.* **D9**, pp. 361-368, 2000.

[31] G. J. Feldman and R. D. Cousins, "A Unified Approach to the Classical Statistical Analysis of Small Signals," *Phys. Rev.* **D57**, pp. 3873-3889, 1998.

[32] W. R. Johnston, J. D. Romano and J. T. Whelan, in preparation, 2002.

[33] A. Lazzarini, R. Schofield, and A. Viceré, LIGO Technical Document, `http://www.ligo.caltech.edu/docs/T/T010101-01.pdf`, 2001.

[34] S. Klimenko and J. Castiglione, "Coherence of power lines," LIGO Technical Document, LIGO-T020109-00-Z, 2002.

[35] S. Klimenko, "Line Monitor," LIGO Technical Document, LIGO-T020108-00-Z, 2002.

[36] A. Searle, private communication, 2002.

[37] S. Klimenko, G. Mitselmakher, A. Sazonov "A cross-correlation technique in wavelet domain for detection of stochastic gravitational waves," e-Print Archive: gr-qc/0208007, 2002.

[38] W. O. Hamilton et al., "Resonant detectors and interferometers can work together" Proceedings of the SPIE conference, Hawaii, 2002.

[39] `http://gravity.phys.lsu.edu/allegro/`

[40] L.S. Finn and A. Lazzarini, "Modulating the experimental signature of a stochastic gravitational wave background," *Phys. Rev.* **D64**, pp. 082002-082021, 2001.

[41] `http://www.geo600.uni-hannover.de/`