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

Document Type **LIGO-T010017-00-Z** 23 February 2001

Proposal to set an upper limit on stochastic sources using LIGO engineering data

Bruce Allen, Warren Anderson, Sukanta Bose, Nelson Christensen, Ed Daw, Mario Diaz, Ronald Drever, Sam Finn, Peter Fritschel, Joe Giaime, Bill Hamilton, Warren Johnson, Erik Katsavounidis, Michael Landry, Albert Lazzarini, Martin McHugh, Adrian Ottewill, Jamie Rollins, Joe Romano, Siong Heng, Bernard Schutz, Alberto Vecchio, Harry Ward, Rainer Weiss, John Whelan, Bernard Whiting, Graham Woan

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: prop.tex— Processed with LaTeX on 2001/02/23

1 Character of proposed activity

We propose to use coincident engineering data from the LIGO Hanford and LIGO Livingston Observatories (LHO and LLO) to set an upper limit on the strength of a stochastic background of gravitational radiation. This is primarily a data analysis activity, but it will also require the development of LIGO Data Analysis System (LDAS) and LIGO/LSC Algorithm Library (LAL) software and LDAS/LAL system integration to satisfactorily process the data. If possible, we might extend the upper limit analysis to include ALLEGRO and/or GEO data.

2 Scientific and technical rationale

Although the LIGO interferometers will not be at design sensitivity and will operate in coincidence for only a week, the engineering data will be the most sensitive ever taken and thus merits analysis and publication. The upper limit analysis for the engineering data will drive hardware and software development in preparation for the analysis of real science data beginning in Fall 2002. It will also test the LDAS data analysis pipeline and the integration of the LAL stochastic background routines within LDAS.

Since any source of correlated environmental noise in separated detectors mimics the correlation arising from a true stochastic background signal, the stochastic sources upper limit analysis will also push development of intra-site (i.e., same site) and inter-site (i.e., different site) correlation techniques and add to our understanding of how the physical environment affects an interferometer. Moreover, the possibility of using ALLEGRO and/or GEO data may provide further understanding of multi-instrument searches and inter-site correlations.

3 Technical Approach

To set an upper limit on the strength of a stochastic background of gravitational radiation, we propose to do the following:

1. Characterize sources of cross-correlated environmental noise, concentrating on (in probable order of importance): (i) power mains, (ii) magnetic field fluctuations (e.g., thunderstorms), and (iii) seismic disturbances (possibly up-converted). We will need to distinguish between steady-state and transient sources of environmental correlations. The transient contamination should not present a major problem for a stochastic source analysis. The steady-state contamination, however, needs to be identified and properly accounted for.

2. Remove spectral lines at the individual detectors (e.g., power mains lines and violin modes) and other identified sources of cross-correlated environmental noise. This will be accomplished by regression, adaptive filtering or other techniques (e.g., notching). We will need to evaluate the different methods, comparing e.g., the loss in signal-to-noise ratio (SNR) due to reduced bandwidth versus incomplete line removal, etc.

3. Calculate the value of the optimally-filtered cross-correlation statistic for each ~ 10 second stretch of data from the LHO-LLO pair, down-sampled to 1024 Hz (Nyquist = 512 Hz) and high-pass filtered (> 50 Hz) to remove seismic noise. We will then average these values (weighted by the inverse of the

variance of the noise in the detectors) over the duration of the engineering run, yielding an estimate x_0 of the cross-correlated gravitational wave signal strength.

4. Use Monte Carlo simulations and a frequentist (or Bayesian) statistical analysis to turn the measurement x_0 into a 90% confidence level upper limit or two-sided confidence interval on $\Omega_{\rm gw}(f)$ in the frequency band $50 \text{ Hz} \leq f \leq 512 \text{ Hz}$.

3.1 Details

Cross-correlated noise: Diagnostics of cross-correlated environmental noise will form a major component of our analysis. This involves examining the spectral nature of the cross-correlations, possibly including time-frequency characterizations of the cross-spectra. Such a study should enable us to identify sources of inter-site correlations. It is expected that the largest such correlation will occur at the power line frequency and its harmonics, and that these signal components would overwhelm the cross-correlation statistic if not dealt with. Thus a significant effort will be made in testing the various line-removal techniques that are currently being developed.

Cross-correlation statistic: We will use the standard optimally-filtered cross-correlation statistic to search for a stochastic background of gravitational radiation, with filter function chosen to maximize the SNR for uncorrelated Gaussian detector noise and a Zel'dovich spectrum for the stochastic signal (i.e., constant fractional energy-density per unit logarithmic frequency interval: $\Omega_{\rm gw}(f) = \Omega_0 = {\rm const}$).

Monte Carlo simulations: Although one expects the measured values of the optimally-filtered cross-correlation statistic to be Gaussian distributed, we still need to perform Monte Carlo simulations to determine the statistical distribution of the cross-correlation statistic in the absence and presence of a simulated stochastic gravitational wave signal. The probability distributions are needed, for example, to construct the confidence belts used to convert the measurement x_0 to a 90% upper limit on $\Omega_{\rm gw}(f)$ (or two-sided confidence interval) as mentioned above. We can simulate the detector noise using the engineering data from LHO and LLO shifted in time relative to one another by any amount greater than the light travel time between the two sites (approximately 10 msec). This will give us two data streams that have no cross-correlated gravitational wave component, but contain (possible) persistent or long duration cross-correlated environmental noise components. After processing the data, we can histogram the measured values of cross-correlation statistic, and calculate various moments (e.g., sample mean, variance, skewness, kurtosis). These moments can then be compared with similar quantities obtained from the Monte Carlo simulations. This should provide some indication of any systematic errors that might bias the calculation of the upper limit.

Other possibilities: Depending on how much progress is made before the engineering run, we may want to extend the above upper limit analysis to include ALLEGRO and/or GEO data. In fact, one might actually be able to set a better upper limit on $\Omega_{\rm gw}(f)$ from ALLEGRO-LLO correlation, given the closeness of the LLO interferometer and the ALLEGRO bar. Such an analysis would require that ALLEGRO be operational during the engineering run, and that we develop the appropriate software (and hardware) tools to process the bar data in a straightforward fashion (e.g., heterodyning), given the different pass-band of the ALLEGRO bar detector.

In addition, the GEO-600 interferometer might also be operational during the engineering run. Although the large separation between the LIGO and GEO interferometers implies that a GEO-LIGO correlation will not improve the stochastic background upper limit that much, we can still increase our understanding of cross-correlated environmental noise by doing such a correlation.

4 Deliverables

Software: LAL modules for stochastic background searches will be written and tested, as will software for the LDAS Data Conditioning API. This code will be able to: (i) Calculate the values of the optimally-filtered cross-correlation statistic for each ~ 10 second stretch of IFO data. (ii) Estimate (standard) cross-correlations and cross-spectral densities between IFO-IFO, IFO-PEM, and PEM-PEM channels at each site and between sites. (iii) Remove spectral lines (e.g., power mains and violin modes) using regression, adaptive filtering, or other techniques. (iv) Remove other dominant environmental disturbances (e.g., magnetic field fluctuations) by regression with appropriate PEM channels. Code will also be written (as part of the LDAS Wrapper API) to integrate the LAL stochastic background modules into LDAS.

In addition, if we decide to do an ALLEGRO-LLO correlation, we will also develop code (within LAL) to calculate the overlap reduction function between interferometers and bars, and to work with heterodyned data.

All LAL modules will be documented as specified in the "LIGO Data Analysis System Numerical Algorithms Library Specification and Style Guide" (LIGO-T990030-07). All LDAS software will be documented as specified in "LDAS System Software Specification for C, C++, and Java" (LIGO-T970211).

Reduced data sets: Trend data containing the values of the IFO-IFO, IFO-PEM, and PEM-PEM correlations in 1 Hz bins (from 50 Hz to 512 Hz) for a selected number of PEMs (e.g., magnetometers, seismometers, accelerometers, RF monitors) will be written to the LIGO data archive, for each ~ 10 second stretch of data. (This corresponds to roughly 10 Gbytes of data for a 5 day engineering run.) Values of the optimally-filtered cross-correlation statistic for each ~ 10 second stretch of IFO data will also be written to the archive.

Publications: A number of papers are expected to result from the upper limit analysis: (i) A LIGO/LDAS document describing in detail the data analysis pipeline used in our upper limit analysis. (ii) An *Astrophysical Journal* or *Physical Review Letter* stating an upper limit on $\Omega_{\rm gw}(f)$ and discussing its implication on stochastic source models, etc. (iii) A *Physical Review D* article detailing the upper limit analysis, which was described only briefly in the *Astrophysical Journal* or *Physical Review Letter*. (iv) A geophysical journal article describing in detail the results of the cross-correlated environmental noise analysis (e.g., power mains, magnetic disturbances, seismic noise, ...). (v) Separate *Physical Review D* articles describing the LLO-ALLEGRO and/or GEO-LIGO upper limit analyses (if these analyses are actually performed).

All papers submitted for publication will follow the procedures specified in the *LSC Publications and Presentations Policy*.

5 Required resources

Since the engineering run will take place in late September 2001, all requisite hardware and software tools must be completed and tested by then.

Hardware: The analysis will be done off-line, off-site using a standard workstation running LDAS on an Intel Linux platform. AIT-2 reduced data tapes will need to be written for the subsequent off-site processing. Suitable local data storage, access bandwidth, and site licensing of IBM's DB2, and any other commercial products (e.g., Sun servers for DB2) will need to be provided by the institution where the analysis is performed.

Several PEMs (e.g., magnetometers, seismometers, accelerometers, RF monitors, and possibly a few others) and diagnostic hardware will need to be installed and validated at the sites.

Software: Several components of LDAS and LAL must be working and integrated before we can perform the upper limit analysis. These include the software modules listed in Section 4, plus the ability to: (i) extract reduced data from the LIGO archive for off-line, off-site processing, and (ii) write trend data to the LIGO data archive and/or metadata database.

Data: We will need access to data from the LIGO archive. The reduced data sets that will be processed off-line, off-site should contain: (i) IFO time-series data for the individual detectors. (ii) PEM time-series data from a selected number of PEMs (e.g., magnetometers, seismometers, accelerometers, RF monitors) for estimating the amount of environmental noise. (iii) Swept-sine frequency response functions for the individual detectors, which provide information about whitening and anti-aliasing filters. (iv) Diagnostic channels indicating the interferometer state and quality of the data—e.g., whether the interferometers are simultaneously in lock, the presence (or absence) of anomalous instrument behavior, etc. (v) Estimates of the noise power spectra of the individual detectors.

Personnel: We expect that the proposed activities will require about 10 FTEs for the duration of the project. We also anticipate needing about 0.5 FTE months of time from each of two LDAS personnel during the verification and validation of all code that will run within the LDAS environment.

Other: We will also need the ability to perform real-time calibration to check the data analysis pipeline—e.g., inject (via hardware) correlated noise at each site, and then check that we can identify and extract it. If we decide to do an ALLEGRO-LLO analysis, we will also want to be able to write ALLEGRO data as part of LLO frames.

6 Work plan

Ten teams of volunteers will work on various tasks needed for the upper limit analysis. A single individual has been designated to coordinate the efforts of each team. The stochastic sources upper limit group will have monthly teleconferences to report on progress toward the milestones. The group has a web page at http://feynman.utb.edu/∼joe/research/stochastic/upperlimits/ and a mailing list stochastic@feynman.utb.edu which is archived to the web page. Details about the individual tasks (including timelines, work plans, milestones, and volunteers) can be found on the group web page.