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

Document Type	**LIGO-T010025-00-Z**	February 12, 2001

Determine upper limits on event rates for inspiralling compact binaries with LIGO engineering data

Bruce Allen, Sukanta Bose, David Churches, **Patrick Brady**, Duncan Brown, Jordan Camp, Nelsen Christensen, Jolien Creighton, Teviet Creighton, S.V. Dhurander, **Gabriela Gonzalez**, Andri M. Gretarsson, Gregg Harry, Vicky Kalogera, Joe Kovalik, Nergis Mavalvala, Adrian Ottewill, Ben Owen, Tom Prince, David Reitze, Anthony Rizzi, B.S. Sathyaprakash, Peter Shawhan, Julien Sylvestre, Linqing Wen, Alan Wiseman.

Distribution of this draft:

Inspiral Upper Limit Group, LSC Spokeperson,
LIGO Laboratory Director, Software Coordinator,
Upper Limit Group Chairs

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,v 1.8 2001/02/13 01:41:12 patrick Exp patrick — Processed with LaTeX on 2001/03/01

1 Scientific problem addressed

Coalescing compact binaries are among the best understood sources of gravitational waves in the frequency band of LIGO and other earth based detectors. We propose to analyze engineering data, acquired using the 2km interferometer at Hanford and the 4km interferometer at Livingston, to determine upper limits on the event rates of inspiralling compact binary systems.

Our target population will be non-spinning, compact-binary systems with the mass of each element in the range $1M_\odot \leq M \leq 20\,M_\odot$. Specifically, we will develop techniques to place upper limits on inspiral rates for binary populations with accurately modeled waveforms. These populations include, but are not limited to, binary neutron star systems. We note that the astrophysical relevance of the upper limits to binary black-hole systems is questionable since post-Newtonian waveforms are inaccurate at frequencies above ~ 50 Hz$(20M_\odot/M_{\rm tot})$. Nevertheless, such binary black-holes are deemed the most likely source for LIGO-I making the development of data analysis techniques urgent and relevant.

We anticipate learning a great deal about the data and the instruments over the course of this project. While this is secondary to developing and testing a pipeline for the analysis, it might be considered scientifically more interesting since the astrophysical limits will likely be several orders of magnitude higher than expected rates.

2 Technical Approach

The search for inspiral events will begin with (i) time series from the two interferometers, (ii) calibration information for each interferometer and (iii) channel(s) to characterize the quality of the gravitational-wave data.

In addition, the search will use servo loop signals, PEM channels, etc, to discriminate good/bad data and veto events observed in the gravitational wave channels. Over 1000 different channels are recorded for each interferometer. As our work plan develops and experience is gained, we will establish the most important channels for this search type.

2.1 Data flow pipeline

1. **On-site, real-time analysis**: The data from each interferometer will be searched for a restricted set of binary inspiral signals at each site in real time. This stage of the search will be executed on the LDAS Beowulf clusters at the two sites. The search codes will be built to run within LDAS as dynamically loaded shared object libraries; the structure of the search codes will follow the specification provided in the wrapperAPI baseline requirements. Events generated by the search codes will be inserted into the database for later analysis. This will be considered the first cut through the data. Some use will be made of vetoes based both on the nature of the inspiral signal itself, and on the auxiliary channels.

   In addition, we expect the DMT to provide triggers which will be inserted into the database. A subset of PEM channels may also be analyzed in real-time to detect noise bursts and these events injected into the data base

2. **Off-line analysis**: The on-site, real time analysis will be designed to test data flow pipelines which might be used during the LIGO-I science run. The majority of the analysis will be done off-line and may require repeating various steps multiple times.

(a) *Filtering of GW and PEM channels*: The data from each interferometer will be searched for the remaining binary inspiral signals. Events generated by the search codes will be inserted into the database. This will be considered the first cut through the data. Some use will be made of vetoes based both on the nature of the inspiral signal itself. For high mass binaries, we may also search for evidence of merger and black-hole ringdown in the data stream.

(b) *Correlation of inspiral and PEM events*: Auxiliary channels, identified to provide vetoes, will be searched for burst events. When such a noise burst is coincident with an event identified by the inspiral search code, that inspiral event will be vetoed if it could be caused by the noise registered in the PEM channel.

(c) *Coincidence between interferometers*: After vetoes have been applied to the inspiral event lists from each interferometer, events coincident within the $\sim$ 10ms time window and by signal parameters will be retained for combined analysis. All other events will be discarded. The different response of the two interferometers may provide an additional veto.

(d) *Combined analysis using both data streams*: The segments of data which produces the list of events retained at this stage will be reanalyzed using coherent multi-detector methods. This will provide the final cut in the analysis pipeline. Events below a suitable threshold will be discarded.

3. **Simulation to determine efficiency**: Independent of the statistical method used to place an upper limit on the rate of inspirals, it will be important to determine the efficiency of the pipeline to detect the target population of sources. This will be achieved by re-analyzing the data with simulated signals injected through the same pipeline discussed above. The fraction of the population detectable by the pipeline is paramount to determining an upper limit.

4. **Statistical analysis**: Several statistical techniques exist to determine the upper limit on inspiral events. All require the same type of output, but each makes more or less use of the resulting output. A method will be chosen to determine the best upper limit without precluding detection of a signal.

2.2 Computational and analysis tools

Cataloging and searching: Simulated and astrophysically plausible events will be injected into the LIGO relational database tables. These cataloged events will be the outputs of astrophysical event filters (e.g. output of inspiral chirp filters that satisfy certain threshold criteria). The outputs will be generated using codes under development by the LSC. Two independent inspiral codes should be available: (i) traditional template based code and (ii) an FCT based code. Both codes will be tested in advance of the data run, but it is likely that only one of them will be used during the real-time portion of the analysis.

Instrumental triggers will also be injected into the LIGO relational database tables. These cataloged events will be outputs of detector characterization methods (e.g. identifying transients in some of the PEM channels). Some amount of experimentation will be needed to determine an adequate set of channels to examine. We anticipate that the E2 data will be very useful for this purpose.

The LIGO relational database will be queried (using the LDAS software) to determine astrophysical and instrumental correlations. During this stage of the search, different queries will be carried out to determine a list of events consistent with binary inspiral. For example, these events should be seen in the GW channel *only*, satisfy some form of χ^2 veto test, be observed with same physical parameters (masses, etc) in each interferometer, etc. The tools for database mining already exist, however this phase of the analysis will benefit greatly from quick-look software which can easily link to the database. Such tools will be developed (under MATLAB for example).

It may be possible to use the LIGO E2 data to construct queries that reduce the false alarm rate without reducing the correct detection rate too much. This can be determined by Monte-Carlo injection of simulated signals into the data stream.

Finally, some subset of these candidate events should be followed up using a coherent two interferometer search technique. That is, data from both interferometers should be combined in the filtering process. This should enhance the detection efficiency.

Some or all of the above steps may be repeated many times in order to tune and augment the search codes before a final off-line analysis of the data.

Simulation: The injection of simulated signals into the GW channels will play an important role in the statistical interpretation of the results from the database queries. This is an off-site analysis task which should be undertaken using the same analysis pipeline, but using data containing the simulated signals. Much of the software needed for injecting signals is already under development within LAL; the rest will be developed and tested as part of this project. This analysis will be carried out off-site, probably at Caltech or other suitable LSC facilities.

Statistical Analysis: The statistical methods to be used in this analysis are reasonably well understood. To accurately determine the statistical significance of the search output, however, an accurate log must be kept of all analysis performed on the data over the course of the project. A format for this log and a mechanism to maintain and make use of it will be developed as part of this project.

3 Deliverables

The proposal will result in a data analysis pipeline suitable for placing upper limits on the rate of binary inspirals. Components of this pipeline include: (i) template based, hierarchical search code suitable for LDAS MPI execution together with LAL compliant functions, (ii) FCT, hierarchical search code (being developed as part of a previous proposal, see LIGO-LIGO-G000217-00-D), (iii) instrument characterization tools suitable for LDAS execution together with LAL compliant functions, (iv) LAL functions to perform coherent searches for binary inspiral signals in multiple interferometers, (v) LAL compliant functions to inject fake signals into the data stream, (vi) implementation of an electronic data-analysis log book, and (vii) visualization (MATLAB and C) tools designed to facilitate analysis of database queries.

Two papers are expected to result from these activities. First, a paper describing the methodology and results of the upper limit analysis. Second, a paper describing in detail the results of the instrument characterization activities with specific emphasis on the implications for binary inspiral detection algorithms.

4 Required resources

Software: This project requires software to carry out searches for binary inspiral waveforms in LIGO data. It is assumed that template based and FCT based search codes will be implemented in LAL and available under LDAS. It is also assumed that the LDAS software pipeline will be available for on-site analysis of the data in the manner proposed for the full-scale science runs in the future.

Hardware: Availability of a Beowulf cluster at each site is required for the first cut analysis. Follow-up analysis will be carried out on LIGO facilities at CACR and UWM.

Data: Data from the LIGO E2 run will be used during the testing and development phase. This will be essential to develop a search strategy in advance of the data taking. Data from LIGO En will be used to determine the upper limit; it is assumed that coincident data from two interferometers will be available and that reduced data sets will be made available in frame format for off-site analysis.

Personnel: We anticipate that the proposed activities will require about 8-10 FTE's for the duration of the project. In addition to the proposers, we anticipate the need for about 0.5 FTE months of time from each of two LDAS personnel during the verification and validation of all code to run within the LDAS environment.

5 Work plan

The working group will have bi-monthly teleconferences to report progress on the milestones. Brady and Gonzalez will coordinate the efforts of the group. Eight teams of volunteers will work on the various aspects of the analysis pipeline; a single individual has been designated to coordinate the efforts of each team. The group has a web page at http://www.lsc-group.phys.uwm.edu/iulgroup and a mailing list iulgroup@gravity.phys.uwm.edu which is archived to the web page. An outline of tasks and milestones is provided in Table 1. There are two phases in the plan. Phase I culminates in a three day verification and validation test – the inch-pebble MDC – scheduled for 16-18 May 2001. Phase II follows this MDC and is planned to culminate in a scientific validation test in late August. The on-line real-time component of the proposed upper limit analysis will take place during the engineering run in September. Off-line analysis will proceed for approximately three months following the data acquisition. The first of the two papers should be ready for submission in late December 2001; the second paper will be submitted around March 2002.

Table 1: Task table and milestones for Inspiral Upper Limit working group. Volunteers have joined teams to complete each major task. A list of the volunteers and more details about the execution of the tasks can be found on the group web page. There are two phases in the plan. Phase I culminates in the first integration test: the inch-pebble MDC. Phase II follows this MDC and is planned to culminate in a scientific validation test in late August.

Inspiral Upper Limits

Project **Inspiral Upper Limits**
Starting Date 01-Jan
Completion Date 26-Aug
Present Date 01-Jan

Legend
◆ Project Milestone
Uncompleted part of task
Completed part of task
Overdue part of task

Task Description	Starting Date	Ending Date	% Comp.	No. of Days
General timeline				
Brainstorming: Phase I	02-Jan	01-Apr	0	90
Development: Phase I	01-Feb	16-May	0	105
Integration into LDAS: Phase I	01-Mar	16-May	0	77
Milestone: End Phase I	16-May	16-May	0	0
Inch-pebble MDC	16-May	18-May	0	3
Brainstorming: Phase II	18-May	16-Jun	0	30
Development: Phase II	18-May	05-Aug	0	80
Integration into LDAS: Phase II	18-Jun	23-Aug	0	67
Milestone: End Phase II	24-Aug	24-Aug	0	0
Scientific Validation Tests	24-Aug	26-Aug	0	3
Template Placement Code (Phase I)				
Development (ASIS)	01-Jan	30-Apr	0	120
Integrate with TIC/FCT	02-Mar	02-May	0	62
Integration into LDAS	03-Apr	16-May	0	44
Templated Inspiral Code (Phase I)				
Development (ASIS)	01-Jan	30-Apr	0	120
Integrate with placement code	02-Mar	02-May	0	62
Integration into LDAS	03-Apr	16-May	0	44
FCT code (Phase I & II)				
Development (G000217-00-D)	01-Jan	30-May	0	150
Integration into LDAS	02-Mar	29-Jun	0	120
Instrument Characterization Tools (Phase I & II)				
Brainstorming	01-Jan	30-Apr	0	120
Code development	02-Mar	29-Jun	0	120
Integration into LDAS	03-Apr	23-Aug	0	143
Multiple interferometer pipeline/code (Phase I & II)				
Identify and compare methods	01-Jan	30-Apr	0	120
Design tools for coincidence/cohere	02-Mar	03-May	0	63
Develop coherent code	03-May	23-Aug	0	113
Develop DB tools	03-Jun	23-Aug	0	82
Database tools (Phase I)				
Database MDC	01-Jan	30-Jan	0	30
C/Matlab ionterface	01-Jan	01-Mar	0	60
Evaluate storage needs	01-Mar	30-Mar	0	30
Develop viewing analysis tools	03-Apr	23-Aug	0	143
Review table definitions	15-Apr	15-May	0	31
Event identification/integration	15-Apr	23-Aug	0	131
Simulation tools (Phase I & II)				
Identify source populations	01-Jan	01-Mar	0	60
Code development	02-Mar	30-May	0	90
Integration into LDAS	03-May	23-Aug	0	113
Statistics (Phase I & II)				
Brainstorming	01-Jan	31-Mar	0	90
Design scheme	02-Mar	29-Jun	0	120
Development	01-May	23-Aug	0	115
Identify logbook needs	01-Jan	19-Feb	0	50
Evaluate logbook facilities	15-Feb	16-Mar	0	30
Implement logbook	16-Mar	14-Apr	0	30