LIGO-G030365-00-Z

ACIGA Data Analysis

Susan M Scott, Antony C Searle, Benedict J Cusack, Andrew J Moylan, David E McClelland
Department of Physics, Faculty of Science,
The Australian National University, Canberra ACT 0200, AUSTRALIA

David Coward, Ron Burman, Eric Howell, David Blair
School of Physics, Faculty of Science,
The University of Western Australia, Crawley WA 6009, AUSTRALIA

Abstract

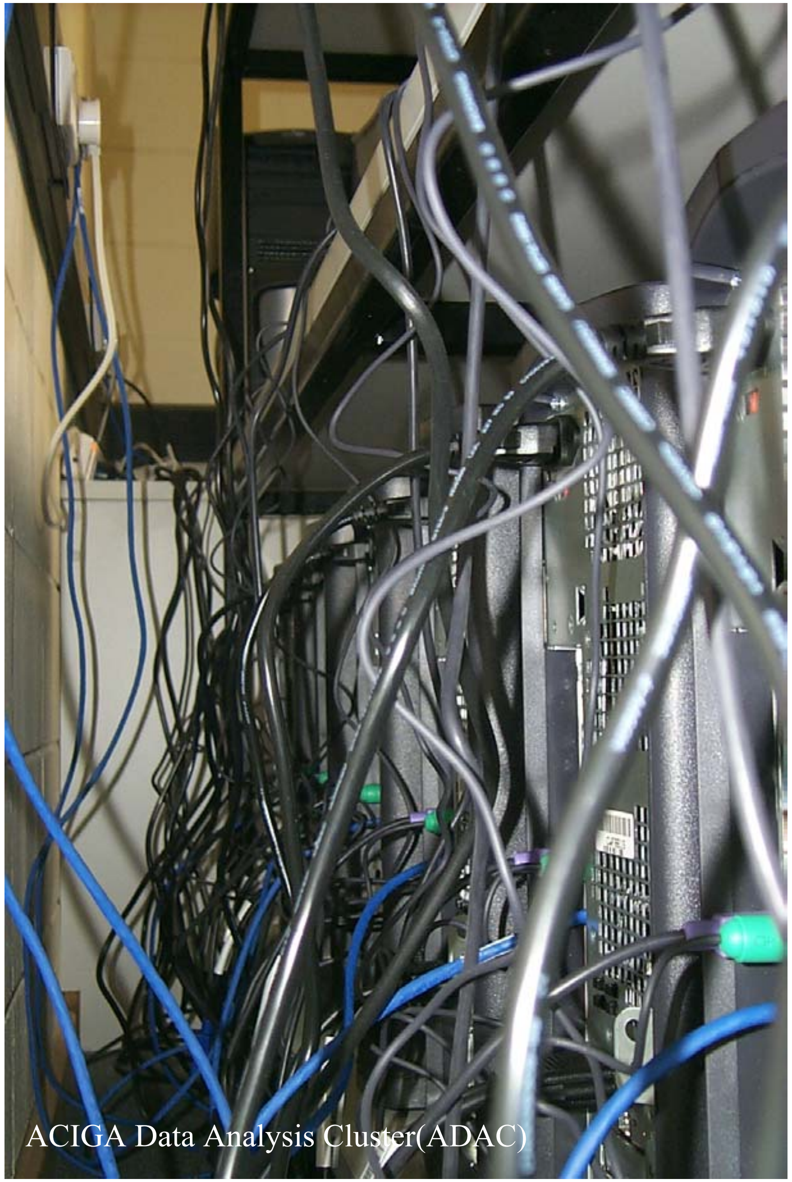
The Data Analysis program of the Australian Consortium for Interferometric Gravitational Astronomy (ACIGA) was set up in 1998 by the author to complement the then existing programs working on suspension systems, lasers and optics, and detector configurations. The ACIGA Data Analysis program continues to contribute significantly in the field; we present an overview of our activities.

LIGO Data Analysis System

ACIGA Data Analysis works in close collaboration with LIGO. We have made significant contributions to the LIGO Data Analysis System (LDAS), and to the Data Conditioning API in particular. These systems underpin LIGO's science results. We will describe contributions of spectral line removal code and the Universal Data Type (UDT). The ACIGA Data Analysis Cluster (ADAC) installed at The ANU is an LDAS system used to characterize and validate our LDAS contributions, analyse LIGO data and environmental data (including locally acquired environmental data).

The LIGO Data Analysis System (LDAS) [1] is the data acquisition, storage, and analysis software and hardware system for the LIGO project. It consists of installations at both observatories and participating institutions including The California Institute of Technology (Caltech), The Massachusetts Institute of Technology (MIT), The Australian National University (ANU), The Pennsylvania State University (PSU), and The University of Texas at Brownsville (UTB). Observatory systems acquire and locally cache data for medium-term storage before it is transferred to the Center for Advanced Computing Research (CACR) at Caltech for archiving in their High Performance Storage System (HPSS). The data may then be accessed by LSC (LIGO Science Collaboration) member institutions and processed by an LDAS installation. Typically, processing is performed in two stages, corresponding to non-parallel and parallel operations. Parallelisable operations are distributed across LDAS-managed Beowulf clusters (collaborative networks of inexpensive consumer computers with unrivalled cost-performance ratios). For example, the parameter space of an exhaustive search may be partitioned into regions, each allocated to a particular computer in the cluster. Some operations are common to all searches—typically pre-processing, or *conditioning*, the data to make it more amenable to searching—and these operations are performed once by high-performance LDAS ‘servers’ before the data is distributed to the Beowulf cluster.

The Data Conditioning API (Application Programming Interface) [2] is the component that performs the pre-processing of data in LDAS. It accepts data from other APIs, typically the FrameAPI which is responsible for ingesting data from file, applies a user-defined *algorithm* consisting of a series of signal-processing ‘actions’, and dispatches the conditioned data to other APIs, typically the Beowulf clusters. As such, it is the component where all generic science tasks occur. The ACIGA Data Analysis team has contributed substantially to the design, development and testing of much of the Data Conditioning API and its signal-processing actions, including Discrete Fourier Transforms (DFTs), linear filtering, resampling, etc. as well as advanced algorithms using system identification to remove artefacts.


ACIGA Data Analysis Cluster(ADAC)

In late 2002 we directly participated in the first science analyses of data taken from LIGO’s S1 Science Run. We have worked closely with the LIGO Stochastic Background search group to develop line removal tools. Line removal was integrated into the stochastic background pipeline, and the impact of correlated spectral lines on the stochastic background search codes was assessed [3]. The analysis culminated in the setting of an upper limit on the strength of a cosmological background of gravitational radiation—results are in preparation.

The ACIGA Data Analysis Cluster (ADAC) is a fully-functional LDAS installation at The ANU Department of Physics. Currently in commissioning, it consists of three dual-processor servers managing a Beowulf cluster of eight single-processor nodes, with a terabyte of local storage and a link to the Mass Data Storage System (MDSS) at The

ANU Super-computer Facility (ANUSF). It will give the team local capability to perform the intensive characterisation on extended datasets required to fully validate complex algorithms like line removal, and to use the industrial-strength algorithms of LDAS for independent ACIGA research.

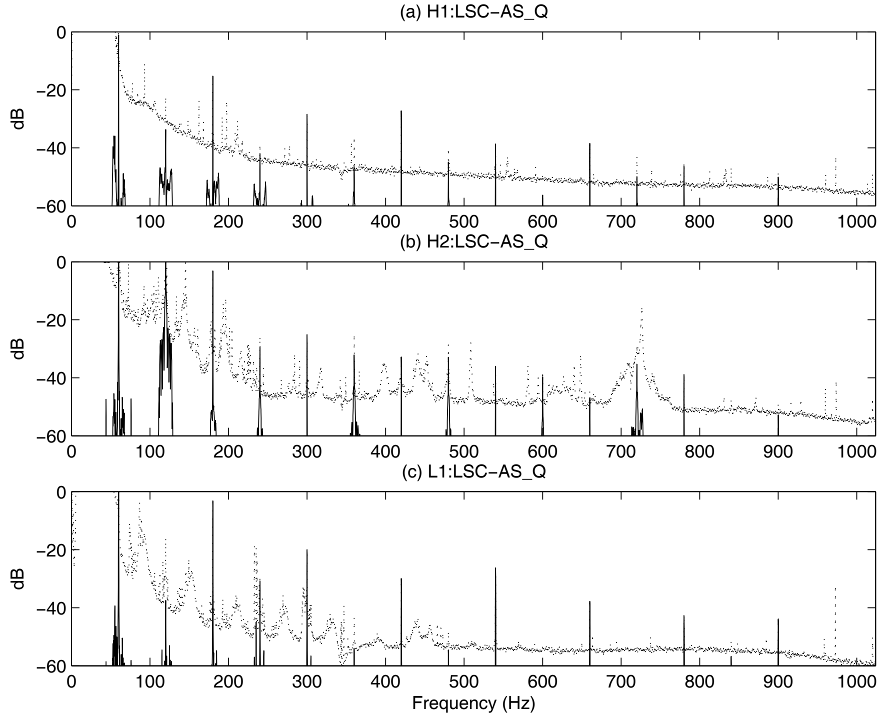
The ACIGA Data Analysis Cluster (ADAC) gives ACIGA a conformant LDAS implementation to analyse LIGO data without duplication of LIGO's investment in the development of data analysis tools. The facility consists of three dual-processor servers managing a Beowulf cluster of eight high-performance nodes, gigabit Ethernet switching, and a terabyte local cache of RAID storage capacity. The cluster has already been used in the characterisation of the performance of line removal tools on actual LIGO data [4].

In the terms of the Memorandum Of Understanding (MOU) between ACIGA and LIGO, ACIGA has access to data taken by the LIGO interferometers during engineering and science runs. Access to all of these terabyte-scale datasets is, however, impractical over the internet. We intend to purchase two LIGO-compatible tape drives for installation at ADAC and Gingin and a multi-terabyte reusable tape cache to facilitate exchange of data between ACIGA and LIGO sites. We also plan the formation of a local archive of LIGO data on the MDSS.

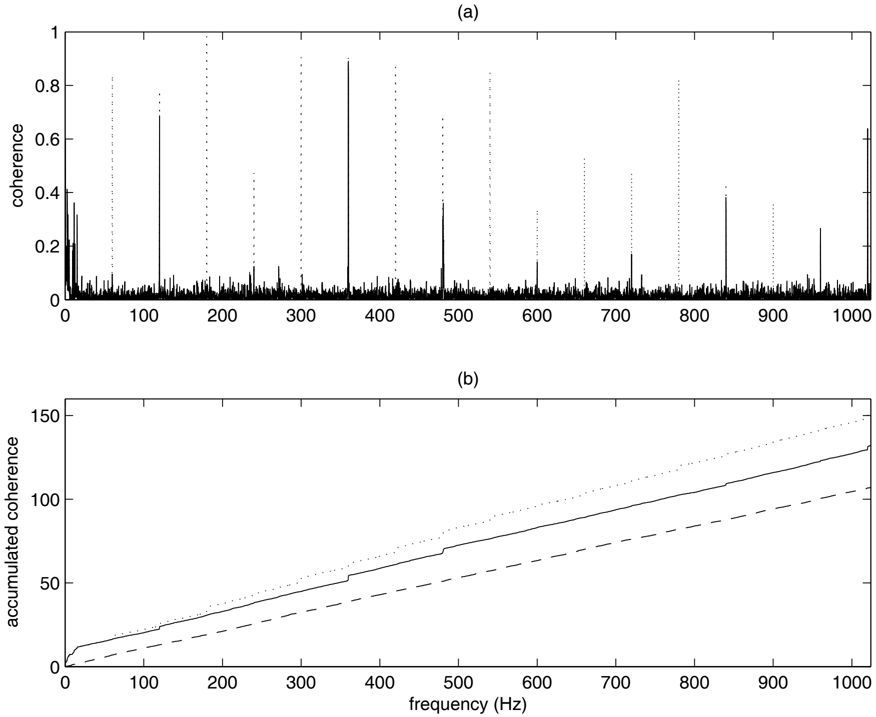
One of the most visible instrumental artefacts in the output of interferometric observatories is the interference due to the alternating-current power supply to the instrument. This manifests itself as a series of narrow spectral *lines* at 60 Hz and its harmonics. Unlike most instrumental noise, the actual generator of the noise can be monitored by sampling the supply voltage and directly measuring the magnetic fields at points throughout the observatory. LDAS stores Physical Environment Monitor (PEM) channels along with the LSC channels. The *line removal* code developed at The ANU ingests these PEM channels and uses system identification techniques [5] to fit an optimal linear model to the interaction of the detector with the measured disturbance. This model can be used to predict the line shape at future times, and hence to subtract it out of the data. Initial characterisation of the algorithm has been very positive, and it has been used by the Data Conditioning API to condition data for ingestion by the stochastic background search code.

[1] http://ldas-cit.ligo.caltech.edu
[2] http://ldas-cit.ligo.caltech.edu/doc_index/datacond_api.html
[3] B Abbott *et al*, *Analysis of First LIGO Science Data for Stochastic Gravitational Waves*, in preparation.
[4] A C Searle, S M Scott, and D E McClelland, *Spectral Line Removal in the LIGO Data Analysis System*, to appear in *Classical and Quantum Gravity*.
[5] L Ljung, *System Identification: Theory for the User* (2nd Ed.) Prentice Hall (1999)

See also the ACIGA Data Analysis research poster: "A stochastic gravitational wave background from general relativistic neutron star formation and bar mode instabilities" by E. Howell et al.



Power spectra of (a) H1:LSC-AS_Q, (b) H2:LSC-AS_Q, and (c) L1:LSC-AS_Q channels (dotted lines) and their respective spectral line models (solid lines). H1 and H2 are, respectively, the LIGO Hanford 4 km and 2 km interferometers, and L1 is the LIGO Livingston 4 km interferometer.



(a) Coherence of H1:LSC-AS_Q and H2:LSC-AS_Q before (dotted line) and after (solid line) line removal. (b) Accumulated coherence of H1:LSC-AS_Q and H2:LSC-AS_Q before (dotted line) and after (solid line) line removal, with the coherence of H1:LSC-AS_Q and L1:LSC-AS_Q (dashed line) providing a noise floor reference.

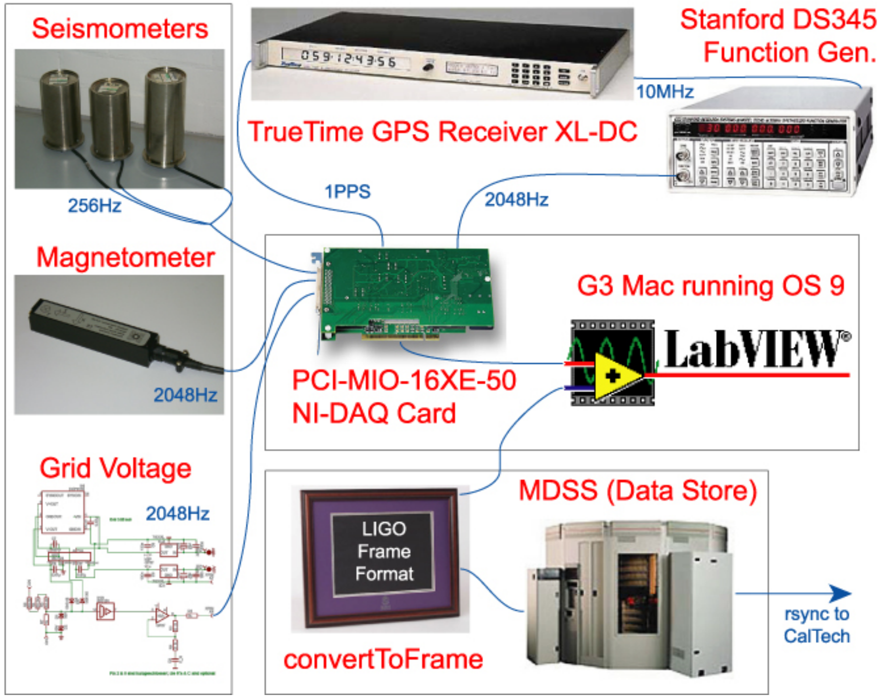
Physical Environment Monitoring

ACIGA participates in the international exchange of environmental data incorporating seismic, magnetic and grid voltage data channels. An around-the-clock environmental data collection system is currently operating at The Australian National University (ANU); ACIGA is the only participant in the southern hemisphere. A schematic outline of this system is shown, as well as results of correlation analyses.

An obvious advantage gained by comparing data from a global network of detectors is immunity from local noise sources. There still exists, however, the possibility of global noise sources coupling parasitic signals into multiple detectors, so an effort is currently underway to understand such noise sources and evaluate their effect on the global array of gravitational wave detectors. In particular, three physical environment variables are being targeted: seismic disturbances, fluctuations in the Earth's magnetic field, and fluctuations in mains voltages. These variables are being monitored continuously at detector sites (LIGO Hanford, LIGO Livingston, VIRGO, GEO previously) and now at The Australian National University (ANU), and various comparison studies have been completed to date.

ACIGA's environment monitoring equipment is located at The ANU's Gravitational Wave Research Facility: three seismometers (for three axes), a three-axis magnetometer, and a simple grid voltage scaling and filtering circuit. A multi-channel analog-to-digital converter and LabView-based data-logging software continuously log and process the data. The ANU's Mass Data Storage System (MDSS) stores the generated data in LIGO frame format ready for collection, via the rsync protocol, from an external party.

The ACIGA data-logging equipment uses a GPS receiver to synchronise samples with the other sites: the receiver's 1 Pulse-Per-Second signal triggers the initial sample at the start of a data taking run, and a function generator (using the receiver's 10MHz signal as a timebase) produces the 2048Hz sampling trigger required by the ADC board. The signal generator limits the accuracy of the timing such that the system will lose one sample roughly every 200 days; the system is manually reset on a regular basis to allow for this, while still maintaining a > 99% duty cycle.

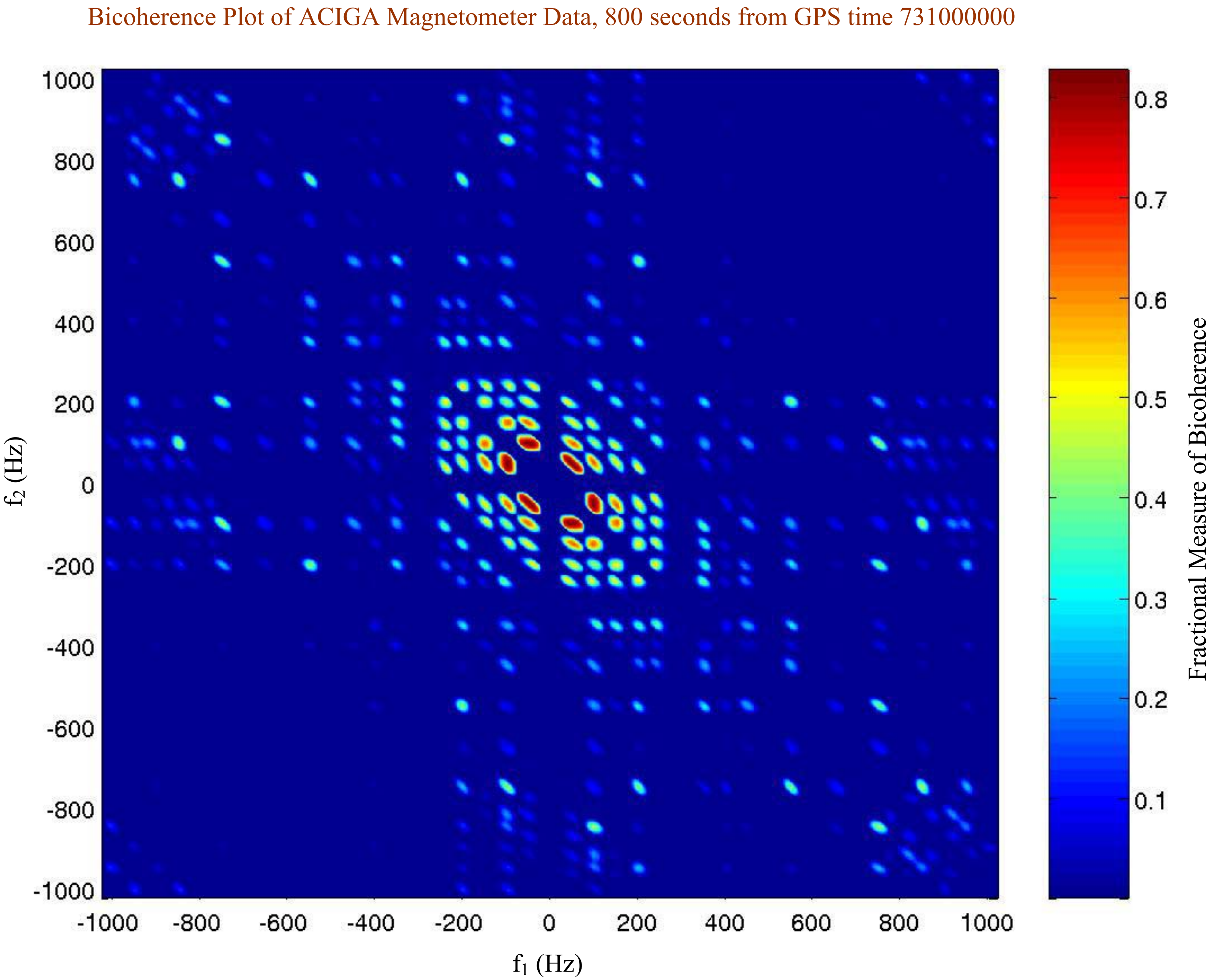


ACIGA's environment monitoring and data-logging setup.

An automated script on LIGO's High Performance Storage System (HPSS) collects frame-formatted data from the MDSS shortly after it is generated, and merges it with data from the other contributors. The HPSS, on which the prototype Network Data Analysis System (NDAS) runs, has provided a test bed for synchronising and merging data from globally separated sources, culminating in the recent merging of S2 data on the system.

Correlation studies using the data generated in the exchange continue. *(See talk from Data Analysis Session: B. Cusack "Global Correlations in Physical Environment Monitors for Gravitational Wave Detection".)* Techniques used include numerical estimates of power-spectral densities, cross-spectral densities and coherences, which are especially useful for identifying correlated lines between sites. Some coherent correlated lines have been detected between magnetometer channels of the Hanford and Livingston sites (see talk for more details), and between the Livingston and ACIGA sites, though this latter correlation occurs at the Nyquist frequency, where both channels have relatively low spectral power, and is thought to be dubious.

More recently, bicoherence studies have been carried out. Bicoherence is the normalised third-order spectrum, and is a function of two Fourier frequencies (see http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=3013 for a set of matlab tools and a brief tutorial on higher-order spectral analysis). An example of a bicoherence plot is shown here, for which ACIGA magnetometer data was used; almost all of the features seen are due to the 50Hz Australian grid voltage and its harmonics. If the data at Fourier frequency coordinates (f_1, f_2) shows a bright spot, this indicates a coherent link between lines at these frequencies in the input time-series data. Additionally, experimentation reveals that there needs to be a third frequency line at $f_1 + f_2$ in order for a bright spot to occur at (f_1, f_2). As such, the spot seems to indicate a coherence between the line $f1$ and the *beat* between lines f_2 and $f_1 + f_2$, or equivalently a coherence between the line f_2 and the beat between lines f_1 and $f_1 + f_2$.

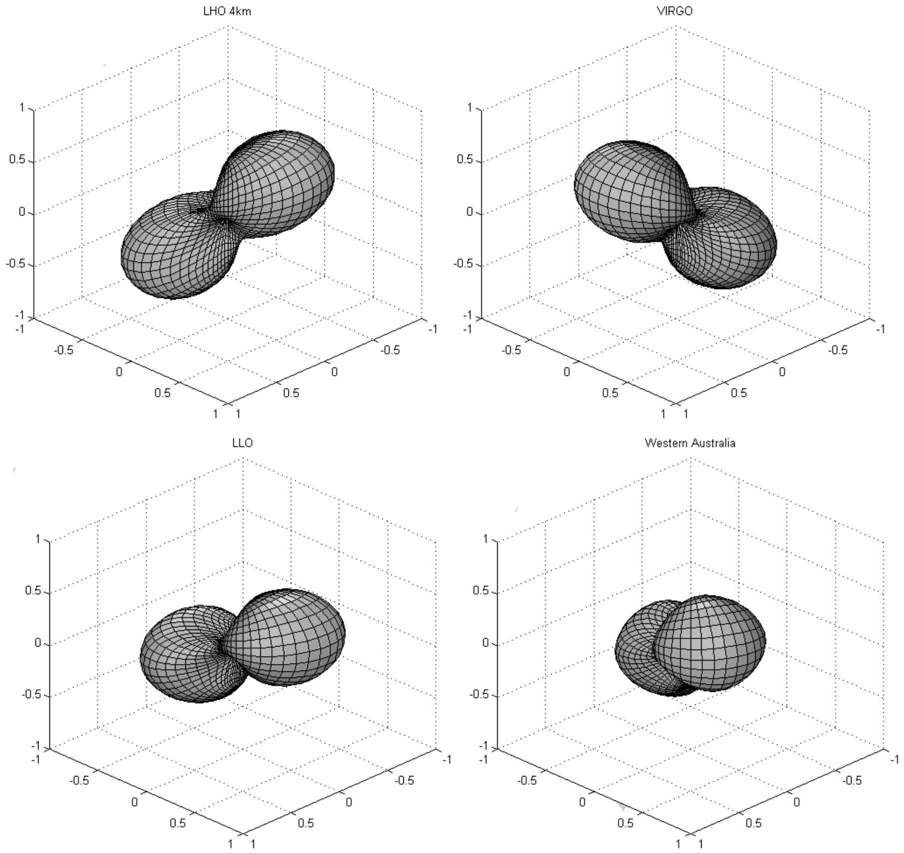


Bicoherence Plot of ACIGA Magnetometer Data, 800 seconds from GPS time 731000000

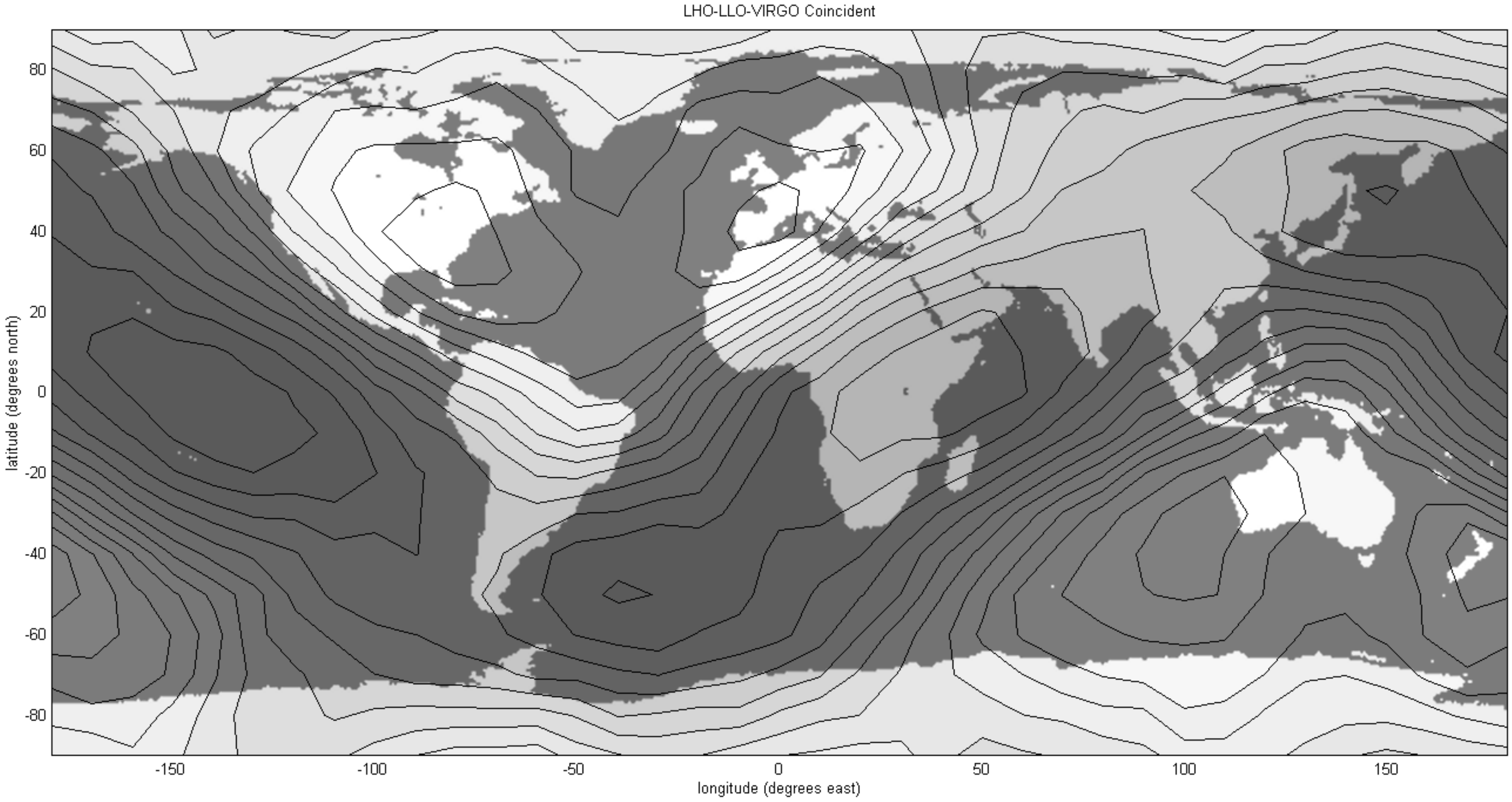
Global Network Optimisation

We have characterized the performance of cooperative global networks of gravitational wave observatories, drawing conclusions about the optimal placement of proposed detectors, including the desirability of a Western Australian observatory.

Ground-based gravitational wave astronomy will eventually be performed with a collaborative network of co-operating instruments. Studies of the optimal methods, both theoretical and practical, for analysing the data from a network of detectors are ongoing. In contrast, The ANU team has developed a simulation to determine, with respect to the detection of binary inspiral events, the optimal location to site a new detector to augment an existing network of detectors, for a variety of data analysis methods. This was partly motivated by the need to quantify the scientific contribution that a proposed Australian detector would make to the international community. It was determined [1] that the siting of the n^{th} detector in a network can have a substantial impact on the sensitivity of the network as a whole, and demonstrated that Western Australian sites—near-antipodal to the LIGO detectors—are optimal for augmenting the existing network. We continue to expand our simulation to include more general networks and other gravitational wave sources, such as (in collaboration with PSU) the detectable fraction of galactic pulsars.



Antenna patterns of existing and proposed interferometers.



Coincident detection rate as a function of the location of a 4th detector added to a LIGO-VIRGO network; brighter is better, contours every 5%.

[1] A C Searle, S M Scott and D E McClelland 2002 *Class. Quantum Grav.* **19** 1465-1470