

Development of a DMT monitor for tracking slow non-stationarities present in LIGO science data.

S. Mukherjee

for the LIGO Science Collaboration.

*Center for Gravitational Wave Astronomy,
The University of Texas at Brownsville, U.S.A.*

Abstract

Presence of slow non-stationarities has been seen in a series of studies with LIGO engineering and science data. It is important to track these slow variations or 'drifts' to diagnose their origin and correct them, as non-stationary data can affect sensitivity of astrophysical searches. A DMT (Data Monitoring Tool) monitor has been built to track slow non-stationarities in LIGO data. The monitor, tested on S4 data, will be operational during LIGO S5 data run. Offline analysis will be performed to unearth details of patterns thus obtained in the output of the monitor.

1. Introduction

The three LIGO detectors (1) have been taking data in science mode since 2002. The details of these science runs can be found in references (2-5). It has been shown in several previous analyses (6-10) that data from the present generation of instruments are non-stationary. Non-stationary features of interferometric data can appear either as sharp transients in time domain, or they may appear as changes in statistical properties (typically measured by change in the second moment) of the noise floor over longer time scales ranging from several seconds to minutes (6). Some of these features also show up as spurious peaks in the frequency domain. Noise floor indicates the part of data that is left behind after one subtracts out the narrowband features (or lines) by any of the well tested methods (11,12,13). Presence of non-stationary features of the latter kind (a.k.a. drifts of the noise floor) can affect the sensitivity of astrophysical searches. An immediate example can be seen in case of the externally triggered burst searches (e.g. Gamma ray burst and gravitational wave association (19-21)) where long stretches of 'off-source' data are compared to 'on-source' data with the underlying assumption that the data are stationary. Deviation from this condition can lead to biased results. It is thus important that these drifts be tracked and analyzed to look for their possible origin and hence be corrected either in hardware or by suitably altering the algorithm by taking in to account the non-stationary

effects. It can thus serve as an important diagnostic as well as a tool to create data flags.

In this paper the authors describe the development of a DMT (*Data Monitoring Tool*)(14, 15) monitor that tracks slow non-stationarity in LIGO data. The DMT is a software specifically developed to facilitate the detector characterization work (22) for LIGO data. The DMT package includes several libraries, simulation tools and graphics abilities that enable a user to perform online as well as offline analysis of LIGO science and engineering run data.

2. DMT and MNFTMon

The monitor described in the present study is based on an algorithm called the *Median based noise floor tracker* (MNFT). The details of this algorithm can be found in reference (6, 7). In short, this algorithm involves the following steps.

- (i) Bandpass and resample given timeseries $x(k)$.
- (ii) Construct FIR (Finite Impulse Response) (17) filter that whitens the noise floor of $x(k)$ producing a resulting timeseries $w(k)$.
- (iii) Remove lines from $w(k)$ using notch filter producing the cleaned timeseries : $c(k)$
- (iv) Track variation in second moment of $c(k)$ using the method of *Running Median* (13).
- (v) Obtain significance levels of the sampling distribution (thresholds) via Monte Carlo simulations.

Every time the running median estimate of the second moment ventures out of the threshold, the data will be deemed non-stationary i.e. it will have significantly departed from Gaussian stationary noise behaviour.

The monitor has been developed using the complete framework of DMT and its base software called 'root' (16). It consists of a main class called NoiseFloorMonitor which is derived from the DatEnv and MonServer classes of the DMT. In addition to this, a set of classes and functions have been developed by the authors to perform some specialized functions for the monitor viz. FIR filtering (17), non-linear filtering, FFT (18) and convolution (using fftw 3.0.1), power spectral density (PSD) (17), random number generation routines etc. These set of functions have been packaged into a library called 'waves'.

In the present version, the monitor is analyzing the main channel (AS_Q) from the two Hanford detectors (H1 and H2) and the Livingston detector (L1). The monitor takes in 60 seconds of data sampled at 16 KHz in one time stride. Figure 1 shows a 600 second segment of raw time series to demonstrate the data type that goes in as the monitor input. Following the steps described above, the data

is first low passed with a cut off frequency of 4KHz and resampled. The spectral noise floor is then estimated using the running median in frequency domain (13) (see figure 2). The lines are rejected of course when the width of the line is smaller than the block length used in calculating the running median. An FIR whitening filter is designed based on this estimate and the data is whitened. In the next step, the lines are subtracted by using a notch filter. This operation is based on an algorithm called the Median based line tracker (MBLT, 13) and is done in the time domain. The power spectrum of the cleaned and whitened data can be seen in figure 3. The second moment is estimated by computing the running median (13) of the squared time series. This is a time domain operation that uses a block length of 4 second segments. Figure 4 displays an actual monitor output using recent LIGO data using full bandwidth.

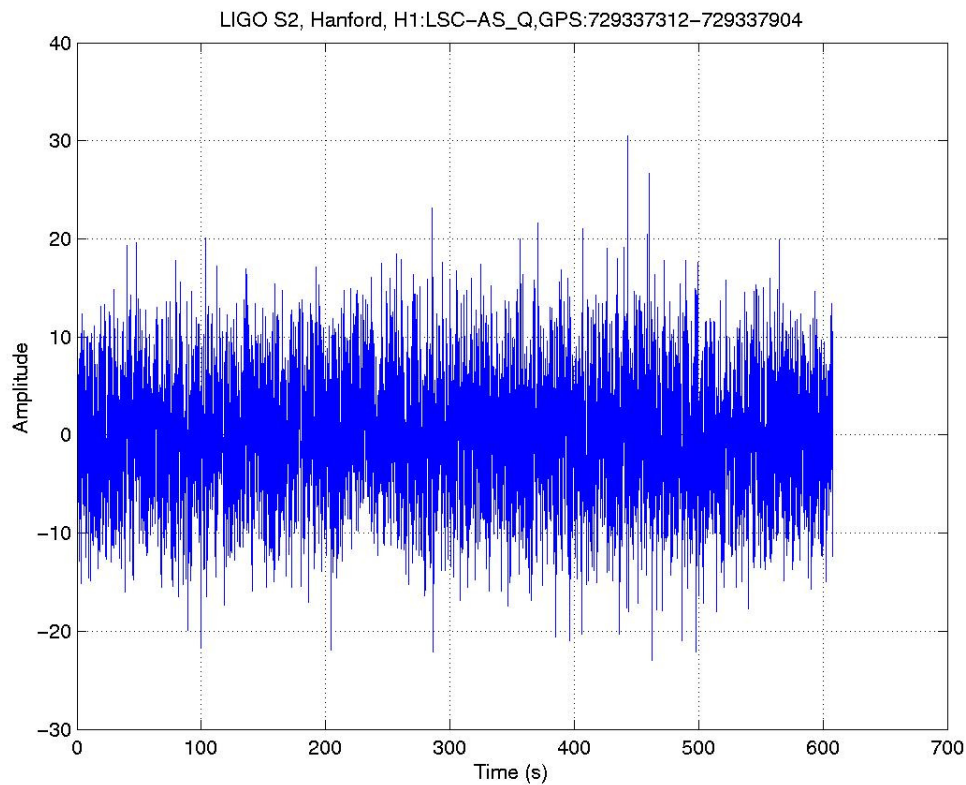


Figure 1. Raw time series data, sampled at 16 KHz, is being used as input to the monitor.

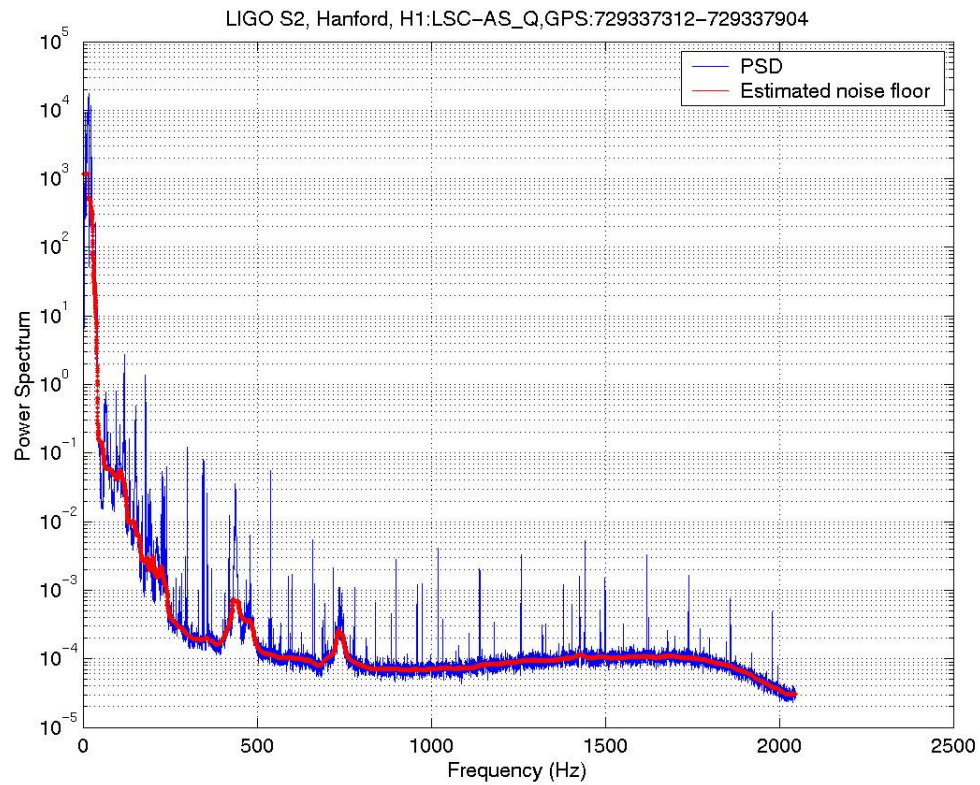


Figure 2. Spectral noise floor is being estimated by using running median. This estimate is used in the construction of an FIR whitening filter. In this figure, the power spectrum is shown in blue (with all lines) and the spectral floor estimate is shown in red. The running median algorithm rejects the lines as 'outliers' and thus gives a smooth estimate of the spectral floor.

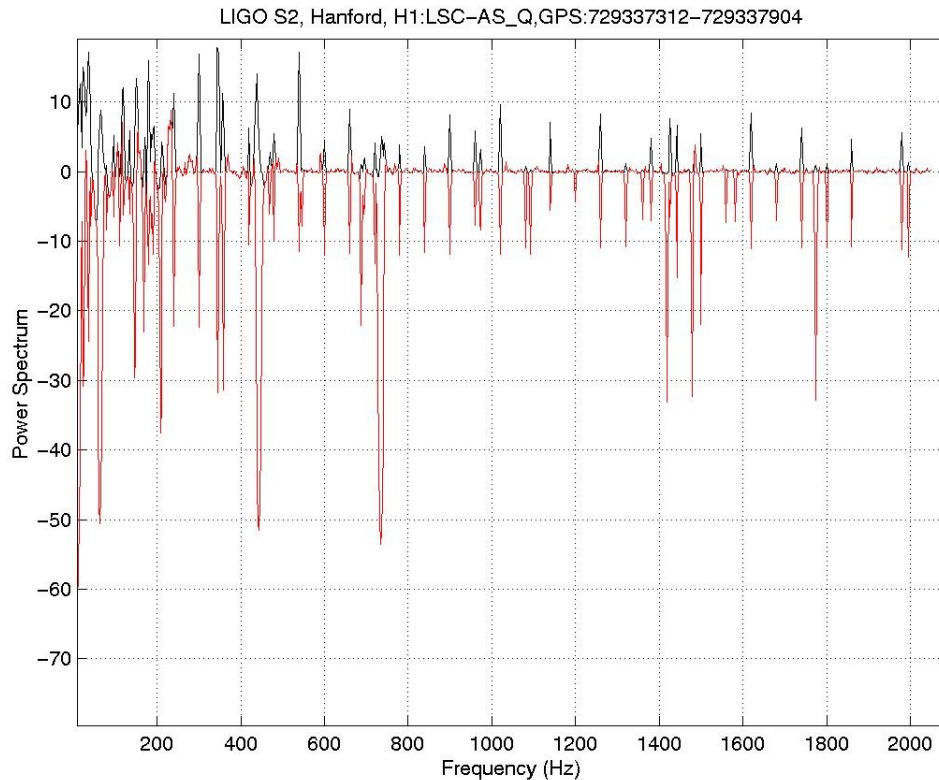


Figure 3. Power spectrum of the whitened and cleaned data. Here, the whitened data with lines is shown in black. The whitened and cleaned spectrum after line subtraction is shown in red.

The output can be viewed on the DMT viewer as a time series that is updated every 60 seconds. Thresholds are also displayed. The code runs online. It takes less than 8 seconds to process 60 seconds of raw data on one of the designated DMT machines at Hanford. The analysis is carried out in four frequency bands : 0-20 Hz, 20-100 Hz, 100-200 Hz and 200-2KHz. This is done to track the behaviour of data in different frequency bands, since the frequency response of different bands can vary. It is thus a useful diagnostic.

On-line testing of the monitor performance has been done by the authors at both Livingston and Hanford sites. The monitor codes are frozen and committed to the Global Diagnostics Systems (23) repository to run for the mini M7 data run (24) before the start of LIGO's fifth science run (S5).

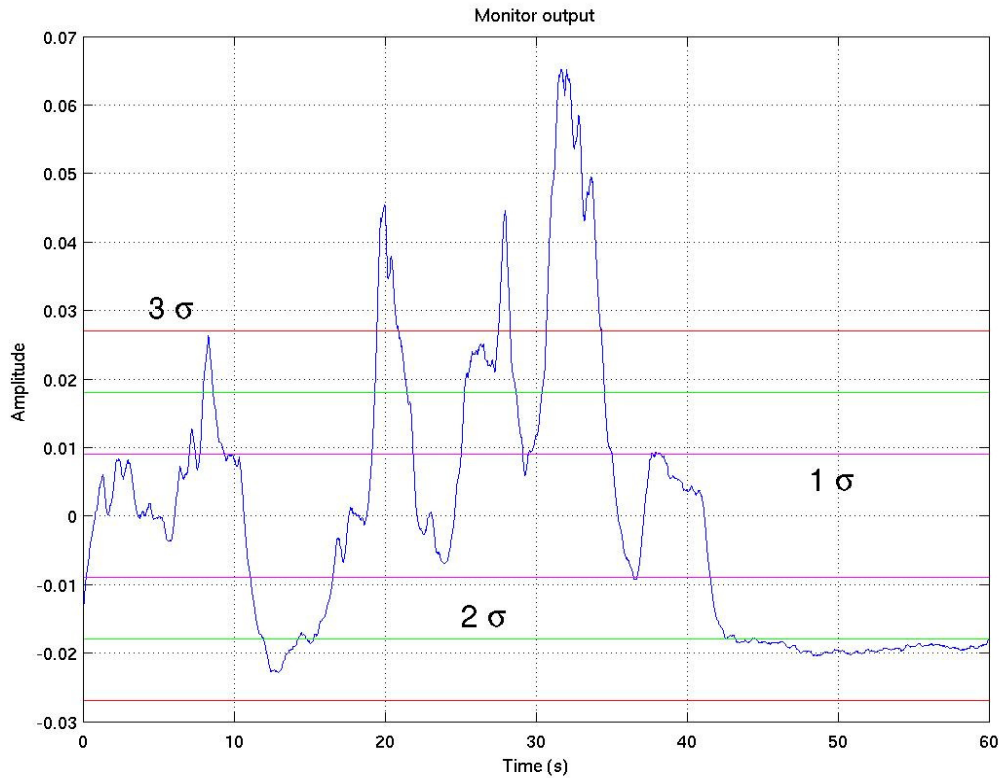


Figure 4. This figure shows the MNFTMon output. The blue curve is the running median estimate of the second moment of the cleaned and whitened data. The different levels of the threshold (corresponding to 1, 2 or 3 σ) are also shown. The data can be inferred to be non-stationary when it crosses the 3 σ threshold. In this example, such threshold crossings have taken place around 20 s and also just before and after 30 s.

Goals of MNFTMon for S5

The monitor will run continuously for the entire duration of S5. The goal of this monitor is to do an off-line analysis of the triggers (i.e. threshold crossings) it has recorded and to study any patterned behaviour in the monitor output. This will lead to putting flags on the data segments that has shown maximum non-stationarity, indicated by frequent threshold crossings and hence having high trigger rates.

As of now, the main AS_Q channels are looked at. However, during the course of S5 run, the monitor functionalities will be extended to include several auxiliary channels as well, such that any patterned behaviour of the output can be traced back to its origin.

Figure of merit resulting from the MNFTMon output

The monitor output can lead to the development of a useful figure of merit (FOM) for the astrophysical burst searches. For example, the signal to noise ratio (SNR) of any signal e.g. 1.4 -1.4 solar mass binary inspiral, or a sine Gaussian can be calculated using a matched filter in presence of Gaussian stationary noise. The output of the monitor can be accurately modeled using Auto Regressive Moving Average (ARMA) models (25,26). The SNR of the same signal can now be calculated using a sufficiently long realization of the modeled noise. The deviation ratio i.e. the ratio of the SNR in presence of Gaussian stationary noise to that in presence of the modeled non-stationary noise, can be plotted as a FOM indicating the sensitivity of the astrophysical search in presence of this kind of non-stationarity. This part of the work is underway.

Acknowledgement

The authors thank John Zweizig and Keith Riles for many suggestions and help during the development of the monitor and the LIGO laboratory for cooperation during the online testing of the monitor. The authors acknowledge partial support for this work from NSF (grants HRD-0206028 and PHY-9981795) and from NASA (grant NAG5-13396).

References

- (1) Barish, B. and Weiss, R. 1999 LIGO and the detection of gravitational waves. *Phys.Today* **52**, 44
- (2) <http://blue.ligo-wa.caltech.edu/scirun/S4>
- (3) <http://blue.ligo-wa.caltech.edu/scirun/S3>
- (4) <http://blue.ligo-wa.caltech.edu/scirun/S2>
- (5) <http://blue.ligo-wa.caltech.edu/scirun/S1>
- (6) Mukherjee, S. 2003 Median based noise floor tracker : robust estimation of noise floor drifts in interferometric data. *Class. Quantum Gravity* **20**, 925-935.
- (7) Mukherjee, S. 2003 Study of non-stationarity in LIGO S2 data : Thresholds, tolerance and models. LIGO Technical document G030374-00-Z, Caltech, Pasadena.
- (8) Mukherjee, S. 2003 Median based noise floor tracker : robust estimation of noise floor drifts in LIGO S1 data. LIGO Technical document G030019-00-Z, Caltech, Pasadena.
- (9) Mukherjee, S. 2003 Robust estimation of noise floor drifts in interferometric data. LIGO Technical document G030166-00-Z, Caltech, Pasadena.
- (10) Mukherjee, S. and Grosso, R. 2005 Median baseline monitor tracking slow non-stationarity in LIGO data. Amaldi conference, Okinawa, Japan.

- (11) Finn, L.S. and Mukherjee, S. 2001 Data conditioning for gravitational wave detectors : a Kalman filter for regressing suspension violin modes. Phys. Rev. D **63**, 062004, 1-19.
- (12) Sintes, A. M. and Schutz, B.F. 1998 Coherent line removal : filtering out harmonically related line interference from experimental data with application to gravitational wave detectors. Phys. Rev. D **58**, 122003.
- (13) Mohanty, S.D. 2002 Median based line tracker (MBLT): model independent and transient preserving line removal from interferometric data. Class. Quantum Gravity **19**, 1513-1519.
- (14) <http://www.ligo.caltech.edu/~jzweizig/DMT-Project.html>
- (15) Zweizig, J. 2003 Data simulation for DMT. LIGO Technical document G030493-00-E, Caltech, Pasadena.
- (16) <http://root.cern.ch>
- (17) Proakis, J.G. and Manolakis, D.G. 1996 Digital Signal Processing Third Ed. Prentice Hall, New Jersey.
- (18) <http://www.fftw.org>
- (19) Finn, L.S., Mohanty, S.D. and Romano, J.R. 1999 Detecting an association between gamma ray and gravitational wave bursts. Phys. Rev. D **60**, 121101
- (20) Mohanty, S.D. 2003 Optimum integration length for computation of the cross correlation coefficient. LIGO Technical document T030113-00-D
- (21) Leonor, I. 2005 Search for short duration GW bursts coincident with S2/S3/S4 gamma ray bursts. LIGO technical document G050432-00-Z
- (22) <http://www-mph.physics.lsa.umich.edu>
- (23) <http://apex.ligo-wa.caltech.edu/gds/>
- (24) <http://blue.ligo-wa.caltech.edu/engrun/M7/>
- (25) Mukherjee, S. 2004 Interferometric data modelling : issues in realistic data generation. Classical and Quantum Gravity, **21**, S1783-1792.
- (26) Hamilton, J. D. 1994 Time series analysis, Princeton University Press.