medNoiseChar: Assessing LIGO Data Stationarity

Lee Samuel Finn

Center for Gravitational Wave Physics, Penn State

LIGO-G030420-00-Z

Introduction

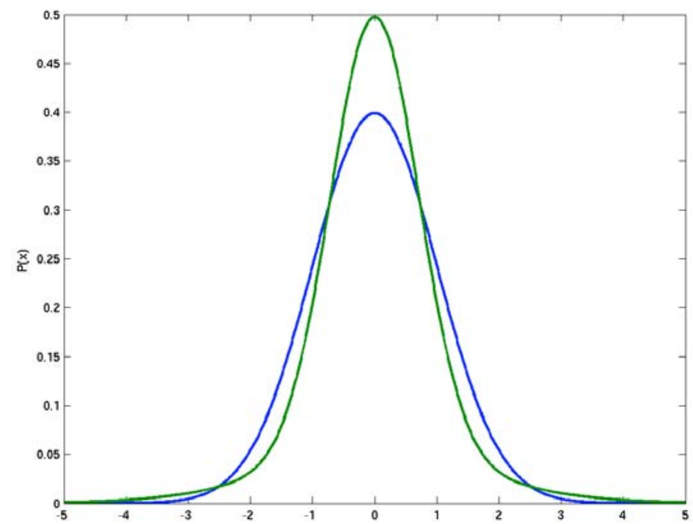
- Problem:
  - Well-behaved detector has stationary noise, calibration
  - Data analysis methods rely on stationarity of noise, calibration
  - How to assess, monitor calibration, noise stationarity?
- medNoiseChar
  - Monitors variation of three measures of detector noise statistics and identifies approximately stationary noise segments
- Outline
  - Characterizing detector noise
  - Monitoring stationarity
  - Identifying stationary epochs
  - A look at S2 stationarity

Basic Idea

- Identify interesting, relevant measures of detector noise statistics
  - » E.g., mean, variance in particular band
- Evaluate statistic periodically
  - » E.g., on consecutive 16s intervals
- Is measurement on current segment consistent with measurement on previous segment? With measurement on past several segments?
  - » Each measurement is an estimate: consistency is a test based on distribution of errors associated with each measurement
- Stationary epoch: longest set of consecutive intervals whose statistics are "consecutively consistent"

Characterizing detector noise

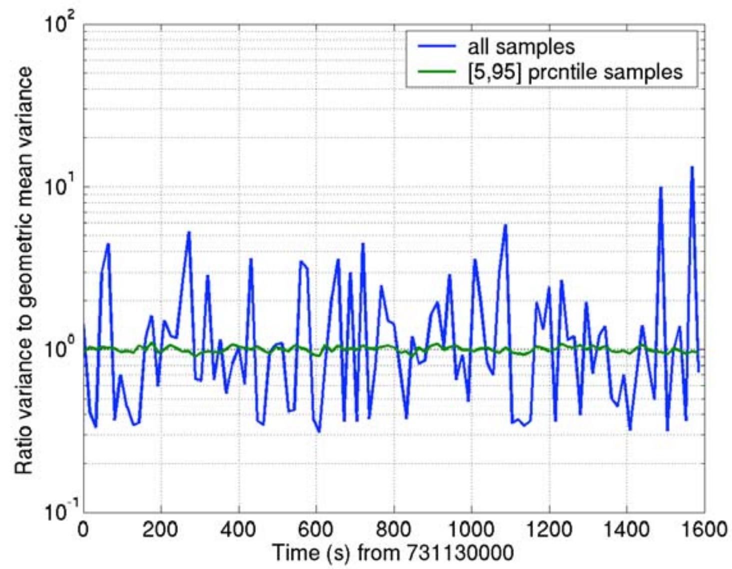
- Central tendency
  - » E.g., mean
- Dispersion
  - » E.g., variance
- "Gaussianity"
  - » E.g., $\chi 2$ fit to Gaussian
- Problem: Detector noise shows strong, rapid variation in sample distribution wings
  - » Timescale: seconds or less
  - » Wings: 2-3σ
  - » Variability: strong, asymmetric departures from Gaussianity apparent in mean, dominate variance



Same mean, variance; different χ^2 fit to Gaussian

Characterizing detector noise

- Focus on distribution mass, not outliers
  - » Rank order N samples per averaging time
  - » Drop first, last pN samples
    - – These are the outliers
  - » Evaluate mean, variance, χ^2 fit to Gaussian of remaining samples
  - » Measures are resistant to changes in outliers
- [5,95] percentile cut
  - » Reduces variability in dispersion on 16 s intervals by > factor 30 on typical segment
- Caveat: *musn't ignore outlier variability*
  - » But that's a different study



	$\sigma_\nu/<\sigma_\nu>$	$\sigma_\nu/<\sigma_\nu>$ [5,95]
Gaussian noise	0.0063	0.0063
S2 data	1.8823	0.0414

LIGO-G030420-00-Z

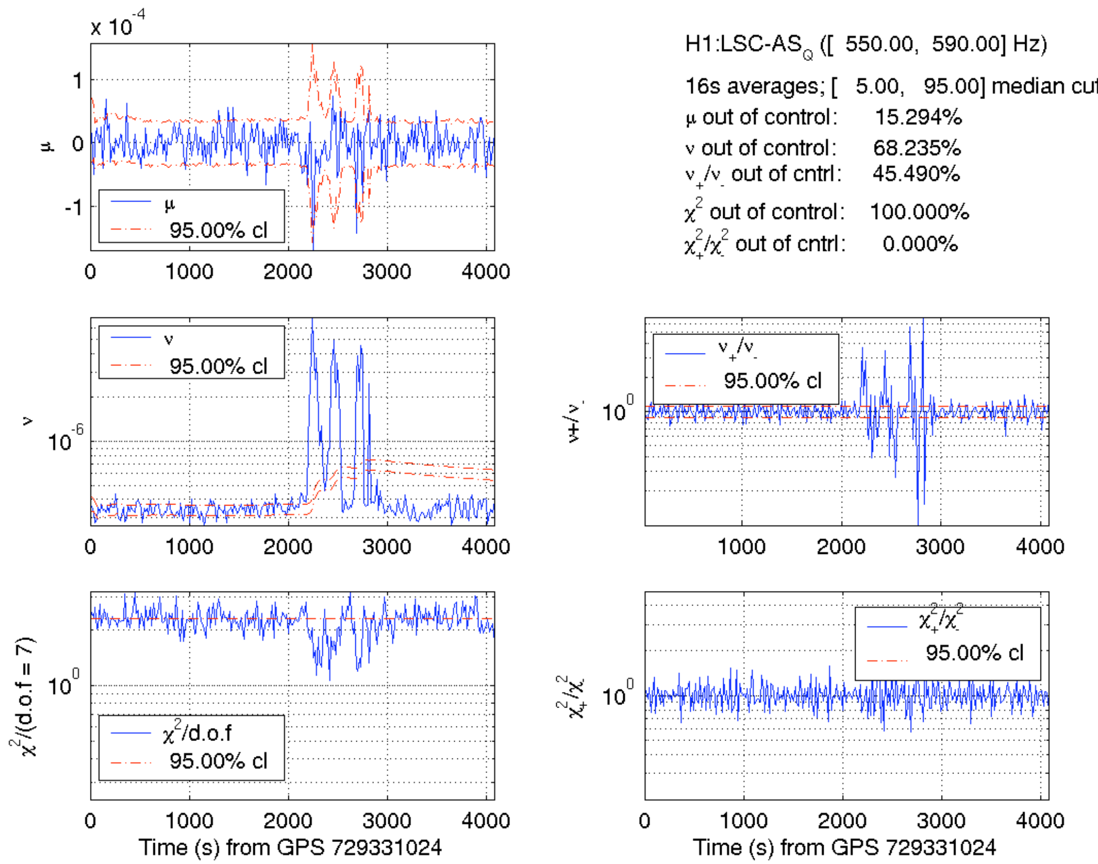
Monitoring stationarity

- Focus on properties of truncated distribution estimated on consecutive data segments
  - » E.g., consecutive 16s intervals
- Mean
  - » N samples: error bars (approximately) proportional to standard error of mean ($\sigma/N^{1/2}$)
  - » Compare to last interval mean, accumulated mean since epoch start
- Variance
  - » Ratio of two variance estimates from same distribution follow *f*-distribution (depends on number of samples in each estimate)
  - » Compare current variance estimate to last estimate, accumulated estimate since epoch start
- "Gaussianity"
  - » Ratio of two χ^2 quality of fit to model distribution follow *f*-distribution (depends on number samples in each estimate)
  - » Compare current fit quality to last

Choices

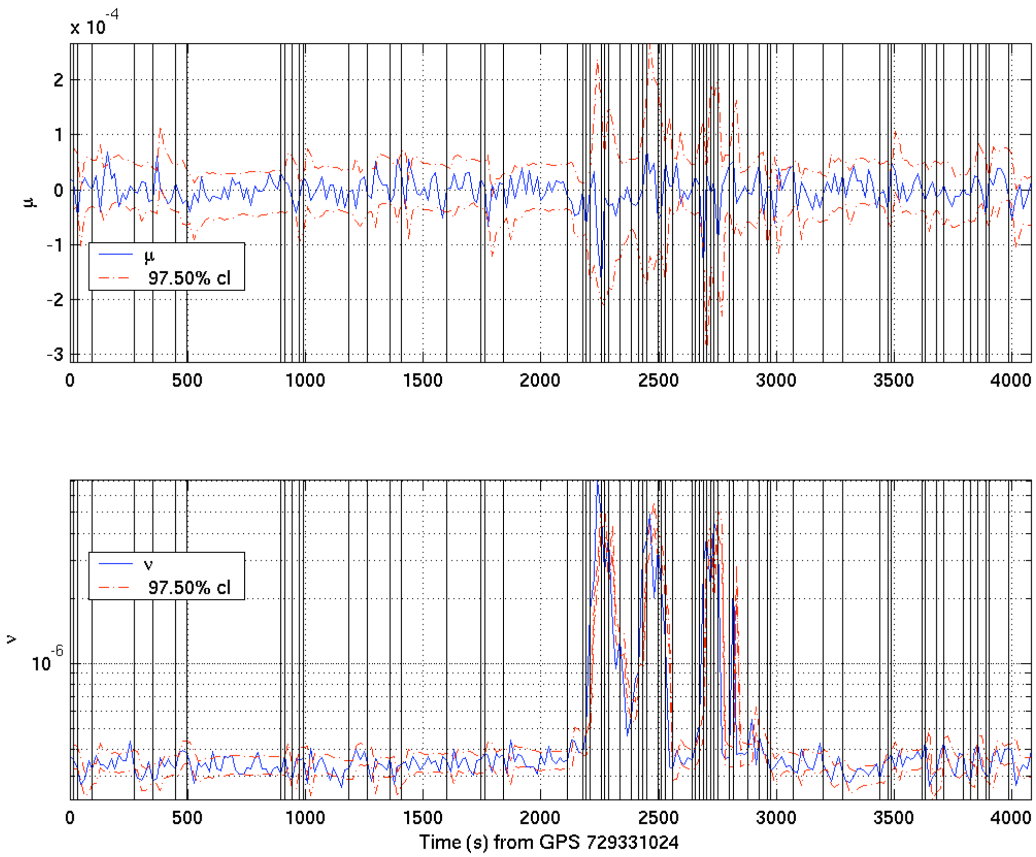
- Band
  - Focus attention on band $[f_l, f_h]$
- Averaging time
  - Interval duration τ on which statistics are accumulated
- Percentile cut
  - Define bulk distribution by throwing out top, bottom p percentile of band-limited samples
- Number of bins for χ^2 fit to Gaussian
  - Bins chosen for constant probability per bin
- Stationarity criteria
  - Stationary epoch ends immediately before first of two consecutive intervals that are inconsistent (probability p) in either mean or variance with epoch cumulative mean, variance

Sample output: monitoring stationarity

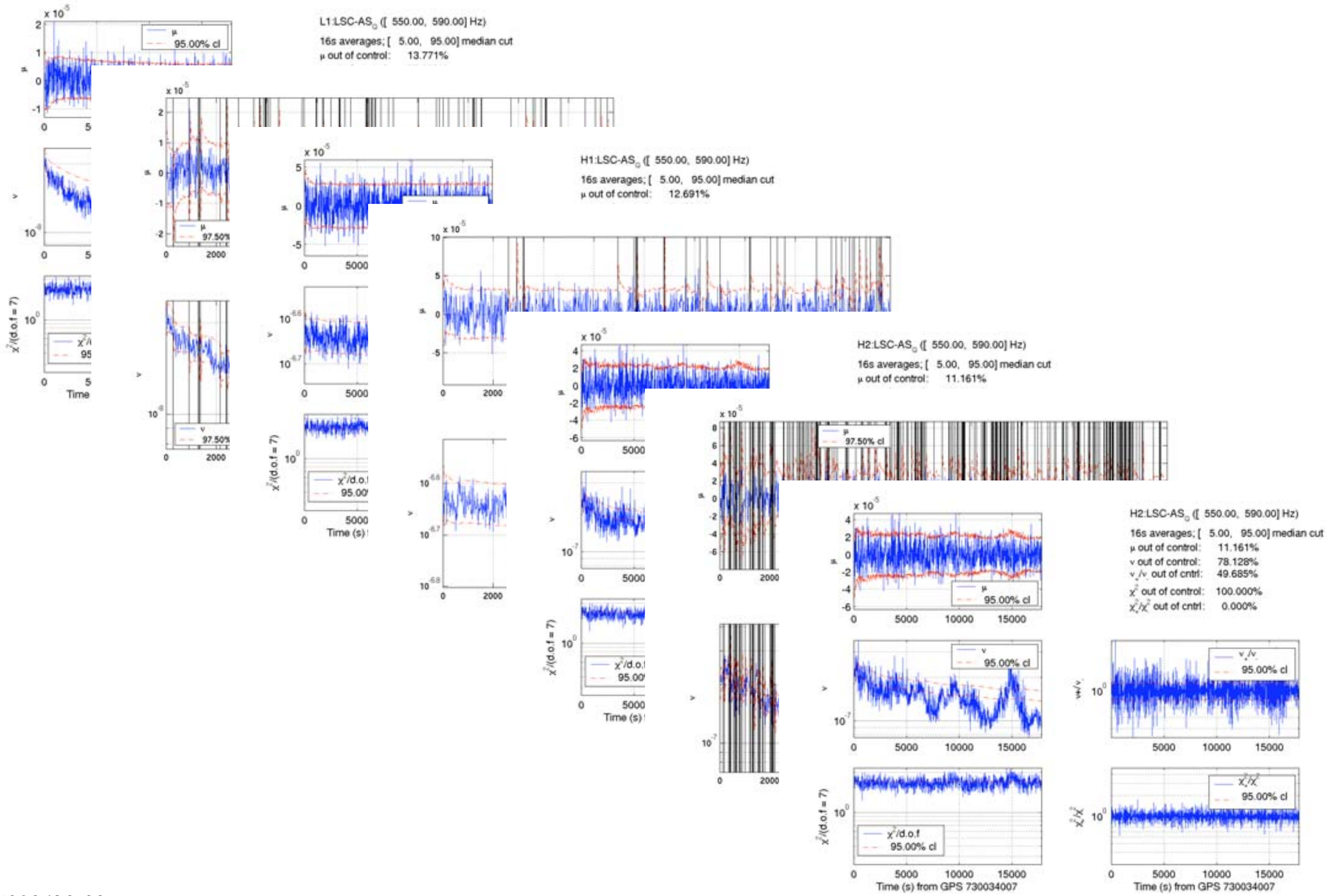


LIGO-G030420-00-Z

Sample output: stationary epochs

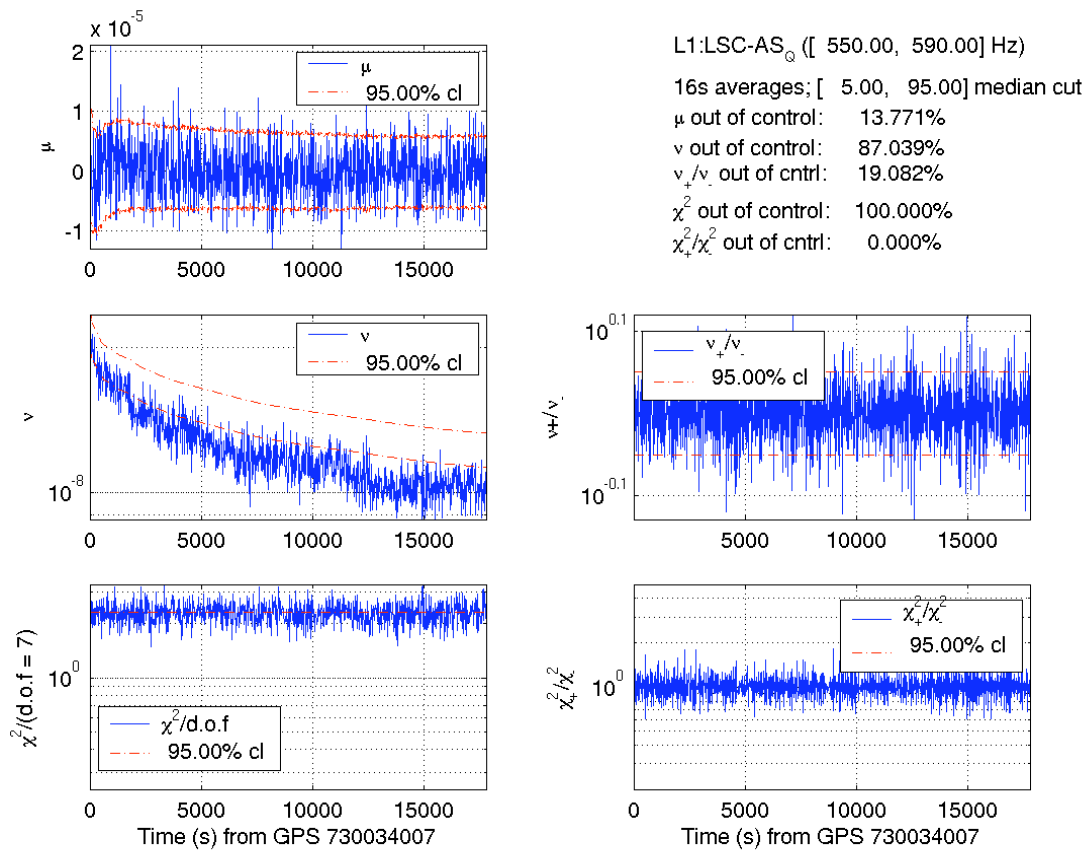


Anecdotal look at S2: Longest triple science mode segment



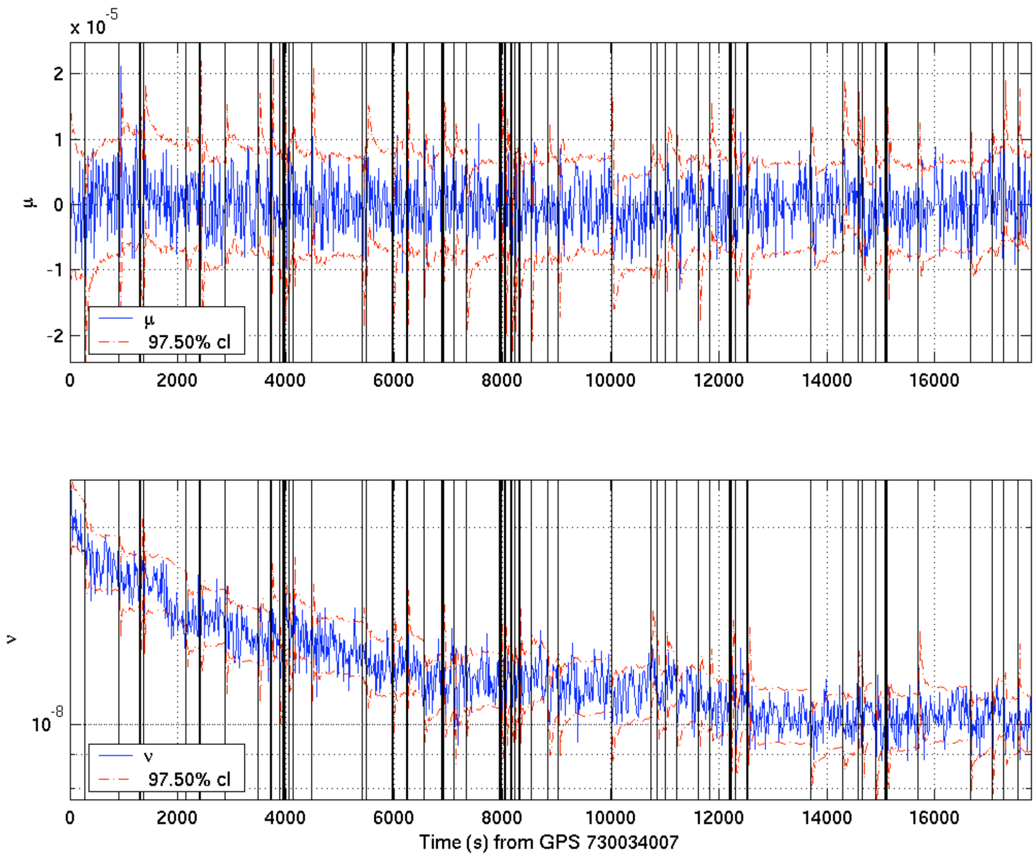
LIGO-G030420-00-Z

L1 Stationarity Monitor: Longest Triple Science Mode

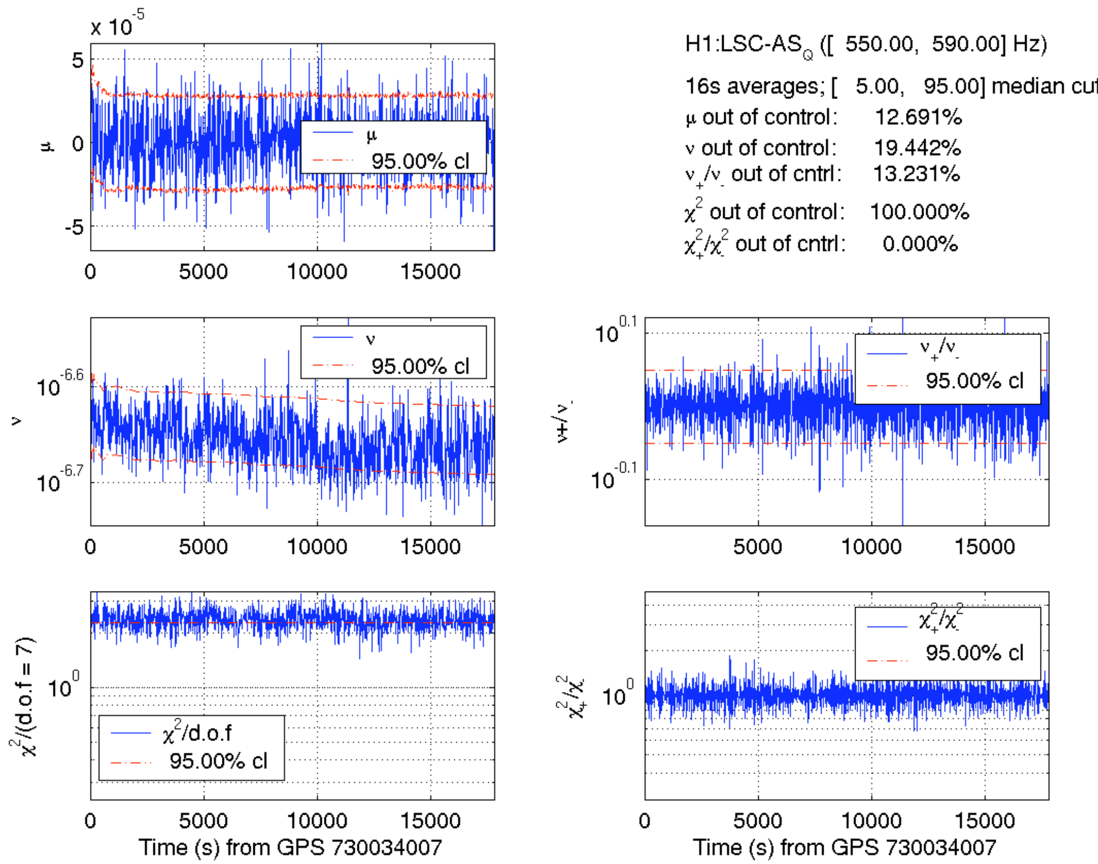


LIGO-G030420-00-Z

L1 Cuts:
Longest Triple Science Mode

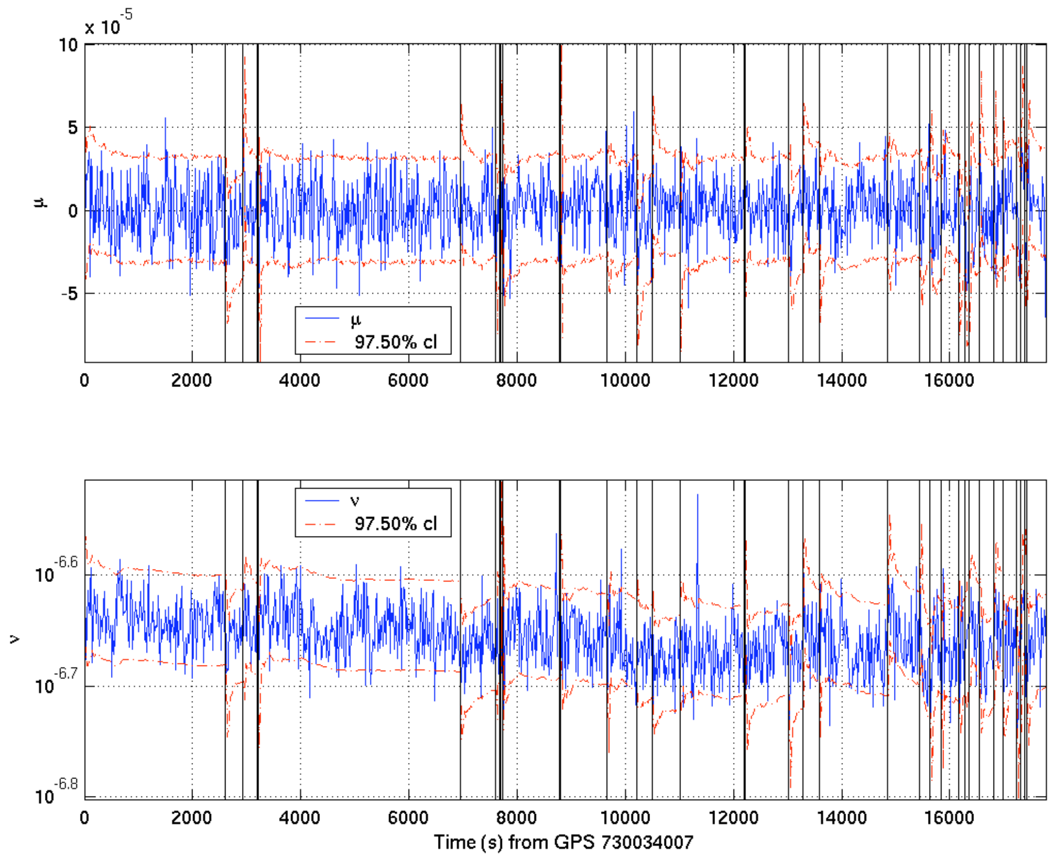


H1 Stationarity Monitor: Longest Triple Science Mode

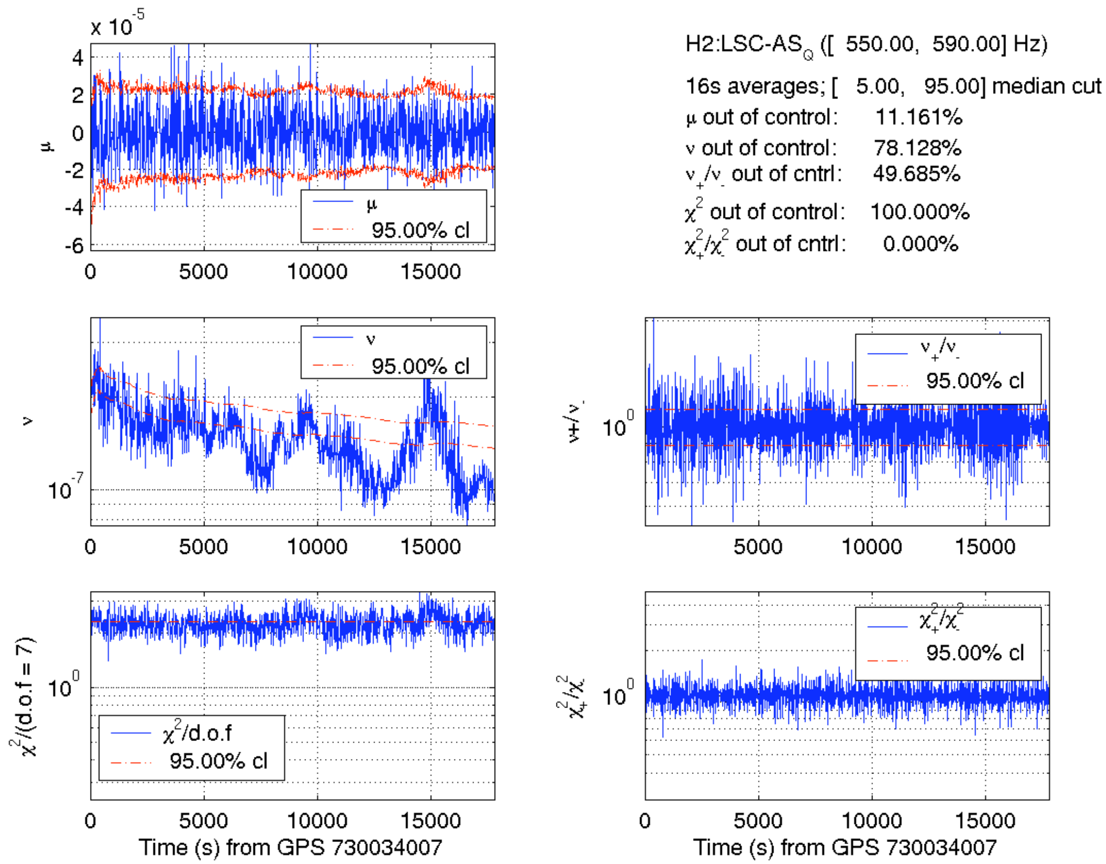


LIGO-G030420-00-Z

H1 Cuts:
Longest Triple Science Mode

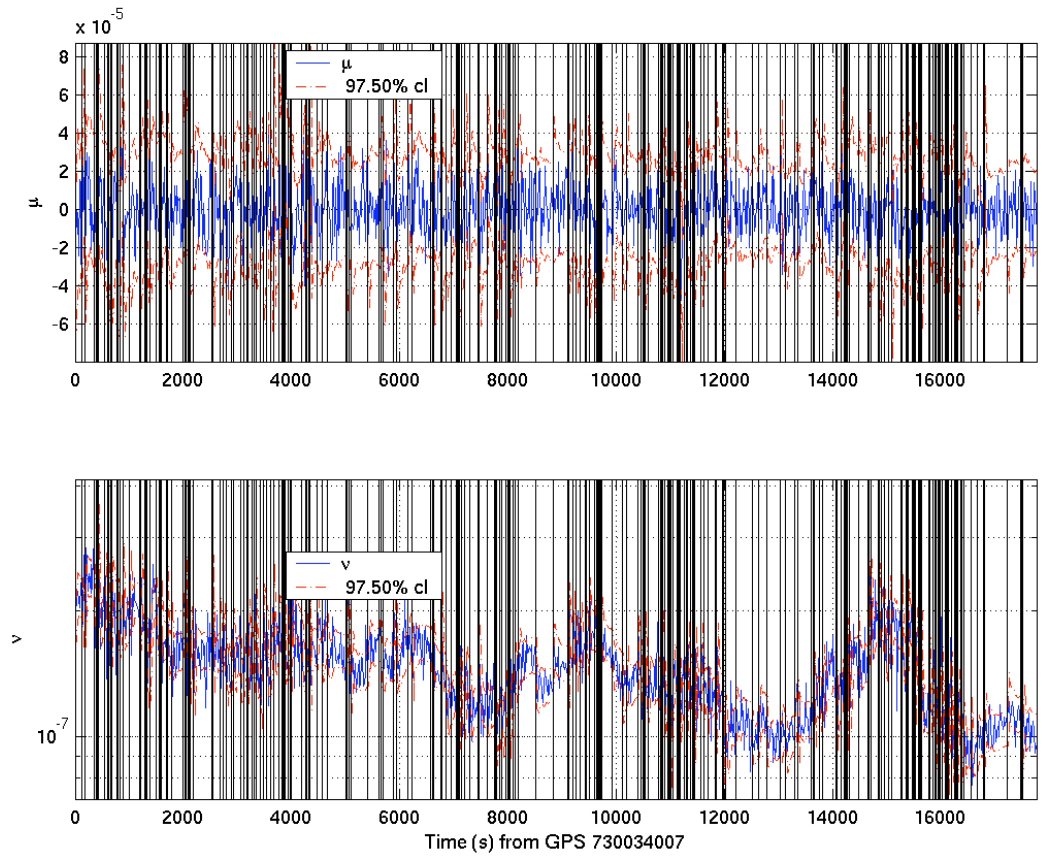


H2 Stationarity Monitor: Longest Triple Science Mode

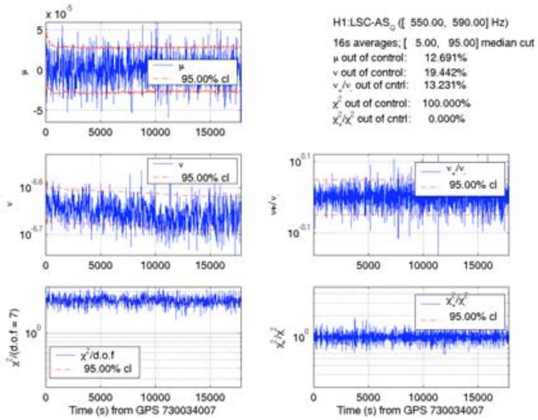


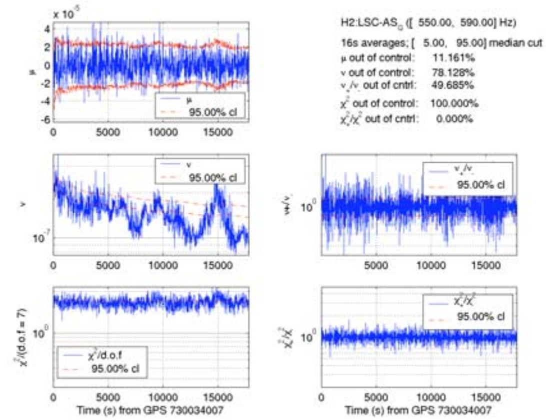
LIGO-G030420-00-Z

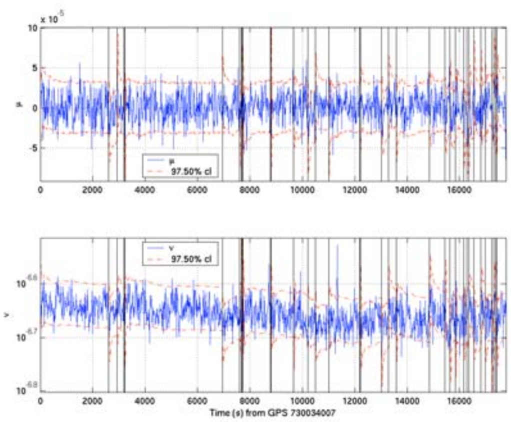
H2 Cuts: Longest Triple Science Mode

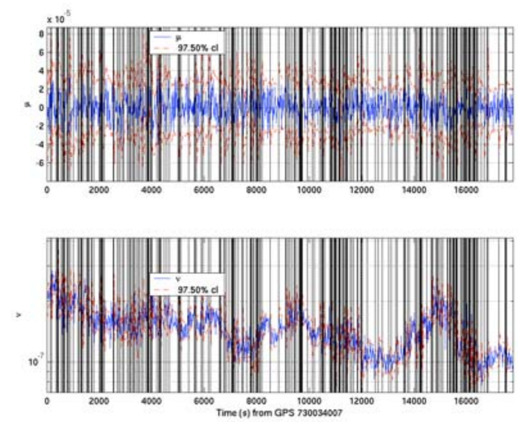


H1 & H2: Contrast



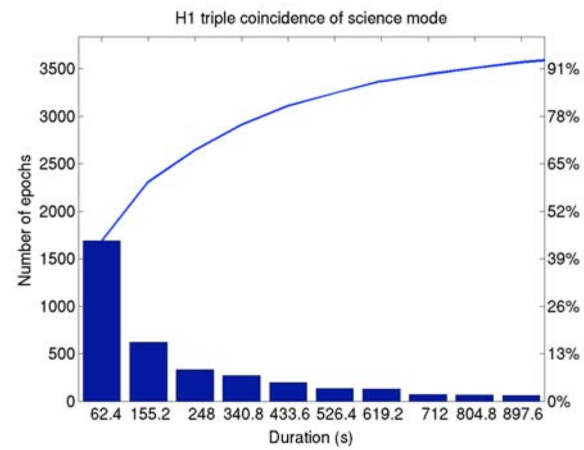


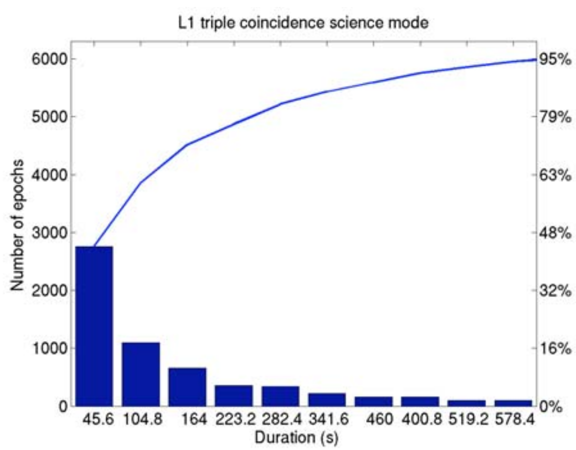


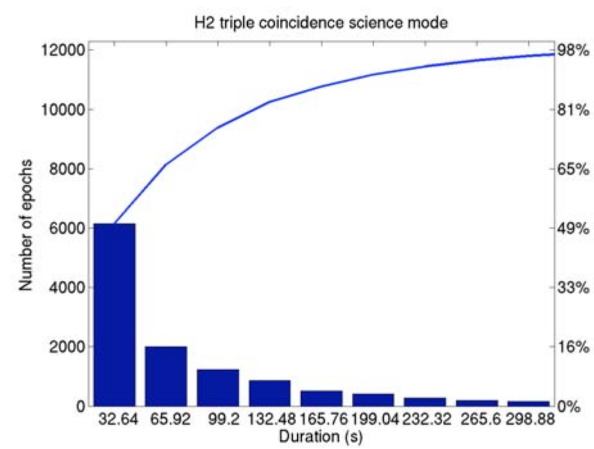


Stationary Epoch Durations

- [550 Hz, 590 Hz]
- 16 s averaging time
- [5, 95] percentile cut
- 97.5% confidence stationarity criteria

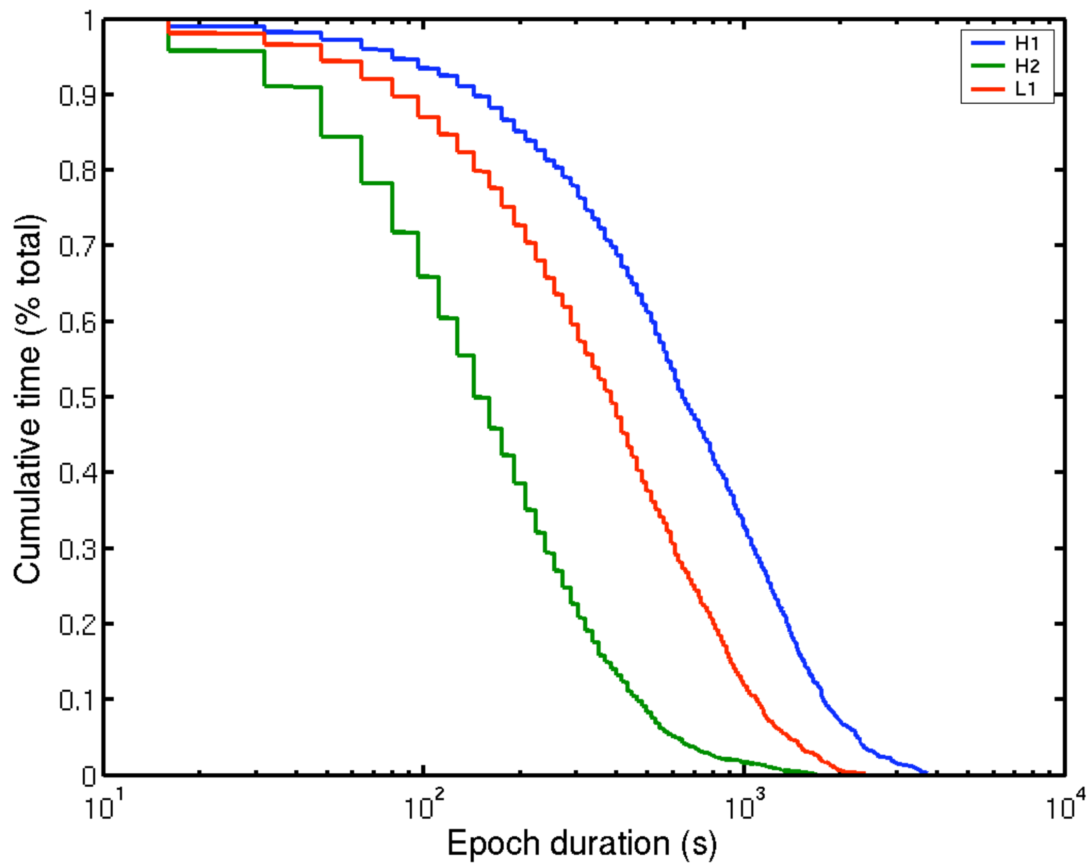






LIGO-G030420-00-Z

Stationary Epoch Durations



Summary

- S2 noise shows ***strongly variable, frequent outliers***
  - » Trimming samples outside of [5,95] percentile on 16s data in [550,590] Hz band reduces range of variance by factor 35 or more
- S2 noise shows ***strong variability on minute*** timescales
  - » 50% of H2/L1/H1 *epochs* shorter than 64/96/144 s: expect median durations of ~750s for Gaussian stationary noise
  - » 50% of H2/L1/H1 *durations* in epochs shorter than 144/384/640 s: expect ~1780 s for Gaussian stationary noise
- Science results seriously affected by noise non-stationarity
  - » Sophisticated analyses can account for variability *if not too rapid*
  - » Greater, more rapid variability, weaker conclusions
- medNoiseChar available as Matlab script and as a standalone Ap that runs on frame data
  - » Same concepts being applied to calibration lines: cf. lineamp being developed by Mike Ashley