Data Analysis and Detector Characterization Leading to Gravitational Wave Astronomy

L. S. Finn

Penn State

LIGO-G000418-00-D

Outline

- Penn State Group
- LSC Activities
- Proposed research

Penn State Group

- Faculty
  - LSC:
    - Finn (Theory/analysis),
    - Gonzalez (Experiment/analysis)
  - Non-LSC:
    - Ashtekar (classical and quantum gravity)
    - Pullin (numerical relativity)
    - Laguna (numerical relativity)
- Theory/Analysis
  - Post-docs
    - Sutton (NSERC Fellow, new to field)
  - Graduate students
    - Van den Broeck, Huckans (both first year)
  - Undergrad. Students
    - Hepler, Hsu, Rotthoff, Shapiro, Winjum

LSC Activities

- Member, Collaboration Council
- Member, Software Coordination Committee
- Co-chair burst upper limit analysis group
  - With P. Saulson
- Team Lead, Data Conditioning API Implementation
- Member, Executive Committee

Proposed activities

- Data conditioning: preparing data for analysis
  - Regression (including adaptive line removal)
  - Lead datacondAPI development
- Data characterization
  - Descriptive statistics
    - Higher-order distribution moments/cumulants
    - KDD istribution estimation
  - Parametric power spectrum estimation

Proposed activities, cont'd

- Data Analysis
  - Analysis in presence of non-Gaussian, stationary noise
    - Likelihood function estimation incorporating higher-order moments
  - Aperture Synthesis
    - Synthesizing a larger detector from several smaller ones
  - Upper limit physics
    - Searching for unmodeled bursts
    - Periodic signals from pulsars

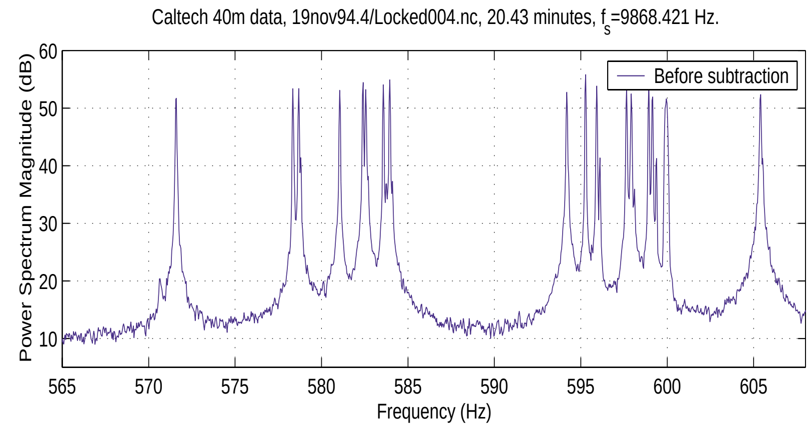
Data conditioning

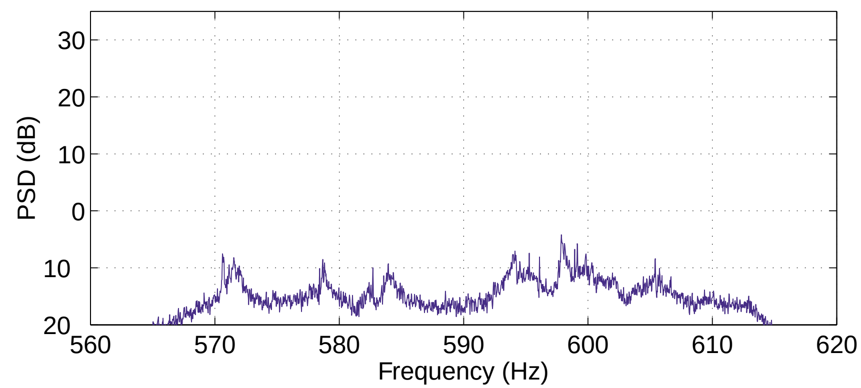
- Analyzed data should be free of instrumental artifacts
  - Eliminate noises that can be eliminated
- Analyzed data should be white
  - Technical advantages in numerical analysis speed, accuracy
- Analyzed data should focus on signal band

- Data conditioning: preparing data for analysis
  - Drop-out correction
  - Regression
    - Violin modes; power-main features; seismic, other disturbances
  - Whitening and power spectrum estimation
  - Basebanding; resampling

Data conditioning and analysis

- How important are data artifacts?
- Focus: LIGO 40 M data
  - 570-610 Hz band
  - Violin modes, 600 Hz power main feature
- Remove artifacts
  - Kalman filter for violin modes
  - Regress power main against magnetometer
  - Note mean square in-band noise significantly reduced

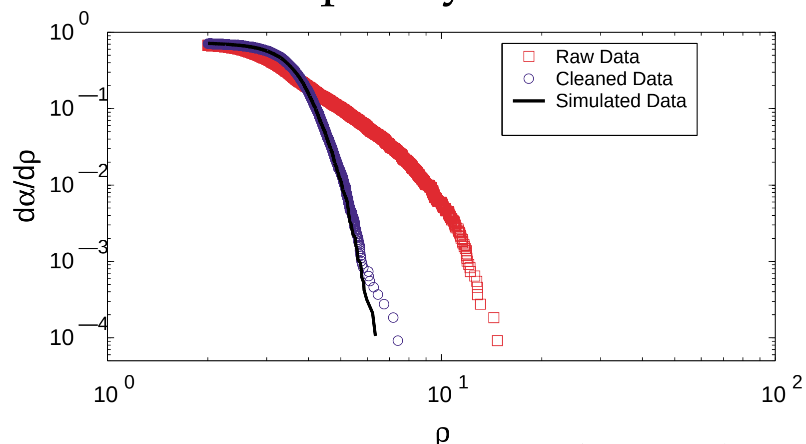




Data conditioning and analysis

- How important are data artifacts?
- Focus: LIGO 40 M data
- Remove artifacts
- Optimal filter for BH formation
  - "Raw" data, data with artifacts removed, simulated Gaussian noise

- Results
  - Data quality …



  - … dramatically through proper conditioning
  - … increasing sensitivity, detection efficiency & confidence

Data Conditioning API Development

- datacondAPI
  - LDAS Subsystem
  - Prepares data for all LIGO production analyses
  - Shared responsibility for statistical characterization
  - Touches all data in 7x24 operations
- Completed
  - Infrastructure
    - Linear filtering, FFT, decimation, basebanding,
  - Descriptive statistics
    - Mean, variance, power spectrum estimation
- Beginning
  - *Regression, line removal*
  - Drop-out/veto management
  - *Signal id tools*
- Development Team
  - ANU: Searle
  - LIGO/CIT: Blackburn, Charlton, Ehrens, Lazzarini, Maros, Salzman
  - PSU: *LSF* (team lead), Rotthoff, Shapiro, Hepler,
  - UT Brownsville: Romano

Data Characterization

- Statistical characterization of stationary noise
  - Tests of stationarity
  - Distribution moments/cumulants
  - Parametric models of power spectra for filtering, instrument characterization
- Non-stationary artifact identification
  - "Burst" identification, classification using "AI" tools

Power Spectrum Estimation

- Principal statistical characterization
  - Also required for analysis
- Welch estimate
  - PSD ~ <|DFT(s.*w)|^2>
- Welch estimate non-parametric
  - But principal spectral features described by simple transfer functions
- PSD as filter modulus
  - Welch vs. MA model
    - y(z) = B(z)e(z)
    - PSD(y)~|B($e^{2\pi i f}$)|^2

Power Spectrum Estimation

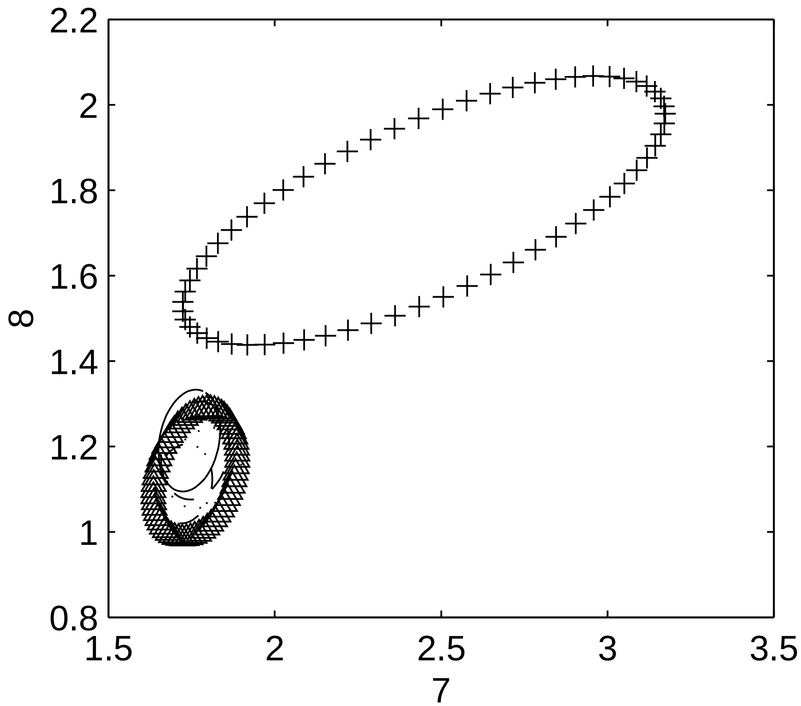
- PSD as filter modulus
  - Welch vs. MA model
    - $y(z) = b(z)e(z)$
    - $PSD(y) \sim |b(e^{2\pi i f})|^2$
- Alternative models
  - ARMA Model
    - $a(z)y(z) = b(z)e(z)$
    - $PSD(y) \sim |b/a|^2$
  - AR Model: $B(z) = 1$
- Advantages
  - Fewer model parameters, more accurate estimates
    - Model params: poles, zeros, gain
  - Model accuracy readily assessed
    - Test $e = ay/b$ for whiteness
  - Model parameters have physical interpretation
    - Characterize instr. State

Burst artifact identification

- What distinguishes noise “bursts” from stationary background?
  - Burst: short duration change in noise character
  - Noise before, after, has identical character
- Search for bursts is search for change
- Segment input signal
  - E.g., 1/8th second sub-intervals
- Develop statistics on sub-intervals
  - E.g., power, max amplitude in sub-bands
- Fit distribution to segment statistics
  - E.g., mixture Gaussian
- Find outliers
  - Segments that are “unusual” in context of overall distribution

Example

- April engineering run
- 6 statistics/segment
  - 5-level wavelet decomp.
    - Essentially a fully-reconstructable octave analysis
  - Max(abs(d)), Max(abs(a))
    - d = wavelet detail
    - a = wavelet approxim.
- Multivariate mixture Gaussian model
  - Every segment drawn from one of N m.-v. Gaussians
- Shown: iso-prob. contours for two statistics

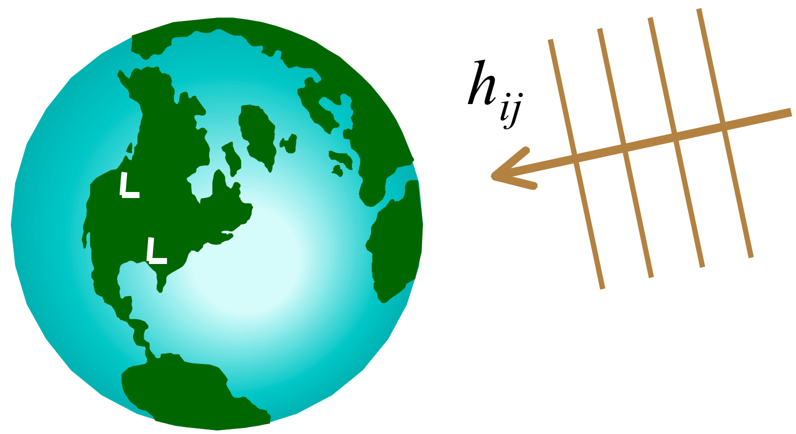


Data Analysis

- Multiple detectors and aperture synthesis
- Data analysis in presence of stationary, non-Gaussian noise
- Analyses
  - Phenomenology
  - Unmodeled burst analysis
    - Using KDD techniques described above
  - Periodic signal analysis
    - Focus on neutron star structure (crust strength)
  - Supernovae: target of opportunity

Aperture synthesis for gravitational wave detectors

- GWave detectors are phase sensitive
  - IFO and bars
- Multi-detector response is phase coherent
  - Phase difference depends on source RA, dec.
- Synthesize larger aperture by interfering response
  - Narrower antenna beam, higher sensitivity



- Network Analysis Development Team
  - Blackburn, Dhurandhar, *Finn*, Lazzarini, Marka, Mendell, Mours, Searle
  - Just getting underway

Data analysis and stationary, non-Gaussian noise

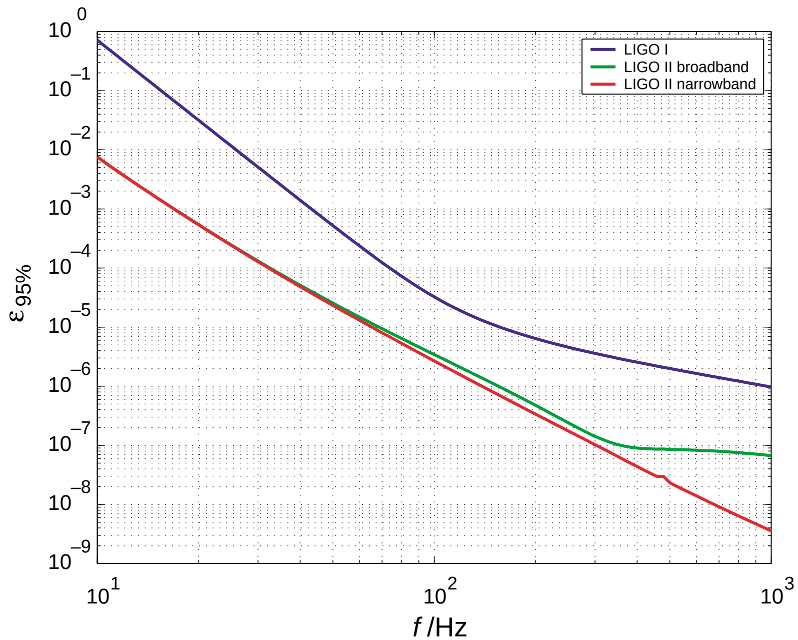
- Likelihood is foundation of statistical analysis
  - Prob of observation under fixed hypothesis
  - Aka sampling function
- Likelihood is estimated
  - E.g., noise mean, variance determines L for Gaussian noise
- Construct likelihood from estimates of higher-order moments
  - Use information-theoretic criterion to make minimum information assumptions on unknowns

Core-collapse supernovae

- Assume:
  - Supernova in galactic neighborhood
    - Neutrinos fix collapse to within 1s
  - Waveform unknown
    - Focus on detector-detector x-correlation
  - 1 KHz signal bandwidth
- Reach:
  - Mass fraction ε converted to gravitational waves
- LIGO I:
  - $\varepsilon_{95\%} < 2\mathrm{x}10^{-4}$ for SN at 55 Kpc
    - Expected rate 1/30y
- Target of opportunity

Periodic Signals

- Focus on pulsars
  - $f_{gw} = 2f_{pulsar}$
  - $h \propto \varepsilon = (\Delta I / I)$
- Reach: upper limit on ε
  - 1 yr observation
  - 10 Kpc distance
  - Declination average
  - Significance: 95%
- Theoretical prejudice
  - $\varepsilon < \sim 10^{-6}$
    - From pure Coulumb lattice crust strength
- Observational constraints
  - $\varepsilon < \sim 10^{-8}$ for *old* (recycled) pulsars



Summary

- Data Conditioning
  - Removing instrumental, environmental artifacts
    - Regression, line removal,
    - datacondAPI
- Data Characterization
  - Stationary noise
    - Distribution estimation
    - Parametric models for PSD estimation
  - Non-stationary noise
    - Burst identification
- Data Analysis
  - Upper limit physics
    - Unanticipated burst sources
    - Pulsars
    - Supernovae
  - Multiple detectors and aperture synthesis
  - Data analysis in presence of stationary, non-Gaussian noise