Materials Presented at LSC Meeting

A. Lazzarini

2000 March 16 - 18

LIGO Livingston Laboratory

LDAS Collaborative Activities

LDAS Collaborations

- LSC - Data analysis software for LDAS pipeline
  - » PSU, UTB, ANU - work on dataConditioningAPI
    - – C++ code directly merged with LDAS API, as specified in dataConditioningAPI requirements & specification documents
    - – Wraps FFTW for PSDs FFTs (ANU)
    - – Time domain filtering, decimation of data, heterodyning (PSU)
    - – Data summaries (trends): P-P, μ, σ, median, etc., by channel (UTB)
      - For ingestion into metadatabase on framed data
    - – Targeting initial end-to-end test as part of a dataConditioning Mock Data Challenge being developed as part of an integrated Lab-LSC software development schedule.

LDAS Collaborations

- LSC - Data analysis software for LDAS pipeline
  - » UWM, Cardiff(GEO) - work on MPI-based template filtering
    - – Conforming MPI-based temlate analysis to LAL style specification
    - – Working with LDAS to identify, iterate on S/W interfaces between LDAS (C++ wrapper and Tcl command layer) and MPI template filktering code (C, per LAL specification)
    - – Targeting a mid-April first pass at LDAS + MPI template analysis on linux cluster

LDAS Collaborations

- Model for working closely with LSC groups is forming well
  » Small working groups to produce specific software
    – Daily email discussions
    – Weekly in-depth teleconferences to review week-to-week progress, address issues & questions
  » LDAS programmer support & participation as needed
  » WWW sites for discussions, sharing code, documents
  » CVS accounts at LIGO/LDAS for submitting, sharing, controlling and maintaining software

LDAS Collaborations

- International - VIRGO
- Frame format revision
  - » Pending documentation in T970130, Rev. D
  - » Pending as Frame Format Release 4
  - » Introduces new structures: Table of Contents, Table
  - » Modifies others, e.g. FrDetector to accommodate multiple interferometers in one frame, anticipates network analysis by allocating IDs other major detectors: GEO, TAMA, Bars, prototypes.
  - » I/O libraries being revised
    - – VIRGO's Fcl (C version)
    - – LIGO's Framecpp (C++ version)

LDAS Collaborations

- International - ANU
- Co-investigators with ANU proposal to Australian Research Council (ARC) for data analysis support to ANU
  - » Shared postdoctoral fellow
  - » Support for combination of work:
    - – Research identified in LSC Whtite Paper on Data Analysis
    - – Close work with LIGO/LDAS to implement C++ APIs
      - Possibly port LDAS to ANU facility

LDAS Collaborations

- GriPhyN (Grid Physics Network) proposal to NSF's Information Technology Research (ITR) Program
  - » Collaboration composed of NSF-funded research projects seeking to enhance computational and database resources
    - – HEP - US collaborations at LHC - CMS & ATLAS
    - – Gravitational physics - LIGO - Lab (Prince, Lazzarini, Williams/CACR) and UWM (Allen)
    - – Astronomy - Sloan Digital Sky Survey (SDSS)
    - – Computer science - Chicago, Berkley, others
  - » Submitted pre-proposal on 30 Dec 1999 ($14.5M)
    - – One of 920+ proposals
    - – 120 invivted by NSF to submit full proposals by 17 April 2000
      - 120 pre-proposals request ~ $600M
      - ITR program has ~ $90M earmarked

LDAS Collaborations

- GriPhyN proposes to prototype a grid of computer, database resources to permit users across US to access data, computing facilities efficiently
- CS R&D:develop/define new concepts for resource management, allocation, scheduling on a global scale
  - » Most of effort to go initially in developing protocols, middleware software that manages resourcs, user requests
  - » “Hugely scaled up version of LDAS”
- Physics R&D: users of technology
  - » Use models, testbeds, benchmarks, ...

Higher Order Statistics to Characterize Non-Gaussian Stationary Noise

A report of work performed by Denis Petrovic (SURF 1999)

Noise Characterization & Higher Order Statistics (HOS)

» Cumulants and polyspectra for zero mean real processes

$$C_{2x}(k) = E\{x(n)\ \ x(n+k)\}$$

$$C_{3x}(k,l) = E\{x(n)\ \ x(n+k)\ \ x(n+l)\}$$

$$C_{4x}(k,l,m) = E\{x(n)\ \ x(n+k)\ \ x(n+l)\ \ x(n+m)\} - (C_{2x}(k)\ \ C_{2x}(l-m) + C_{2x}(l)\ \ C_{2x}(k-m) + C_{2x}(m)\ \ C_{2x}(k-l))$$

» S_{nx}- n-th order spectrum, C_{nx}- n-th order cumulant

$$S_{nx}(f_1, f_2, \ldots, f_{n-1}) = \boldsymbol{F}\{C_{nx}(\tau_1, \tau_2, \ldots, \tau_{n-1}\}$$

» Third order cumulants and bispectrum (third order polyspectra) of a zero mean real random processes

$$C_{3X}(k,l) = E\{x(n)\ \ x(n+k)\ \ x(n+l)\}$$

$$S_{3x}(f_1, f_2) = E\left\{\hat{x}(f_1)\ \ \hat{x}(f2)\ \ \hat{x}^*(f_1 + f_2)\right\}$$

$$\hat{x}(f) = \boldsymbol{F}\{x(t)\}$$

Properties of HOS

- Properties of polyspectra
  - For order (n>2) generally complex functions
  - Zero for order (n>2) if the process is Gaussian
  - Power spectrum - first in a series of polyspectra
- Motivations for using HOS techniques
  - Identify non-Gaussian signals by suppressing additive Gaussian noise
  - Extract information due to deviations from Gaussianity
  - Detect and characterize non-linear properties in a signal
  - Implement a real-time HOS filter to characterize LIGO data
- - Bicoherence as a metric of nonGaussian characteristics

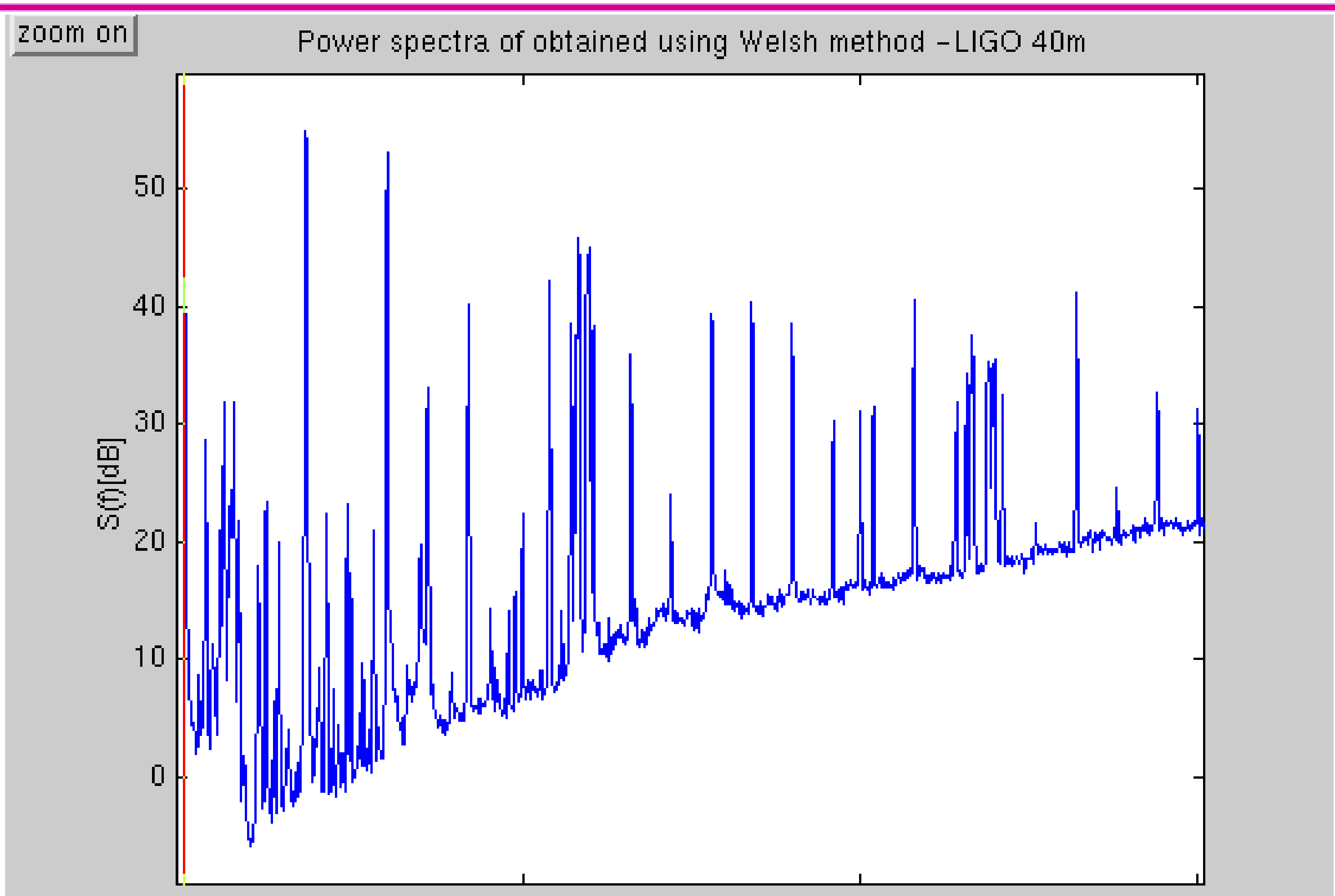
$$bic(f_1, f_2) \equiv \frac{|S_{3x}(f_1, f_2)|}{\sqrt{S_2(f_1)\ S_2(f_2)\ S_2(f_1 + f_2)}}$$

LIGO-G000021-00-E

Application of HOS to 40m data

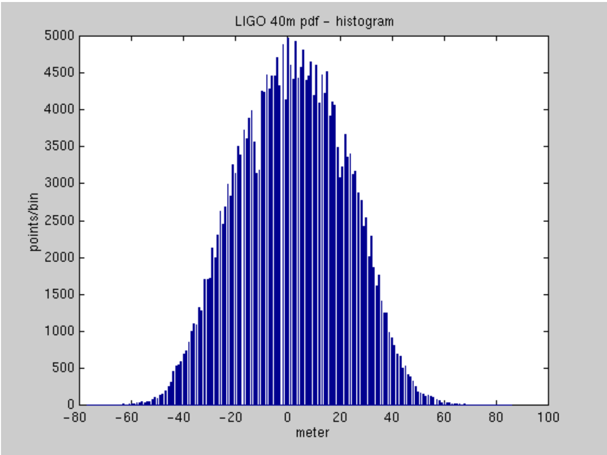
- Use LIGO 40m power spectrum to identify characteristic features
  - LIGO 40m power spectrum is rich in harmonics
  - Presence of the sidebands around some of the harmonics
- Generate simplified models of a signal with these“features”
- Generate bispectral representation of the signal
- Identify and understand characteristic patterns in bispectrum
- Generalize to look for similar features in LIGO data

40m whitened strain spectrum (Nov 1994)

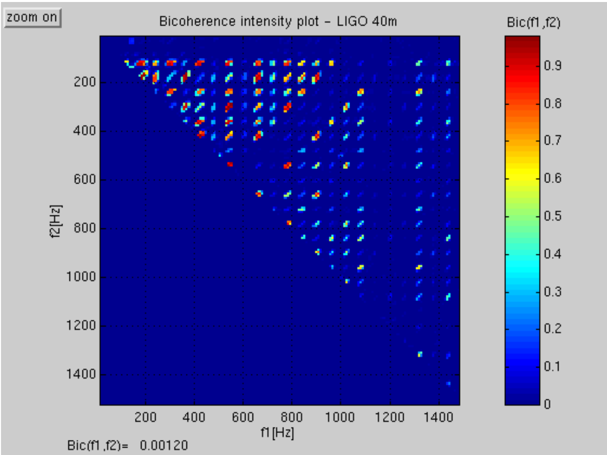


Probability Distribution Function and Bicoherence of LIGO 40m data

- PDF of LIGO 40m data suggests that the process is Gaussian
- Presence of peaks in LIGO 40m data bicoherence



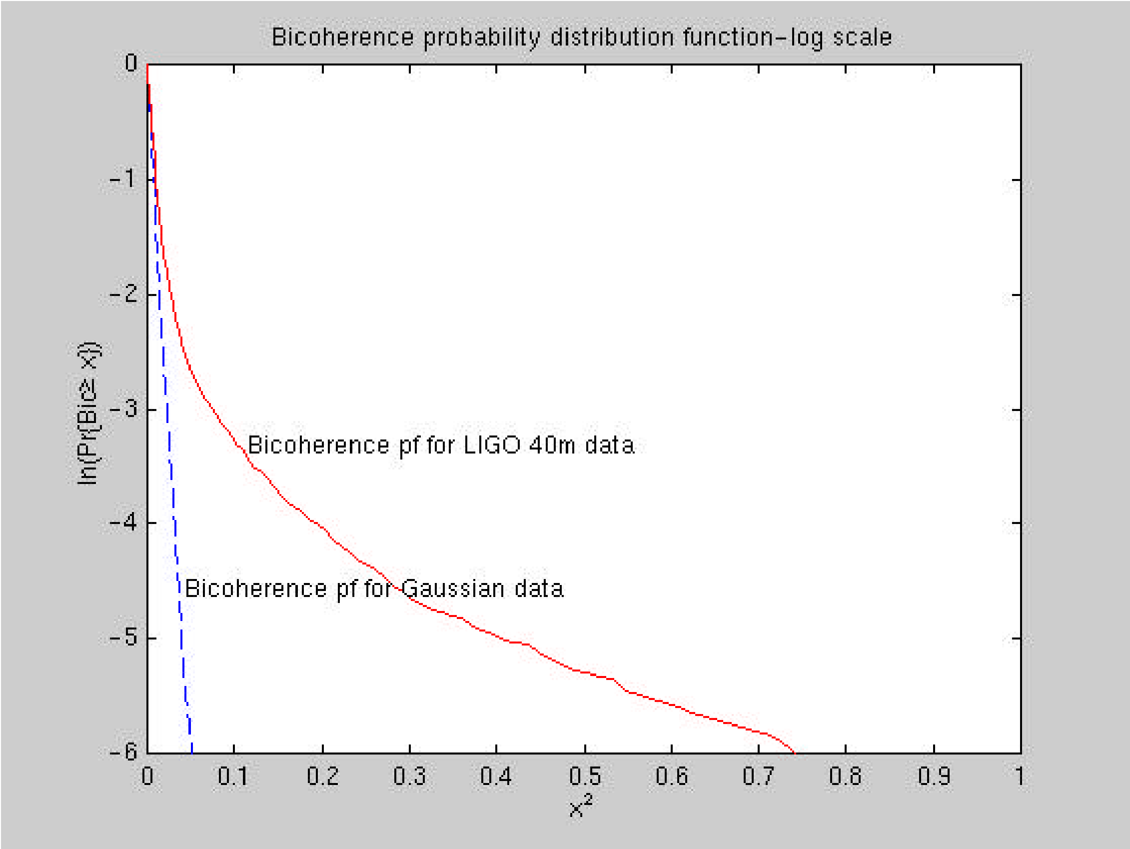
Histogram of strain signal



Bispectrum of strain signal

Non-Gaussian characteristics of 40m strain

- Histogram of the magnitudes of bic(f_1,f_2) exhibits substantial deviation from modeled Gaussian data



LIGO-G000021-00-E

Example Studies

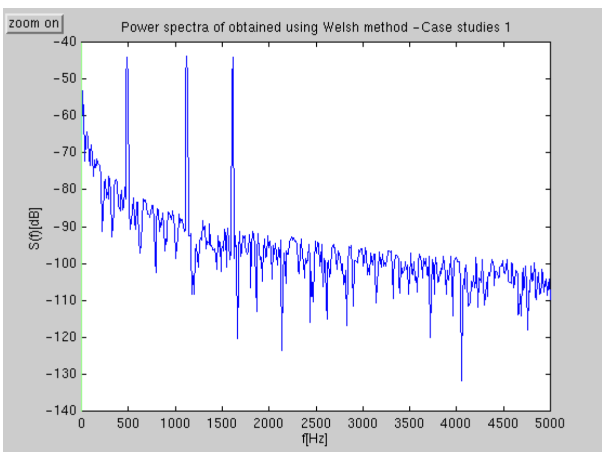
- Create time series data with known linear and non-linear characteristics to "calibrate" bispectrum, bicoherence

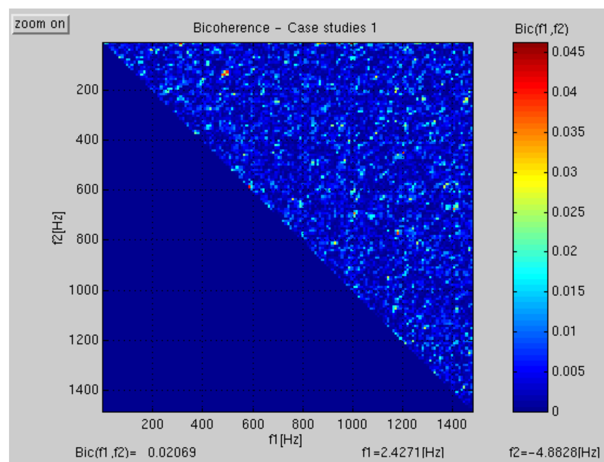
$$x(t) = x_1(t) + x_2(t) + B\ x_3(t) + Cx_1(t)\ x_2(t) + n_{Gaussian}(t)$$
$$x_1(t) = A\ Cos(2\pi f_1 t + \varphi_1)$$
$$x_2(t) = A\ Cos(2\pi f_2 t + \varphi_2)$$
$$x_3(t) = A\ Cos(2\pi f_3 t + \varphi_3)$$

  » A,B,C arbitrary constants;
  » ϕ_1, ϕ_2, ϕ_3 uniform RVs $[-\pi,\pi]$;
  » $n_{Gaussian}(t)$ colored (e.g., 1/f) noise

Example Studies:1

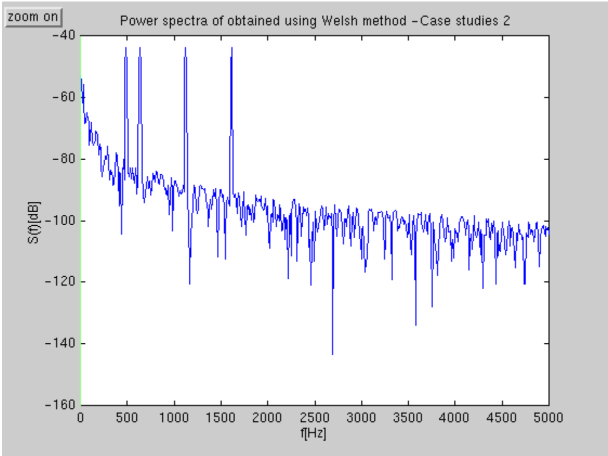
- Case 1: B=1, C=0, $f_3=f_1+f_2$
  - » Accidental occurrence of sum-frequency
  - » No phase coherence - ϕ_3 is independent of ϕ_1 , ϕ_2
  - » Signal consists of three independent harmonics
  - » - No features present in bicoherence spectrum

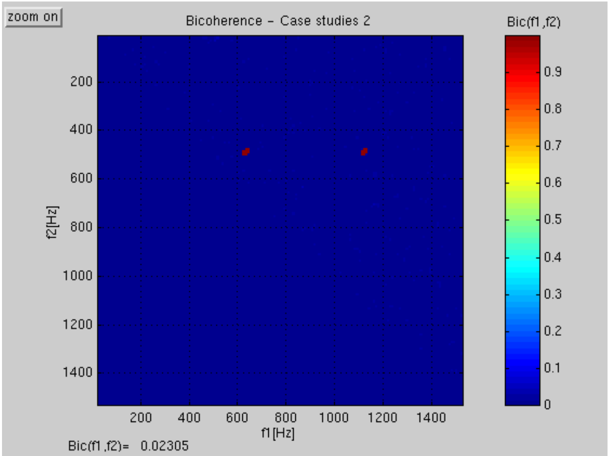




Example Studies:2

- Case 2: B=0, C=1
  - » Nonlinear mixing of “strong” lines creates of sum & difference frequency lines at $f_3=f_1 \pm f_2$ with phase coherence -
    - – ϕ_3 is related to ϕ_1 , ϕ_2
  - » Peaks in bicoherence occur as a result of the non-linear coupling between x_1 and x_2.





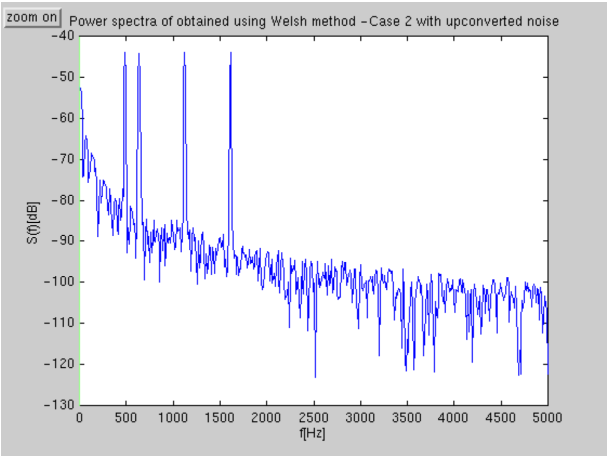
LIGO-G000021-00-E

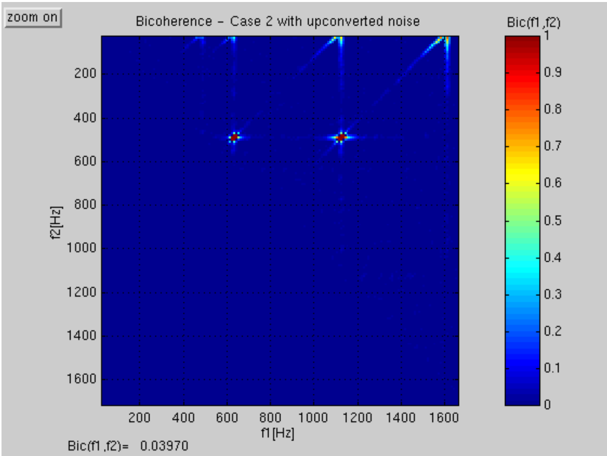
Example Studies:3

- Case 3: Upconverted broadband noise

$$x(t) = \left(1 + n_{Gaussian}(t)\right)\ \left(x_1(t) + x_2(t) + B\ \ x_3(t) + Cx_1(t)\ \ x_2(t)\right)$$

  - » Noise upconversion produces characteristic "wings" at base of peaks

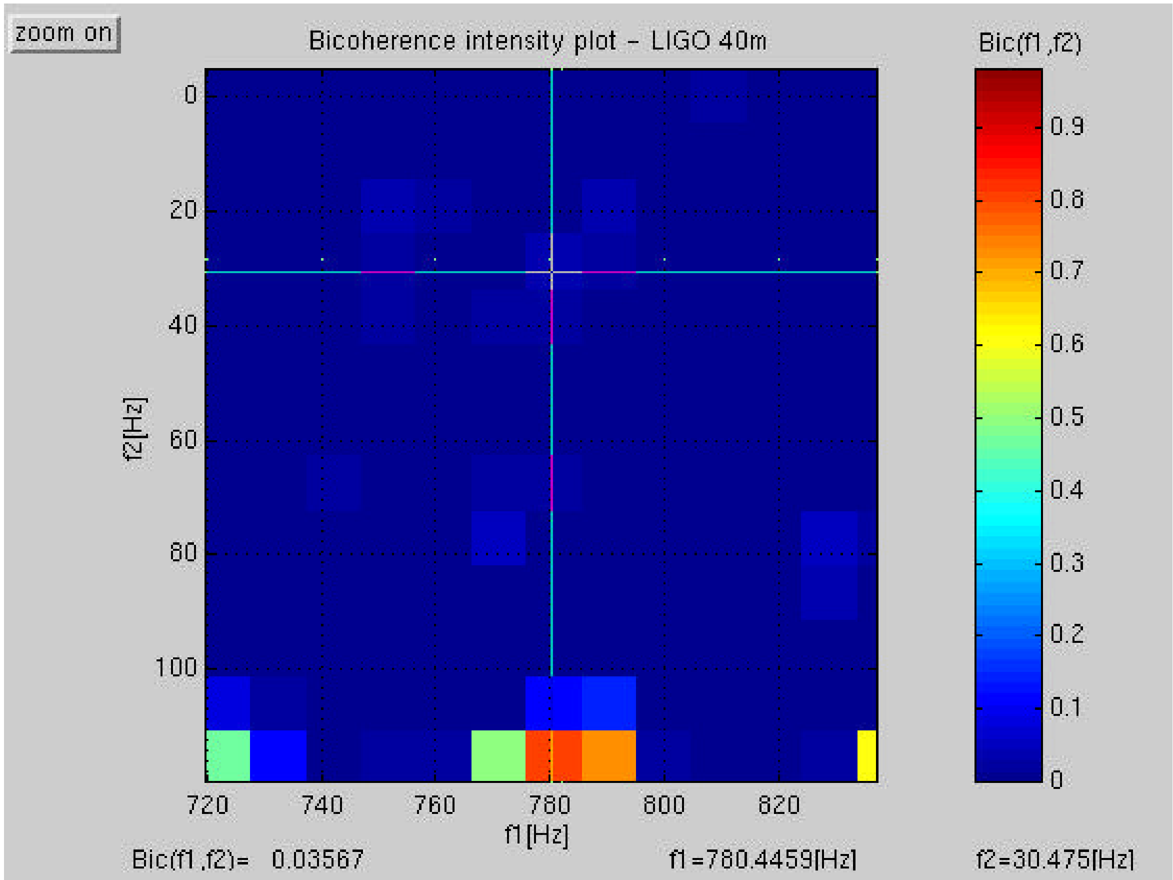




LIGO-G000021-00-E

Example study: 3

- Similar features
- are present in LIGO 40m data
- Bicoherence can detect upconversion of low frequency components in power spectrum



LIGO-G000021-00-E

Harmonic process - bispectrum

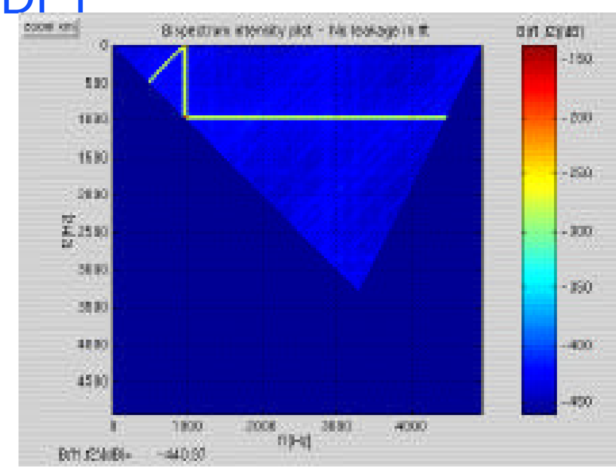
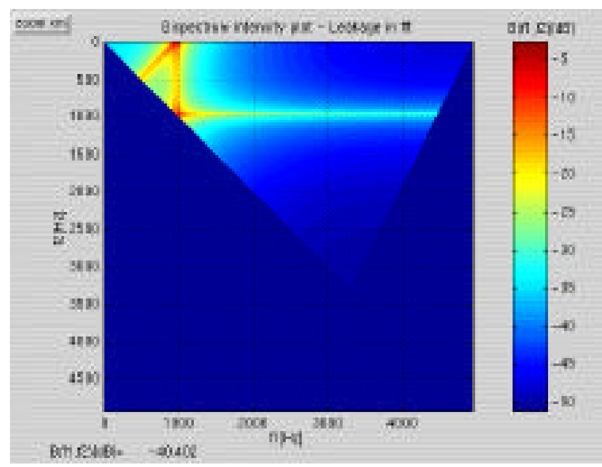
- Bispectrum of a harmonic process is theoretically zero
- Estimated bispectrum is an approximation of the formula

$$S_{3X}(f_1, f_2) = E\left\{\hat{x}(f_1)\ \hat{x}(f_2)\ \hat{x}^*(f_1 + f_2)\right\}$$

$$\hat{x}(f) \approx \delta(f - f_0)e^{i\phi} + \delta(f + f_0)e^{-i\phi}$$

$$\hat{x}(f) = \boldsymbol{F}(x(t)); \quad f_0 \; - \; frequency\ of\ a\ periodic\ signal$$

- Estimated bispectrum largely depends on spectral characteristics of DFT

LIGO-*G000021-00-E*

Fast Chirp Transform

Fast Chirp Transform , "FCT"

- Concept developed by Prince, prototyped by Jenet
  - » Submitted to PRD for publication, dated 2000.01.31
    - – *"Detection of Variable Frequency Signals Using a Fast Chirp Transform"*
  - » Motivated by desire for "on the fly" template generation
    - – Need in EM pulsar $\{f, \dot{f}, \ddot{f}, \ldots\}$ searches with EM dispersion
  - » Need to explore mass parameter space for which templates are not (yet) known -- allows "parameterized waveform" searches
  - » Need to estimate false alarms by looking at nonphysical regions in parameter space
- Work not part of ASIS scope
- Proposal to LSC being prepared to solicit interest in addressing validation, implementation, evaluation
  - » Majid, Jenet, summer student, interested LSC members

FCT Concept

- Generalize FT: $\chi_{FT}(t) = \int df \; h[f] \; e^{2\pi i f t} \longrightarrow \chi_{CT}(t) = \int df \; h[f] \; e^{i\phi(f)}$
- Express phase as series in *f:* $\phi(f) = 2\pi f\tau + \delta\phi(f); \delta\phi(f) = \sum_{m>1} k_m [f\tau_m]^m$
- Discretize to FFT, FCT:

$$\chi_{FFT}(k) = \sum_{j=0}^{N_0-1} h[j] \; e^{2\pi i (\frac{jk}{N_0})} \longrightarrow \chi_{FCT}(k,\{l_p\}) = \sum_{j=0}^{N_0-1} h[j] \; e^{2\pi i \left[\frac{jk}{N_0} + \sum_{p>1} l_p \left(\frac{j}{N_0}\right)^p\right]}$$

- 2-parameter example--quadratic chirp:

$$\chi_{FCT}(k,l) = \sum_{j=0}^{N_0-1} h[j] \; e^{2\pi i \left[\frac{jk}{N_0} + l\left(\frac{j}{N_0}\right)^2\right]} = \sum_{p=0}^{N_1-1} \sum_{j=j_{\min}(p)}^{j=j_{\min}(p+1)-1} h[j] \; e^{2\pi i \left[\frac{jk}{N_0} + l\left(\frac{j}{N_0}\right)^2\right]}$$

  - » Require for each interval $\{p,p+1\}$ that term $l(j/N_0)^2$ changes less than some acceptable maximum value, ε:

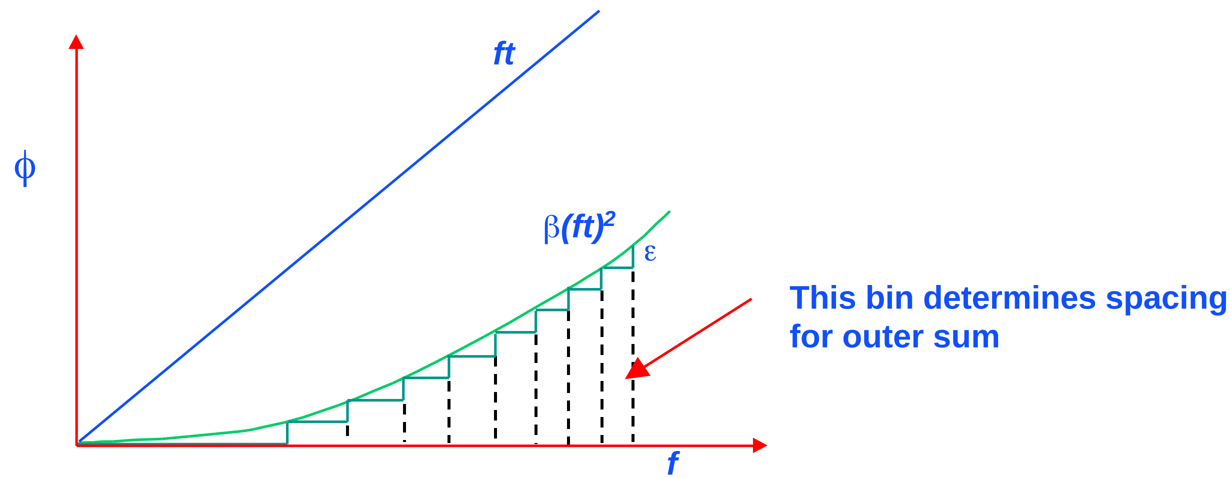
$$l\left[\left(\frac{j_{\min}(p+1)}{N_0}\right)^2 - \left(\frac{j_{\min}(p)}{N_0}\right)^2\right] < \varepsilon$$

FCT Concept

- With this constraint move nonlinear term outside inner sum:

$$\chi_{FCT}(k,l) \approx \sum_{p=0}^{N_1-1} \overbrace{e^{2\pi i l \left(\frac{j_{\min}(p)}{N_0}\right)^2}}^{\textbf{Function of } p \textbf{ alone}} \underbrace{\sum_{j=j_{\min}(p)}^{j=j_{\min}(p+1)-1} h[j]\, e^{2\pi i \left[\frac{jk}{N_0}\right]}}_{\textbf{Function of } j \textbf{ alone}}$$

  - » Limits on outer sum determined by how fast nonlinear term changes:



FCT Concept

- Convert nonlinear term to (an approximate) linear term (I.e., keep first term in Taylor series):

$$\left(\frac{j_{\min}(p)}{N_0}\right)^2 \longrightarrow \frac{p}{N_1};$$

  - » This determines value of N_1:

$$N_1 = \left\lceil \frac{N_0}{\Delta_{\min}} \right\rceil; \quad \Delta_{\min} \longrightarrow \underset{over\ \ j < N_0}{Min\ \ \Delta} \left\{ \frac{\left(\frac{j+\Delta}{N_0}\right)^2 - \left(\frac{j}{N_0}\right)^2}{\varepsilon} = 1 \right\}$$

  - » This ensures maximum error in nonlinear->linear approximation $\leq \varepsilon$

FCT Concept

- The final (approximate) expression is the “FCT”:

Function of *p* alone

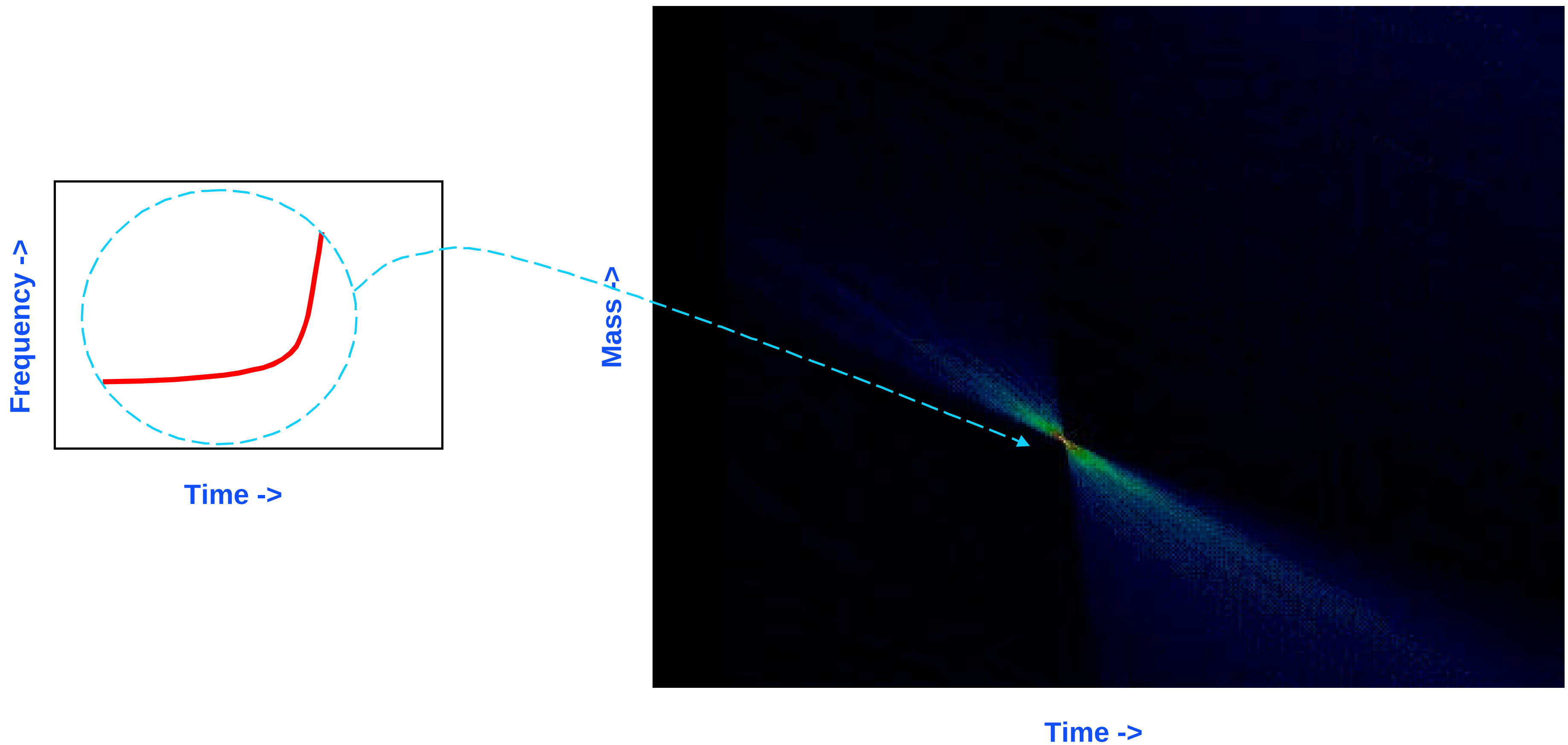
$$\chi_{FCT}(k,l) \approx \sum_{p=0}^{N_1-1} e^{2\pi i l \left[\frac{p}{N_1}\right]} \sum_{j=j_{\min}(p)}^{j=j_{\min}(p+1)-1} h[j]\, e^{2\pi i \left[\frac{jk}{N_0}\right]}$$

Function of *j* alone

FCT Implementation

- **Draft proposal prepared to LSC to get work started -- “mock data challenge”**
- **Full (inefficient) 2D FFT implemented for PN chirps, pulsars, quadratic chirps**
  - » Prototype works as advertised (!)
- **Issues that need to be addressed:**
  - » Optimization
  - » Parallelization within MPI (does its implementation “break” the present MPI-based pipeline baseline in any way?)
  - » Validation against existing techniques
  - » Identifying its role within the analysis plan
    - – Transforms template search to index-counting bookkeeping task for conventional FFTs
    - – Allows natural extension into regime where templates not known, do not exist
      - Monitor non physical parameter space to estimate false alarms,...

FCT for PN 1.4+1.4 Mass Chirp



LIGO-G000021-00-E

FCT for Quadratic Chirp + Gaussian Noise

LIGO-G000021-00-E