Search implementation for the gravitational wave stochastic background applied to the S1 LIGO I Science Run

Albert Lazzarini

for the LIGO Scientific Collaboration

06 May 2003

LIGO Science Seminar at Caltech

SGWB Working Group WWW SITE: http://feynman.utb.edu/~joe/research/stochastic/upperlimits/

Outline of Talk

- Stochastic GW background
- LIGO S1 run summary
- Search technique, implementation
- Details of analysis
- Results
- Conclusions

The Stochastic GW Background

- The stochastic GW background arises from an incoherent superposition of unresolved sources of gravitational radiation bathing Earth.
  - » Measure: Ω_{GW} -- energy density in Universe associated with GWs
  - » $\Omega_{gw}(f)$ -- frequency distribution of energy

$$\Omega_{\rm gw}(f) \equiv \frac{f}{\rho_{\rm c}} \frac{d\rho_{\rm gw}}{df}$$

- ***Cosmological sources***
  - » GW can probe the very early universe
  - » Inhomogeneities near Planck time, inflation
  - » Phase transitions
  - » Cosmological defects
- *Astrophysical sources*
  - » *NS/NS, WD/WD, periodic sources, SNe*
    - *$\Omega_{gw}(f) \leq 10^{-8}$ in LIGO band (Maggiore, gr-qc/0008027)*

Relationship between cosmological quantities and measurable quantities

- Power spectrum, $S_{gw}(f)$: $\lim_{T\to\infty} \frac{1}{T} \int_0^T |h(t)|^2 \, dt = \int_0^\infty S_{\rm gw}(f) \, df$

- $\Omega_{gw}(f)$ in terms of $S_{gw}(f)$: $S_{\rm gw}(f) = \frac{3H_0^2}{10\pi^2} \, f^{-3} \Omega_{\rm gw}(f)$

- Strain for constant Ω: $h(f) = S_{\rm gw}^{1/2}(f) = 5.6 \times 10^{-22} \, h_{100} \sqrt{\Omega_0} \left(\frac{100\text{Hz}}{f} \right)^{3/2} \text{Hz}^{1/2}$

What is known about the stochastic background?

Observational Technique	Observed Limit	Frequency Domain	Comments
Cosmic Microwave Background	$\Omega_{\rm gw}(f)\ h_{100}^2 \leq 7\times 10^{-11}(H_0/f)^2$	$H_0 < f < 30\ H_0$	Allen & Koranda, Phys. Rev. D50 (1994) 3713.
Radio Pulsar Timing	$\Omega_{\rm gw}(f)\ h_{100}^2 \leq 2\times 10^{-9}(f/f_0)^2$	$f \geq f_0 = 2\times 10^{-9}$ Hz	Lommen, astro-ph/0208572
Big-Bang Nucleosynthesis	$\int_{f>10^{-8}\ \rm Hz} d\ln f\ \Omega_{\rm gw}(f)\ h_{100}^2 \leq 10^{-5}$	$f > 10^{-8}$Hz	Kolb & Turner, The Early Universe, Addison Wesley 1990
Interferometers	$\Omega_{\rm gw}(f)\ h_{100}^2 \leq 3\times 10^5$	100 Hz $\lesssim$ f $\lesssim$ 1000 Hz	Garching-Glasgow [34]
Room Temperature Resonant Bar (correlation)	$\Omega_{\rm gw}(f)\ h_{100}^2 \leq 3000$	$f = 900$ Hz	Glasgow [33]
Cryogenic Resonant Bar (single)	$\Omega_{\rm gw}(f)\ h_{100}^2 \leq 300$ $\Omega_{\rm gw}(f)\ h_{100}^2 \leq 5000$	$f = 907$ Hz $f = 1875$ Hz	EXPLORER [35] ALTAIR [36]
Cryogenic Resonant Bar (correlation)	$\Omega_{\rm gw}(f)\ h_{100}^2 \leq 60$	$f = 907$ Hz	EXPLORER+NAUTILUS [37, 38]

TABLE I: Summary of upper limits on $\Omega_0 h_{100}^2$ over a large range of frequency bands. The upper portion of the table lists limits derived from astrophysical observations. The lower portion of the table lists limits obtained from prior direct measurement.

Stochastic GW Background Detection

- Cross-correlate the output of two (*independent*) detectors with a suitable filter kernel: $s_i(t) = h_i(t) + n_i(t)$

$$Y = \int_0^T dt_1 \int_0^T dt_2\, s_1(t_1)\, Q(t_1 - t_2)\, s_2(t_2)$$

$$\approx T \int_{-\infty}^{\infty} df\ \tilde{s}_1^*(f)\, \tilde{Q}(f)\, \tilde{s}_2(f)$$

- Requires:
  - (i) Two detectors must have overlapping frequency response functions i.e., $s_1(f)s_2(f) \neq 0,\ \{f\} \subset \varnothing$
  - (ii) Detectors sensitive to same polarization state (+, x) of radiation field, h_{gw}.
  - (iii) Baseline separation must be suitably "short": $L < \lambda_{GW}(f) \Rightarrow \frac{2\pi f L}{c} < 1$
- Limits of detection (1 year integration):
  - » LIGO I: $\Omega_{gw} \leq 10^{-5}$
  - » Advanced LIGO: $\Omega_{gw} \leq 5 \times 10^{-9}$

Stochastic GW Background Detection

- Correlation kernel weighting function -> optimal filter

$$\mu \equiv \langle Y \rangle = \frac{T}{2} \int_{-\infty}^{\infty} df\, \gamma(|f|) S_{\rm gw}(|f|)\, \tilde{Q}(f)$$

$$\sigma^2 \equiv \langle Y^2 \rangle - \langle Y \rangle^2 \approx \frac{T}{4} \int_{-\infty}^{\infty} df\, P_1(|f|)\, |\tilde{Q}(f)|^2\, P_2(|f|)$$

- SNR is maximized for: $\tilde{Q}(f) \propto \dfrac{\gamma(|f|) S_{\rm gw}(|f|)}{P_1(|f|) P_2(|f|)} = \lambda \dfrac{\gamma(|f|) \Omega_{\rm gw}(|f|)}{|f|^3 P_1(|f|) P_2(|f|)}$

$$\mathrm{SNR} \approx \frac{3H_0^2}{10\pi^2} \sqrt{T} \left[\int_{-\infty}^{\infty} df\, \frac{\gamma^2(|f|) \Omega_{\rm gw}^2(|f|)}{f^6 P_1(|f|) P_2(|f|)} \right]^{1/2}$$

Overlap Reduction Factor, γ(f)

- Overlap reduction function, γ(f), is a function of detector geometries, orientations and detector separations

$$\gamma(f) = \frac{5}{8\pi} \sum_{k=+,\times} \int d\hat{\Omega}\; e^{2\pi i f \hat{\Omega}\cdot\vec{x}/c} \left[\mathbf{d}_1 : \mathbf{e}_1^k(\hat{\Omega})\right]\left[\mathbf{d}_2 : \mathbf{e}_2^k(\hat{\Omega})\right]$$

$$e^{+}_{ab}(\hat{\Omega}) = \hat{m}_a\hat{m}_b - \hat{n}_a\hat{n}_b$$

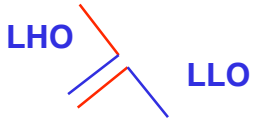
$$e^{\times}_{ab}(\hat{\Omega}) = \hat{m}_a\hat{n}_b + \hat{n}_a\hat{m}_b$$

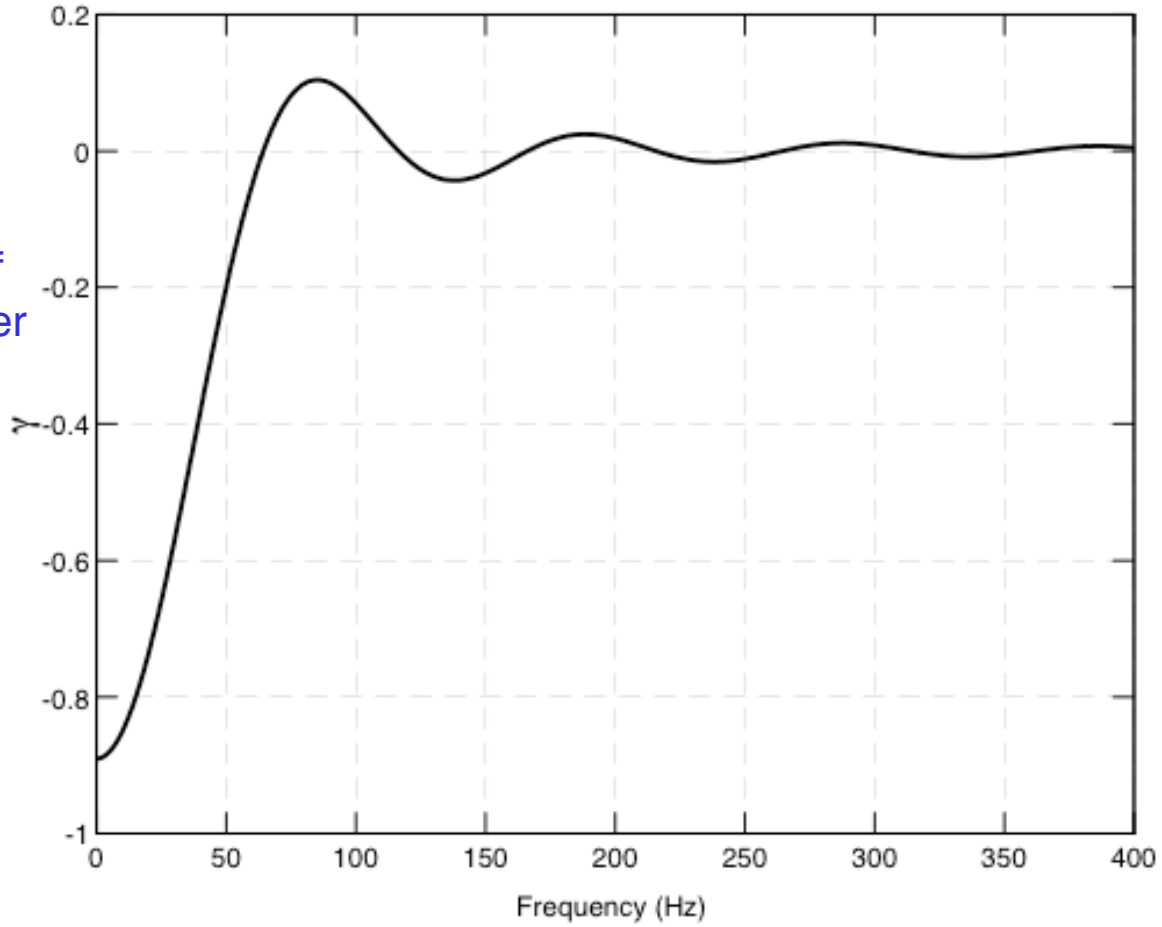
$$\gamma(f) = \gamma_1\left(\frac{2\pi Lf}{c}\right)\,\mathbf{d}_1 : \mathbf{d}_2 + \gamma_2\left(\frac{2\pi Lf}{c}\right)\,(\hat{n}_{12}\cdot\mathbf{d}_1)\cdot(\mathbf{d}_2\cdot\hat{n}_{12}) + \gamma_3\left(\frac{2\pi Lf}{c}\right)\,(\hat{n}_{12}\cdot\mathbf{d}_1\cdot\hat{n}_{12})(\hat{n}_{12}\cdot\mathbf{d}_2\cdot\hat{n}_{12})$$

$$\begin{bmatrix}\gamma_1(x)\\ \gamma_2(x)\\ \gamma_3(x)\end{bmatrix} = \begin{bmatrix} 5 & -\frac{10}{x} & \frac{5}{x^2} \\ -10 & \frac{40}{x} & -\frac{50}{x^2} \\ \frac{5}{2} & -\frac{25}{x} & \frac{175}{2x^2}\end{bmatrix} \cdot \begin{bmatrix} j_0(x)\\ j_1(x)\\ j_2(x)\end{bmatrix}$$

Overlap Reduction Factor, □(f)

- WA - WA == 1
- LA - WA shown
  - □(0) ~ -1 because of WA-LA interferometer orientations:

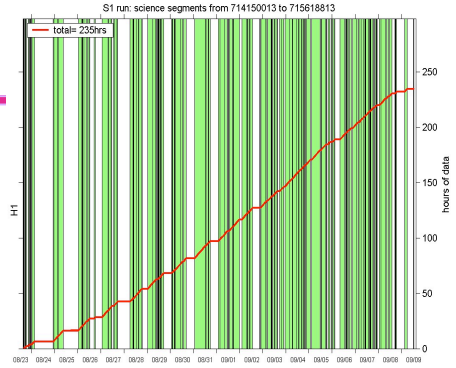




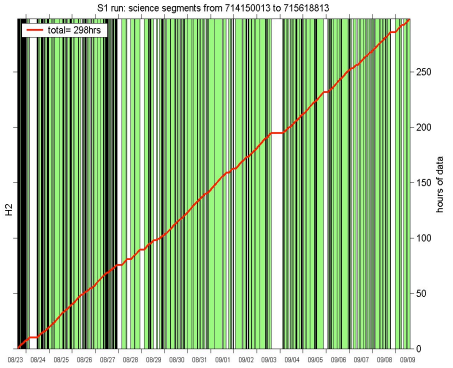
In-Lock Data Summary from S1

Red lines: integrated up time Green bands (w/ black borders): epochs of lock

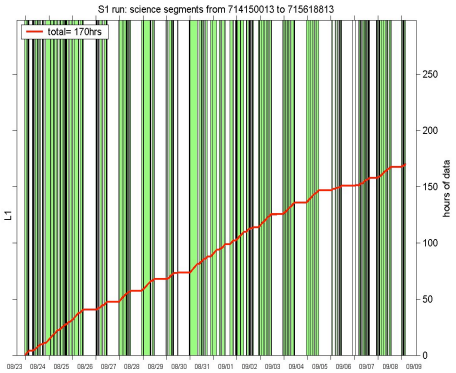
H1: 235 hrs



H2: 298 hrs



L1: 170 hrs



- **August 23 – September 9, 2002: 408 hrs (17 days).**
- **Individual interferometers:**
  - **H1 (4km): duty cycle 57.6% ; Total Locked time: 235 hrs**
  - **H2 (2km): duty cycle 73.1% ; Total Locked time: 298 hrs**
  - **L1 (4km): duty cycle 41.7% ; Total Locked time: 170 hrs**
- ***Double coincidences:***
  - ***L1 & H1 : duty cycle 28.4%; Total coincident time: 116 hrs***
  - ***L1 & H2 : duty cycle 32.1%; Total coincident time: 131 hrs***
  - ***H1 & H2 : duty cycle 46.1%; Total coincident time: 188 hrs***
- **Triple Coincidence: L1, H1, and H2 : duty cycle 23.4% ;**
  - **Total coincident time: 95.7 hrs**

For this analysis:

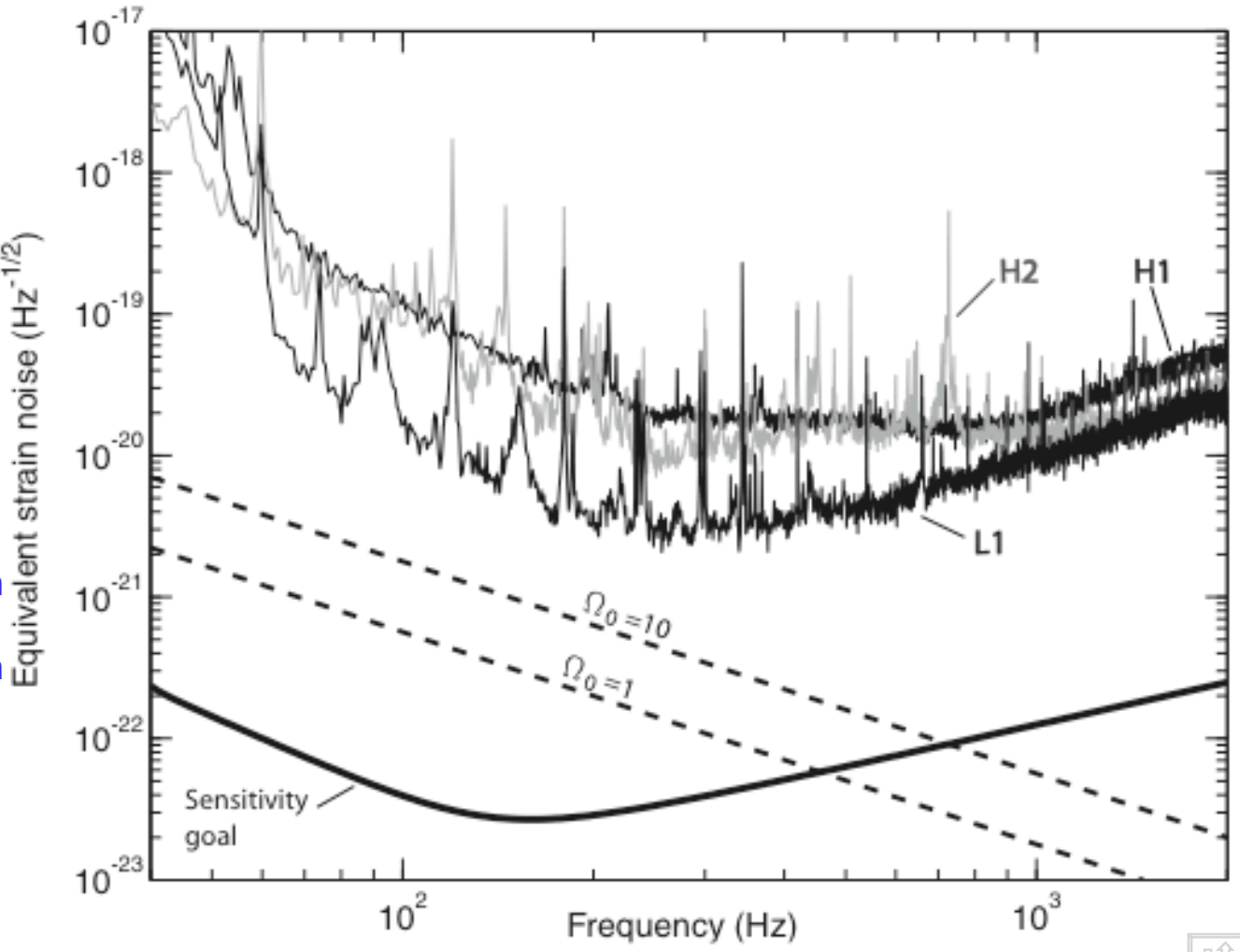
L1-H1:
Valid data: 75 hrs
Quiet Data: 75 hrs
Calibrated Data: 64 hrs
Net uptime: 15.7%

L1-H2:
Valid data: 81 hrs
Quiet Data: 66 hrs
Calibrated Data: 51 hrs
Net uptime: 12.5%

H1-H2:
Valid data: 134 hrs
Quiet Data: 119 hrs
Calibrated Data: 100 hrs
Net uptime: 24.5%

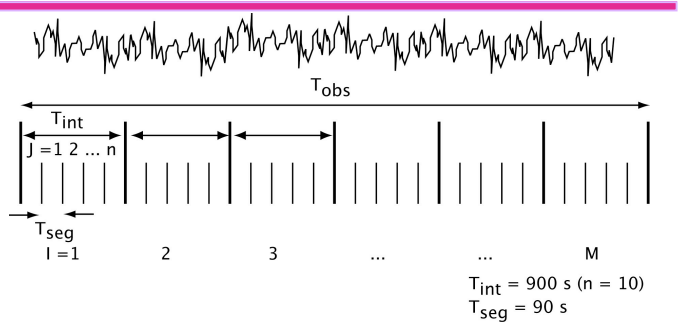
S1 Sensitivities

- Spectra taken just before run
- Cross correlation technique allows one to "dig" signal below noise floor in individual instruments
- Dashed lines show expected 90% confidence bounds one could set:
  - »100 hrs of observation with H2km +L4km (□=10)
  - »150 hrs of observation with H2km + H4km (□=1)
  - »Limits from theoretical SNR equation



Implementation of analysis

- Analysis performed in data intervals of 900s (15 min)
  - »For each 900s interval, ***I*** :
    - –Average power spectrum, mid point calibration used for entire interval
    - –Ten 90s segments are analyzed separately
    - –Provides 10 independent estimates, statistics of estimates
  - »For each 90s segment ***J***, estimate:
    - –resample data to 1024 samples/s (512 Hz Nyquist)
      - 90% of SNR comes below f~300Hz
    - –FFT, window data
    - –Calculate estimate Y_{IJ}
    - –Average n = 45 frequency bins to obtain cross-correlation spectra with Δf = 0.25 Hz
  - »Average 10 values of Y_{IJ} to obtain interval average, Y_I, sample variance, s_I^2



$$\tilde{\bar{g}}_i[l] := \sum_{k=0}^{2N-1} \Delta t \; \bar{g}_i[k] e^{-i2\pi kl/2N}$$

$$\bar{g}_i[k] = \begin{cases} w_i[k]\, g_i[k] & k = 0, \ldots, N-1 \\ 0 & k = N, \ldots, 2N-1 \end{cases}$$

$$Y_{IJ} = \Delta f \sum_{l=0}^{2N-1} \tilde{\bar{g}}_1[l]^* \, \widetilde{Q}'[l] \, \tilde{\bar{g}}_2[l]$$

$$\overline{Y}_I \equiv \frac{1}{10} \sum_{J=1}^{10} Y_{IJ}$$

$$s_I := \sqrt{\frac{1}{9} \sum_{J=1}^{10} (Y_{IJ} - \overline{Y}_I)^2}$$

Implementation of analysis

- Results for 900s intervals combined to obtain run averaged answer ("point estimate").

$$\overline{Y} = \frac{\sum_I \sigma_I^{-2} Y_I}{\sum_I \sigma_I^{-2}}$$

$$\overline{Y}/T \equiv \widehat{\Omega}_{\rm eff} = \Omega_0\, h_{100}^2 + \Omega_{\rm inst}$$

$$\widehat{\Omega}_{\rm eff} - 1.65\,\hat{\sigma} \le \Omega_{\rm eff} \le \widehat{\Omega}_{\rm eff} + 1.65\,\hat{\sigma}$$

$$\Omega_0\, h_{100}^2 \le \max\{\overline{Y}/T,\, 0\} + 1.28\,\widehat{\sigma}$$

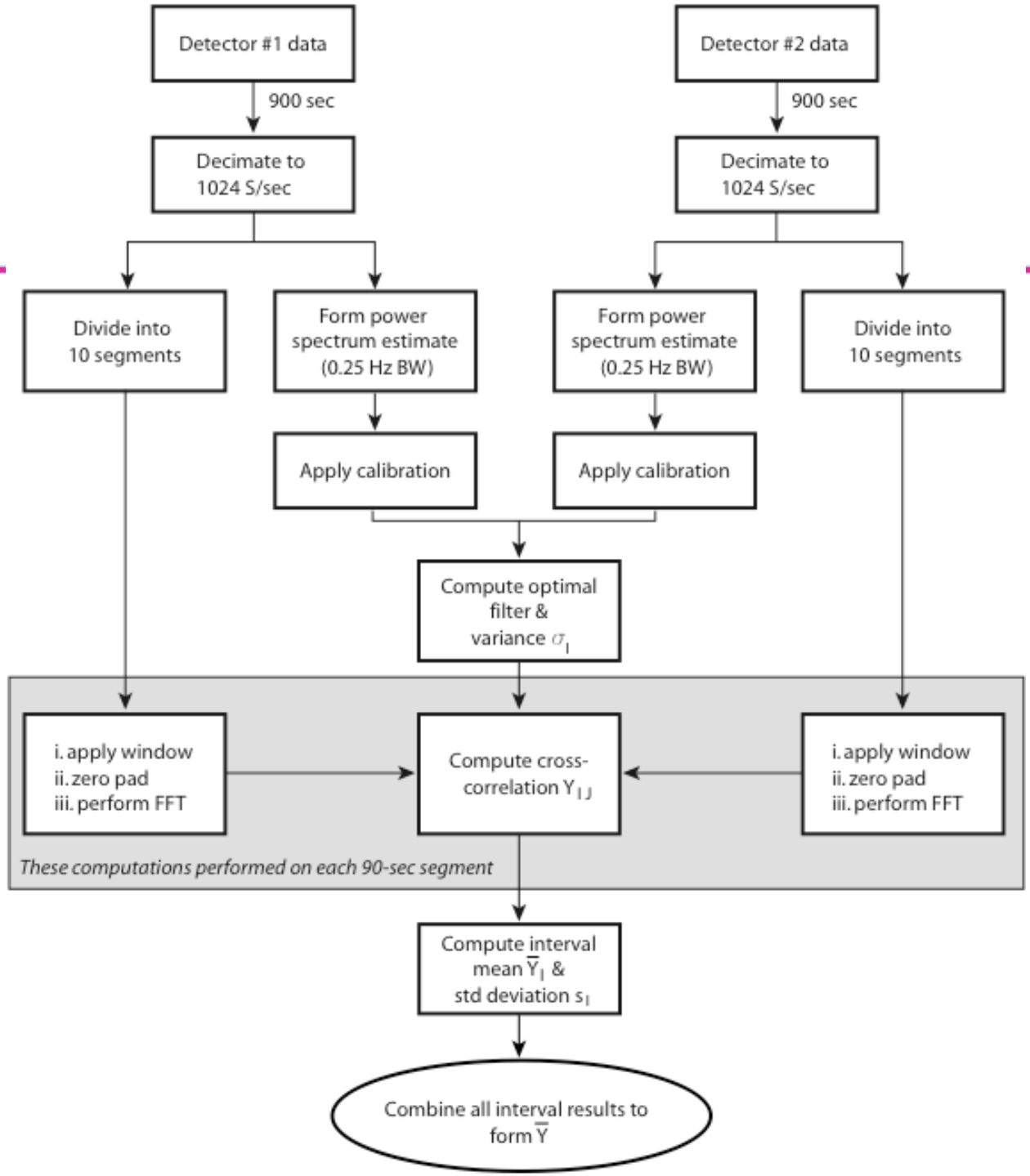
- Weights σ_I^2 are obtained from the power spectra for each interval, *I*.
  - » Measures data quality -- how quiet the interferometer pair was during interval *I*

$$\sigma^2 \equiv \langle Y^2 \rangle - \langle Y \rangle^2 \approx \frac{T}{4} \int_{-\infty}^{\infty} df\; P_1(|f|)\, |\tilde{Q}(f)|^2\, P_2(|f|)$$

- Overall statistical error for the estimate derived from individual interval variances

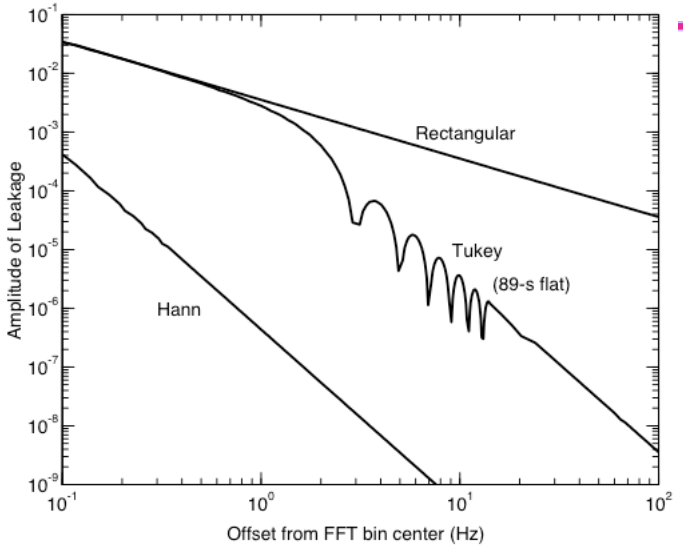
$$\begin{aligned} \hat{\sigma} &:= \sqrt{\mathrm{Var}(\overline{Y}/T)} \\ &= \frac{1}{T} \frac{\sqrt{\sum_I \sigma_I^{-4} \mathrm{Var}(\overline{Y}_I)}}{\sum_J \sigma_J^{-2}} \\ &= \frac{1}{T} \frac{1}{\sqrt{10}} \frac{\sqrt{\sum_I \sigma_I^{-4} s_I^2}}{\sum_J \sigma_J^{-2}} \,. \end{aligned}$$

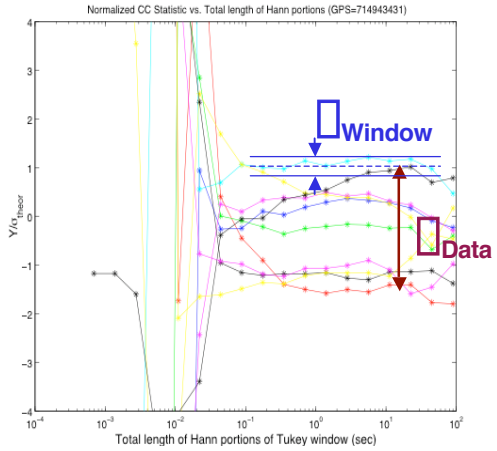
Pipeline analysis flow

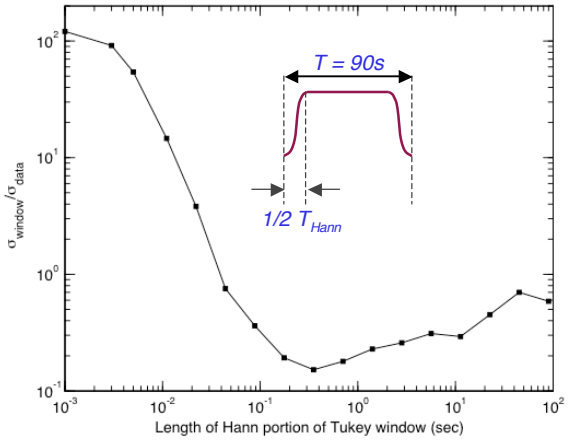


Windowing Effects

- Sources of spectral leakage:
  - » Analysis of data in finite segments
  - » Redness of spectrum (seismic wall)
  - » Narrow band features
    - – Will be removed by notching, need to make sure their effect is contained in a few frequency bins
- Studied different window widths
  - » Used a Tukey window
  - » Flat top plus smooth fall-off at ends:
  - » 0.5s + 89s + 0.5s

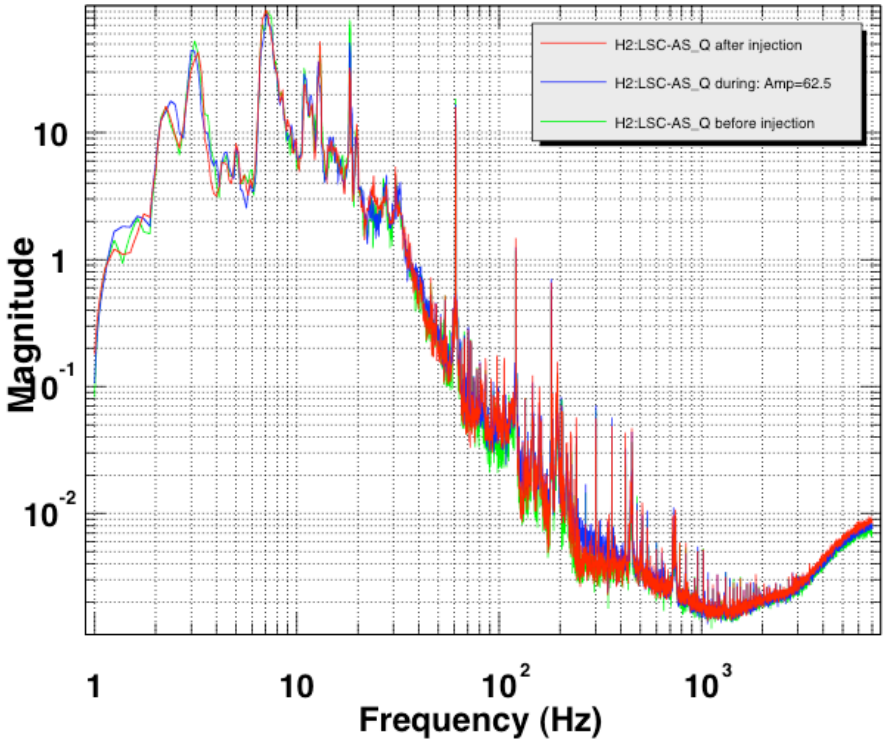




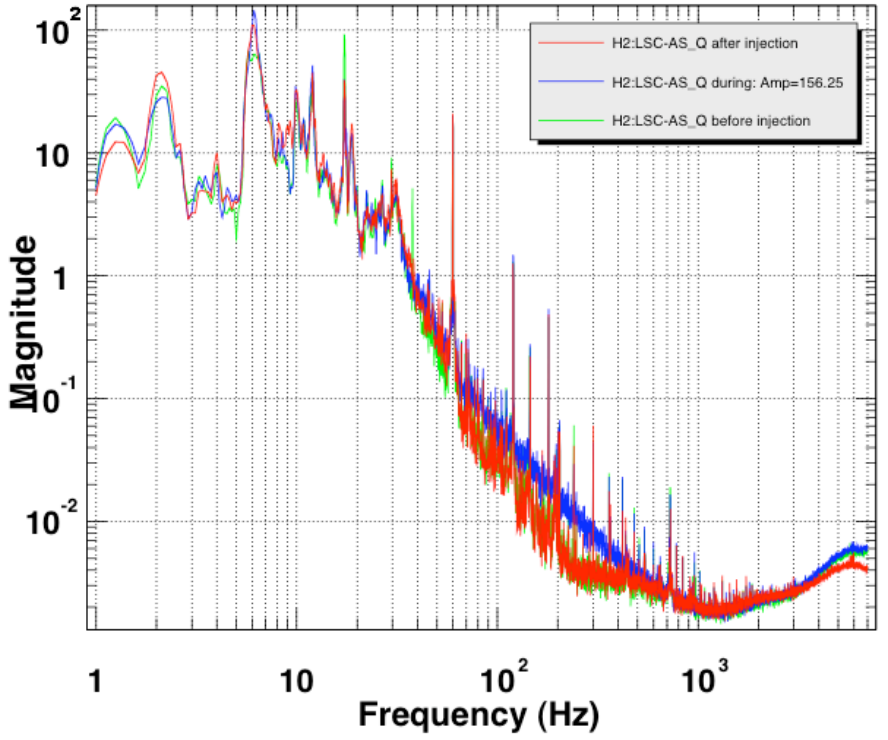


End-to-end pipeline validation SW and HW injection of simulated stochastic backgrounds

- Generate time series of correlated random noise with same properties as SGWB
- Injected in hardware during S1 run at several amplitudes
- Inject post run for further verification



*T0=10/09/2002 06:53:57 Avg=10 BW=0.187493

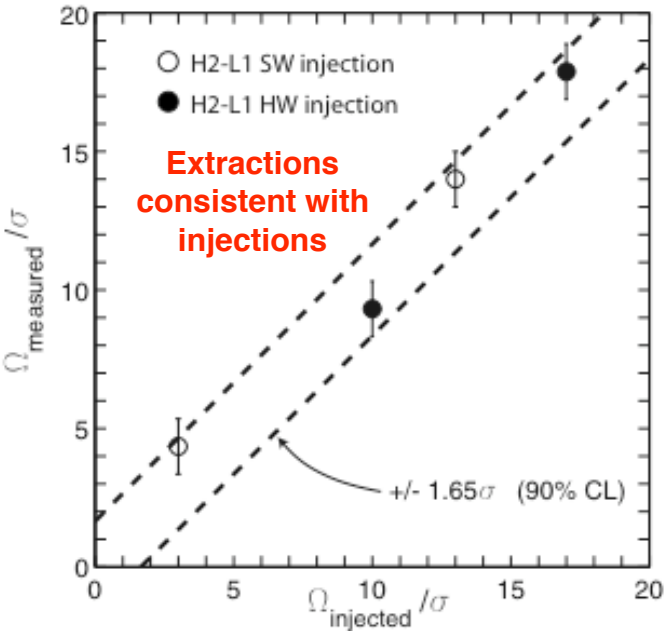


*T0=10/09/2002 04:05:22 *Avg=10 BW=0.187497

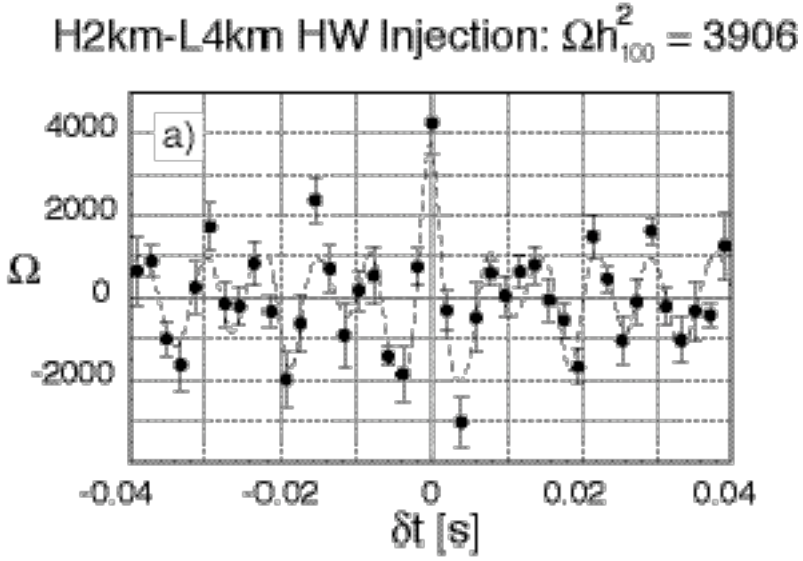
Extraction of injected signals

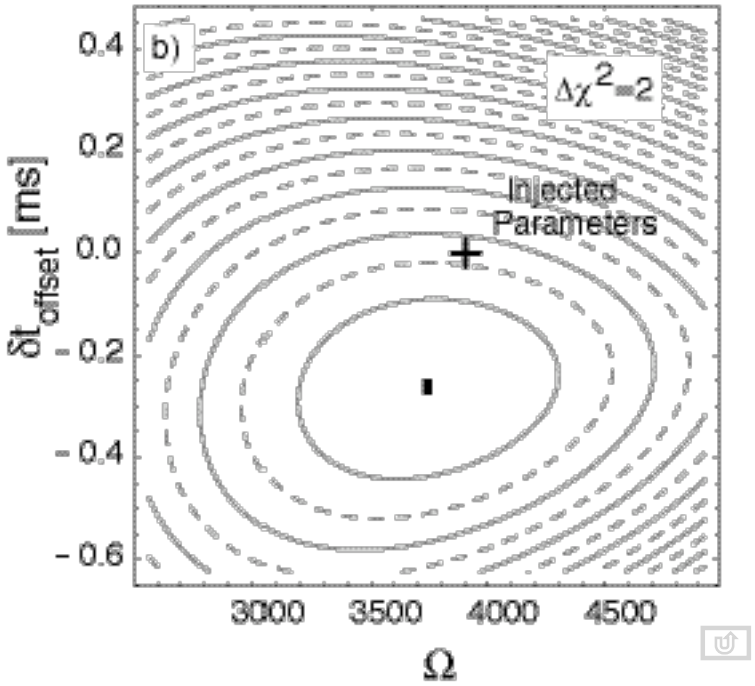
- □² with 2 D.O.F: time offset, amplitude
- Theoretical curve determined by power spectra, $P_1(f)$, $P_2(f)$
- Points obtained by performing a time-shift analysis of data

$$Y(\tau) = \int_0^T dt_1 \int_0^T dt_2\, s_1(t_1+\tau)\, Q(t_1-t_2)\, s_2(t_2)$$
$$\approx T \int_{-\infty}^{\infty} df\ \tilde{s}_1^*(f)\, \widetilde{Q}(f)\, e^{i2\pi f\tau}\, \tilde{s}_2(f)\ . \qquad (6.2)$$

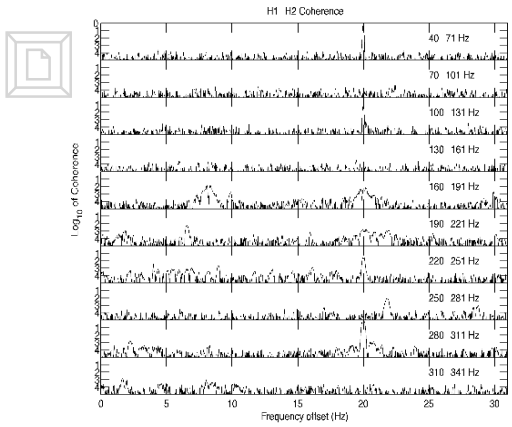


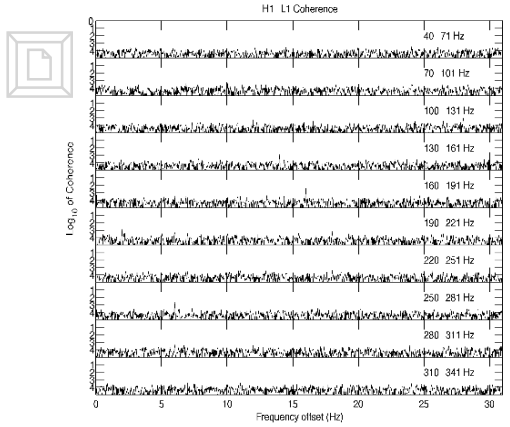
- **Analysis sensitive to relative timing of data streams**
  - **270 □s offset consistent with post-run GPS measurements**

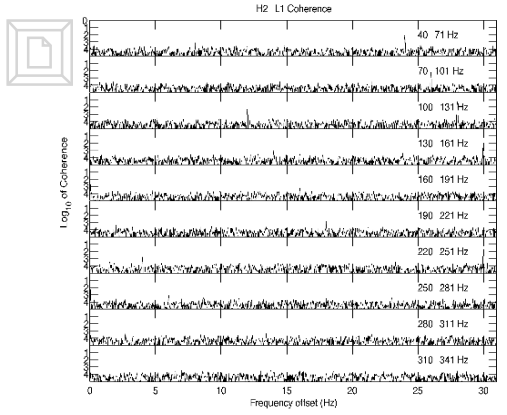




Survey of cross-spectral coherences, $\gamma_{12}(f)$

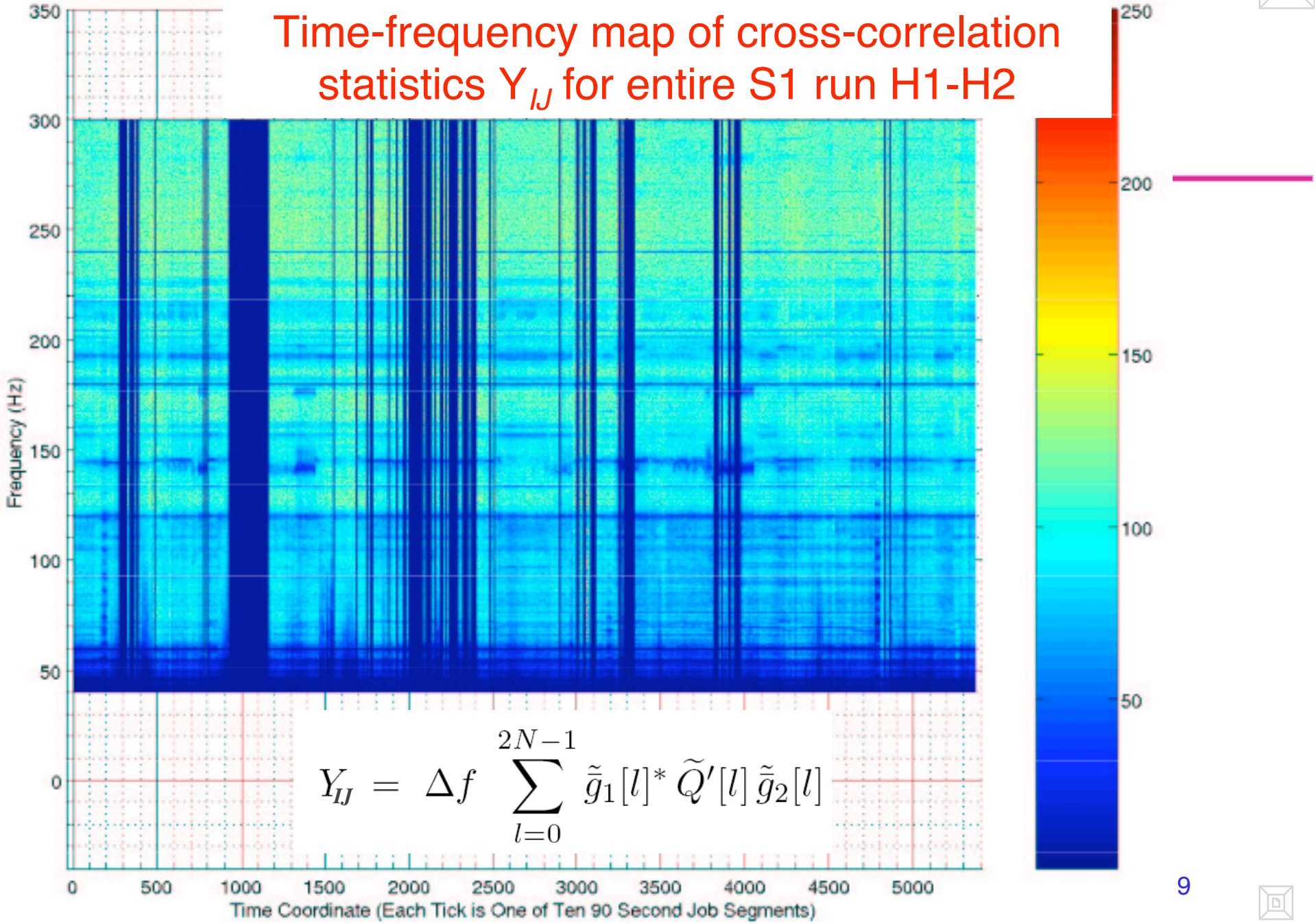






- Narrowband coherences persist over long integration times
  - » n x 60 Hz (H1-H2) line harmonics
  - » n x16 Hz (all pairs) GPS-synchronized clocking electronics for data acquisition system
  - » 250 Hz (?)
- Lines excised by excluding them from the integration over frequency to obtain estimates, Y_{IJ}

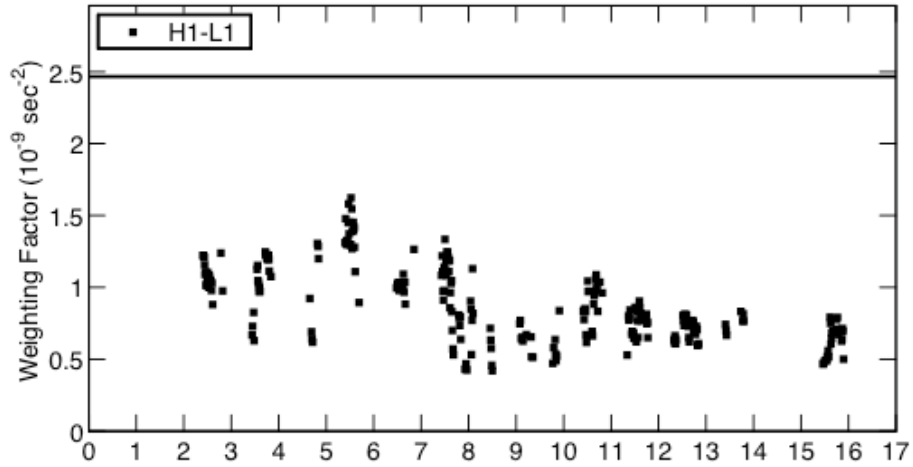
Time-frequency map of cross-correlation statistics Y_{IJ} for entire S1 run H1-H2

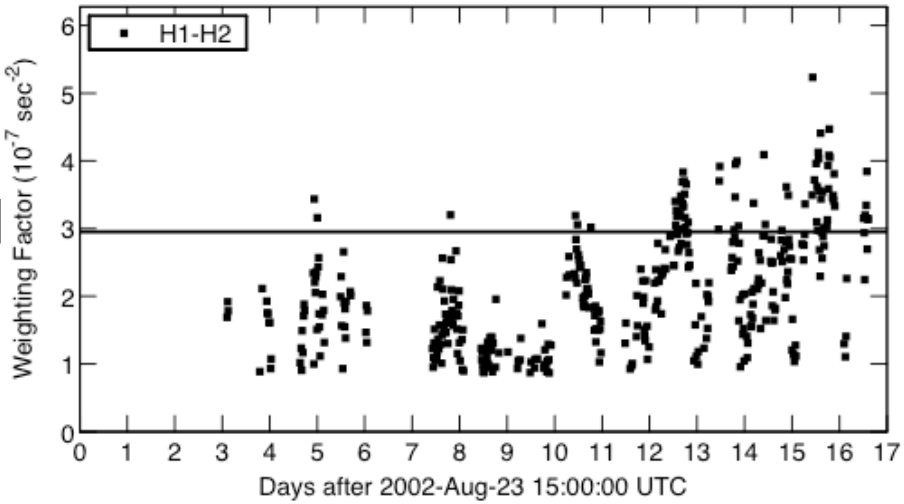


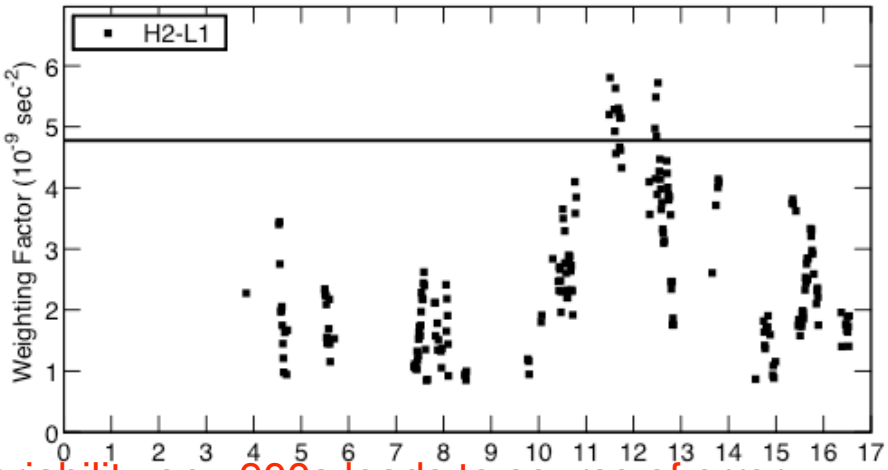
$$Y_{IJ} = \Delta f \sum_{l=0}^{2N-1} \tilde{\bar{g}}_1[l]^* \, \widetilde{Q}'[l] \, \tilde{\bar{g}}_2[l]$$

Data Quality --

Variability in noise floor during S1



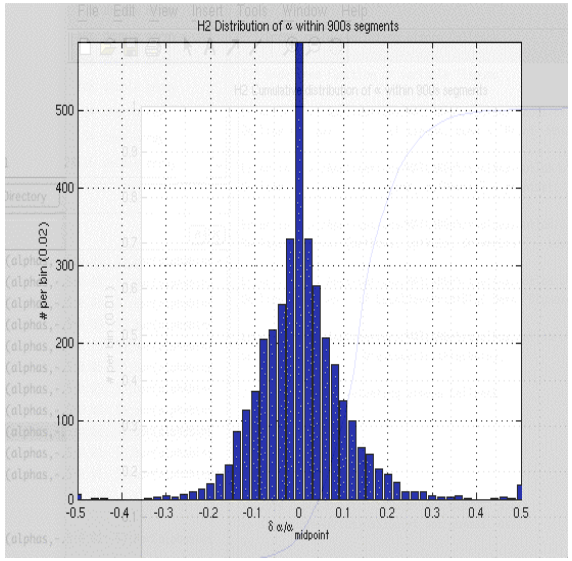


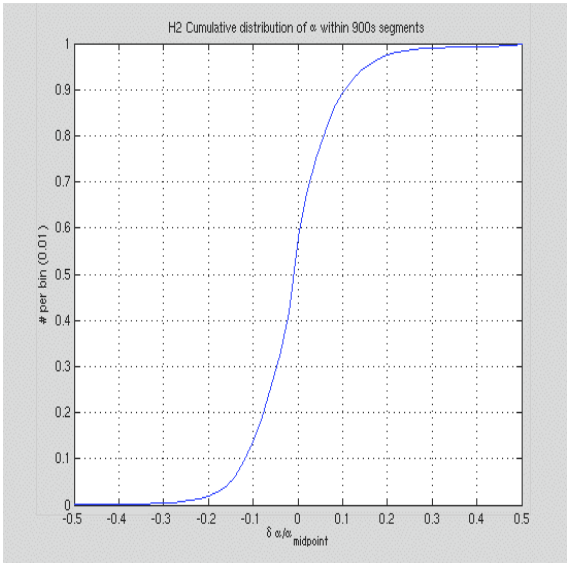


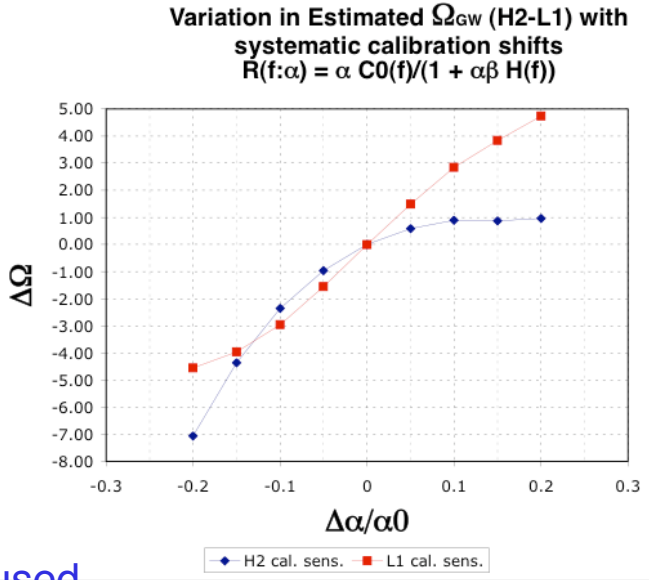
- Variations in statistical weights, ω_i^2, with which individual measurements are combined
- Horizontal lines correspond to "representative" power spectra shown earlier
- Variability shown is on 900s scale, taken into account by weights

- Variability on <900s leads to source of error in estimate beyond statistical errors
  - H1-H2: $\sigma_{\text{PSD nonstationarity}} \sim 0.1$
  - H1-L1: $\sigma_{\text{PSD nonstationarity}} \sim 0.3$
  - H2-L1: $\sigma_{\text{PSD nonstationarity}} \sim 9.3$
- Effect estimated by rerunning analysis with a finer granularity

Data Quality -- Variations in calibrations during S1 Example :H2-L1 (H2 interferometer)







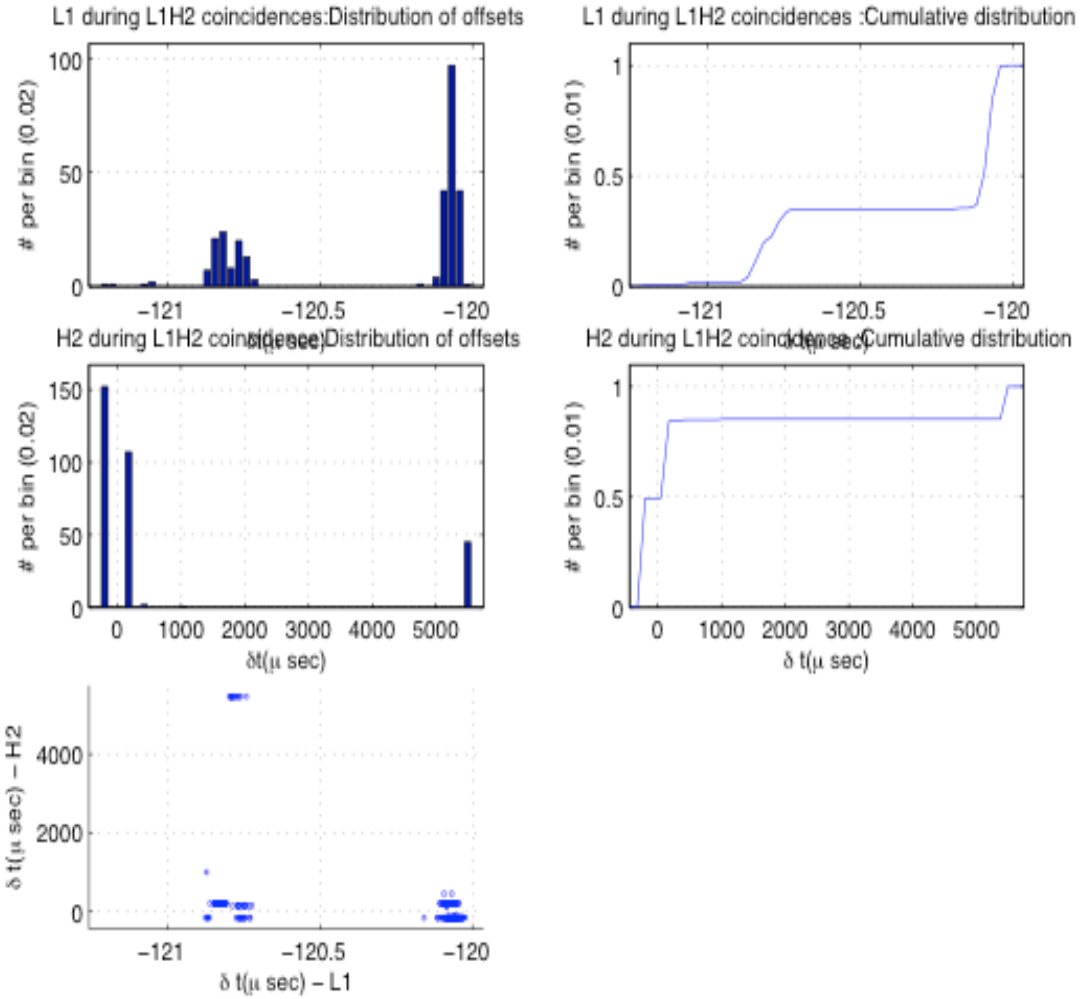
$$R(f:t) = \frac{\square(t)\cdot C_0(f)}{1+\square(t)\cdot\square\cdot H_0(f)}$$

$$\tilde{X}(f) = R(f)\cdot\tilde{g}_{ADC}(f)$$

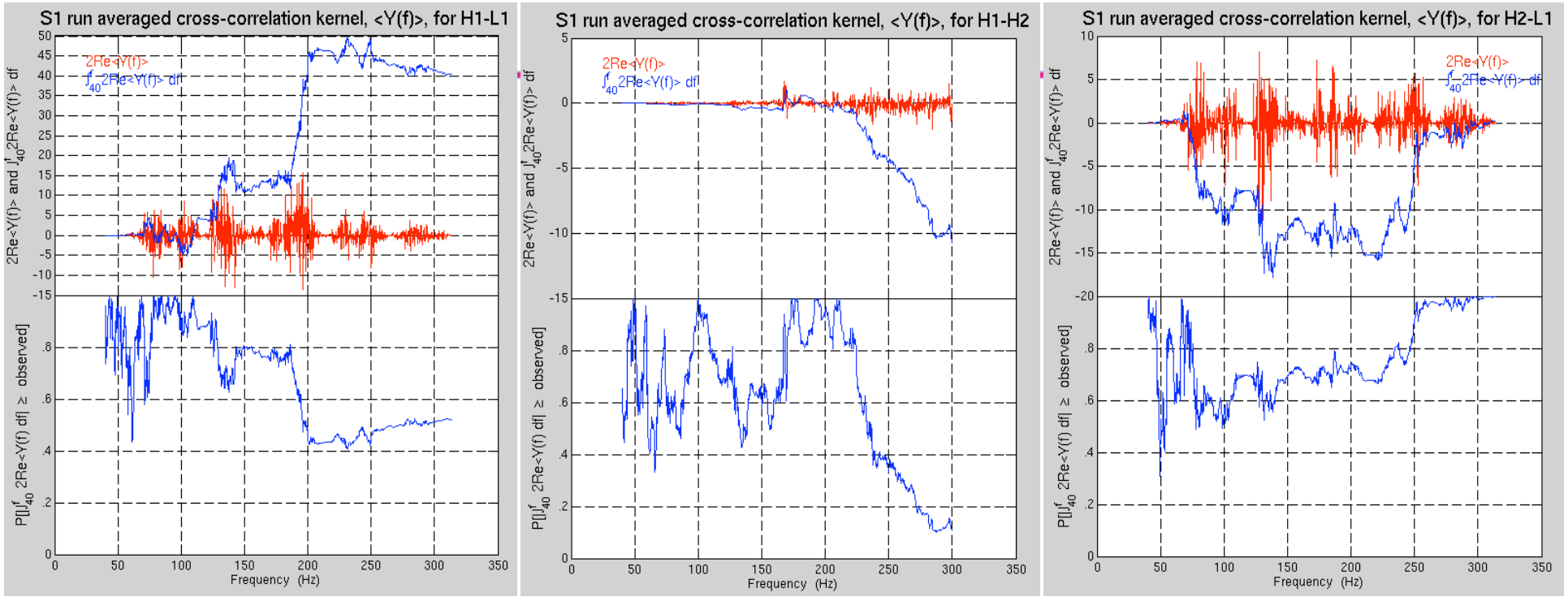
- 900s midpoint calibrations used
  - » 60s trends in calibration were acquired
- Variability on <900s leads to source of error in estimate beyond statistical errors
  - H1-H2: $\square_{calibration} \sim 0.2$
  - H1-L1: $\square_{calibration} \sim 0.4$
  - H2-L1: $\square_{calibration} \sim 1.2$
- Re-ran analysis on finer granularity to assess effect of varying R

Data Quality -- Variations in timing during S1 Example :H2-L1

- Timing offsets of acquisition system changed during S1
  - »Hardware reboots
- Variability over run leads to source of systematic error in estimate beyond statistical errors
- Scale factor, α_{time}, in estimate
  - H1-H2: $\alpha_{time} \sim 1$ (not signifcant)
  - H1-L1: $\alpha_{time} \sim 1$ (not signifcant)
  - H2-L1: $\alpha_{time} \sim 1.05$
    - ± 200 μs shift
- Include in final result as a scaling up of estimate



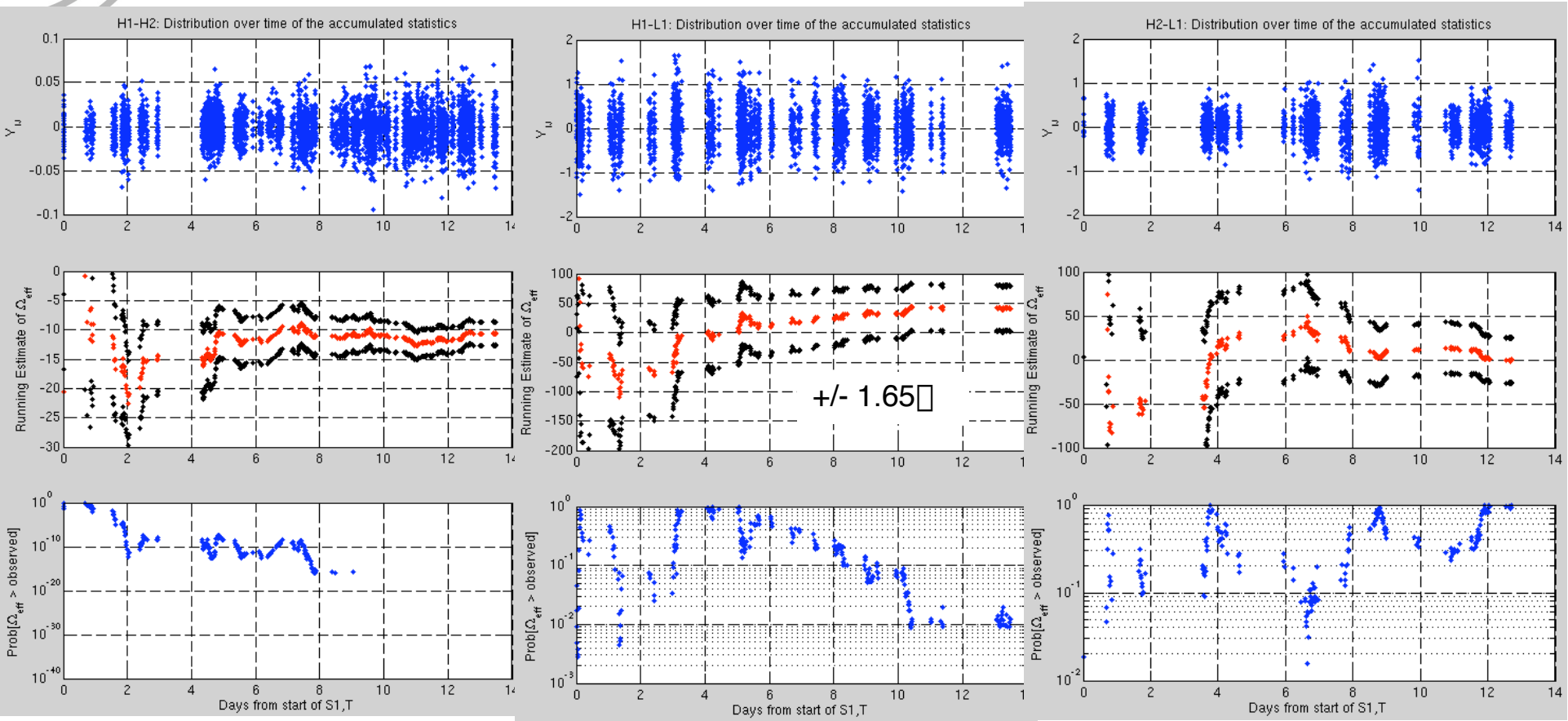
Frequency, time dependence of cross correlation kernels



- Run-averaged kernels $\langle \tilde{\bar{g}}_1[l]^* \, \tilde{Q}'[l] \, \tilde{\bar{g}}_2[l] \rangle$
- Integrals correspond to estimates of $\Omega_{\rm eff}$

$$\overline{Y}/T \equiv \widehat{\Omega}_{\rm eff} = \Omega_0 \, h_{100}^2 + \Omega_{\rm inst}$$

Time evolution over S1 of estimates

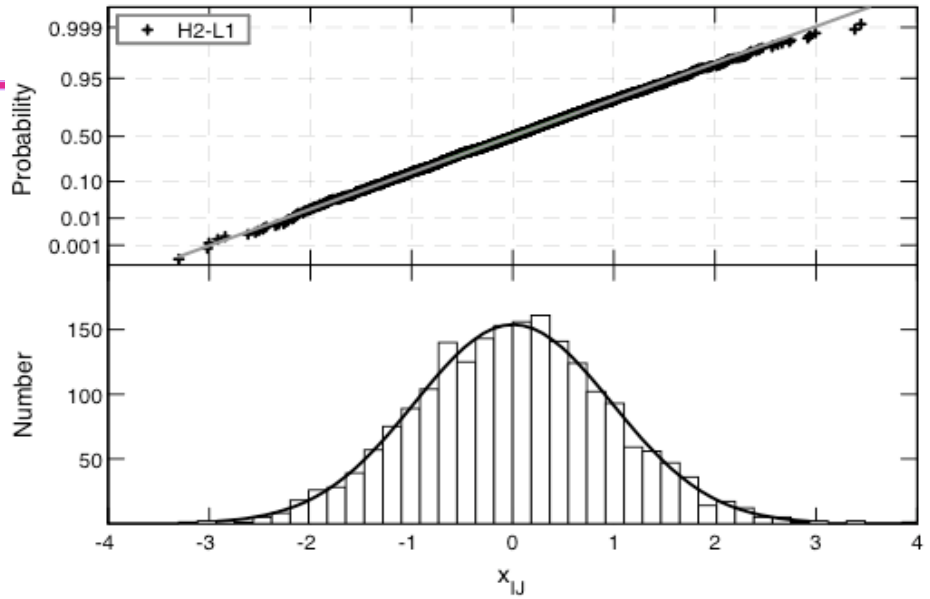


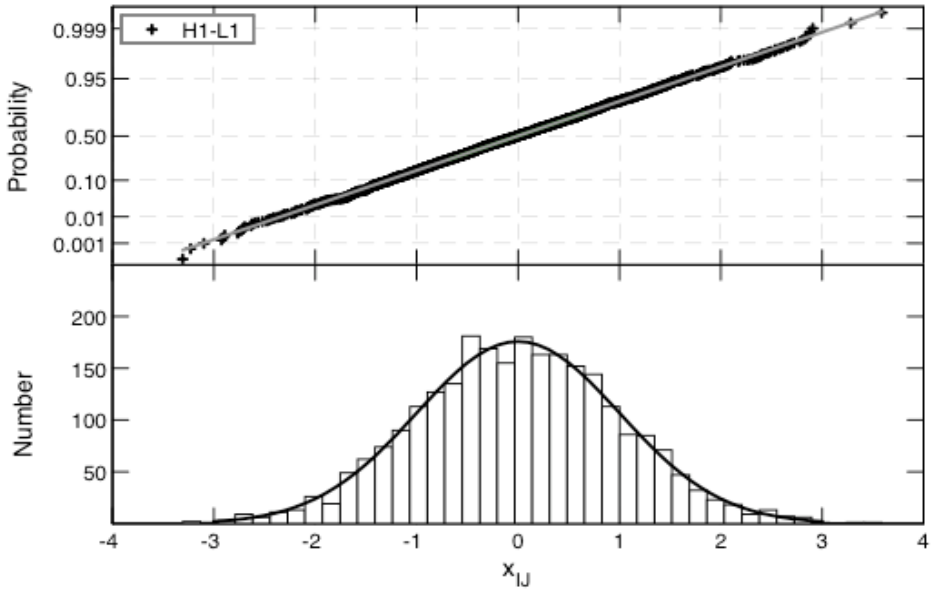
- Running estimates of $\Omega_{\rm eff}$ over run
- End points correspond to estimates of $\Omega_{\rm eff}$: $\overline{Y}/T \equiv \widehat{\Omega}_{\rm eff} = \Omega_0\, h_{100}^2 + \Omega_{\rm inst}$
- Bottom panels show probability of observing running estimate at each time, T, if underlying process is zero-mean Gaussian noise

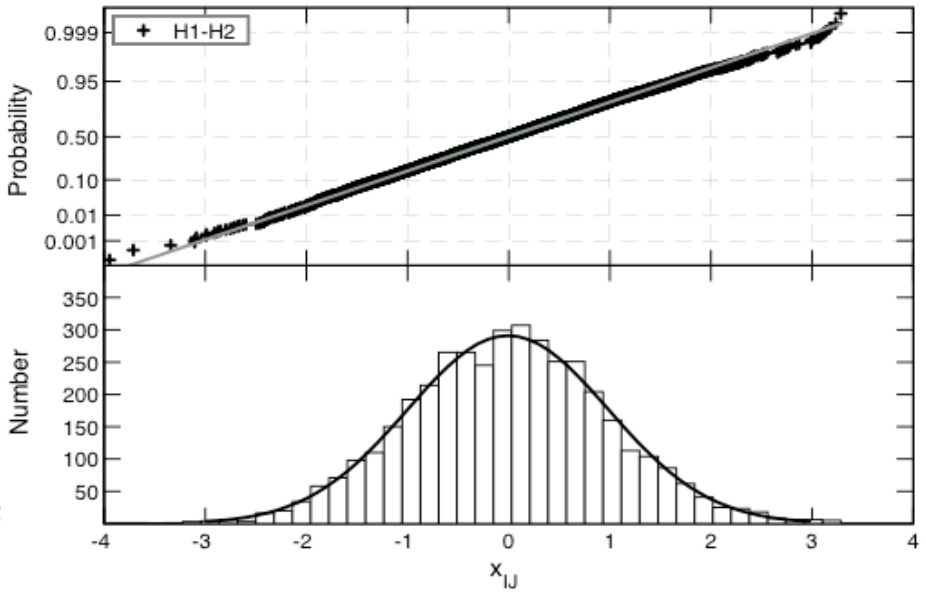
Statistics of estimates

$$X_{IJ} = \frac{Y_{IJ} - \bar{Y}}{\sigma_I}$$

- Normal deviates of 90s estimates from average values are Gaussian RVs:
  - » $\langle X_{IJ} \rangle = 0$
  - » $\sigma^2 = 1$



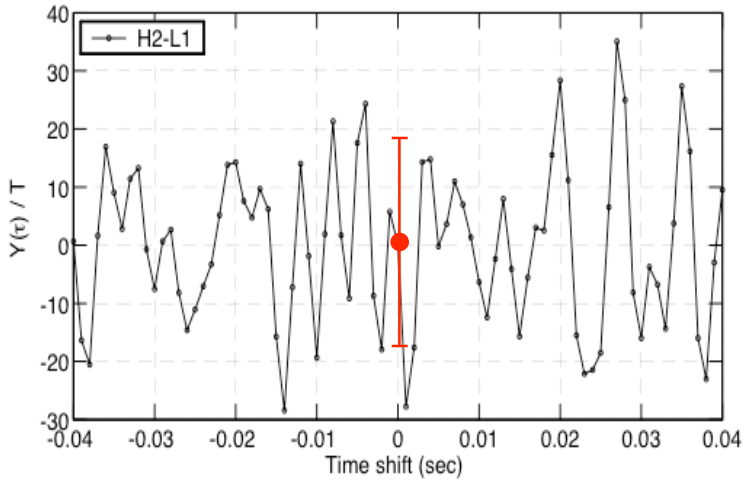


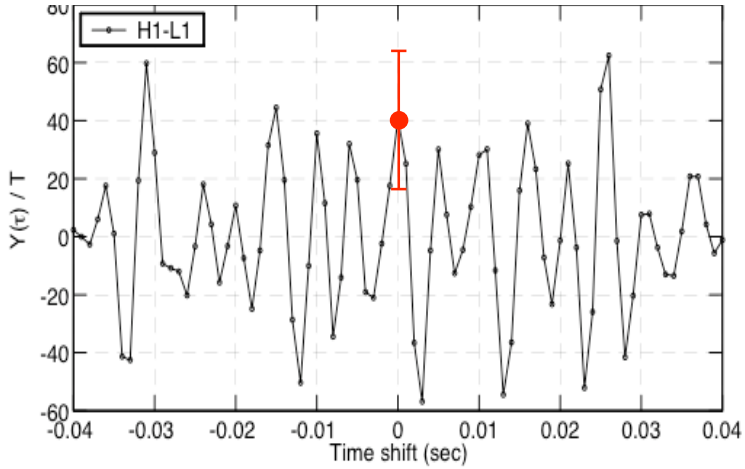


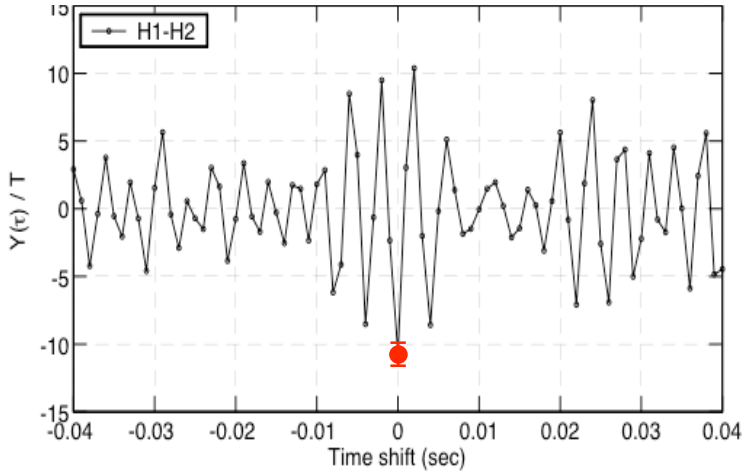
Time shift analysis of final results

$$Y(\tau) = \int_0^T dt_1 \int_0^T dt_2\, s_1(t_1+\tau)\, Q(t_1-t_2)\, s_2(t_2)$$
$$\approx T \int_{-\infty}^{\infty} df\ \tilde{s}_1^*(f)\, \widetilde{Q}(f)\, e^{i2\pi f\tau}\, \tilde{s}_2(f)\ . \qquad (6.2)$$

- H2-L1, H1-L1 consistent with random excursions from point to point
- H1-H2 exhibits time variations not consistent with random noise
  - » Influence of residual instrumental correlations present
  - » Time series not consistent with SNR=10 signal due to □=const.





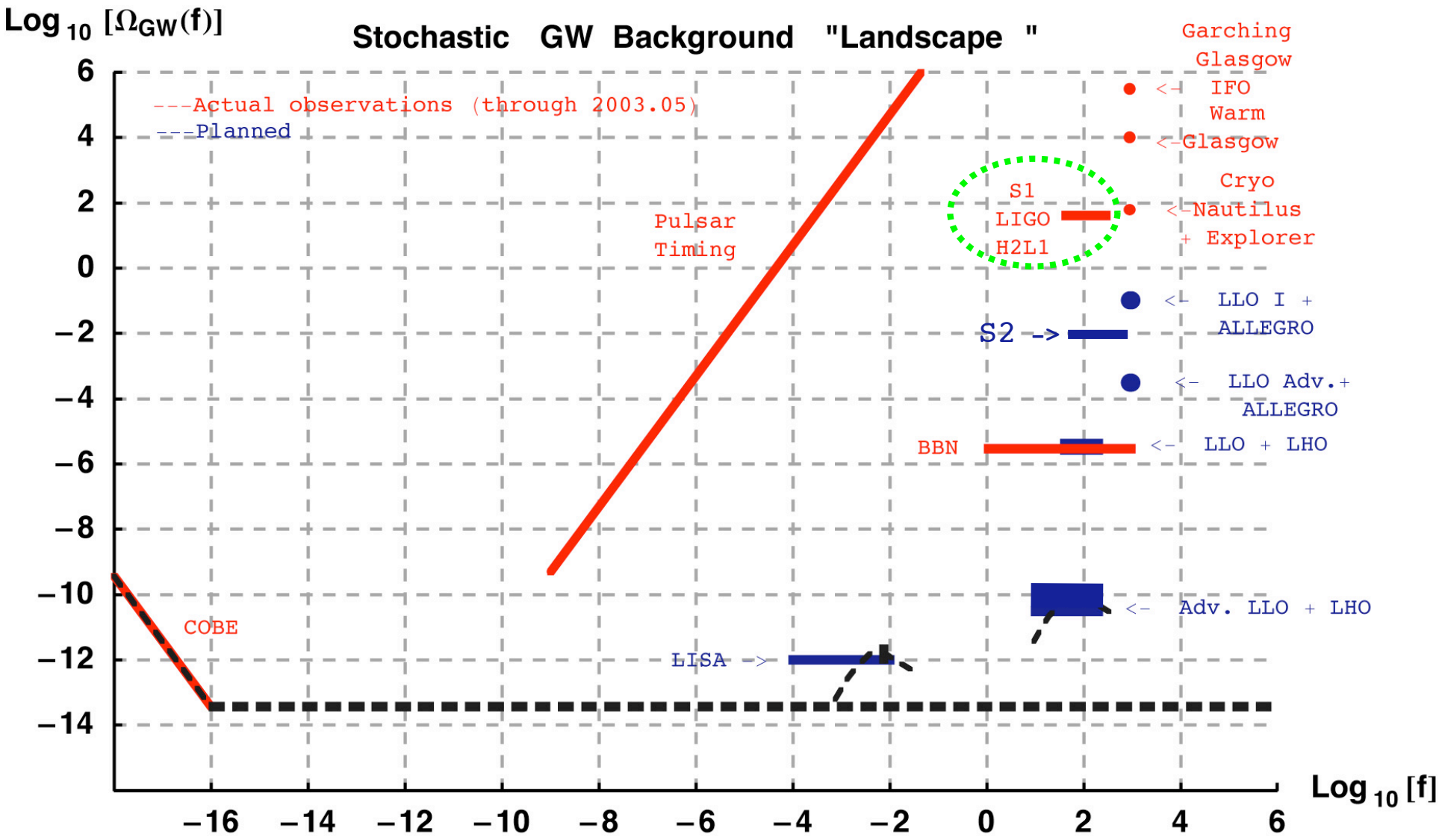


S1 results

	H1-H2	H1-L1	H2-L1
Point Estimates, □eff	-10.7	41.0	0.2
Statistical, □stat	1.2	23.0	15.4
Calibration variations, □cal	0.2	0.4	1.2
Non stationarity, □PSD	0.1	0.3	9.3
Total error (quadrature)	1.2	23.0	18.0
Timing error (scale factor)	1.0	1.0	1.1
Final result	-10.7	41.0	0.2
Symmetric 90%CL: +/- 1.65 □	2.0		
Upper 90% CL □eff +1.28 □		70.4	23.2
Calibration uncertainty, +/-	20%	20%	20%

- H1-H2: $\Box_0 (h_{100})^2 + \Box_{\text{instrumental}} = (-11 \pm 2) \pm 20\%$
- H1-L1: $\Box_0 (h_{100})^2 < 70 \pm 20\%$
- ***H2-L1: $\Box_0 (h_{100})^2 < 23 \pm 20\%$***

Stochastic Gravitational Wave Background "Landscape"



Summary and conclusions

- H2-L1 provides the best upper limit from S1
  - ✓Measurement BW Δf = 274 Hz: 40Hz , f < 314 Hz
  - ✓Within 2x of expectation at beginning of run (~10 vs ~23)
    - ✓64 hrs vs 100 hrs -> 1.25x
    - ✓calibration variation, overall uncertainty
    - ✓non-stationarity of noise floors
    - ✓timing offsets
  - ✓2x - 3x better than previous (narrowband, Δf = 1 Hz) direct measurement with bars @ 1 kHz

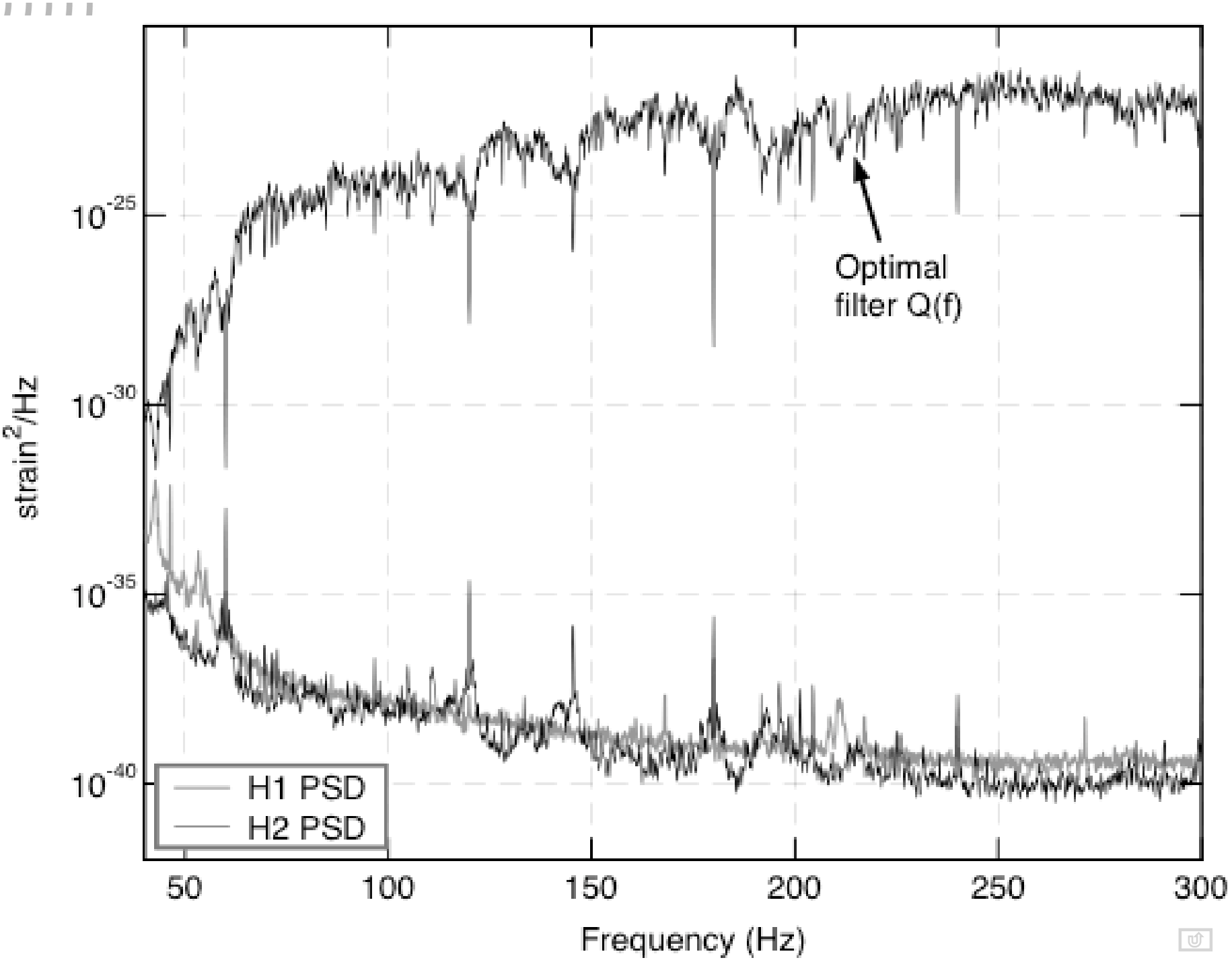
Summary and conclusions

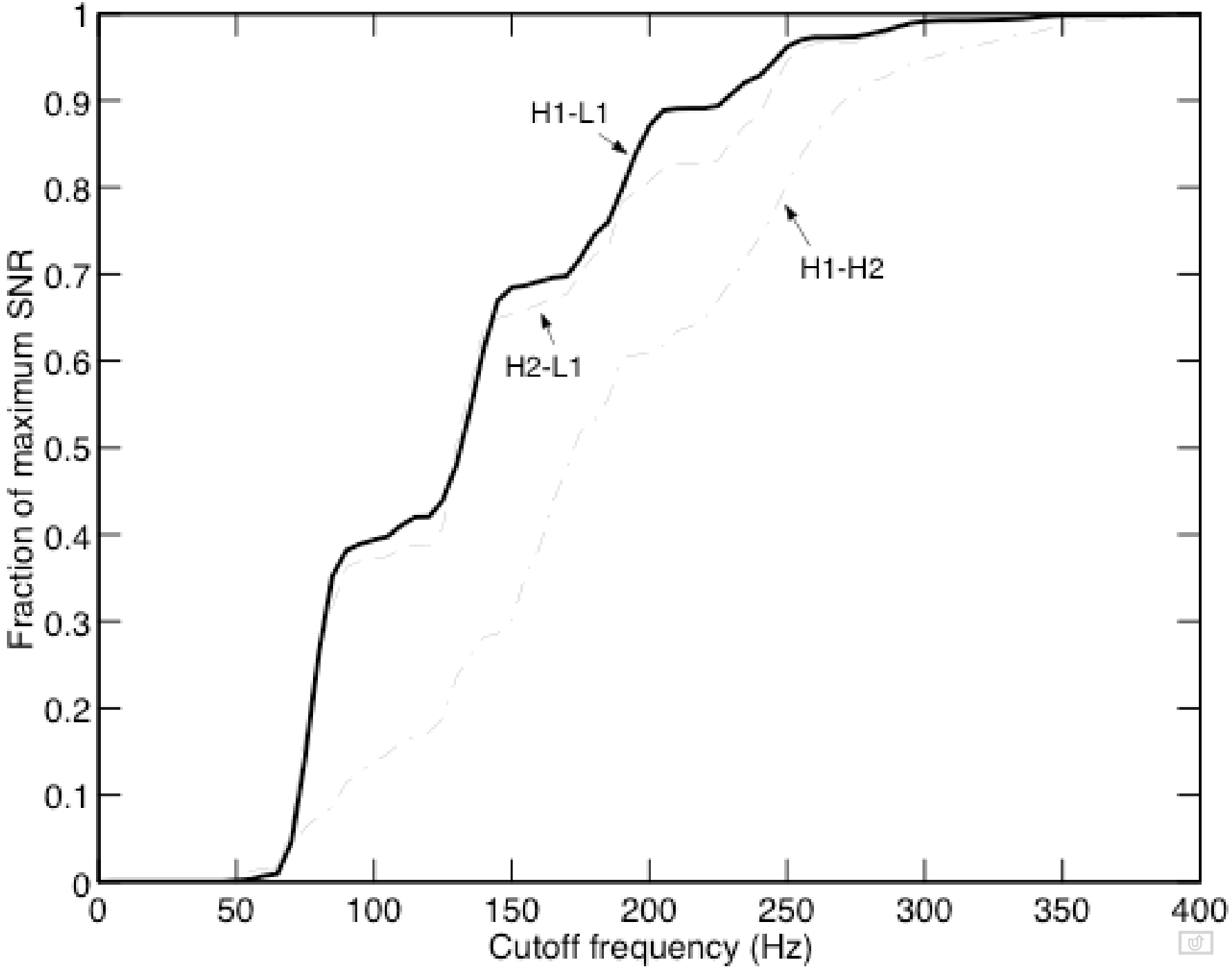
- H1-H2 was not usable, even though it would have had 10x sensitivity
  - » Statistical error is as expected, ~ 1
  - » Bias (-10□) was not foreseen (could have been expected)
- Correlations exhibit instrumental features
  - » WA-LA: narrowband
    - – GPS synchronization of data acquisition systems
    - – 250 Hz feature of presently unknown nature
  - » WA-WA : narrowband and broadband
    - – 60 Hz mains and harmonics, upconversion broadening
    - – Acoustic couplings within corner station between detection systems
      - Broadband 200 - 300 Hz

Summary and conclusions (2)

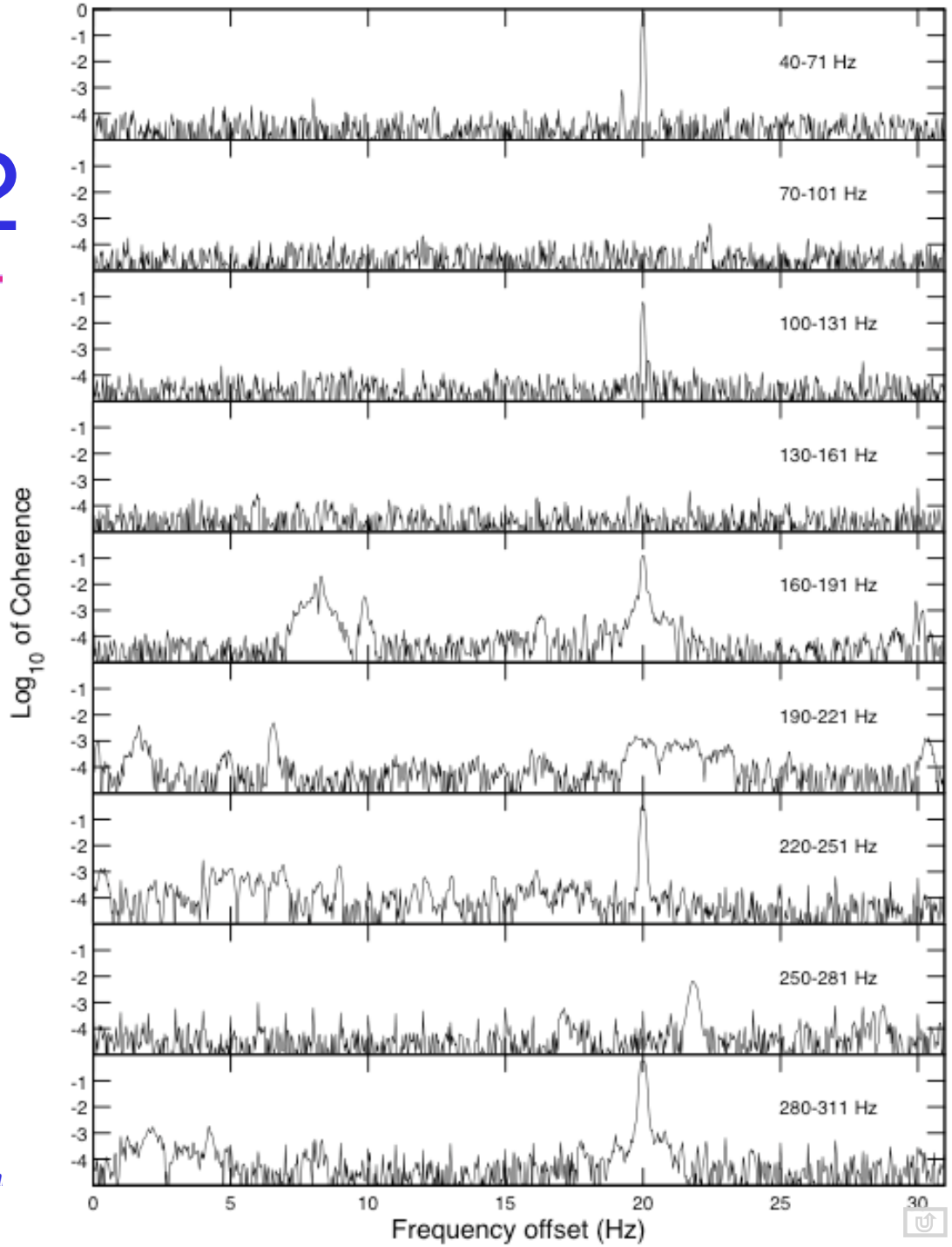
- Plans for S2 (and beyond)
  - » Take calibration variability, non-stationarity into account on the finest possible time scales
  - » Improve on calibration uncertainty
  - » Identify and eliminate or remove correlated noise sources for H1-H2
  - » Review, improve data analysis pipeline (e.g., high-pass filtering, line removal, frequency range, windowing, ...)
  - » Set up infrastructure for ALLEGRO-LLO, GEO-LIGO correlations
    - – ALLEGRO - LLO correlation may allow identification of instrumental biases vs signal
      - • ALLEGRO can be rotated, allowing for an improved analysis (Finn & Lazzarini, PRD 15 October 2001)
  - » Restructure analysis code for more efficient time-shift analysis, simulations
  - » ***Expected S2 upper-limit: $\Omega_0 (h_{100})^2 < 10^{-2}$ for H2-L1***
  - » ***Ultimate LIGO I (1 yr integration): $\Omega_0 (h_{100})^2 < 10^{-5}$ H2-L1***

FINIS


10^{-25}
10^{-30}
10^{-35}
10^{-40}
strain^2/Hz
Optimal
filter Q(f)
H1 PSD
H2 PSD
50
100
150
200
250
300
Frequency (Hz)


Fraction of maximum SNR
Cutoff frequency (Hz)
H1-L1
H1-H2
H2-L1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
50
100
150
200
250
300
350
400

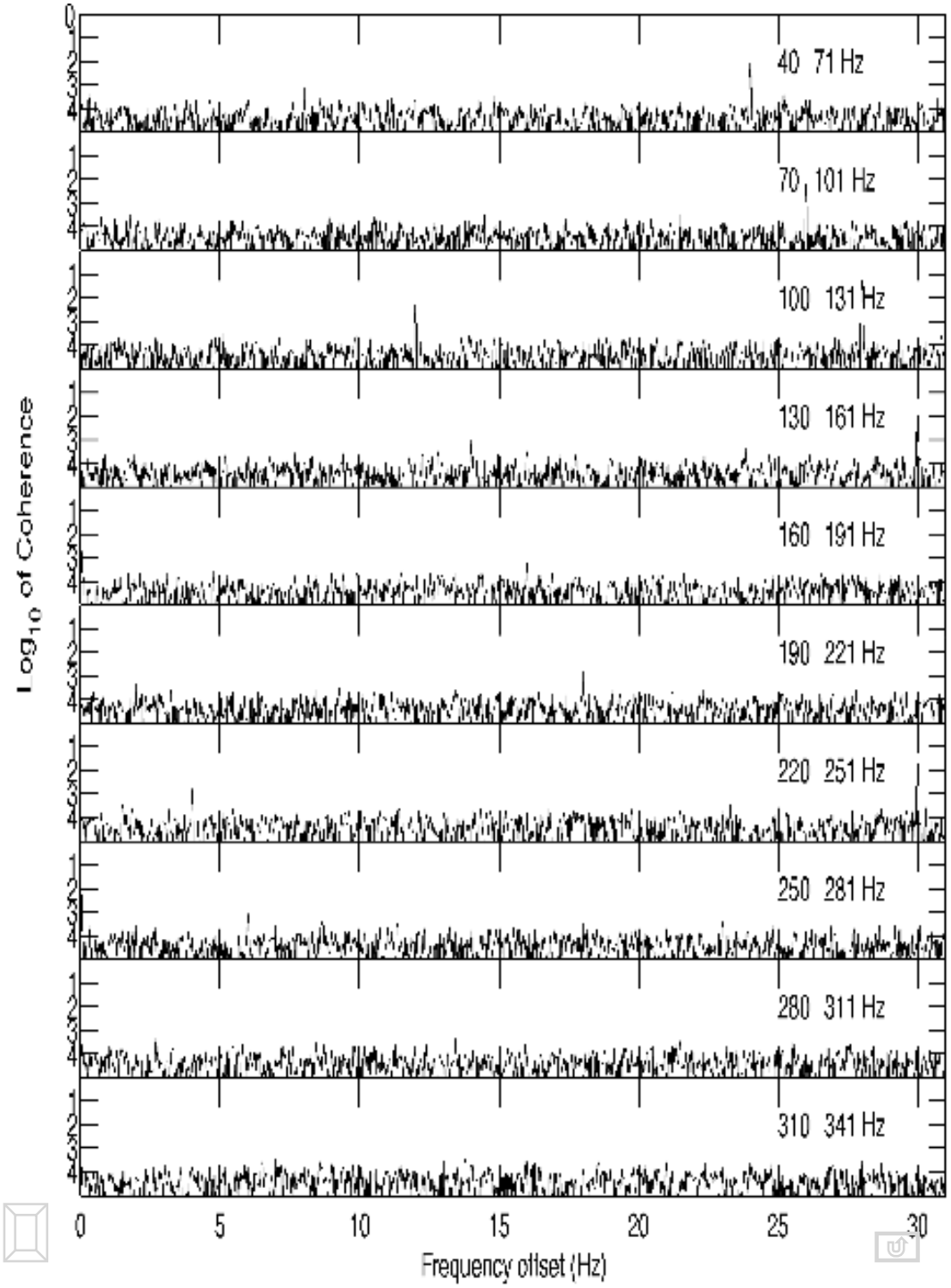
H1 - H2



LIGO-G030242-01-E
06 May 2003

Lazza

H2 - L1



LIGO-G030242-01-E
06 May 2003