First LIGO Search for Binary Inspirals

Peter Shawhan

(LIGO Lab / Caltech)

For the Inspiral Upper Limits Working Group
of the LIGO Scientific Collaboration

Penn State Center for Gravitational Physics and Geometry

March 31, 2003

Thanks to Gaby González and Albert Lazzarini for sharing visual materials

Outline

The First LIGO Science Run

Inspiral Search Fundamentals

Practical Matters

Rate Limit Calculation

The Future

The First Science Run — S1

August 23 – September 9, 2002 (17 days)

GEO ran simultaneously with LIGO

Collected data around the clock

Observatories manned by operators and scientific monitors

Operators keep interferometers working properly

Scimons watch data quality, work on "investigations"

Control-room tools:

Fully computerized control system

Data visualization software

Electronic logbook

Many computer/video screens!

State of LIGO Interferometers During S1

All three interferometers in “recycled” optical configuration

- Livingston 4 km — **L1**
- Hanford 4 km — **H1**
- Hanford 2 km — **H2**

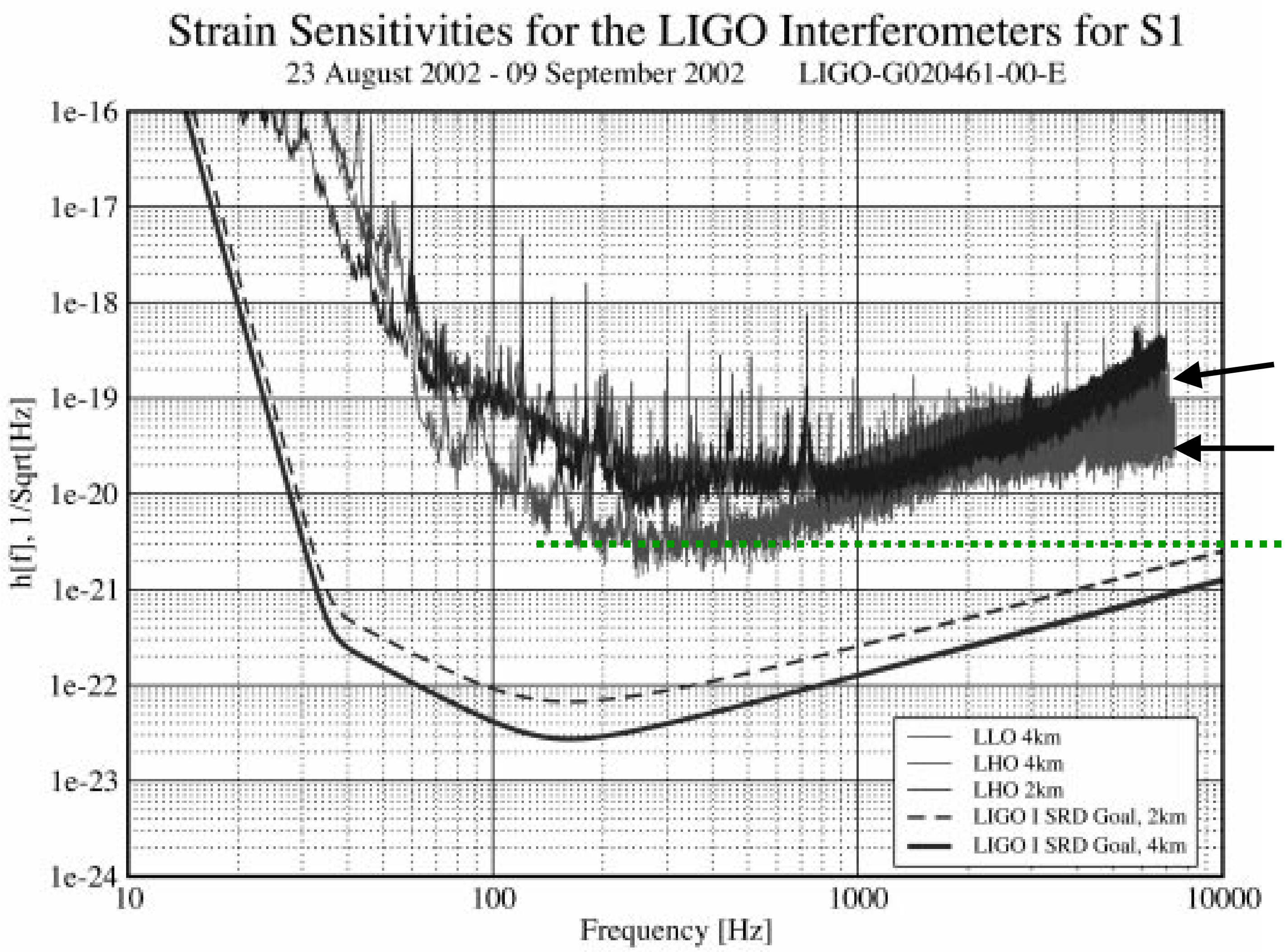
H2 was at full laser power, others at reduced power

All three used “common-mode servo” and Earth-tide compensation

Limitations:

- Ground noise at Livingston generally made it impossible to lock the interferometer during workdays
- Very little of auto-alignment system was operational ⇒ drifts
- Occasional extended difficulties with locking – due to alignment sensitivity?

Strain Sensitivities During S1



H1 & H2

L1

3×10^{-21} at ~300 Hz

Ranges for
Binary Neutron Star Inspirals

For an *optimally oriented* 1.4+1.4 $M_\odot$ binary system, to yield SNR=8 :

L1 ~175 kpc

H1 ~38 kpc

H2 ~35 kpc

Notes:

Averaging over orientations reduces these by a factor of sqrt(5)

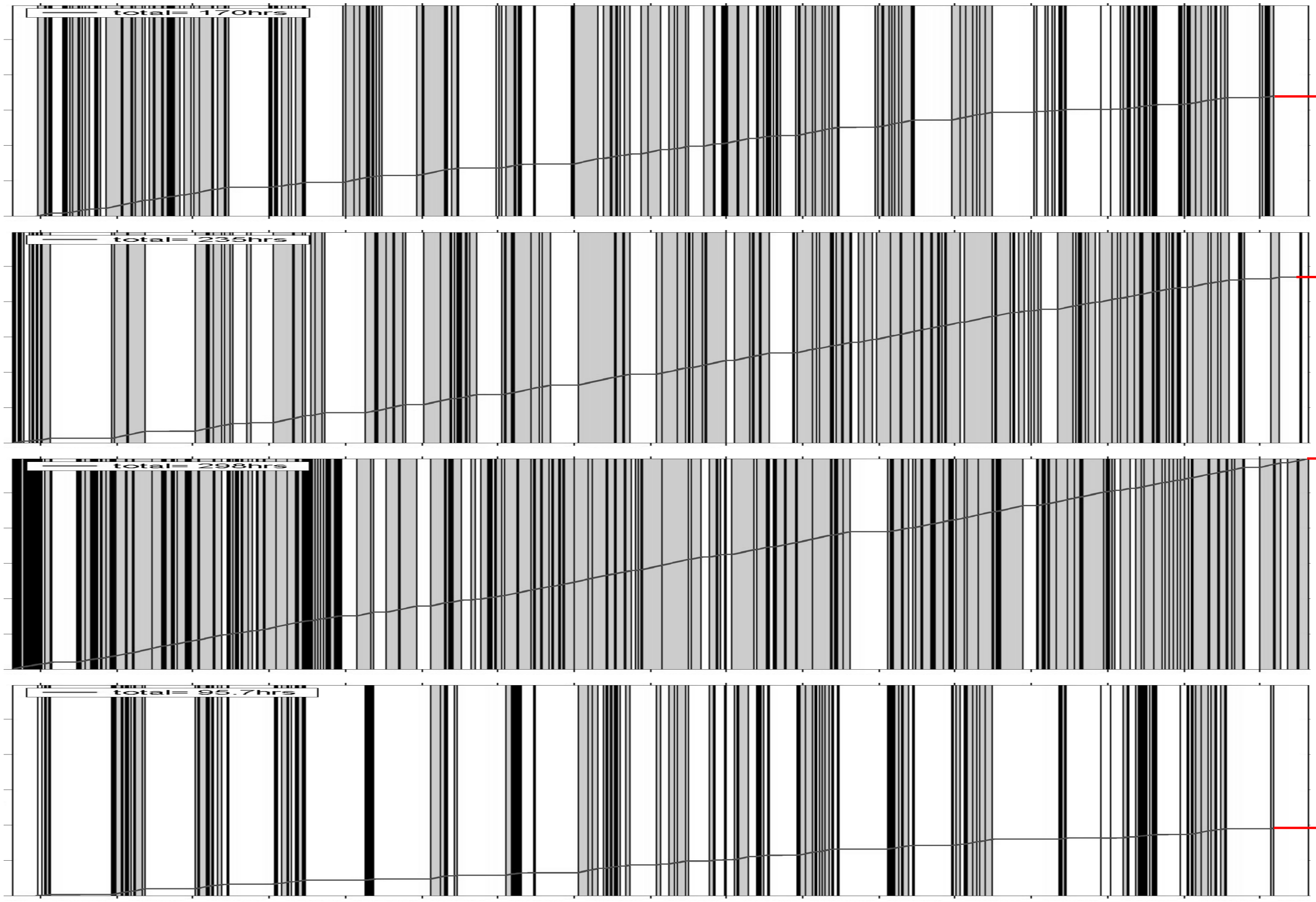
Range is nearly proportional to total mass of binary system

$\Rightarrow$ **L1 could detect almost all binary inspirals in Milky Way, and many in Magellanic clouds**

$\Rightarrow$ **H1 & H2 could detect most inspirals in Milky Way**

$\Big\}$ *if* noise is Gaussian and stationary, so that SNR=8 is enough

S1 Data Statistics



L1

170 hours

H1

235 hours

H2

298 hours

All 3

96 hours

17 days = 408 hours

S1 Data

Data stream includes a large number of channels

- The “gravitational-wave channel”, LSC-AS_Q
- Auxiliary interferometer sensing & control channels
- Environmental monitoring (seismometers, accelerometers, microphones, magnetometers, etc.)
- Control settings

AS_Q and aux Interferometer channels sampled at 16384 Hz

- Digital servo system
- Data volume: 5.8 MB/sec from Hanford, 2.9 MB/sec from Livingston

Full data set written to disk at observatories

- Then copied to tapes
- Full data set sent to Caltech and U. of Wisconsin–Milwaukee (UWM)
- Reduced data set generated and sent to MIT

Data Analysis Organization

Data Analysis is the job of the LIGO Scientific Collaboration

Four LSC "upper limit" working groups were formed

- Organized around signal types: burst, inspiral, continuous-wave, stochastic
- Most data analysis is done in the context of one of these groups
- Interact via weekly teleconferences, email lists, electronic notebooks, occasional face-to-face meetings

Inspiral Upper Limit Working Group

Led by Patrick Brady (UWM) and Gabriela González (LSU)

Others who contributed to this analysis:
Bruce Allen (UWM), Duncan Brown (UWM), Jordan Camp (Goddard), Vijay Chickarmane (LSU), Nelson Christensen (Carleton), Jolien Creighton (UWM), Carl Ebeling (Carleton), Valera Frolov (LLO), Brian O'Reilly (LLO), Ben Owen (Penn State), B. Sathyaprakash (Cardiff), Peter Shawhan (CIT)

Outline

The First LIGO Science Run

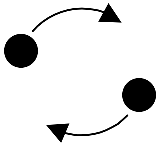
Inspiral Search Fundamentals

Practical Matters

Rate Limit Calculation

The Future

Gravitational Waves from Binary Inspirals

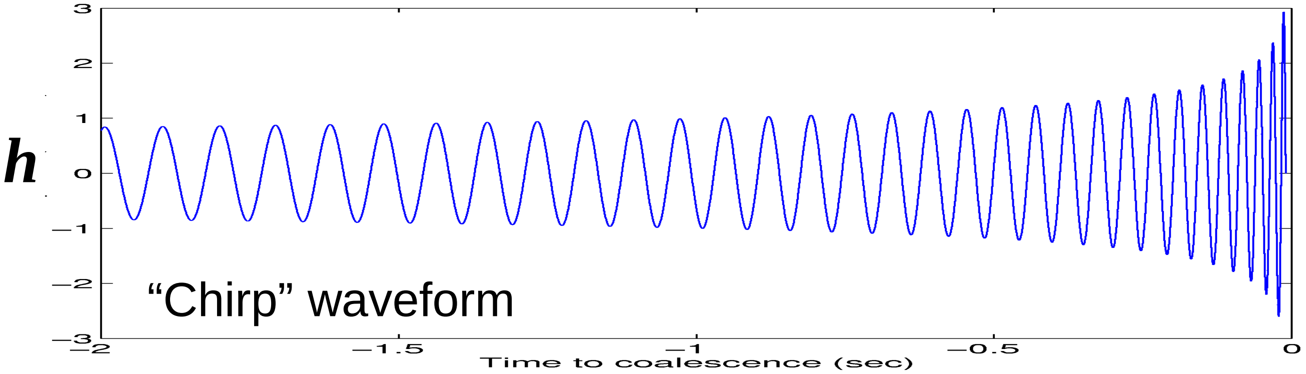


Binary in tight orbit emits gravitational waves

Loss of angular momentum causes orbit to decay

- Decay rate accelerates as orbital distance shrinks

Waveform is well known if masses are small



“Chirp” waveform

Enters LIGO sensitive band ~seconds before coalescence

Binary neutron star systems are known to exist !

- e.g. PSR 1913+16

Overview of the S1 Inspiral Search

Use matched filtering to search for the known waveforms of binary inspirals

Do filtering in frequency domain

- Weight frequencies according to noise spectrum

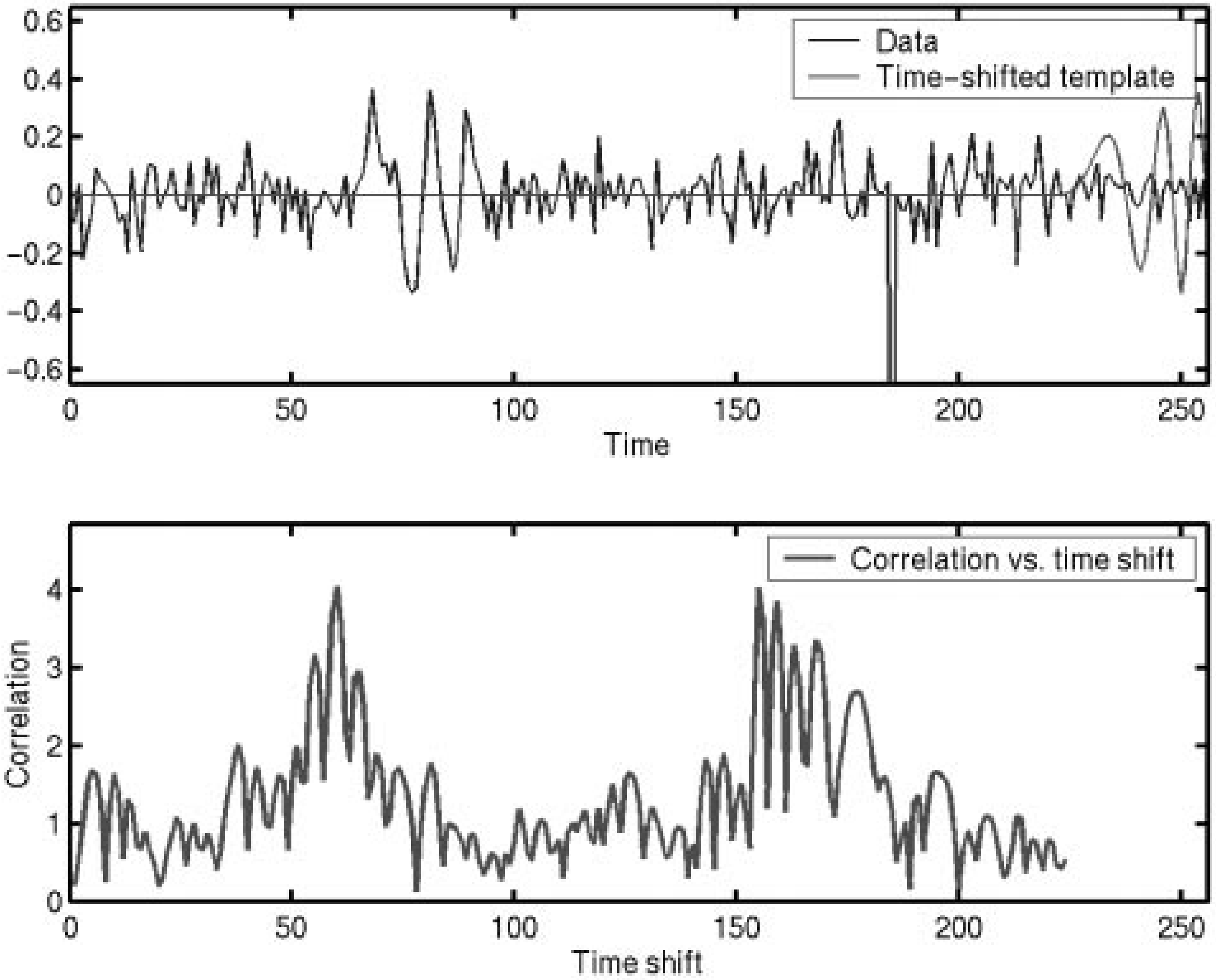
Lay out a "bank" of templates to cover parameter space

- Allow mass of each binary component to be between $M_\odot$ and 3 $M_\odot$
- Includes binary neutron star systems, nominally 1.4 + 1.4 $M_\odot$

Make sure that candidate signals have the expected distribution of signal power as a function of frequency

- Do a chi-squared test

Illustration of Matched Filtering



Optimal Filtering Using FFTs

Transform data to frequency domain : $\tilde{h}(f)$

Generate template in frequency domain : $\tilde{s}(f)$

Correlate, weighting by power spectral density of noise:

$$\frac{\tilde{s}(f)\ \tilde{h}^*(f)}{S_h(|f|)}$$

Then inverse Fourier transform gives you the filter output at all times:

$$z(t) = 4\int_0^\infty \frac{\tilde{s}(f)\ \tilde{h}^*(f)}{S_h(|f|)}\ e^{2\pi i f t}\ df$$

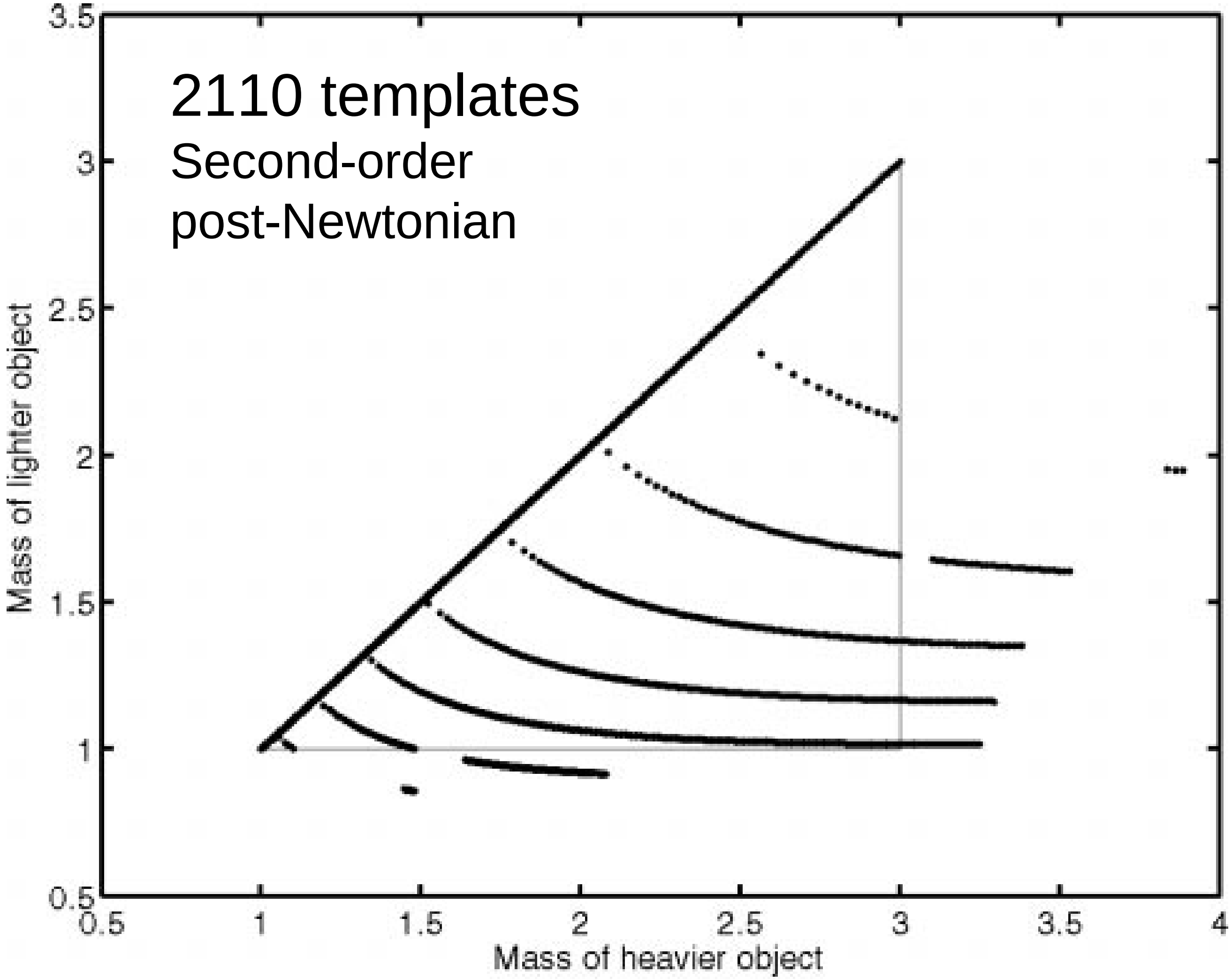
Find maxima of $|z(t)|$ over arrival time and phase

Characterize event by signal-to-noise ratio, ρ

Template Bank

Calculated based on L1 noise curve

Templates placed for maximum mismatch of $\delta = 0.03$



2110 templates

Second-order post-Newtonian

Chi-Squared Test

Any large transient in the data can lead to a large filter output

A real inspiral has signal power distributed over frequencies in a particular way

Divide template into p parts, each expected (on average) to contribute equally to ρ, and calculate a χ^2 :

$$\chi^2(t) = p \sum_{l=1}^{p} \left| z_l(t) - z(t)/p \right|^2$$

(We use p = 8)

"Veto" events with large χ^2

Allow for large signals which may fall between points in the template bank

$$\chi^2(t) \leq 5\left(p + \rho^2\delta^2\right)$$

Data Processing

The search was performed using routines in the LIGO Algorithm Library (LAL), running within the LIGO Data Analysis System (LDAS)

- Template bank is divided up among many PCs working in parallel ("flat" search)
- Most of the processing for this analysis was done on the UWM LDAS system, which has 296 PCs →

Each LDAS job processed 256 seconds of data

- Consecutive jobs overlapped by 32 seconds
- Events which exceeded an SNR threshold of 6.5 and passed the chi-squared veto were written to the LDAS database

Can we really detect a signal?

We used LIGO’s hardware signal injection system to do an end-to-end check

- Physically wiggle a mirror at the end of one arm
- Measure the signal in the gravitational-wave channel

Injected a few different waveforms at various amplitudes

- Example: 1.4+1.4 $M_\odot$, effective distance = 7 kpc 🔈

Signal was easily found by inspiral search code

- The 1.4+1.4 $M_\odot$ template had the highest SNR (= 92)
- Reconstructed distance was reasonably close to expectation
- Yielded a χ^2 value well below the cut

Outline

The First LIGO Science Run

Inspiral Search Fundamentals

Practical Matters

Rate Limit Calculation

The Future

Real Detectors…

… are not on all the time

- ⇒ Only process the good data (requires bookkeeping)
- ⇒ Need to decide how to use the data from each detector

… have time-varying noise

- ⇒ Discard data when detector was not very sensitive
- ⇒ Estimate noise from the data

… have a time-varying response

- ⇒ Need calibration as a function of time

… have "glitches"

- ⇒ Chi-squared veto
- ⇒ Veto on glitches in auxiliary interferometer channels

Making Choices about the Analysis Pipeline

Need to avoid the possibility of human bias when deciding:

- Which interferometers to use
- What data to discard
- Chi-squared veto cut
- Auxiliary-channel vetoes

Can’t make these decisions based on looking at the data from which the result is calculated !

Set aside 10% of triple-coincidence data as a “playground”

- Make all decisions based on studying this sample
- Hope it is representative of the full data set
- Avoid looking at the remaining data until all choices have been made

Final result is calculated from the remaining data

Data Set Selection

We choose to use L1 and H1 only

- H2 was the least sensitive, and glitchier than the others

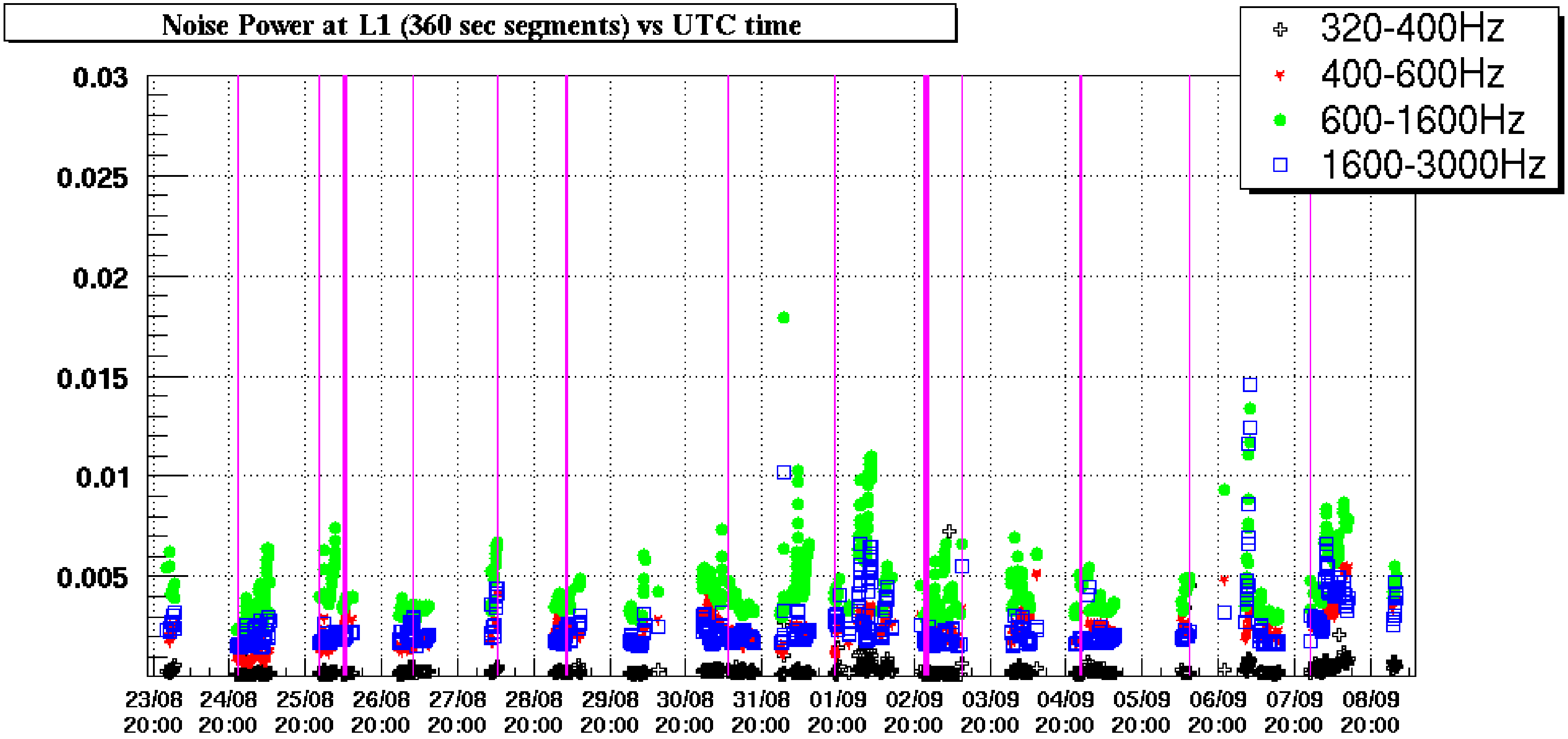
Even when locked, interferometer was not always stable

- Settling down at the beginning of a lock
- Periodic tuning of alignment to maximize light stored in arms
- Operators marked "science mode" data while running – guarantees that no control settings were being changed

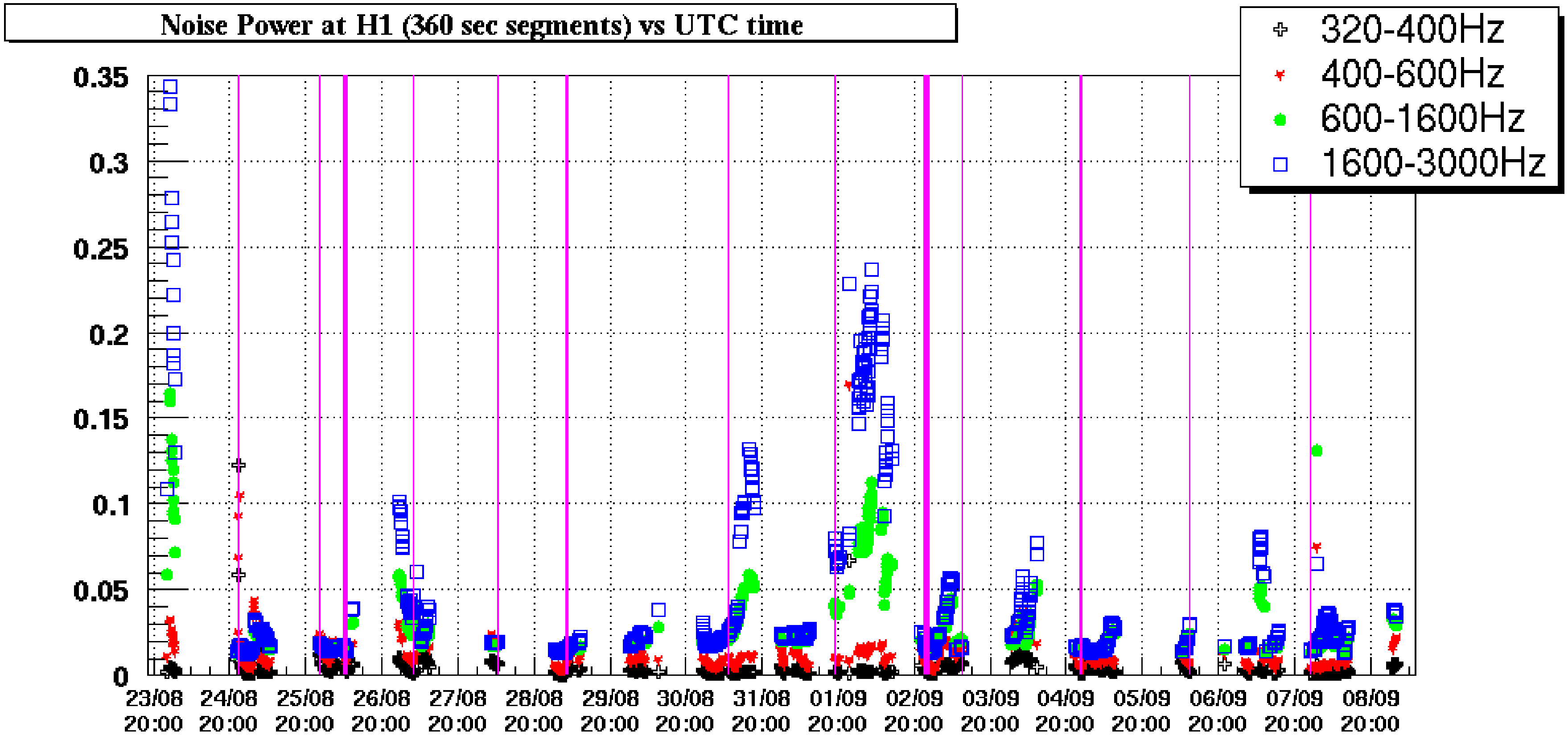
We choose to discard science-mode data when noise is larger than normal — "epoch veto"

- Noise power calculated in four frequency bands
- Entire "segment" of data is discarded if any band exceeds a threshold
- Cuts 23% of L1 data, 31% of H1 data

Epoch Veto Bands for L1



Epoch Veto Bands for H1



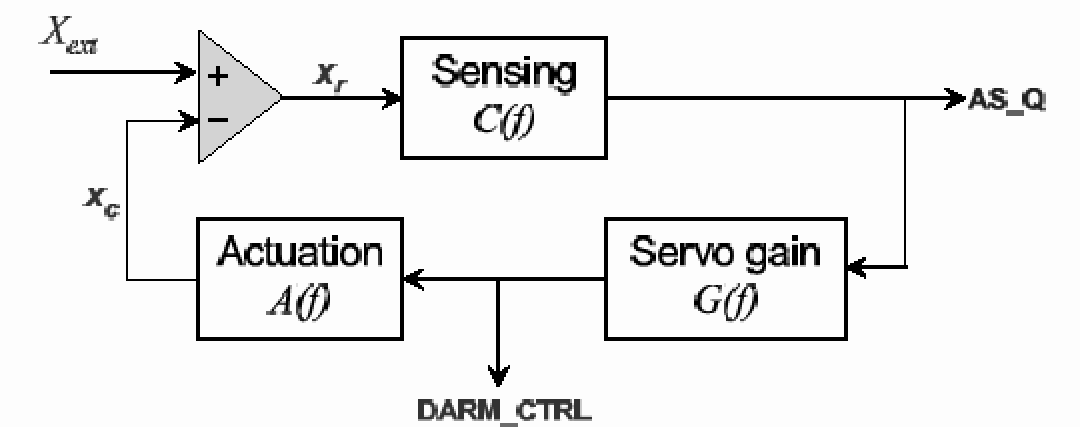
Noise Estimation

- **Crucial, since it enters into the calculation of SNR**
- **Power spectral density (PSD) of noise is calculated from the data which is input to each LDAS job**
  - Calculated by averaging PSDs from 7 overlapping 64-sec time intervals
  - This includes any signal which may be in the data, but that’s OK
- **Optimal filtering in frequency domain requires us to assume that the PSD is constant for the whole job**
  - This isn’t necessarily true !

Calibration

Optical sensing is inherently frequency-dependent

Servo system introduces additional frequency dependence



Occasionally measure complete transfer function

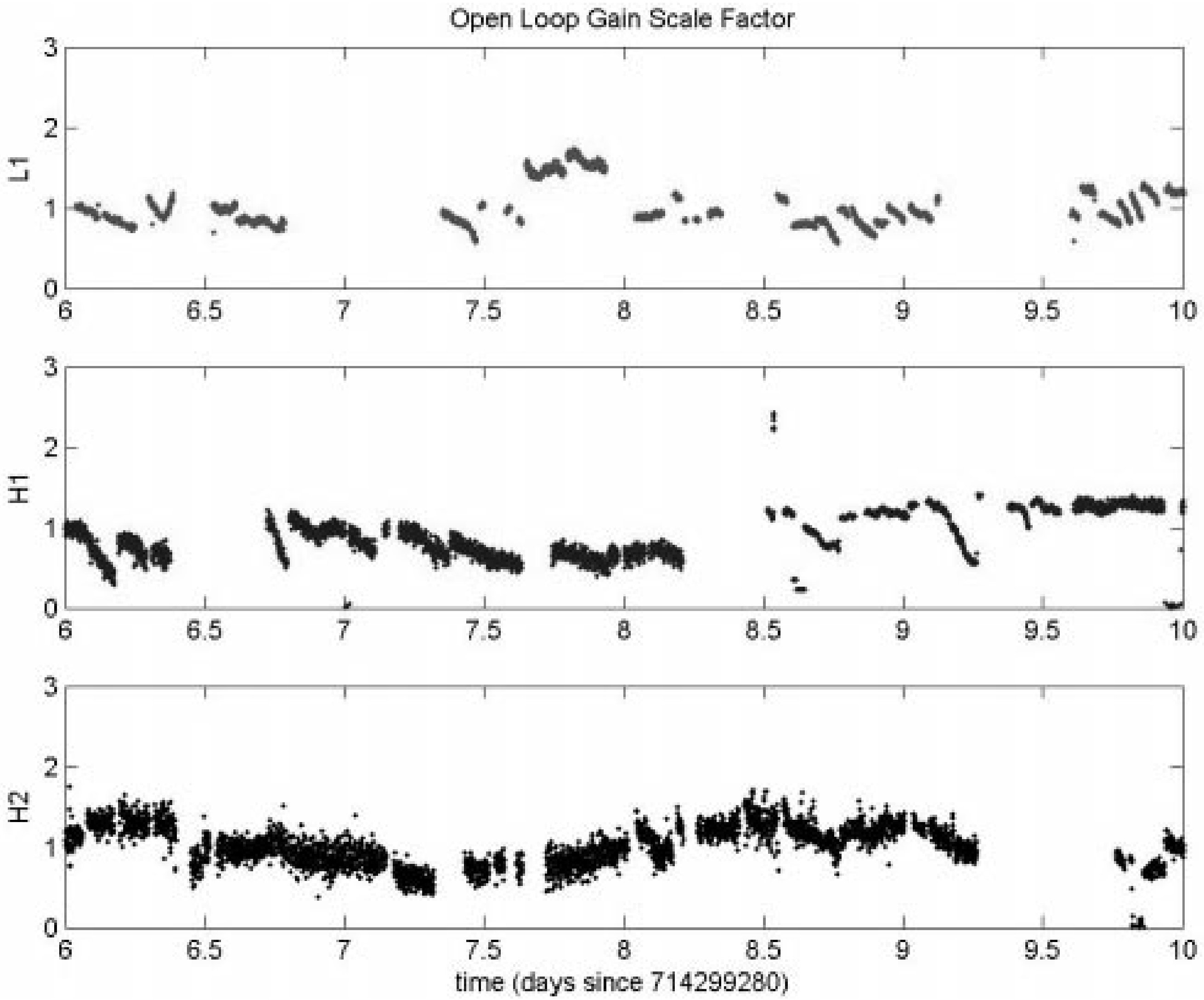
Continuously inject “calibration lines”

- Sinusoidal wiggles on an end mirror, at a few frequencies
- Allow us to track variations in the optical response over time

Effect of Changing Optical Gain

Affects phase as well as amplitude—important for matched filtering

Calibration Stability



Auxiliary-Channel Vetoes

There are “glitches” in the gravitational-wave channel

- Transients larger than would be expected from Gaussian stationary noise
- Seen, at some level, in all three interferometers
- Chi-squared veto eliminates many, but not all

We checked for corresponding signatures in other channels

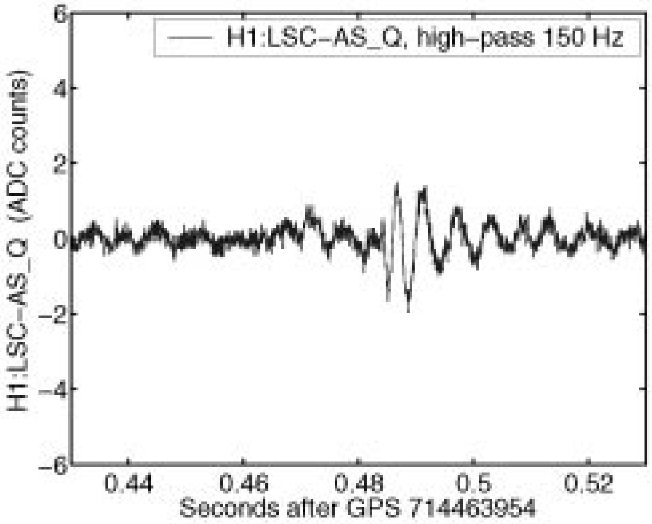
- Environmental channels (accelerometers, etc.)
- Auxiliary interferometer channels

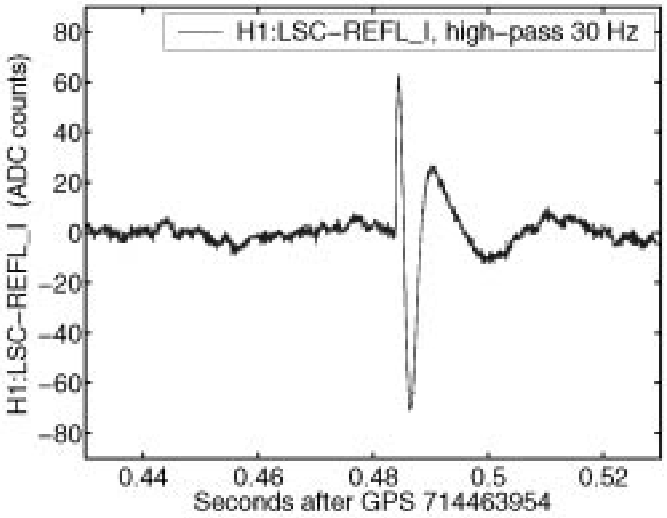
Tried a few glitch-finding algorithms

$\left.\begin{array}{l}\text{absGlitch}\\ \text{glitchMon}\end{array}\right\}$ Part of the LIGO Data Monitoring Tool (DMT)

- Inspiral search code (!)

Big Glitches in H1





← Found by inspiral search code with SNR=10.4

These occurred ~4 times per hour during S1

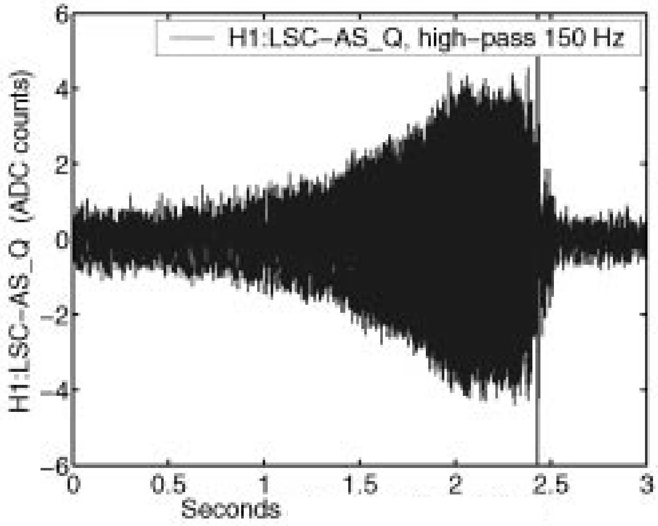
"REFL_I" channel has a very clear transient for almost all such glitches in H1

Use glitchMon to generate veto triggers

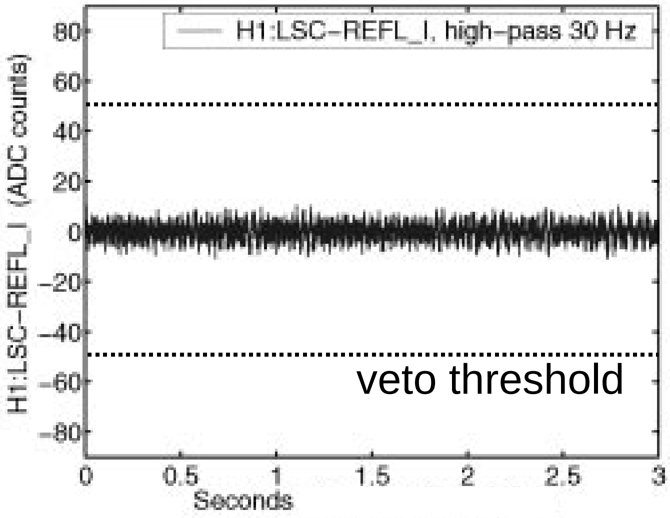
Veto Safety

Have to be sure a real gravitational wave wouldn't couple into the auxiliary channel strongly enough to veto itself !

Check using hardware signal injection data



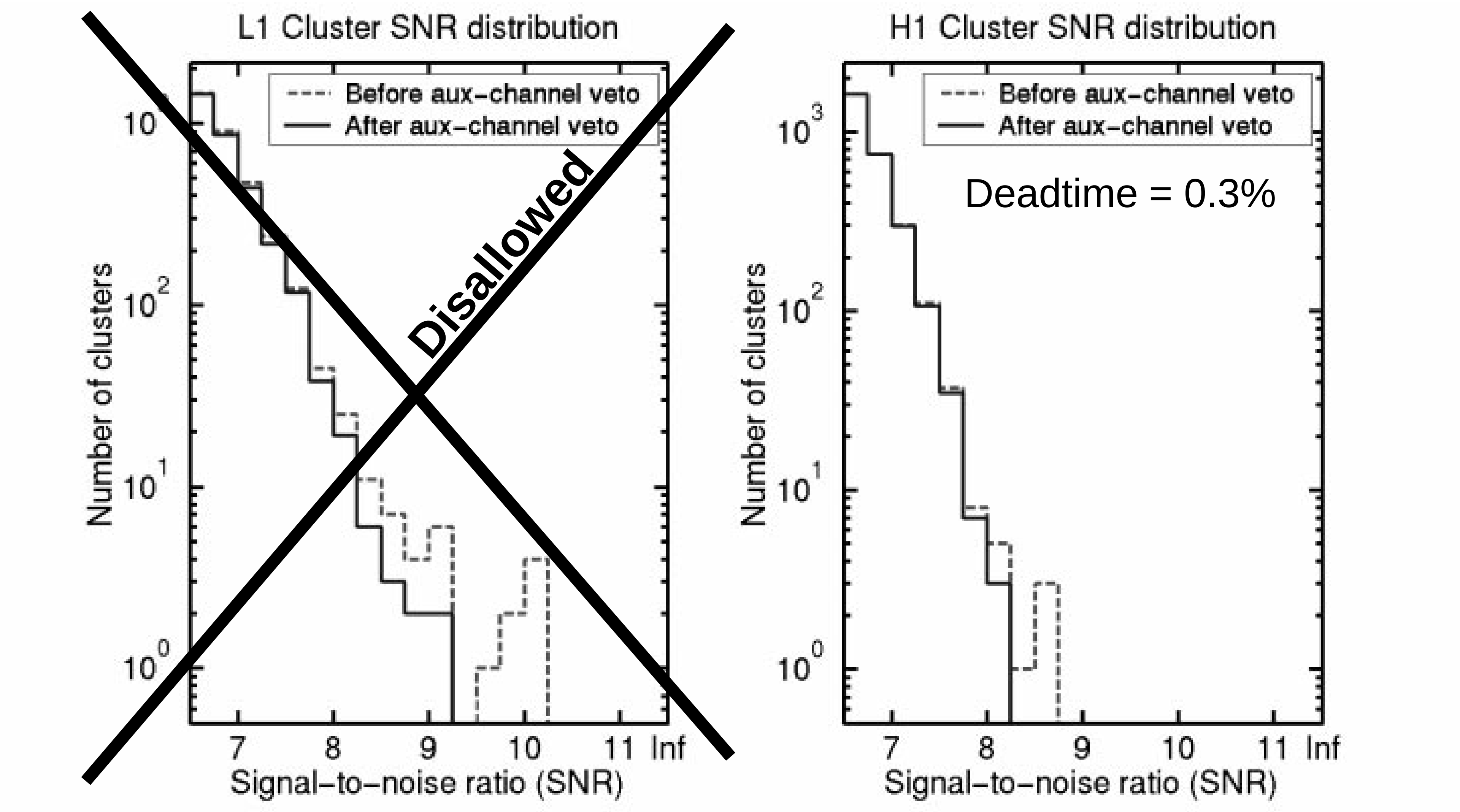
No sign of signal in REFL_I



Best veto channel for L1 ("AS_I") was disallowed because there was a small but measurable coupling

Effect of Vetoes

on Playground Data



Disallowed

Deadtime = 0.3%

Outline

The First LIGO Science Run

Inspiral Search Fundamentals

Practical Matters

Rate Limit Calculation

The Future

Strategy

Expected rate in Milky Way is very low

Perhaps only 10^{-6} per year for binary neutron stars !

Simultaneous observation with multiple detectors gives us a chance to make a (surprising) discovery

Look for coincident event(s) in excess of random background rate

Random background rate can be estimated with time-shift analysis

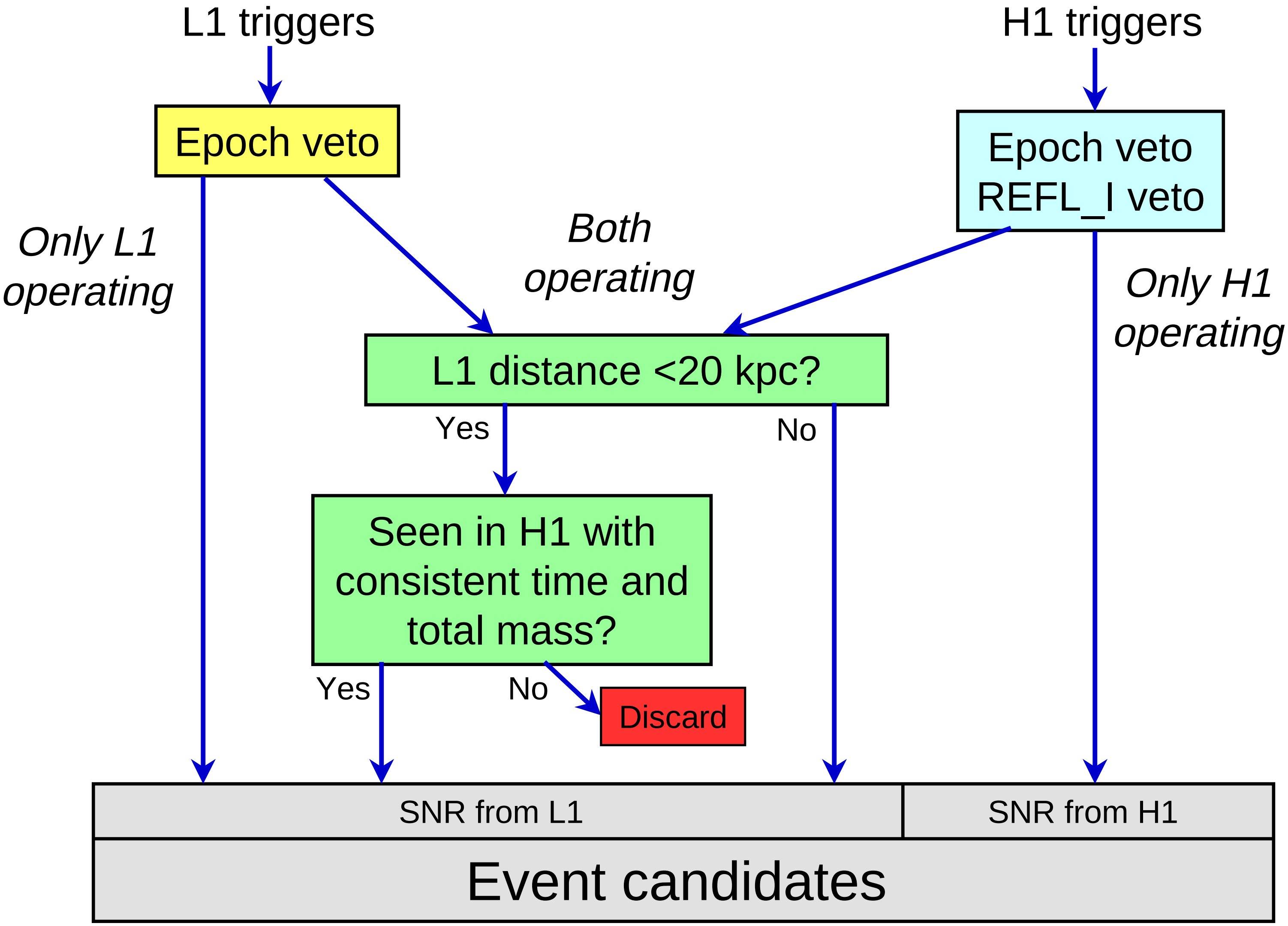
Realistically, analysis will probably yield an upper limit

Can use single-interferometer data to increase observing time

L1 *or* H1 : 289 hours   vs.   L1 *and* H1 : 116 hours

Judging from playground data, this should yield a tighter upper limit

Analysis Pipeline



L1 triggers

H1 triggers

Epoch veto

Epoch veto
REFL_I veto

Only L1 operating

Both operating

Only H1 operating

L1 distance <20 kpc?

Yes

No

Seen in H1 with consistent time and total mass?

Yes

No

Discard

SNR from L1

SNR from H1

Event candidates

Statistical Method

Add together SNR distributions from all 4 categories

No reliable way to estimate the background for single-interferometer events

Would not claim a detection based on this summed-SNR method

Hard to know *a priori* where one should set SNR threshold

$\Rightarrow$ Use the "maximum-SNR statistic" to set upper limit

Useful since candidate events are so sharply peaked at low SNR

Yields a frequentist upper limit

$$R < \frac{2.3}{\varepsilon \, T} \quad \text{at 90\% C.L.}$$

Efficiency of analysis pipeline above observed max SNR

Observation time

Calculating the Efficiency of the Analysis Pipeline

Use a Monte Carlo simulation of sources in the Milky Way and Magellanic Clouds

- Mass and spatial distributions taken from simulations by Belczynski, Kalogera, and Bulik, Ap J **572**, 407 (2002)
- Inspiral orientation chosen randomly
- Distribution of Earth orientation is same as for S1 data

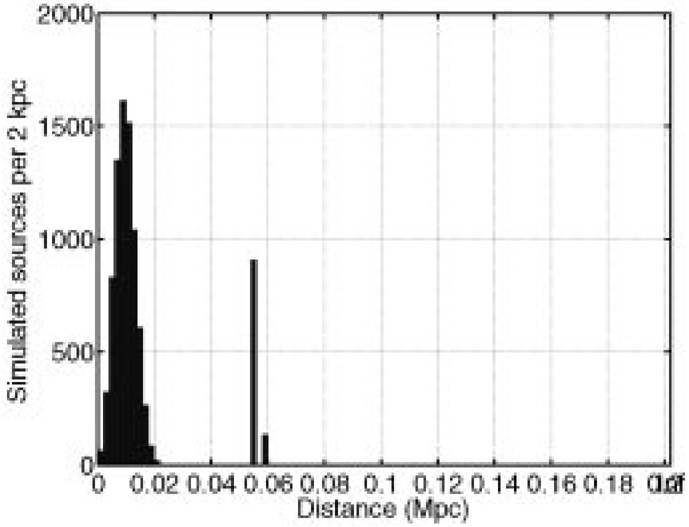
Add simulated waveforms to the real S1 data

Run the full analysis pipeline

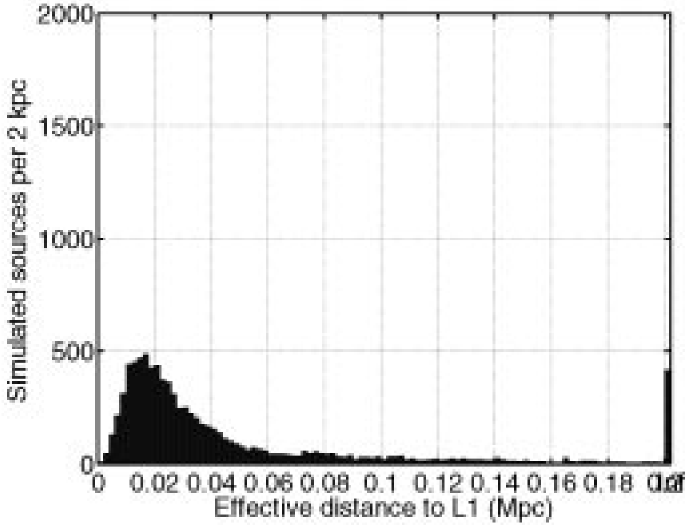
See what fraction of simulated events are found

Distributions from the Simulation

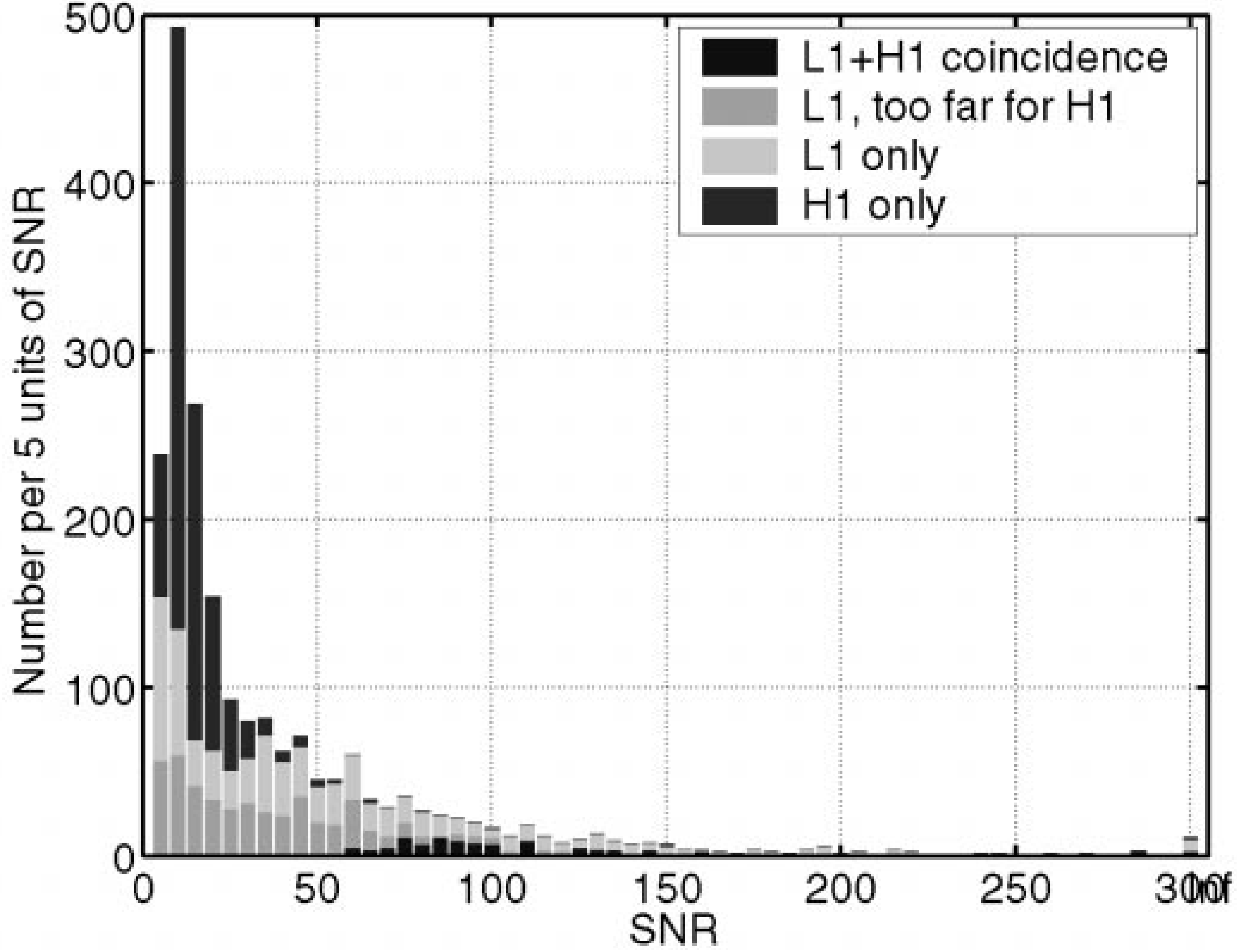
Actual Distance



Effective Distance



SNR Distribution from Simulation



Preliminary Result

(as presented at AAAS Meeting)

Analyzing full dataset yields a maximum SNR of 15.9

- This event seen in L1 only, with effective distance = 95 kpc
- Several others with SNR>12 (inconsistent with Gaussian stationary noise)
- No candidates were seen in coincidence in L1 and H1

Pipeline efficiency for Monte Carlo (require SNR$\geq$15.9) : 0.35

Observation time = 295.3 hours

$\Rightarrow$ *R* < 170 per year at 90% C.L. *

*** Note: This is not the final result**

- It was calculated without using the epoch veto
- An incorrect mass distribution was used for the simulation
- Final result will be somewhat different

Plans to Finish This Analysis

Currently re-doing simulation

Still some systematics to evaluate

- Calibration uncertainty
- Uncertainties in power spectrum estimation
- Modeling of sources in galaxy

A paper has been drafted

- Focuses on method as well as giving the result
- Has been reviewed by LSC internal review committee
- Presented at LSC Meeting two weeks ago
- Hope to submit it in a month or so

We *must* finish this soon and move on to later data

Outline

The First LIGO Science Run

Inspiral Search Fundamentals

Practical Matters

Rate Limit Calculation

The Future

The S2 Run

Now in progress !

Began February 14, runs through April 14

Detector sensitivities are much better than for S1

Duty factors are similar to S1

L1: 38%

H1: 72%

H2: 55%

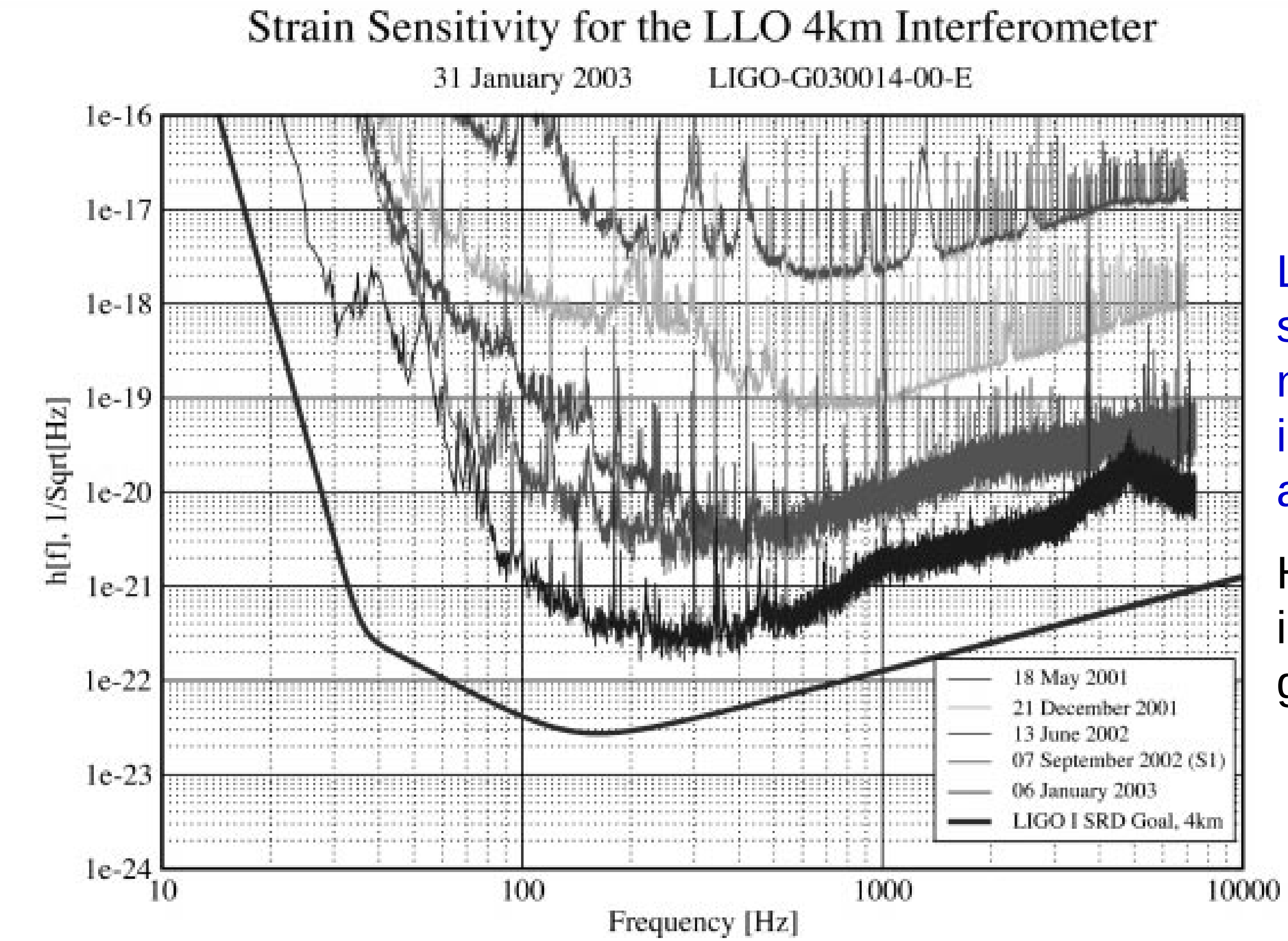
Improvements since S1:

Better alignment control, especially for H1

Better monitoring in the control rooms

Inspiral search code is being run in near-real-time for monitoring purposes

Sensitivity Improvements



L1 can now see binary neutron stars in Andromeda and M33 !

H1 & H2 have improved greatly too

Future Directions for Inspiral Searches

Study additional veto techniques

- Some obvious glitches survive the chi-squared veto
- The chi-squared veto does not use “off-chirp” information

Do coherent analysis of data from multiple detectors

- Restructure analysis pipeline

Search for higher-mass binaries

- Challenge to get accurate waveforms

Search for low-mass MACHO binaries

- Primordial black holes in halo of our galaxy ?

Implement hierarchical search algorithms

Summary

The S1 run provided good data

We had good efficiency for sources throughout our galaxy

We've learned a lot about the details of doing a full analysis

Mechanics of data processing

Calibration, vetoes, multi-detector strategy, statistical methods, …

Much better data is being collected now

S2 only yields a modest increase in number of binary NS inspiral sources

The real payoff will come when we reach the Virgo Cluster

⇒ This is only the first of many inspiral searches !