Data analysis for impulsive signals using interferometers

Peter R. Saulson
Syracuse University
Spokesperson, LIGO Scientific Collaboration

Rilke on gravitational wave detection

Though we are unaware of our true status,
our actions stem from pure relationship.
Far away, antennas hear antennas
and the empty distances transmit …

Pure readiness. Oh unheard starry music!
Isn't your sound protected from all static
by the ordinary business of our days?

from *The Sonnets to Orpheus*, First Part, number XII
by Rainer Maria Rilke, translated by Stephen Mitchell

Outline

- Where I come from: LIGO Scientific Collaboration
- What we are doing in the search for gravitational waves
- How should one go about looking for poorly-modeled burst signals at low SNR?
- A survey of the methods we are trying
- Our results to date
- A glimpse into the future

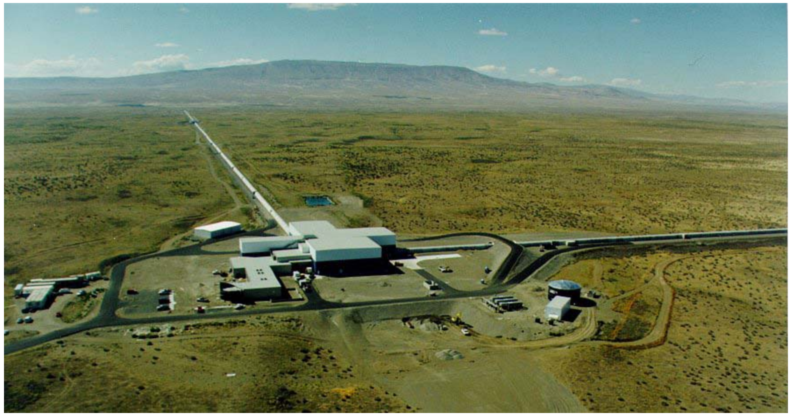
Note to students:

The search with interferometers has a few novel features, but almost all of them have been strongly inspired (if not taken directly) from the search with resonant detectors.

Differences are often a matter of degree only.

As resonant detector bandwidths grow, even those differences may shrink.

LIGO and GEO



Four interferometers contribute data to LSC analyses:

•4 km and 2 km interferometers at LIGO Hanford Observatory

•4 km interferometer at LIGO Livingston Observatory

•GEO600

N.B.: No GEO data available for S2, but back on air for S3.

LIGO

LIGO Scientific Collaboration

LSC

NSF

UTB TSC

Southeastern Louisiana University

California State University Dominguez Hills

Leland Stanford Junior University

Northwestern University

GEO 600

LSU

Universitat de les Illes Balears

Southern University

Hobart and William Smith Colleges

The University of Adelaide Australia

California Institute of Technology

University of Florida

NAOJ 国立天文台

The University of Birmingham

Loyola University New Orleans

University of Wisconsin UW Milwaukee

Carleton College

Syracuse University

Cardiff University

IGR

Universität Hannover

Washington State University

The Physics of the Universe PPARC

PAH

The University of Western Australia

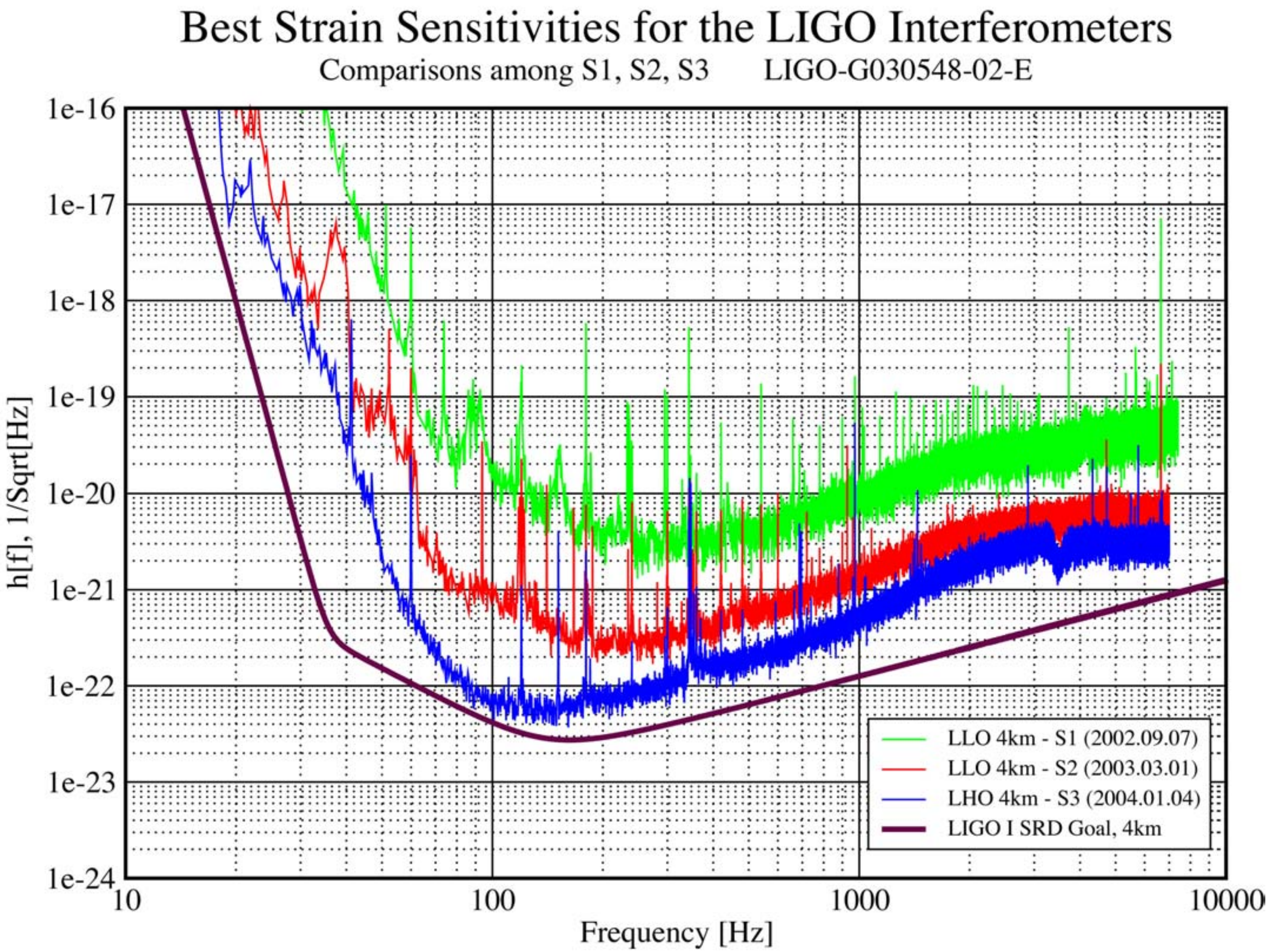
Michigan

Penn State

Massachusetts Institute of Technology

University of Glasgow

Sensitivity



Data Runs

We have carried out three Science Runs (S1--S3) interspersed with commissioning.

S1 run:

17 days (August / September 2002)

Four detectors operating: LIGO (L1, H1, H2) and GEO600

Triple-LIGO-coincidence (96 hours)

Four S1 astrophysical searches published (*Phys. Rev. D* 69, 2004):

- » **Inspiraling neutron stars 122001**
- » **Bursts 102001**
- » **Known pulsar (J1939+2134) with GEO 082004**
- » **Stochastic background 122004**

Data Runs

S2 run:

59 days (February—April 2003)

Four interferometers operating: LIGO (L1, H1, H2) and TAMA300 plus Allegro bar detector at LSU

Triple-LIGO-coincidence (318 hours)

Many S2 searches under way

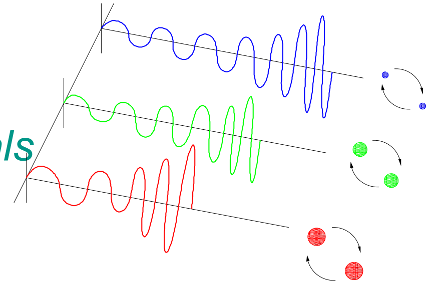
S3 run:

70 days (October 2003 – January 2004) – Analysis ramping up…

We search for four classes of signals

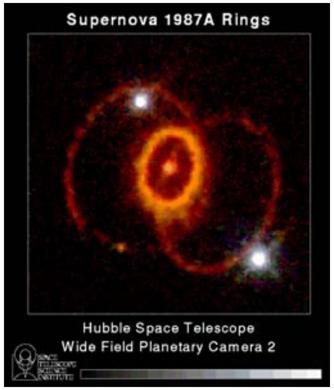
- Chirps

  "sweeping sinusoids" from compact binary inspirals



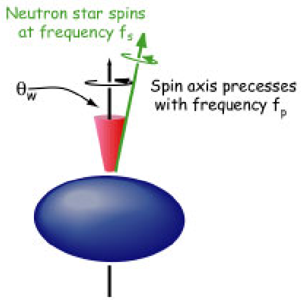
- Bursts

  transients, usually without good waveform models



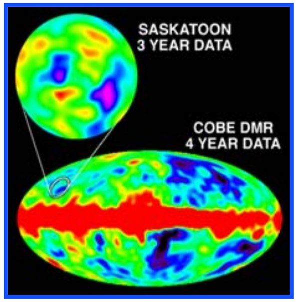
- Periodic, or "CW"

  from pulsars in our galaxy



- Stochastic background

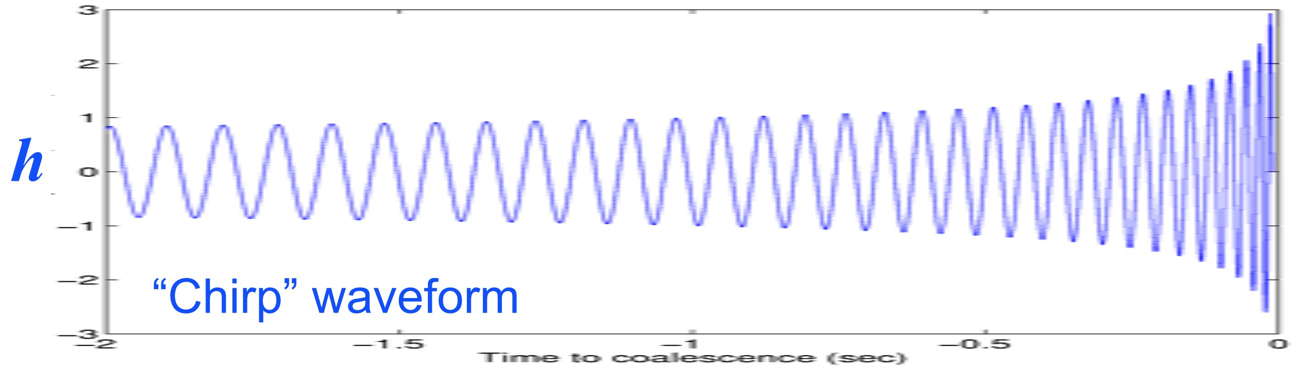
  cosmological background, or superposition of other signals



LIGO-G040435-00-Z

Inspiral Gravitational Waves

Compact-object binary systems lose energy due to gravitational waves. Waveform traces history.



In LIGO frequency band (40–2000 Hz) for a short time just before merging:
anywhere from a few minutes to <<1 second, depending on mass.

Waveform is known accurately for objects up to ~3 $M_\odot$

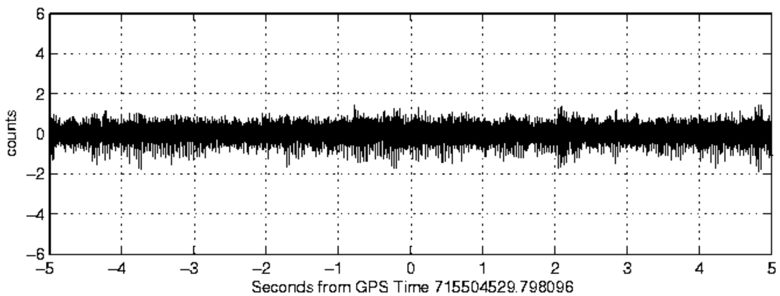
"Post-Newtonian expansion" in powers of (Gm/rc^2) is adequate

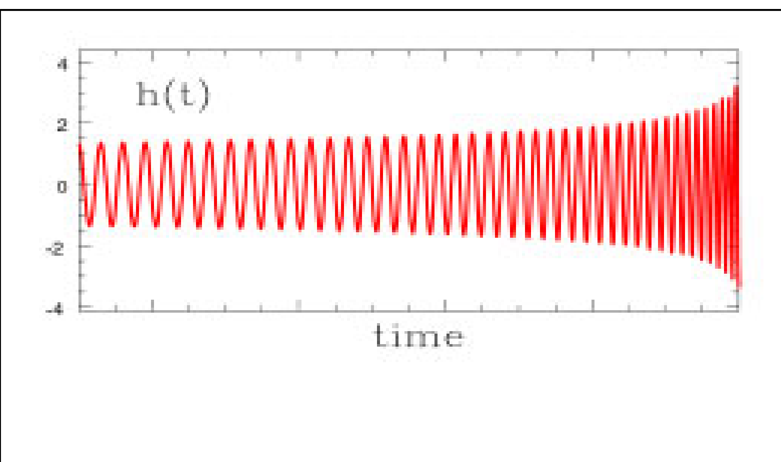
➔ Use *matched filtering.*

Matched filtering:

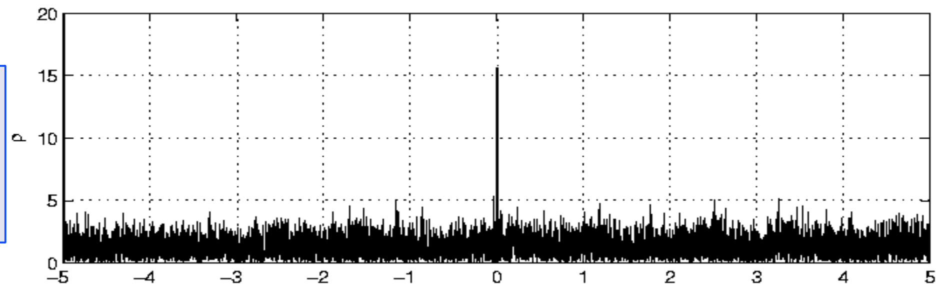
cross-correlation of data with known signal waveform

(whitened)GW Channel
+ simulated inspiral





SNR



Coalescence Time

Cross-correlation

$$s_1(t) = h(t - t_1) + n_1(t)$$

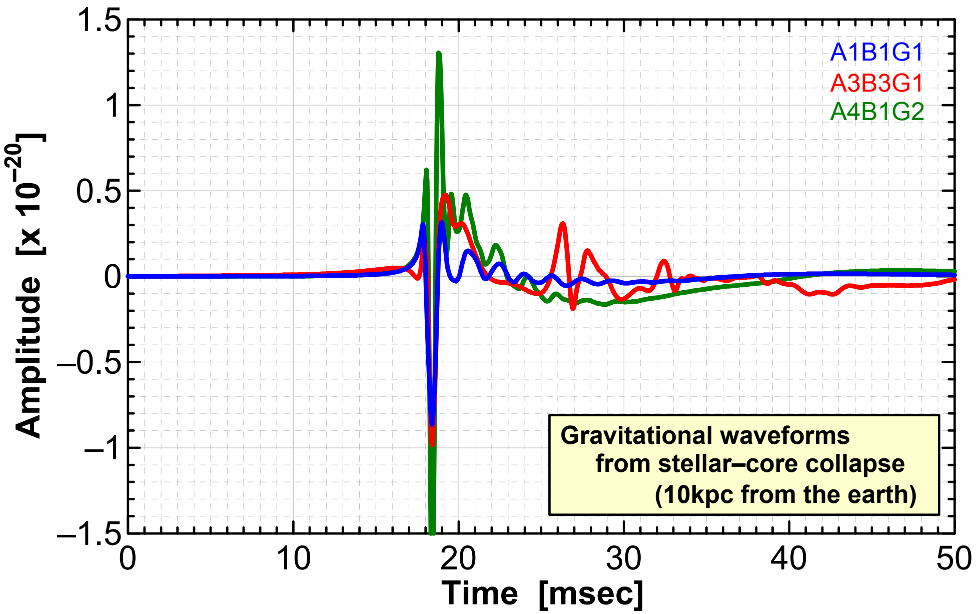
$$h_{known}(t - t_2)$$

$$C(t, t_w, t_{off}) = \int_{t-t_w/2}^{t+t_w/2} s_1(t') h_{known}(t' + t_{off}) dt'$$

$$\Rightarrow \int_{t_w} h^2(t) dt + \text{filtered noise}$$

Burst Signals

- Catastrophic events involving solar-mass compact objects can produce transient "bursts" of gravitational radiation in the LIGO frequency band:
  - core-collapse supernovae
  - merging, perturbed, or accreting black holes
  - gamma-ray burst engines
  - cosmic strings
  - others?
- Precise nature of gravitational-wave burst (GWB) signals typically unknown or poorly modeled.
  - Can't base such a broad search on having precise waveforms.
  - Search for generic GWBs of duration ~1ms-1s, frequency ~100-4000Hz.



possible supernova waveforms
T. Zwerger & E. Muller, Astron. Astrophys. *320* 209 (1997)

The burst challenge

Interpretation: Broadband observations can reveal details of a signal's waveform, and thus of the dynamics of the source.

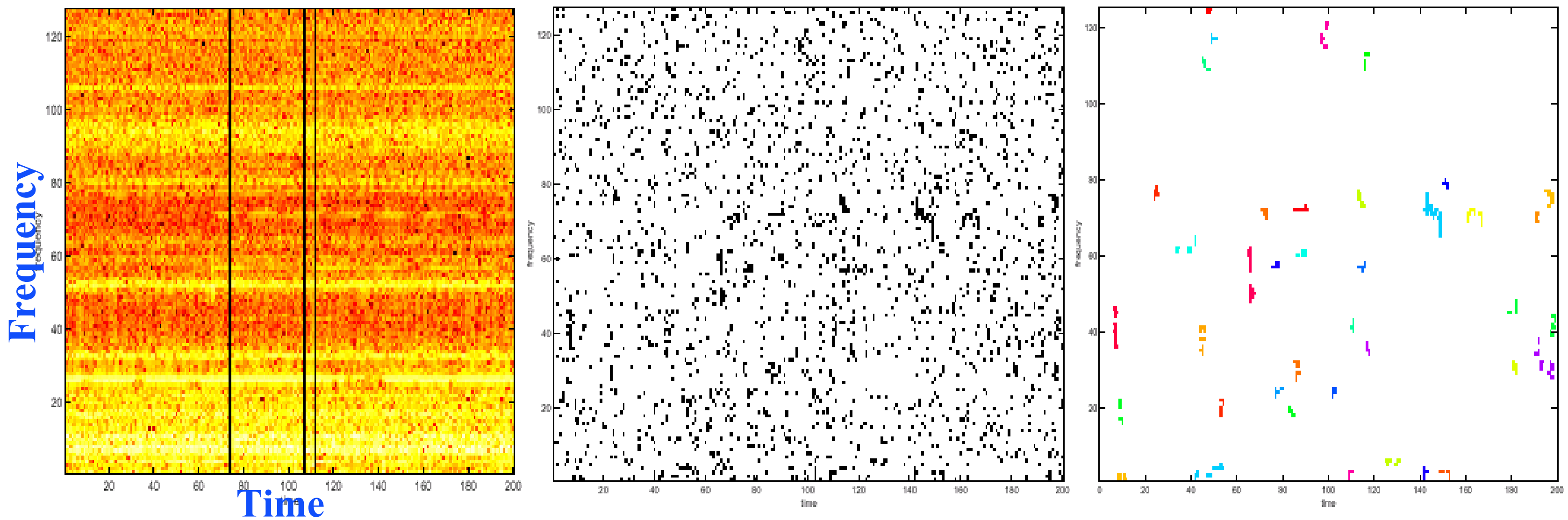
Detection: If the waveform is known accurately, then we can use the waveform as a *matched template* to optimally detect the signal.

But without knowledge of the waveform, we have neither the optimal ability to detect the signal, nor a good criterion for choosing among many possible sub-optimal ways to look for the signal.

The LSC Burst Group has experimented with a variety of *Event Trigger Generators* to search for signal candidates in LIGO data.

Burst search example: Time-frequency methods

One way to search for bursts is by looking for transients in the *time-frequency plane*. Here, we illustrate the TFCLUSTERS algorithm.



- Compute t-f spectrogram, in short-duration time bins.
- Threshold on power in a pixel; search for clusters of pixels.
- Find coincident clusters in outputs of all interferometers.

Burst detection methods

LSC Burst group has used several variations on time-frequency methods.

Also,

- » Time-domain search for statistical *change points*
- » Simple time-domain filtering
- » Matched filtering for special cases where waveform is known
  - – Black hole "ringdowns"
  - – New search for cosmic string cusp and kink events.

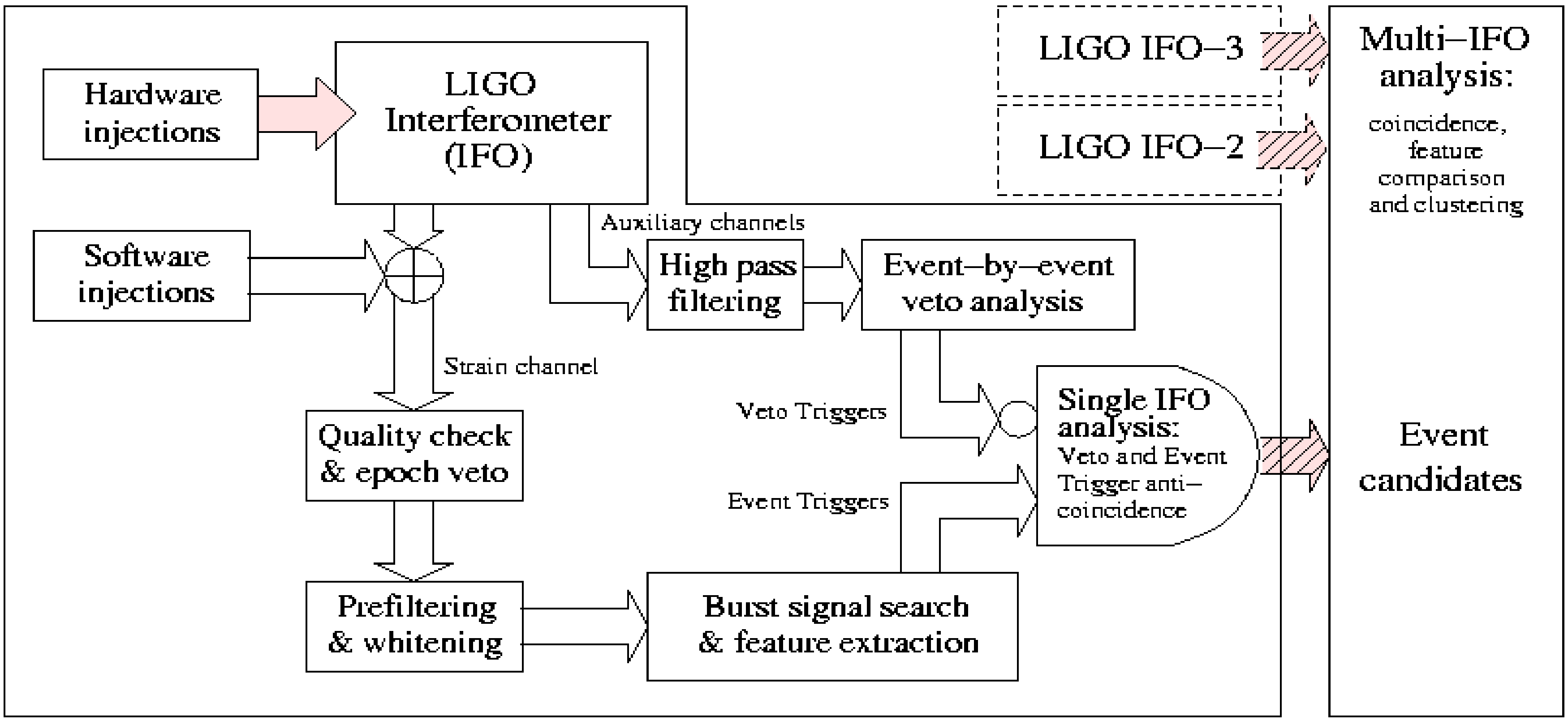
Test the list of coincident event candidates for *coherence* between signals from three LIGO interferometers.

Coherence test is also the basis of a search for signals associated with times of gamma-ray bursts or other astrophysical events ("triggered search".)

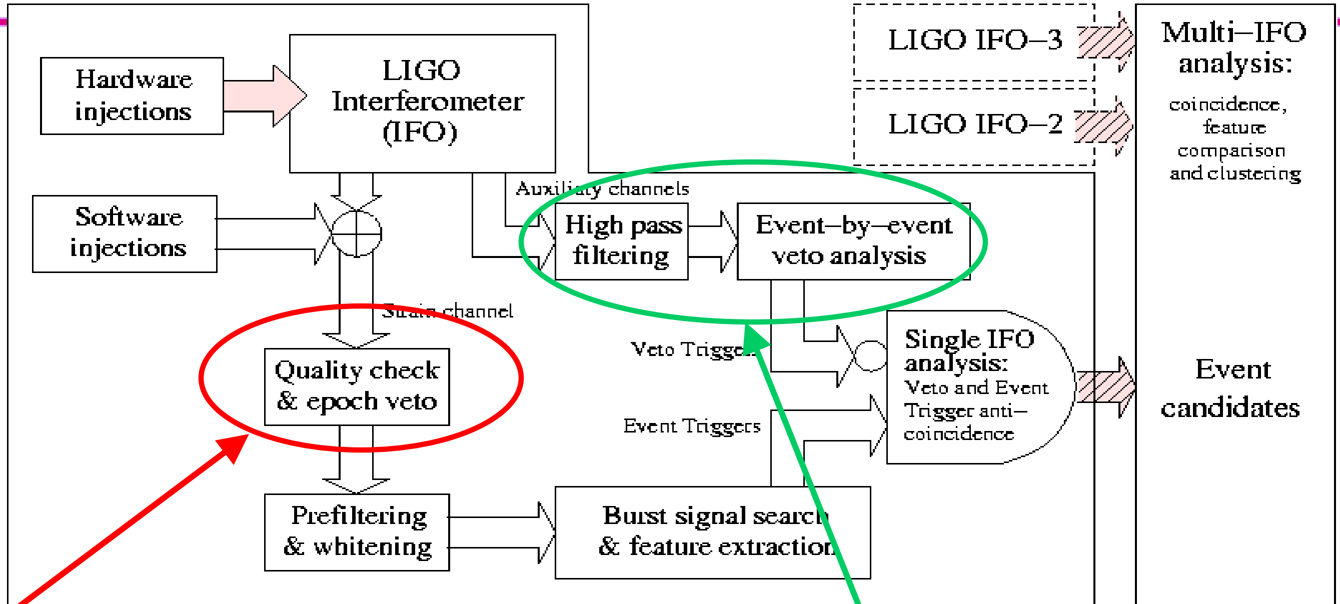
For analysis we need…

- Identification of detector events
  - » Event Trigger Generation and coincidence
- Estimation of expected contribution from background
  - » Time shift analysis
- Estimation of efficiency
  - » Simulations
- Determination of live-time T
  - » Triple-time subject to vetos
- Systematic error estimation and propagation
  - » Calibrations, background estimation, efficiency

A basic pipeline for the untriggered burst search



Data checks



Data Quality:

Identify data that do not pass quality criteria

- Instrumental errors
- Band Limited RMS
- Glitch rates from channel
- Calibration quality

Veto Analysis:

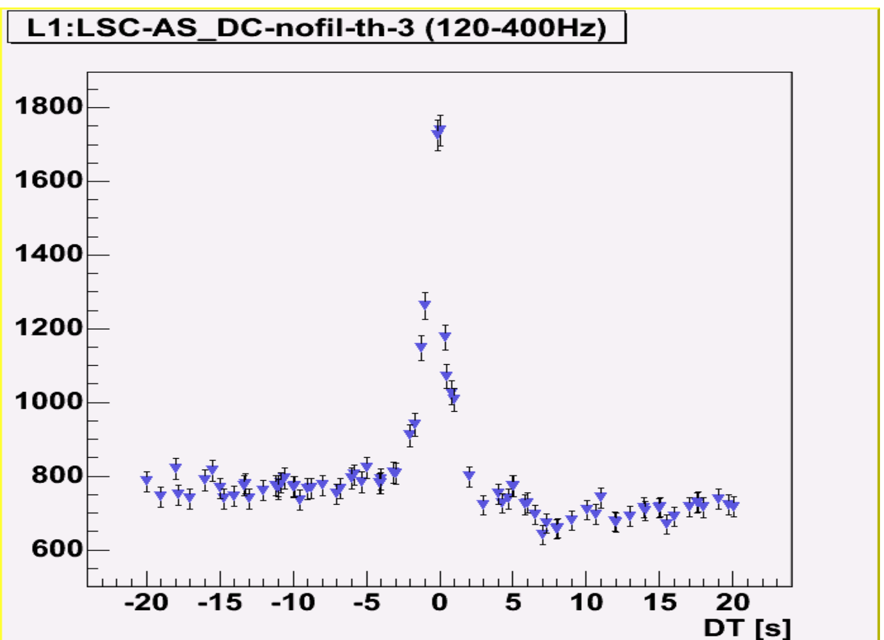
» Goal: reduce singles rates without hurting sensitivity

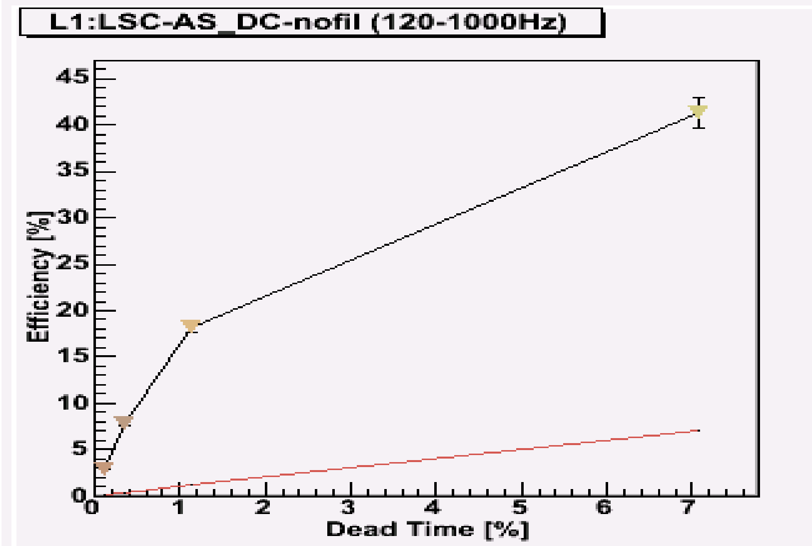
» Establish correlations

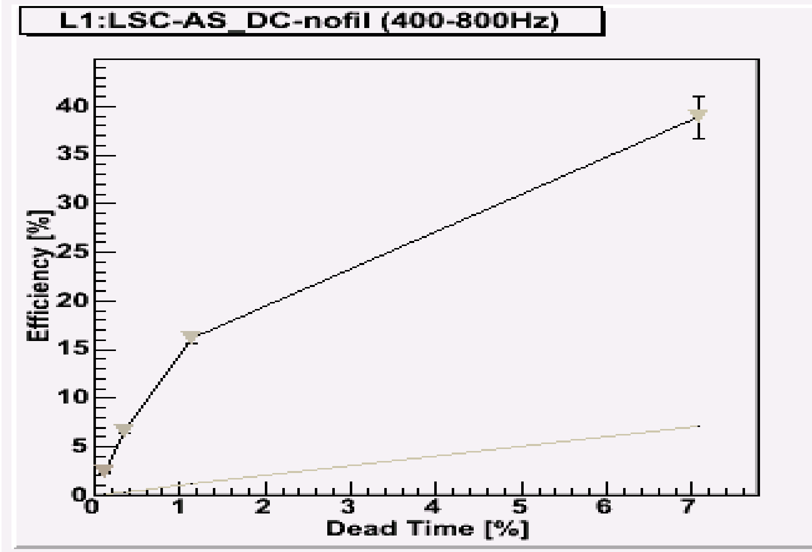
» Study eligibility of veto

Veto Analysis

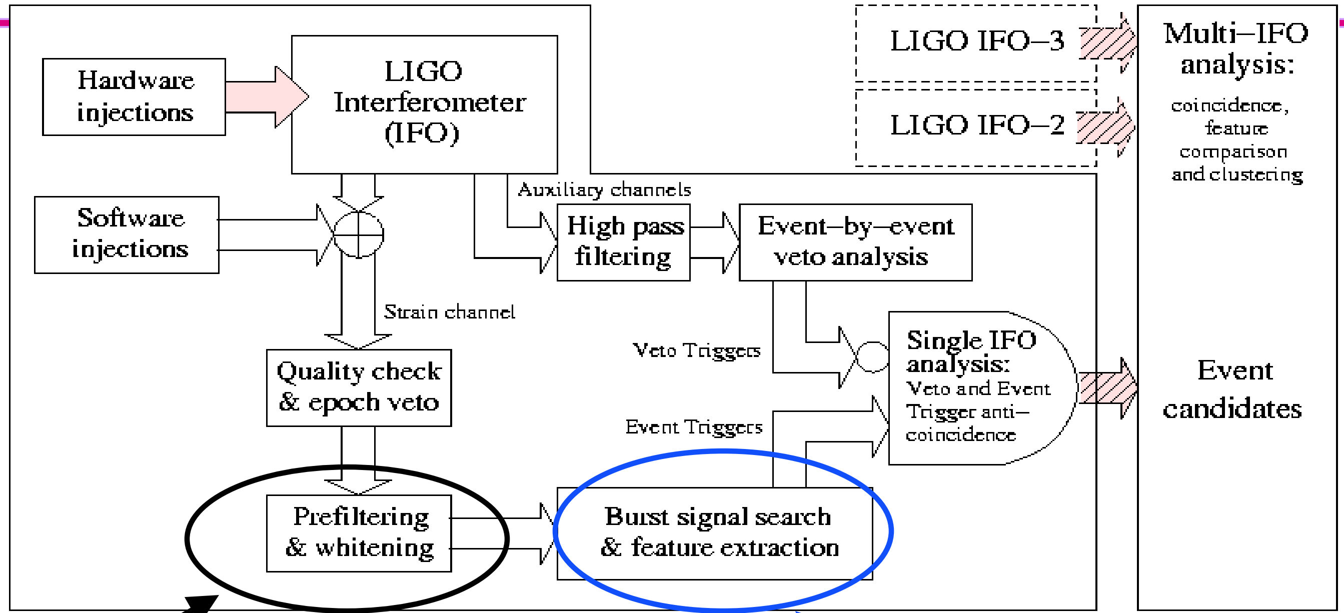
- **Strategy**: Selection of auxiliary channels with glitches that correlate better with burst triggers
  - » Choice among: Interferometer channels, Wavefront Sensors, Optical Levers, environmental sensors (microphones, accelerometers, seismometers, …)
- **Method**: Coincidence analysis and time-lag plots







Event Trigger Generation



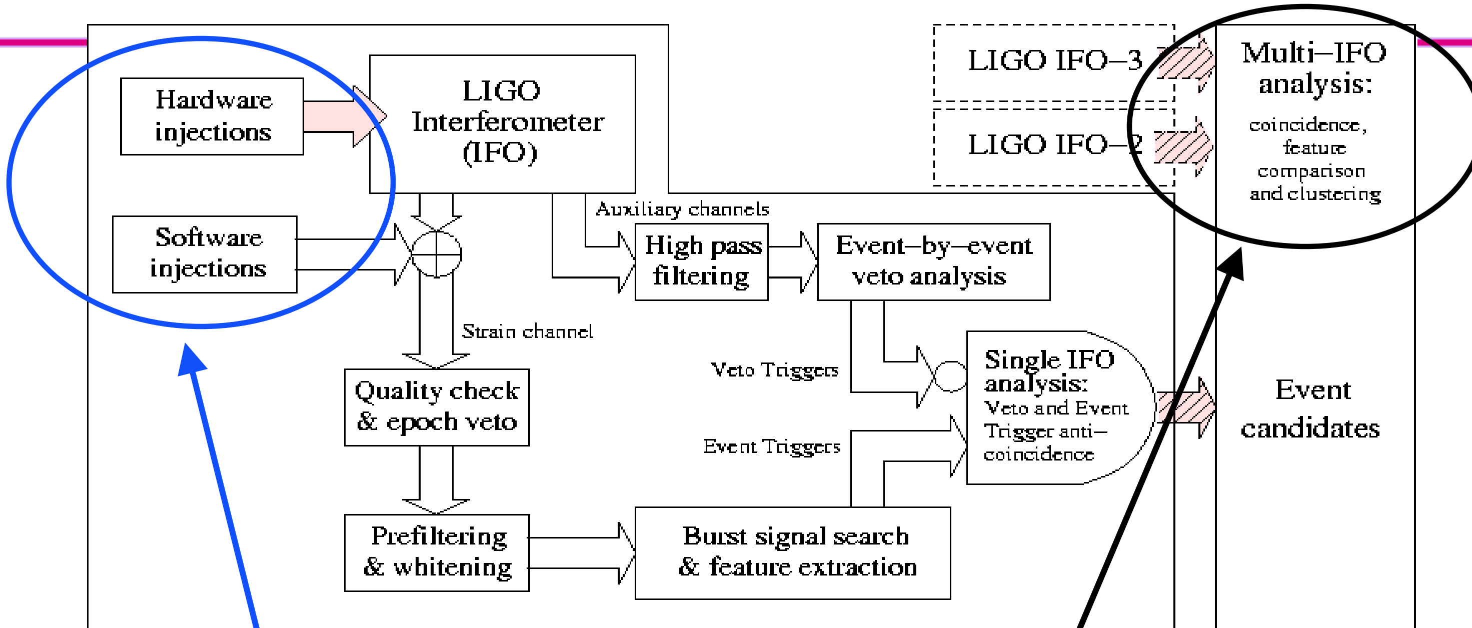
Data Conditioning:

Different ETGs need different kinds of prefiltering for optimal performance.

GW Event Trigger Generators:

Search for unusual transient features in the data.

Other pipeline features



Simulations:

» Use to optimize ETGs

» Employ standard waveforms to measure efficiencies of the search

Coincidence Analysis:

» Time and frequency coincidence

» Waveform consistency: perform a fully coherent analysis on candidate events

All tuning done in playground dataset

We used a playground dataset for all tuning of thresholds, vetoes, simulation methods.

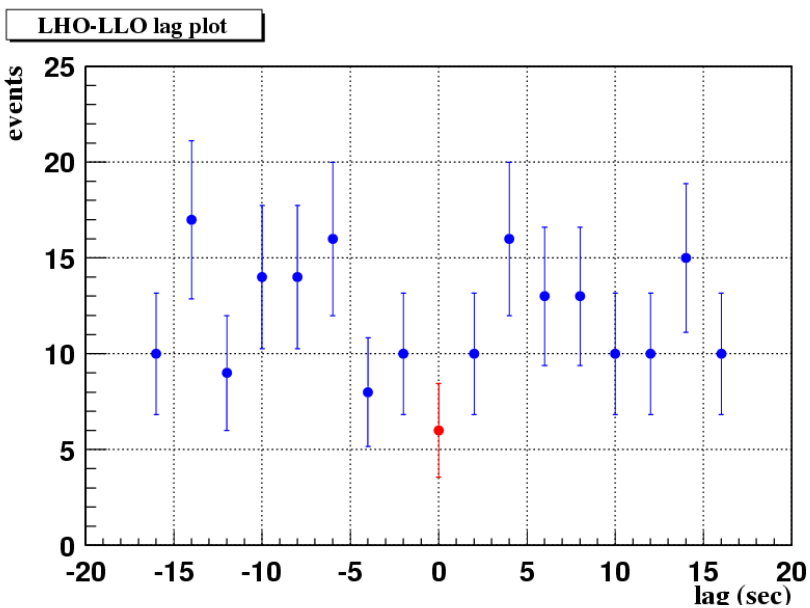
The playground data is about 10% of the run's duration.

After tuning, applied the method to the remaining data, from which analysis results are determined.

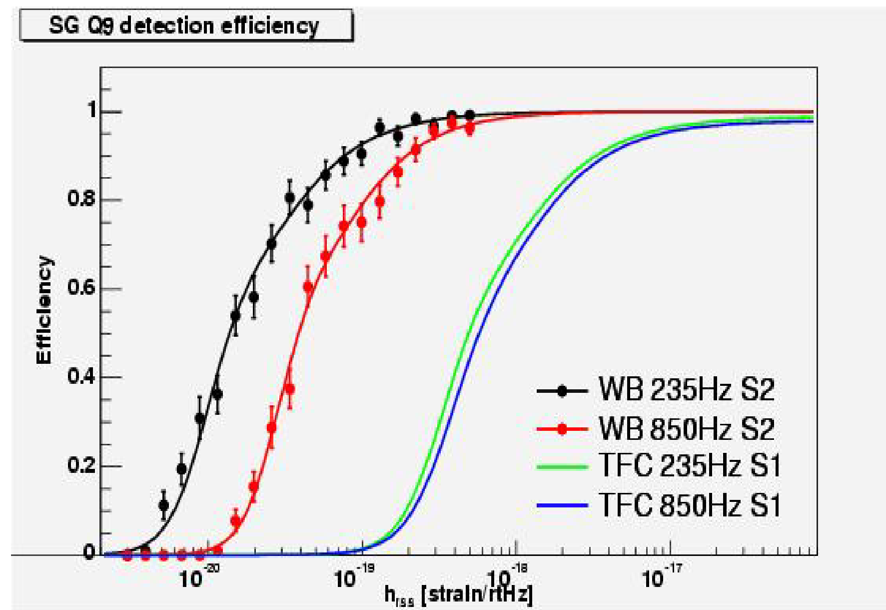
Coincidences, random coincidences, and efficiency

True coincidences at zero lag, and estimate of random coincidences from non-physical time lags.

Determination of search efficiency from artificial addition (in software) of trial signals to data.



S1

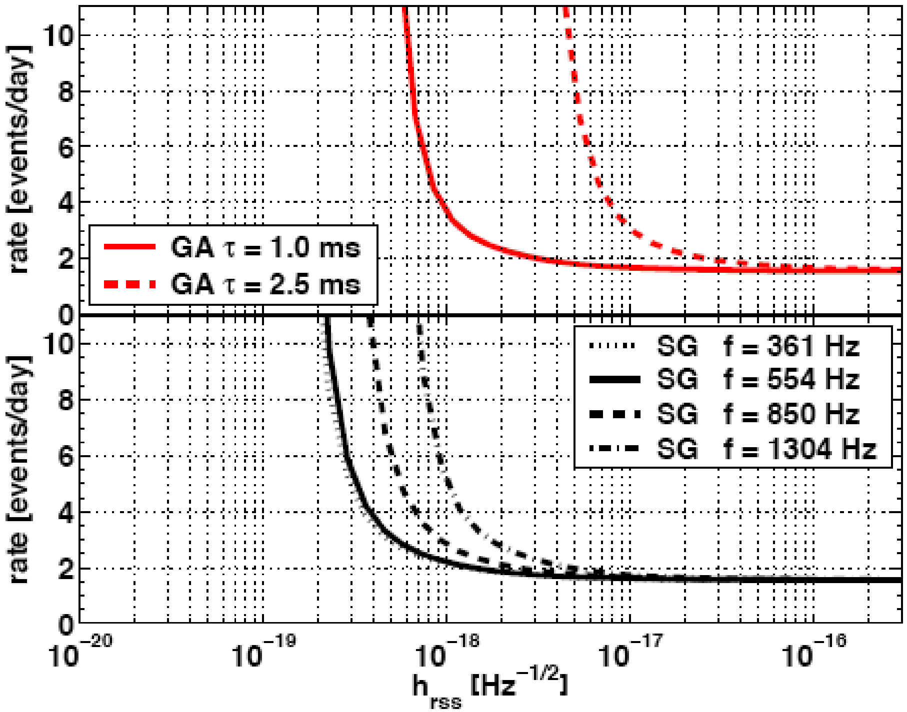


Comparison of S1 and S2

S1 Burst Search Results

- End result of analysis pipeline: number of triple coincidence events
- Use time-shift experiments to establish number of background events
- Use Feldman-Cousins to set confidence upper limits on rate of foreground events:
  - » TFCLUSTERS: <1.6 events/day

Burst model: Gaussian/Sine gaussian pulses



- Determine detection efficiency of the end-to-end analysis pipeline via signal injection of various morphologies.
- Assume a population of such sources uniformly distributed on a sphere around us: establish upper limit on rate of bursts as a function of their strength
- Obtain rate vs. strength plots

h_{rss}: natural measure of strength of unmodeled bursts

- Without a signal model, it isn't obvious what feature of signal strength is the most useful measure of *what we could have seen* in a search that yields an upper limit.
- We have tested search efficiency against a variety of waveforms.
- A rather waveform independent measure of search sensitivity is the root-sum-square amplitude, h_{rss}:

$$h_{rss} = \sqrt{\int_0^{\infty} |h(t)|^2 dt} = \sqrt{\int_{-\infty}^{\infty} |\tilde{h}(f)|^2 df}$$

A look ahead at the S2 burst search

Results from S1 published. S2 results are almost done, but most are not quite ready for sharing.

In what follows, I'll share some of the methods used in the S2 search:

- What they do
- How well they work

Two burst searches under way:

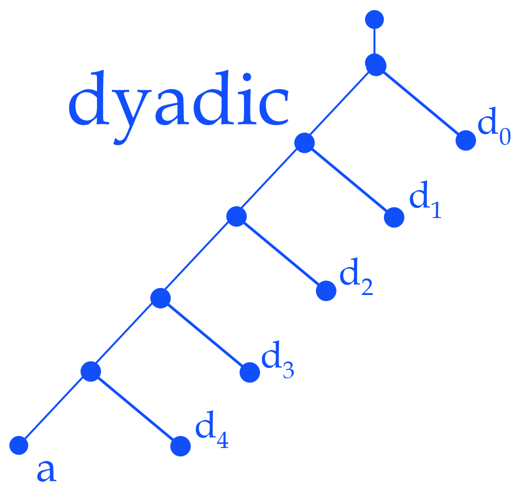
- » Untriggered search, like S1, but with two innovations:
  - – New ETG, WaveBurst.
  - – Waveform coherence test, the r-statistic
- » Coherent search near the time of GRB030329.

WaveBurst: search for bursts using wavelets

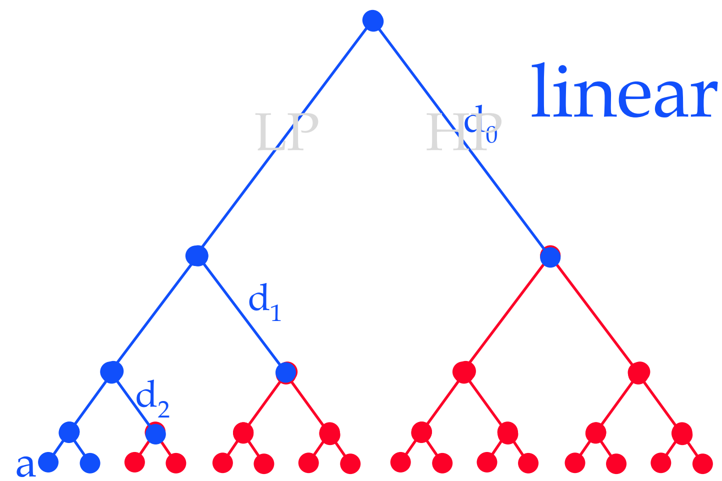
- use wavelets
  - » flexible tiling of the TF-plane by using wavelet packets
  - » variety of basis waveforms for bursts approximation

    Haar, Daubechies, Symlet, Biorthogonal, Meyers.
- use rank statistics
  - » calculated for each wavelet scale
  - » robust
- use local T-F coincidence rules
  - » works for 2 and more interferometers
  - » coincidence at pixel level applied before triggers are produced

Wavelet Transform

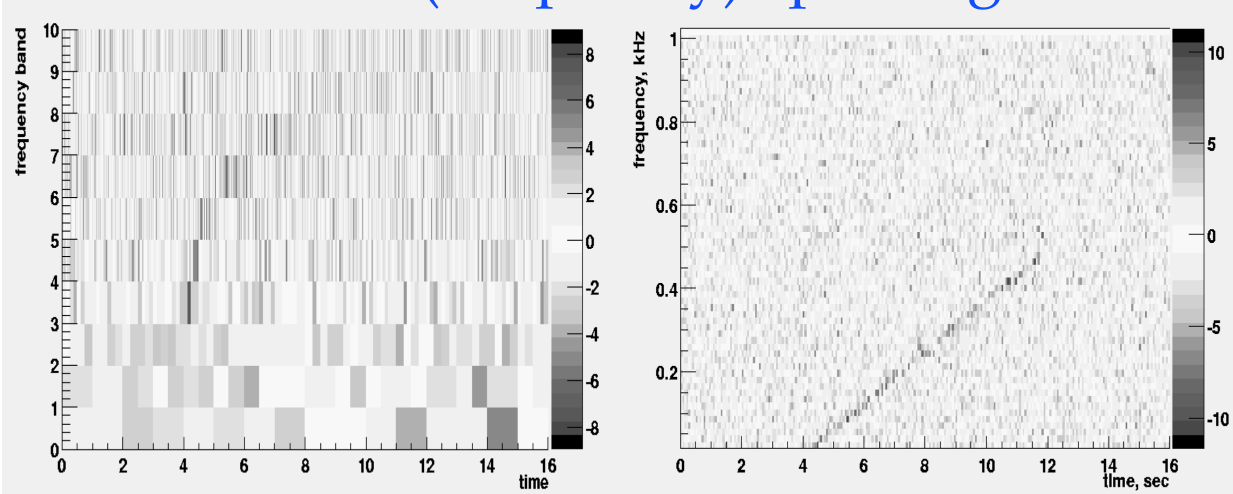


a. wavelet transform tree

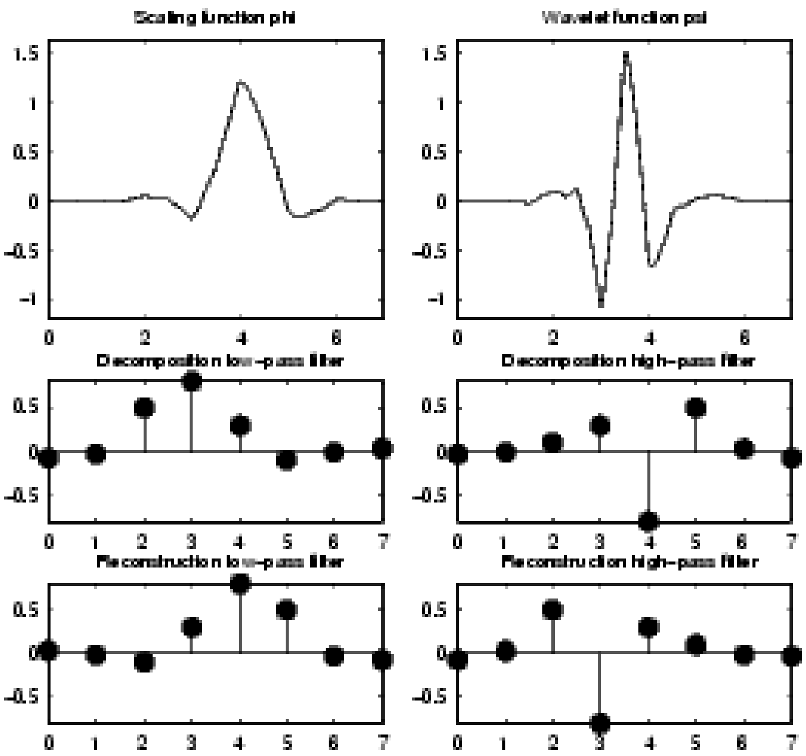


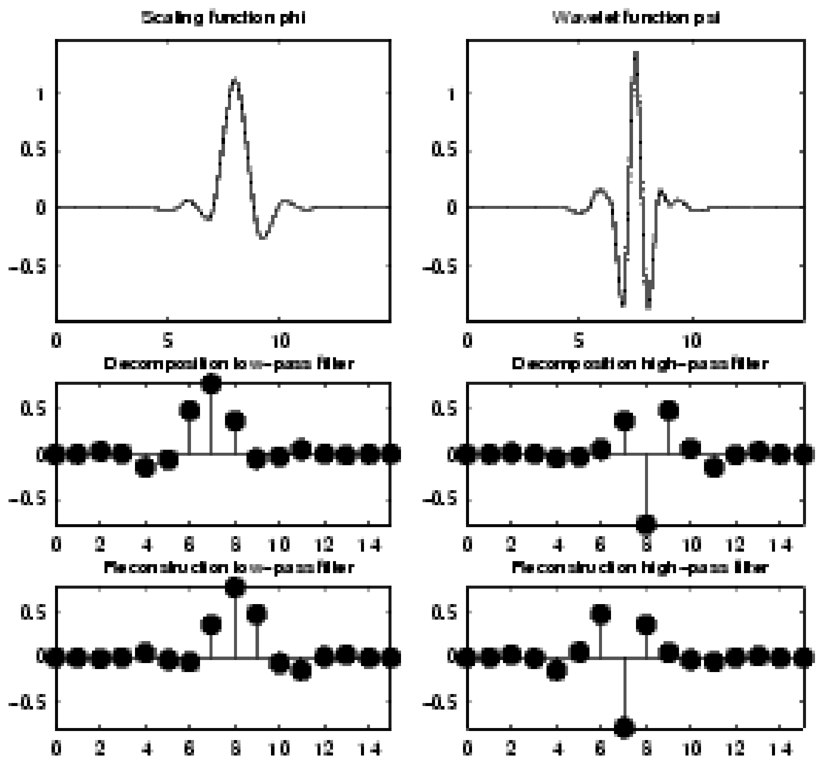
b. wavelet transform binary tree

time-scale(frequency) spectrograms

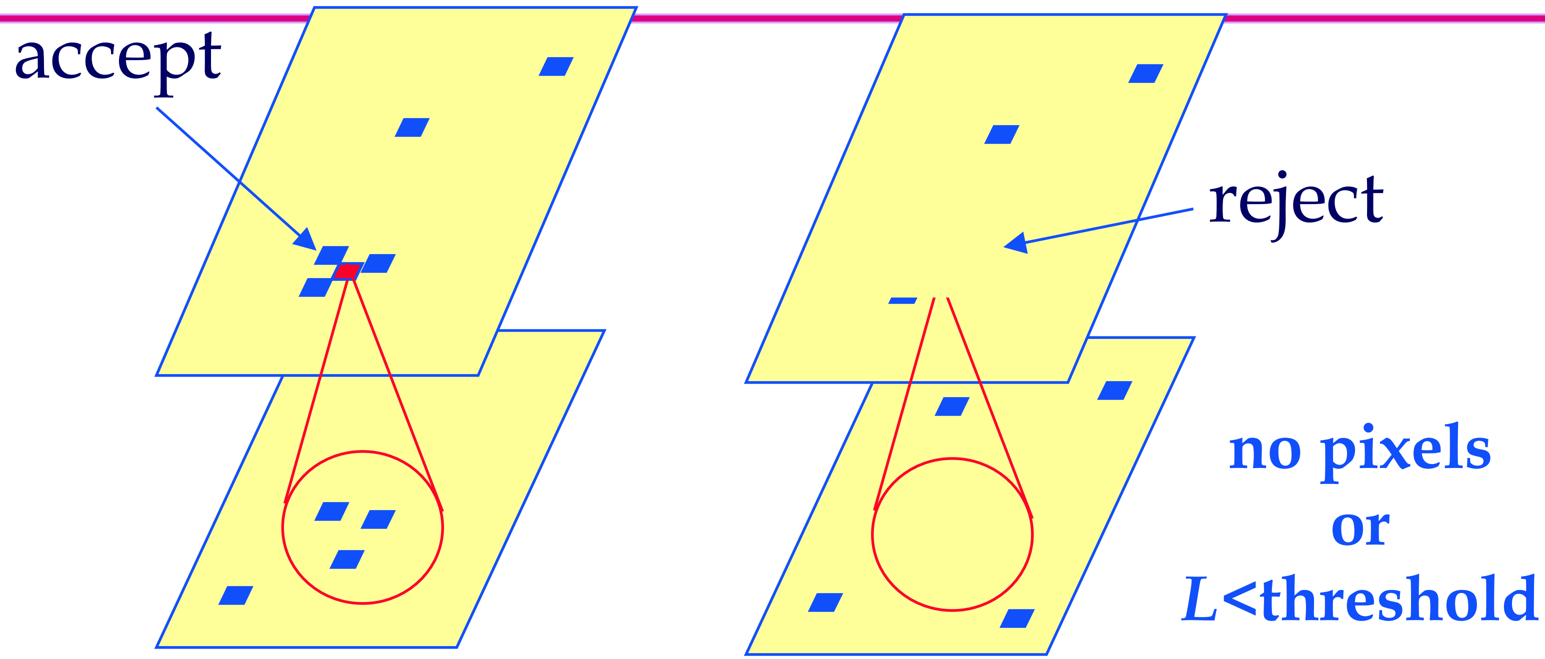


Symlet wavelet





Coincidence



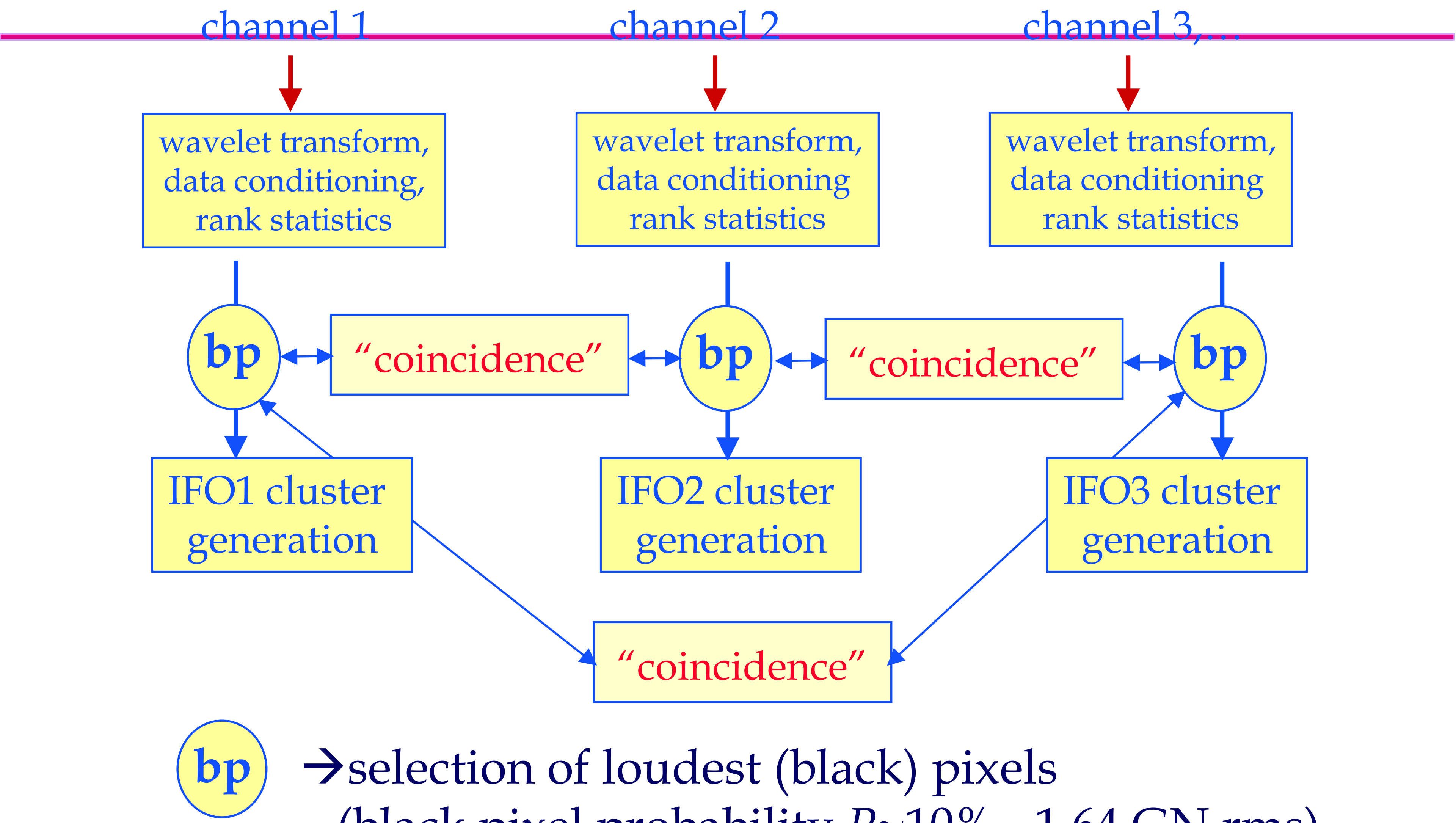
- Given local occupancy P(t,f) in each channel, after coincidence the black pixel occupancy is

$$P_C(t,f) \propto P^2(t,f)$$

for example if P=10%, average occupancy after coincidence is 1%

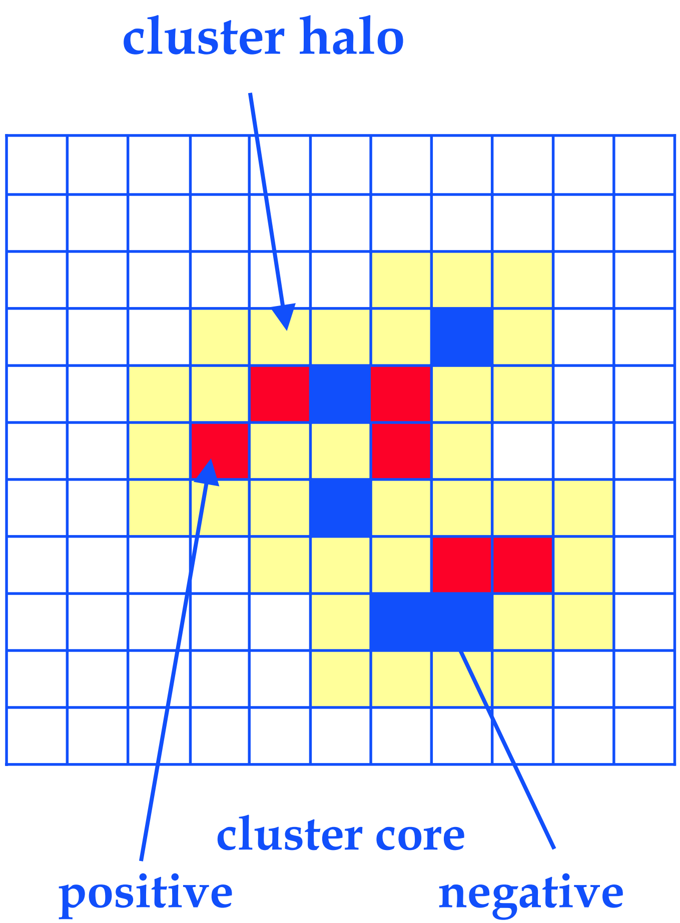
- can use various coincidence policies → allows customization of the pipeline for specific burst searches.

Analysis pipeline



→selection of loudest (black) pixels
(black pixel probability *P*~10% - 1.64 GN rms)

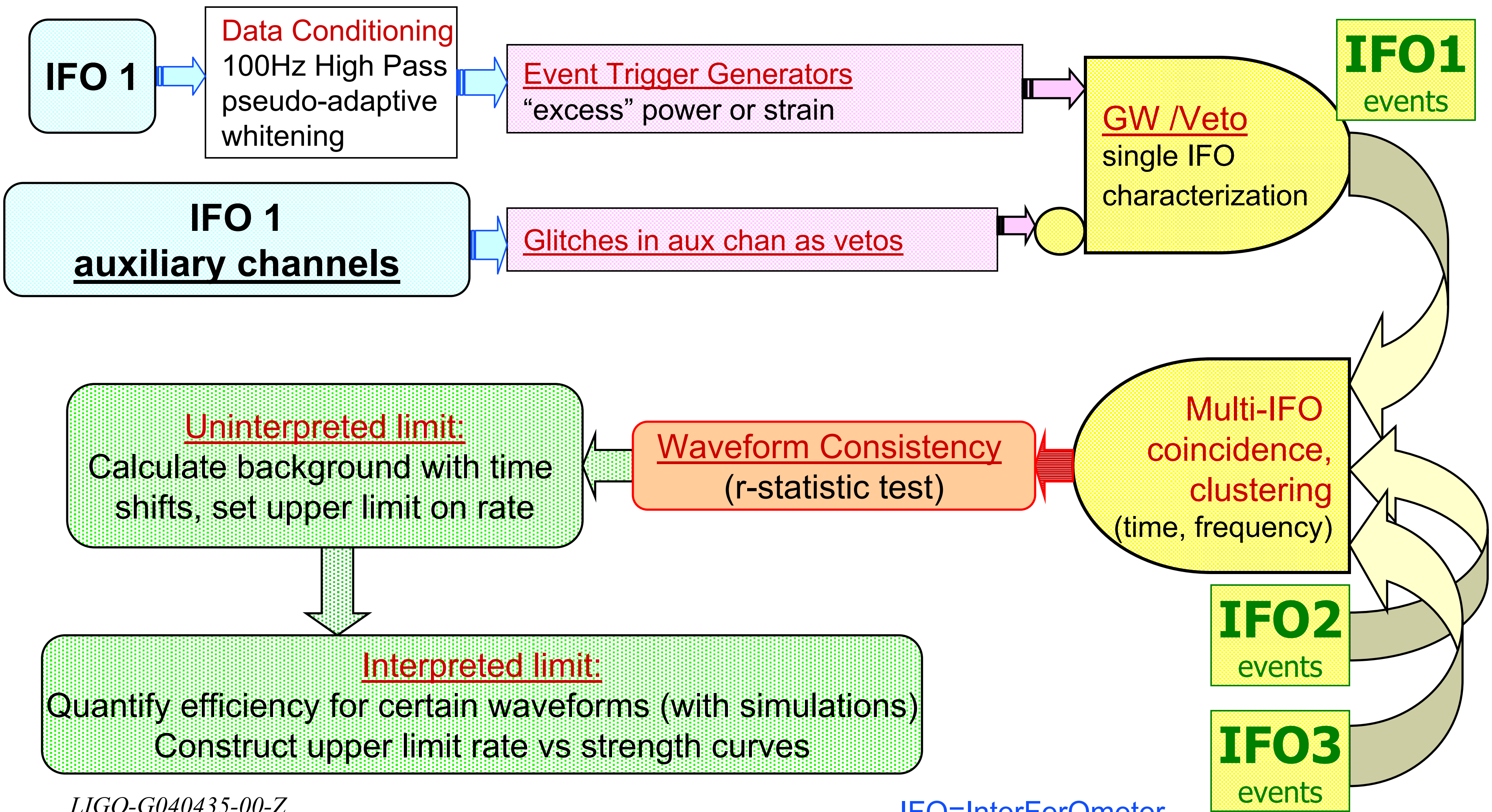
Wavelet Cluster Analysis



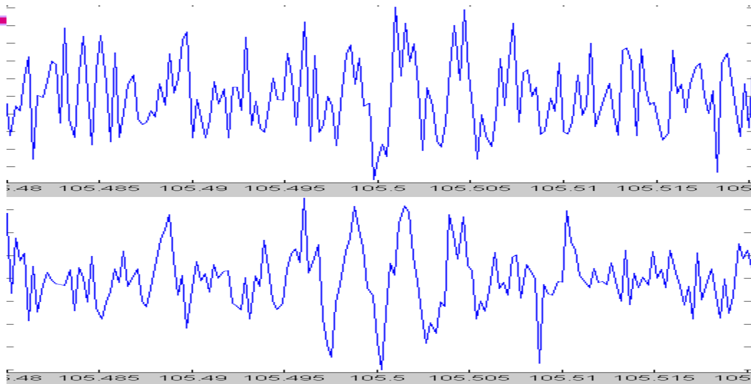
Cluster Parameters

size – number of pixels in the core
volume – total number of pixels
density – size/volume
amplitude – maximum amplitude
power - wavelet amplitude/noise rms
energy - power x size
asymmetry – (#positive - #negative)/size
confidence – cluster confidence
neighbors – total number of neighbors
frequency - core minimal frequency [Hz]
band - frequency band of the core [Hz]
time - GPS time of the core beginning
duration - core duration in time [sec]

r-statistic: an End-of-Pipeline Waveform Consistency Test



Correlation Analysis



$$s_1(t) = h(t - t_1) + n_1(t)$$
$$s_2(t) = h(t - t_2) + n_2(t)$$

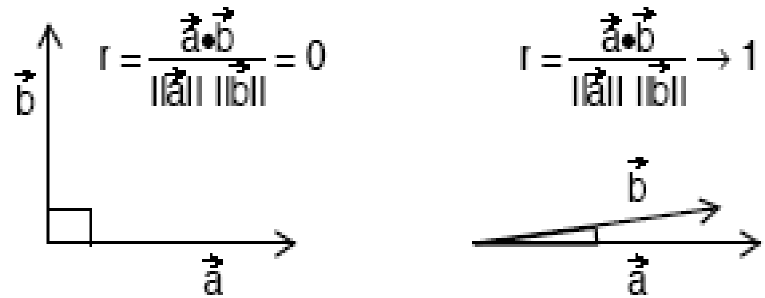
$$C(t, t_w, t_{off}) = \int_{t-t_w/2}^{t+t_w/2} s_1(t') s_2(t' + t_{off}) dt'$$
$$\approx \int_{t_w} h^2(t) dt + \int_{t_w} n_1(t) n_2(t) dt$$

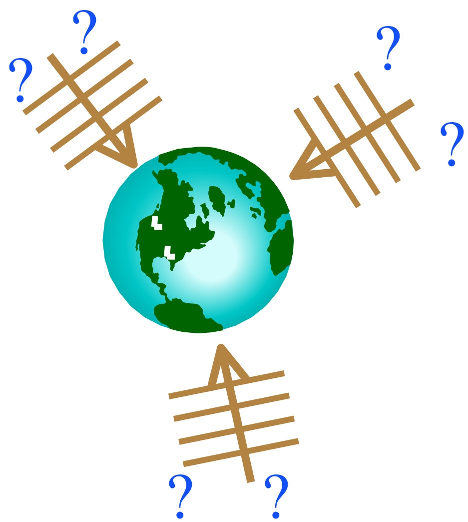
h_{rss}^2

$< > = 0$

The r-statistic test

- Correlated time-series “point” in same direction
  - » Pearson r statistic: cosine angle between time-series vectors
  - » Note: Insensitive to relative amplitude scale
- Don’t know when signals arrive at geographically distinct detectors
  - » Evaluate r over different physical time-lags
- Don’t know signal duration
  - » Evaluate r over range of potential signal durations
- r is function of two detectors, not three (or more)
  - » Evaluate geometric mean of significance for all detector pairs





Reference: Cadonati gr-qc/0407031

r-statistic Test for Waveform Consistency

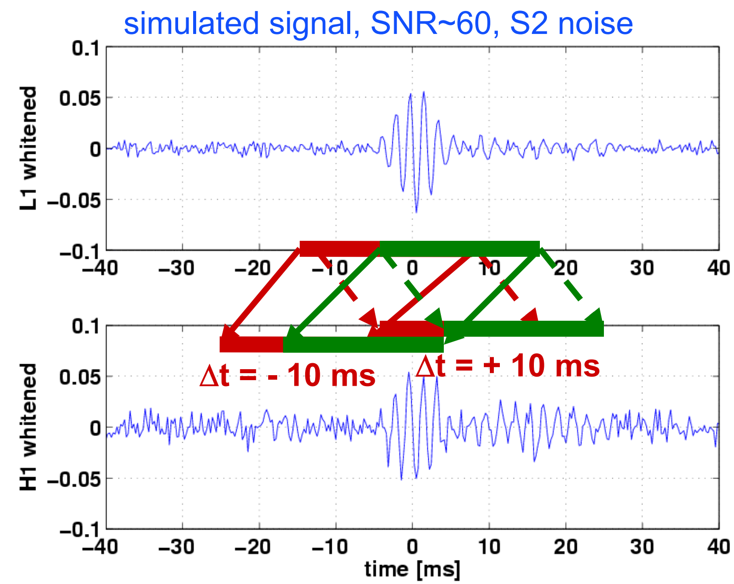
For each triple coincidence candidate event produced by the burst pipeline (start time, duration ΔT) process pairs of interferometers:

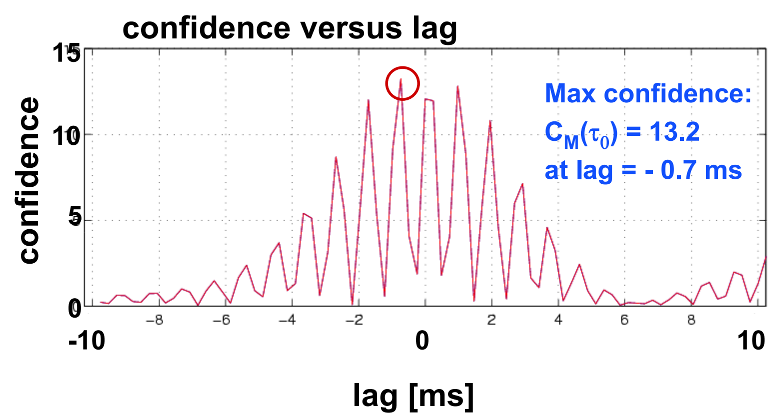
After Data Conditioning:

Partition the trigger in intervals (50% overlap) of duration $\tau =$ integration window (20, 50, 100 ms). For each interval, time shift up to 10 ms and build an r-statistic series distribution.

$$r_k = \frac{\sum_i (x_i - \bar{x})(y_{i+k} - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_{i+k} - \bar{y})^2}}$$

If the distribution of the r-statistic is inconsistent with the no-correlation hypothesis: find the time shift yielding maximum correlation confidence $C_M(j)$ (j=index for the sub-interval)



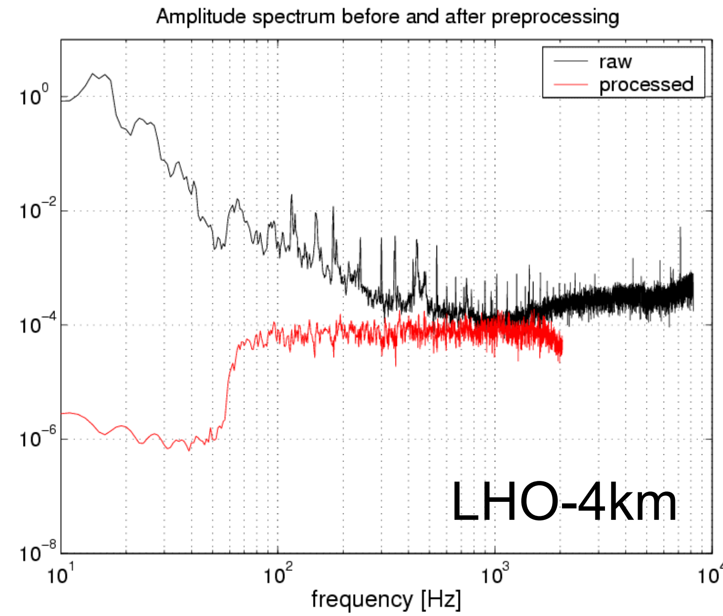


Some Technical Details...

Aggressive pre-processing: band-limit to 100-1600 Hz

Remove all predictable content (effective whitening/line removal): train a linear predictor error filter over 10 s of data (1 s before event start),

$\Rightarrow$ emphasis on transients, avoid non-stationary, correlated lines.



Waveforms are declared “consistent” (event passes the test) if the correlation confidence is above threshold in all three pairs of interferometers.

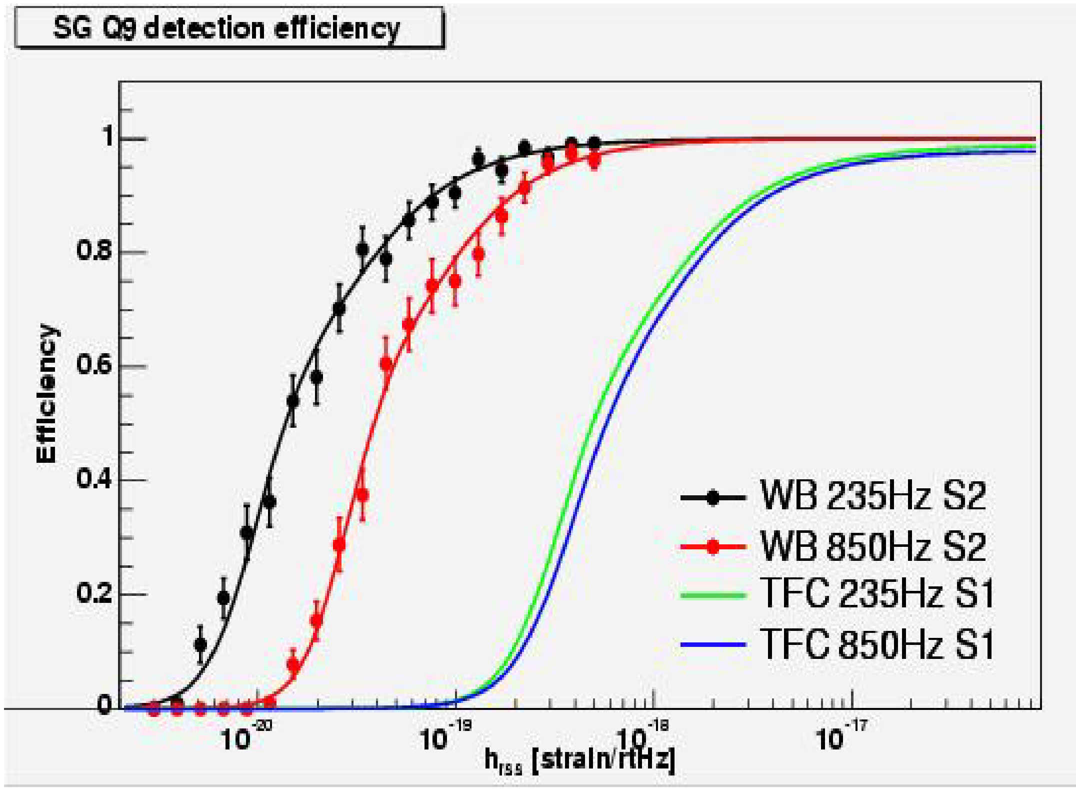
Correlation confidence: $\Gamma = \max(C_M^{L1H1} + C_M^{L1H2} + C_M^{H1H2})/3$

β is the threshold: $\Gamma > \beta$

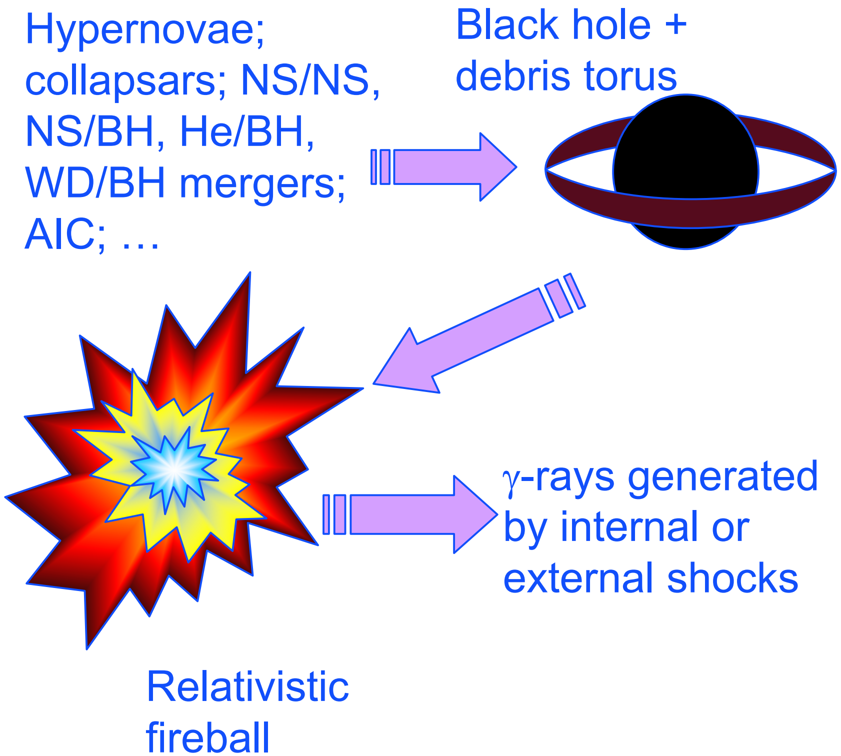
Efficiency of the S2 burst search

- "All time, all sky" search for short (<1 s timescale) bursts
  - » No external triggers! Generate own from observations
- Two step analysis
  - » Trigger generation with WaveBurst
  - » r-statistic test on event candidates
- Substantial improvement over S1 sensitivity
  - » About x10 due to reduction in interferometer noise
  - » Remainder of improvement from more effective analysis

Preliminary



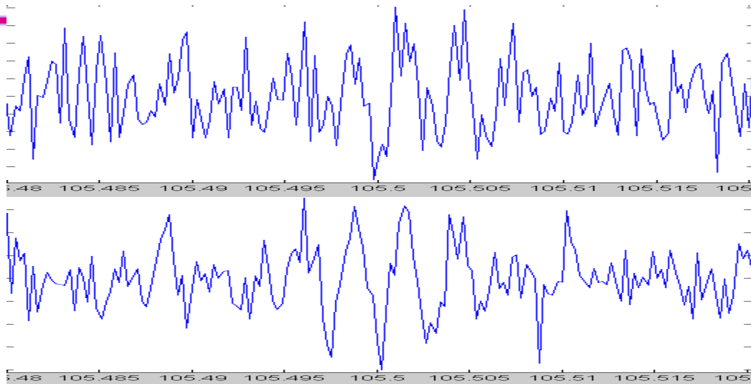
Triggered searches: γ-ray bursts & gravitational waves



For more on GRBs: P. Meszaros, Ann. Rev. Astron. Astrophys. *40* (2002)

- GRB 030329
  - » Detected by HETE-2, Konus-Wind, Helicon/KoronasF
  - » Especially close: $z = 0.1685$; d_L=800Mpc (WMap params)
  - » Strong evidence for *supernova origin of long GRBs*.
  - » H1, H2 operating before, during, after burst
- Radiation from a broadband burst at this distance?
- Exercise analysis

Correlation Analysis



$$s_1(t) = h(t - t_1) + n_1(t)$$

$$s_2(t) = h(t - t_2) + n_2(t)$$

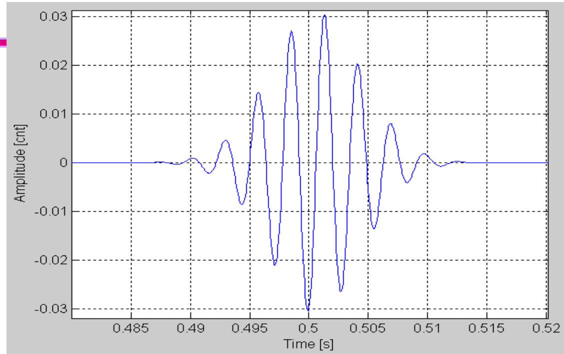
$$C(t, t_w, t_{off}) = \int_{t - t_w/2}^{t + t_w/2} s_1(t') s_2(t' + t_{off}) dt'$$

$$\approx \int_{t_w} h^2(t) dt + \int_{t_w} n_1(t) n_2(t) dt$$

h_{rss}^2

< > = 0

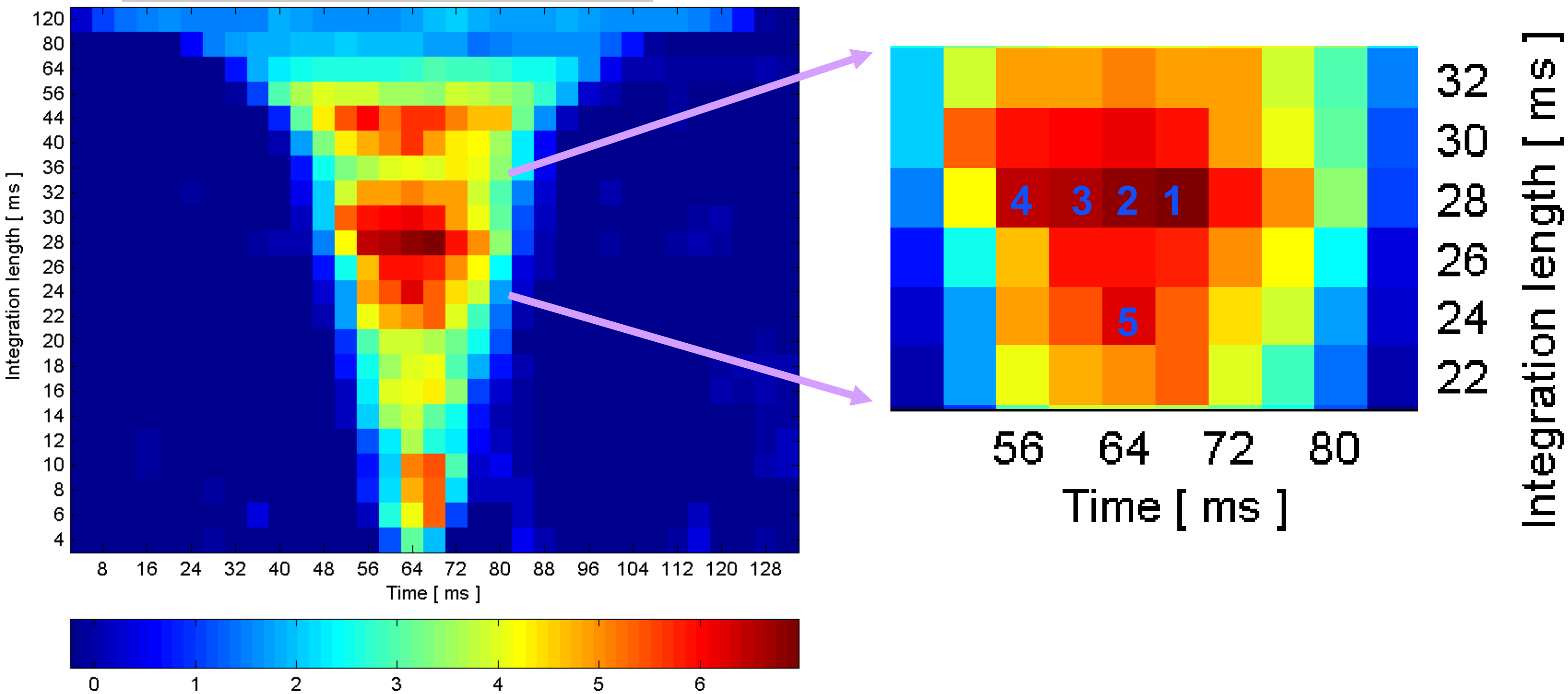
Event Identification – Simulated Signal



Color coding: “Number of variances above mean”

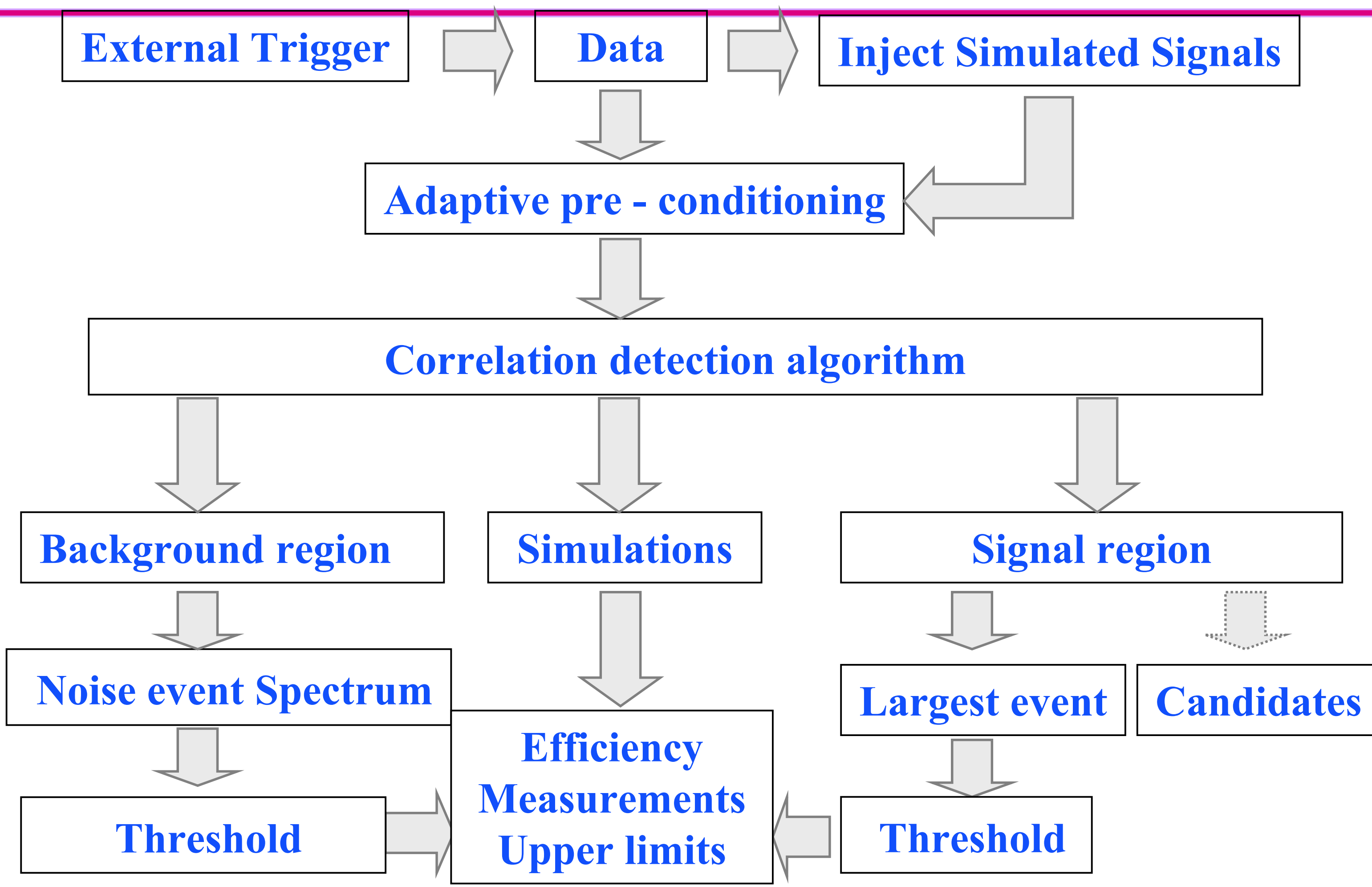
<u>Event strength</u> [ES] calculation:

Average value of the “optimal” pixels

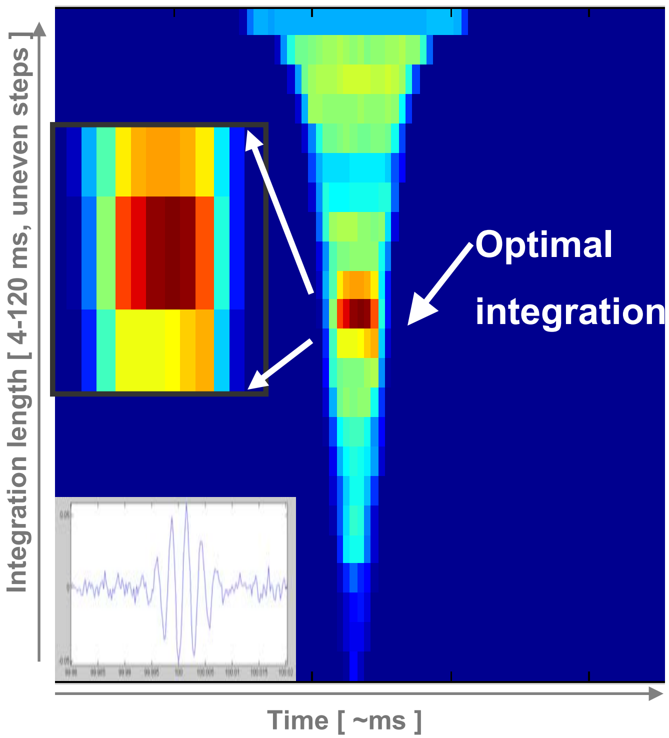


LIGO-G040435-00-Z

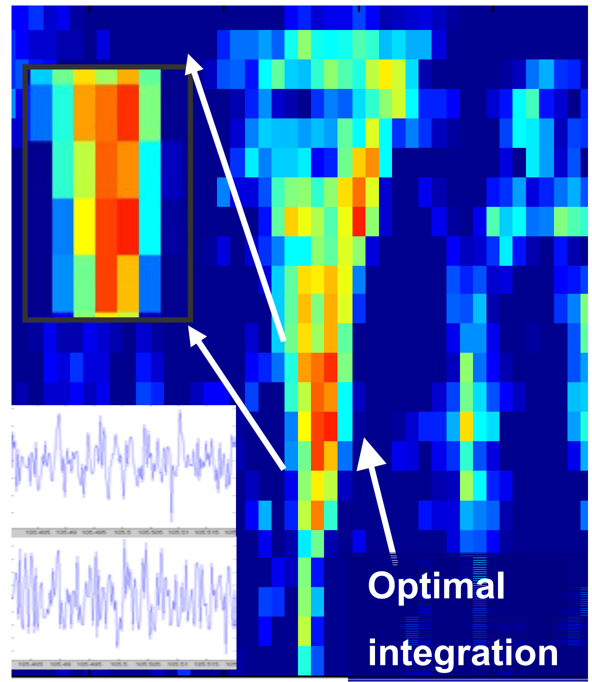
Analysis Flow Chart



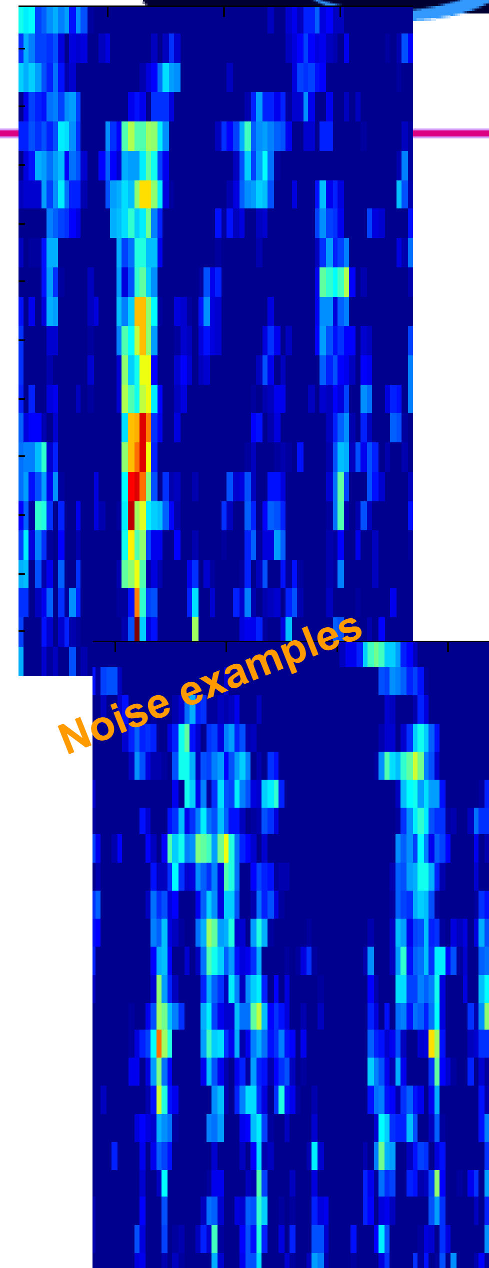
Sensitivity Determination



“Huge” Sine-Gaussian
F = 361Hz, Q = 8.9
h_{RSS} ~ $6x10^{-20}$ [1/√Hz]



“Small” Sine-Gaussian
F = 361Hz, Q = 8.9
h_{RSS} ~ $3x10^{-21}$ [1/√ Hz]
(~ detection threshold)



Analysis Methodology: Externally Triggered Search

- Cross-correlate h_{H1}, h_{H2} near time GRB trigger time …
  - Near: τ in [-120s,+60s]
- … and look for cross correlation exceeding threshold
  - Signal correlated, noise uncorrelated
- Random noise correlations will lead to threshold crossing in fraction α of observations
  - Higher threshold, less likely false positive
  - Estimated by analysis on noise away from GRB trigger
- Set threshold by tolerable false rate α.
  - This analysis: α = 10%

$$c(\tau, \Delta t, T) = \int_{-T}^{+T} h_{H1}(t_0 + \tau + \Delta t) h_{H2}(t_0 + \tau)\, dt$$

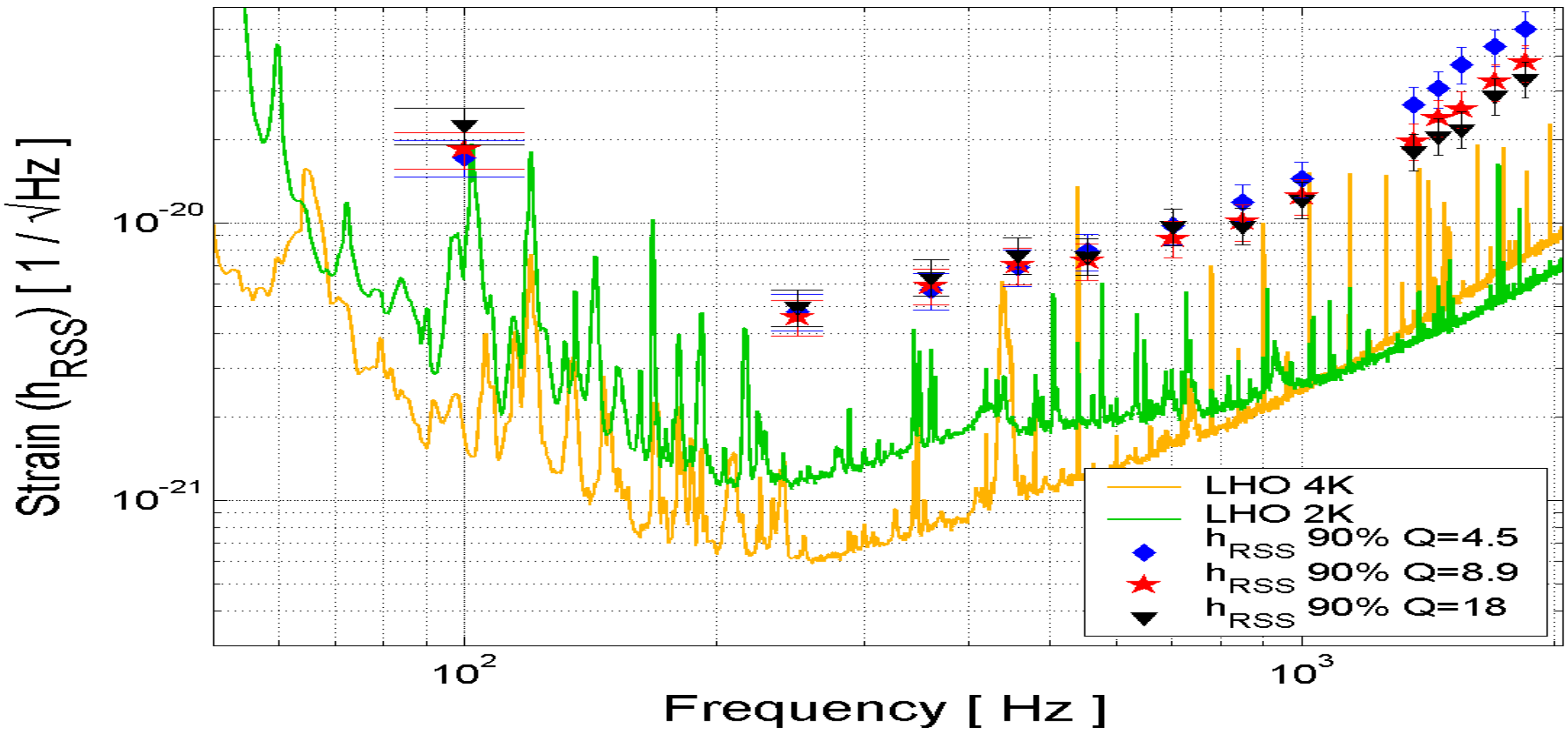
Δt in [-5,+5] ms for timing uncertainties
T in [4,128] ms for short GW bursts
t_0 is trigger time

- *Methodology applicable to GW burst search associated with* any *externally generated trigger* (e.g., SN, neutrino burst, etc.)

For more see: Mohanty et al. gr-qc/0407057
For other x-corr style analyses cf. L. Finn, S. Mohanty, J. Romano, Phys. Rev. D **60** 121101 (1999), P. Astone *et al.*, Phys. Rev. D **66** 102002 (2002)

GRB030329 preliminary result

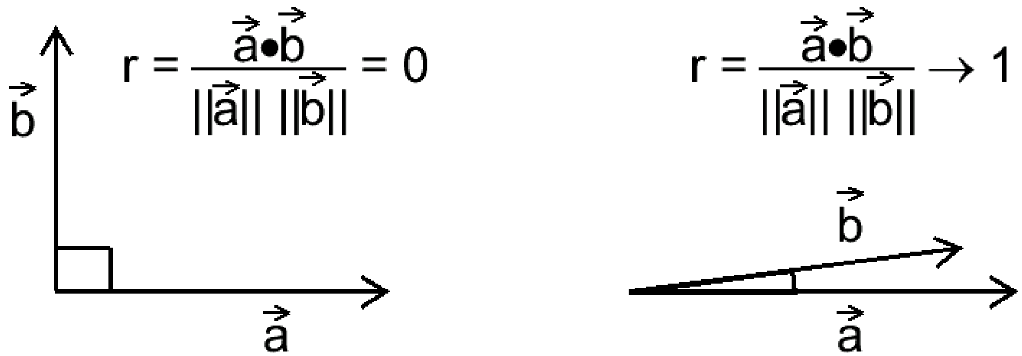
- No event exceeded analysis threshold
- Using simulations an upper limit on the associated gravitational wave strength at the detector at the level of h_{RSS}~$6x10^{-20}$ $Hz^{-1/2}$ was set
- Radiation from a broadband burst at this distance? $E_{GW} > 10^5 M_{\odot}$



A fully coherence-based burst search in S3?

- CorrPower: search algorithm for coherent power estimation
- Multiple operational modes:
  - » enhanced r-statistic on pre-selected candidate events
  - » external trigger search similar to GRB030329
  - » continuous search on the whole dataset (compromising on integration times, online?)
- Effective unification of coherent triggered and untriggered
- Presently undergoing tests
  - » Efficiency at low false-alarm rate
  - » Computing speed

What is CorrPower?



(a) (b)

$$(a+b)^2 = a^2 + b^2 + a\bullet b$$

$$P_{corr} = a\bullet b$$

Checking our instruments and search pipelines

- Calibration lines are always present.
- We make online tests of data quality, with human review afterwards.
- At several times, we inject fake signals by shaking mirrors, to check that they are properly recovered by search pipelines.
- Those *hardware injections* are supplemented by many additional signals added in software to check efficiency of search pipelines.
- We carry out intensive reviews of
  - » search software, and of
  - » complete analysis results.

How would we check a burst detection candidate?

- Understand statistical significance of result
  - » Convincing statistics essential
  - » Other event trigger generators?
  - » Robustness to ETG tuning
- Data validation
  - » Elog check
  - » Suspicious GPS times, times within second, time within ETG stride
  - » Possible unintended injection
  - » Check reduced data against original data frames
- Software validation
- Investigate possible external causes for events
  - » Environmental channels
  - » Auxiliary interferometer channels

How would we check a burst detection candidate? (II)

- Consistency of individual events with GW expectations
  - » Signals have the same sign?
  - » H1-H2 amplitudes consistent?
  - » Waveform reconstruction?
  - » t-f spectrograms?
  - » Consistency of time delays?
  - » If multiple events, consistent with plausible sky distribution (e.g., not too many events coming in along L1 minimum…)?
- Seek any corroborative indicators
  - » Other LIGO searches
  - » Other GW detectors
  - » Other astronomical observations (SNe, GRB, neutrino, …)

Near term prospects

- The LIGO interferometers approached within a factor of 2 – 3 of design sensitivity during S3 run.
- Commissioning now to try to do better, and to improve duty cycle.
  - Hanford 4km now closer than x2 to design sensitivity
  - Livingston site now has extra seismic isolation, should achieve good duty cycle during Fall 2004.
- Expect a 6-month long run during 2005, at or near design sensitivity, with good duty cycle.