Gravitational-Wave Data Analysis

Peter Shawhan

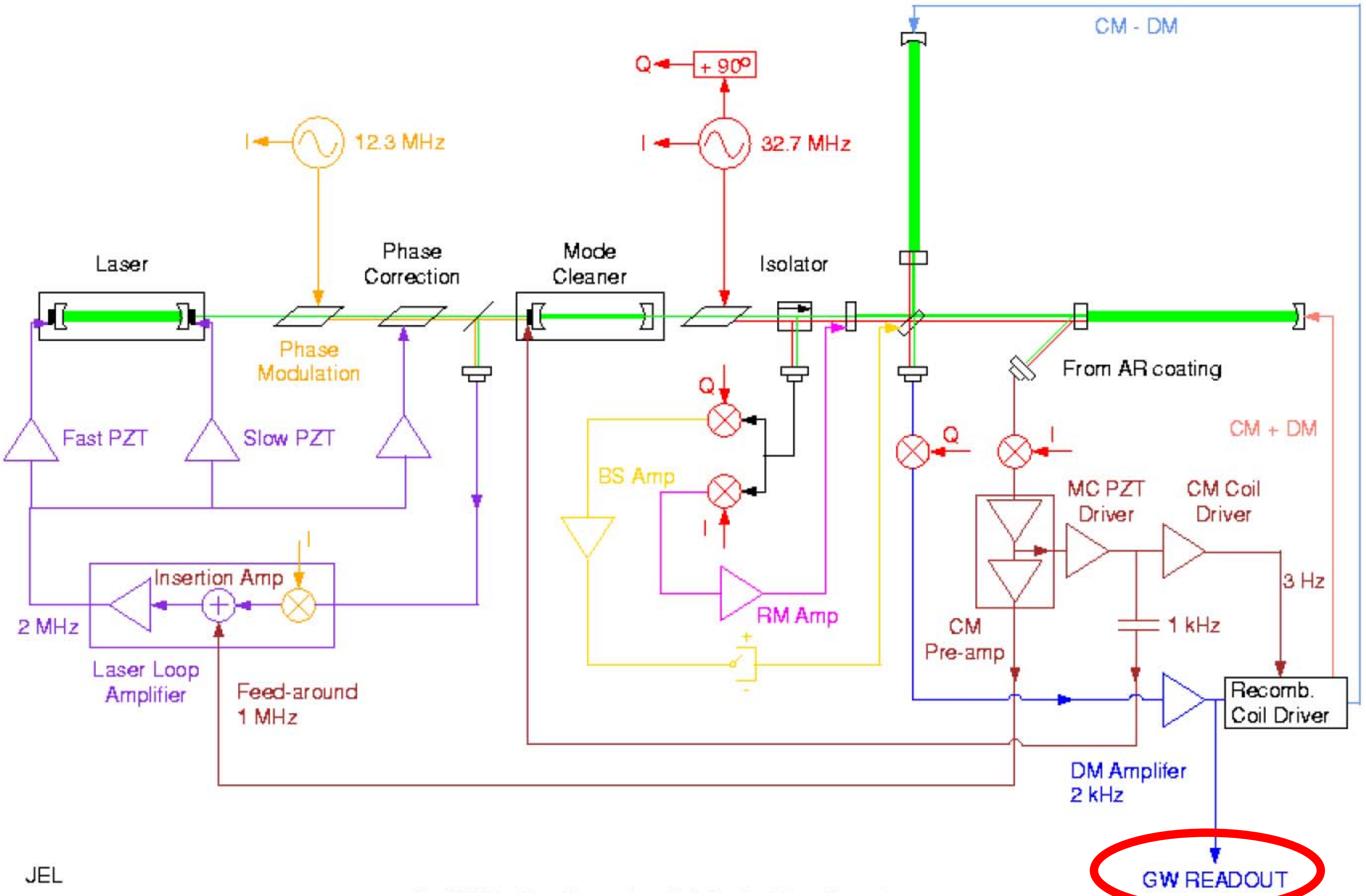
University of Maryland

Physics 798G
April 12, 2007

Outline

- **Gravitational-wave data**
- **General data analysis principles**
- **Specific data analysis methods**
  - Classification of signals
  - Methods for each class of signals
- **Idiosyncracies of real detectors**
- **The gravitational-wave community**

Length Sensing and Control



JEL

Gravitational-Wave Data

Instantaneous estimate of strain for each moment in time

i.e. demodulated channel sensitive to arm length difference

(Or, for resonant detector: displacement sensed by transducer)

Digitized time series recorded in computer files

LIGO / GEO sampling rate: 16384 Hz

VIRGO sampling rate: 20000 Hz

Synchronized with GPS time

Common “frame” file format (*.gwf)

***Many* auxiliary channels recorded too**

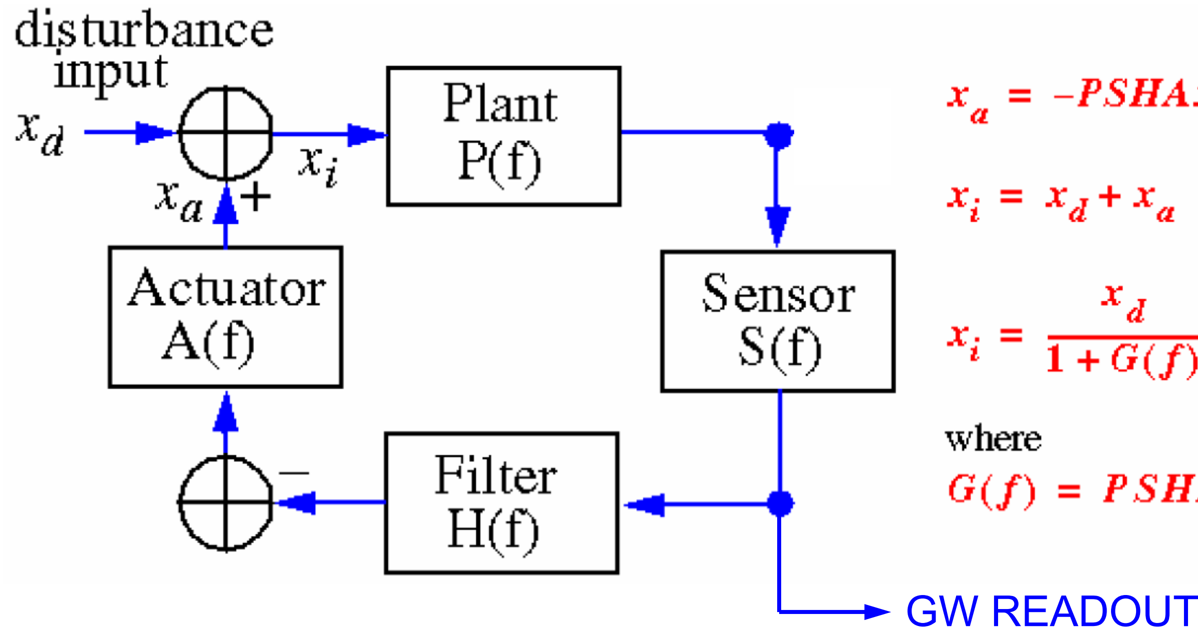
Interferometric sensing and control

Environmental sensors (accelerometers, microphones, magnetometers,…)

Interferometer configuration and facilities housekeeping data

Total data volume: a few megabytes per second per interferometer

Calibration



$$x_a = -PSHAx_i$$

$$x_i = x_d + x_a$$

$$x_i = \frac{x_d}{1 + G(f)}$$

where

$$G(f) = PSHA$$

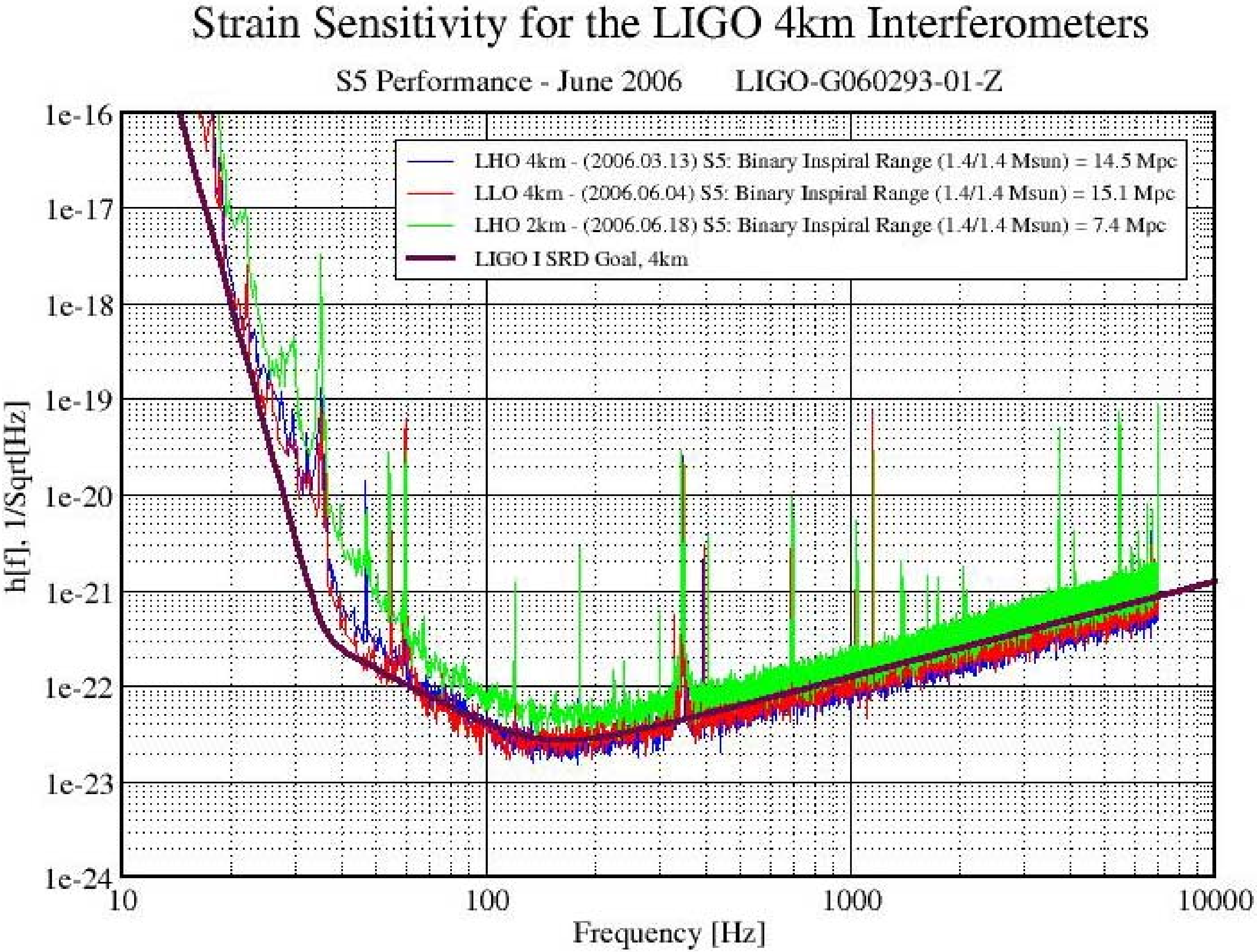
Monitor P(f) continuously with "calibration lines"

Sinusoidal arm length variations with known absolute amplitude

Apply frequency-dependent correction factor to get GW strain

$$h = (\text{GW READOUT}) \times \frac{1+G(f)}{P(f)\,S(f)}$$

Gravitational-Wave Strain Data



Outline

- **Gravitational-wave data**
- **General data analysis principles**
- **Specific data analysis methods**
  - Classification of signals
  - Methods for each class of signals
- **Idiosyncracies of real detectors**
- **The gravitational-wave community**

Basic Principles

Gravitational wave sources are rare and/or intrinsically weak

- Need highly sensitive detectors
- A detectable signal will most likely be near threshold of detectability

Claiming the first detection will be a big deal

- Past detection claims failed to be confirmed
- Want to set a high standard of evidence

Require consistency among multiple detectors

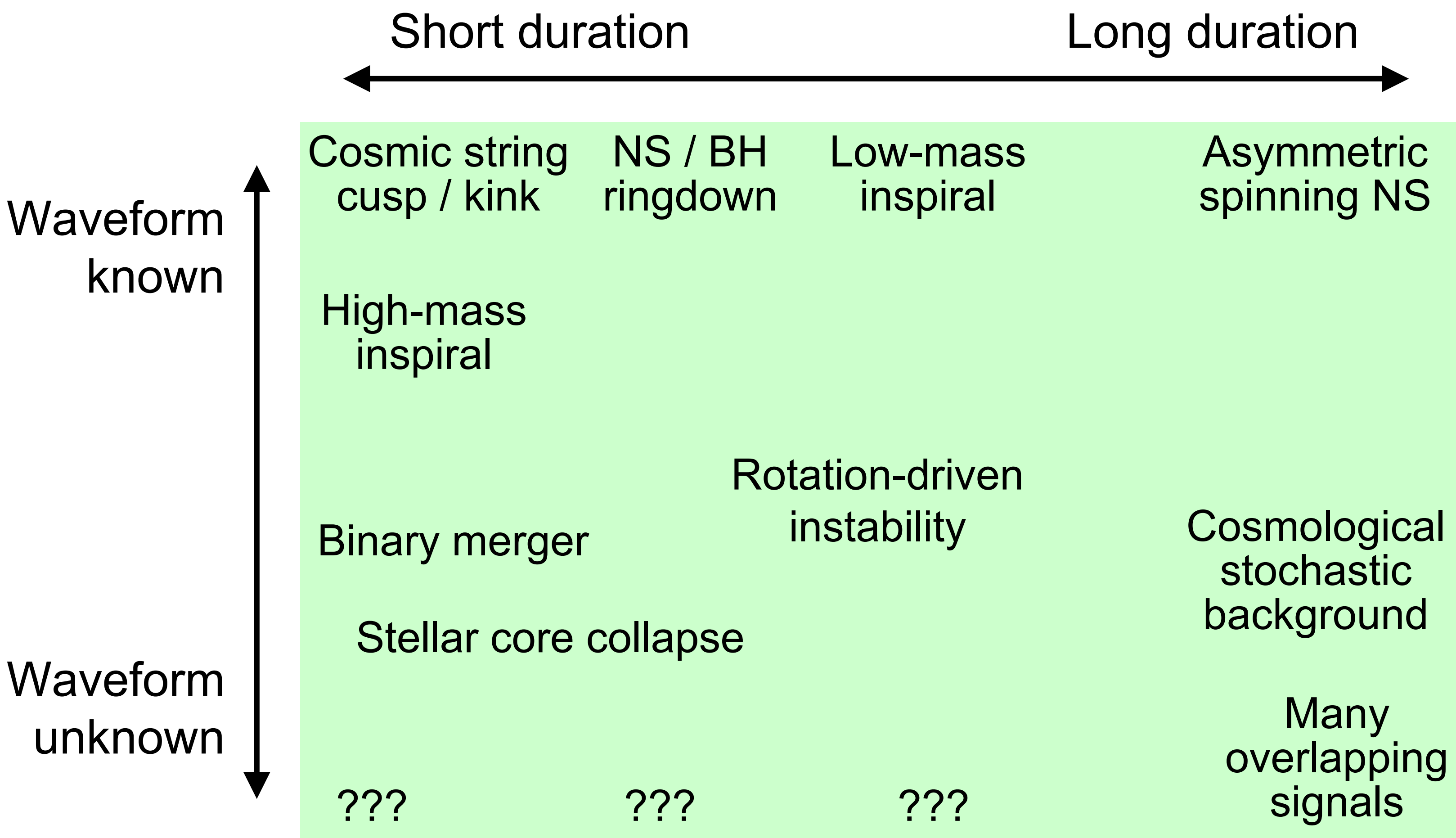
- Individual detectors may glitch
- Require coincidence or cross-coherence of some sort
- Allow for relative time delay, different antenna response, sensitivities

Estimate false alarm rate (“background”) **using time-shifted data**

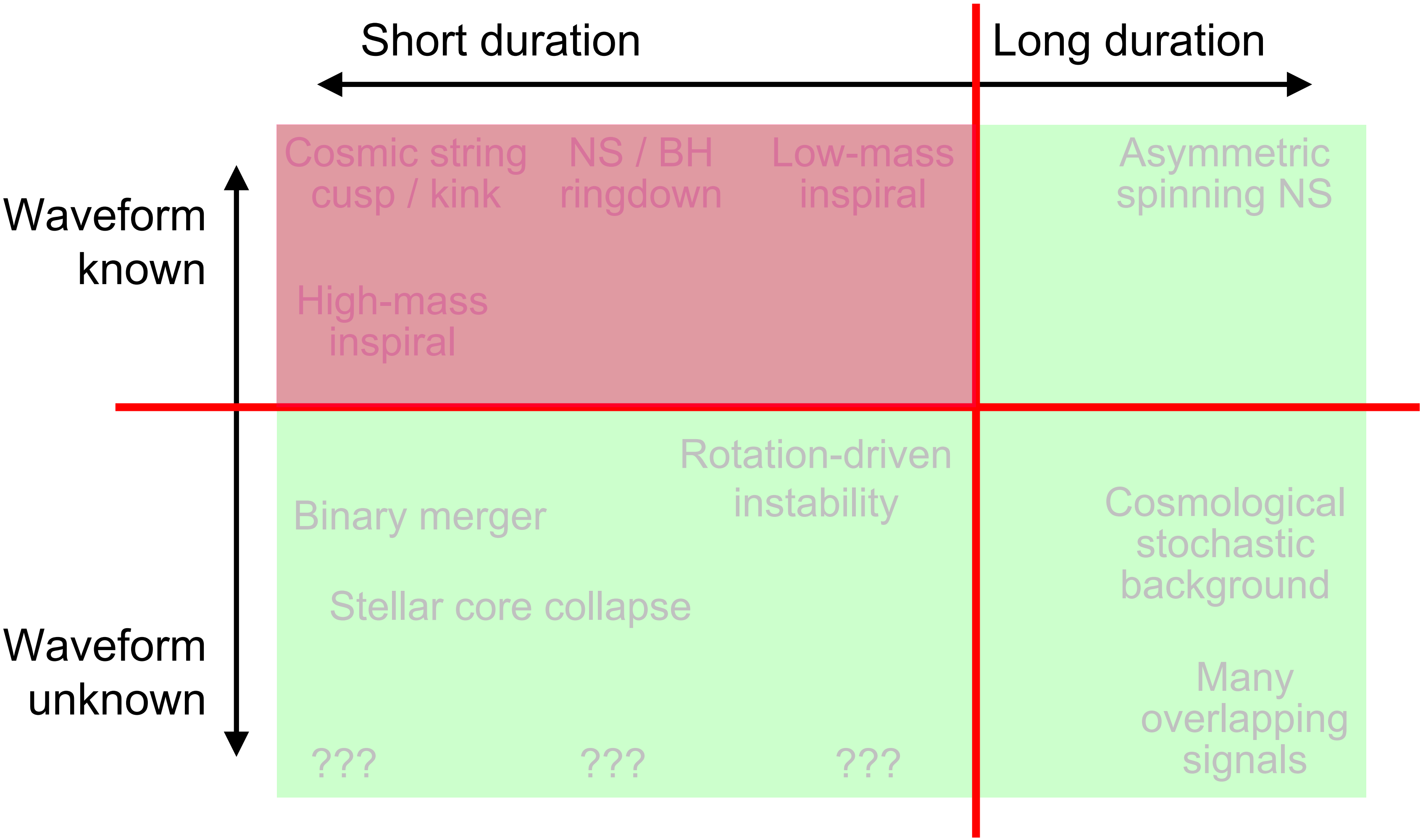
Outline

- **Gravitational-wave data**
- **General data analysis principles**
- **Specific data analysis methods**
  - Classification of signals
  - Methods for each class of signals
- **Idiosyncracies of real detectors**
- **The gravitational-wave community**

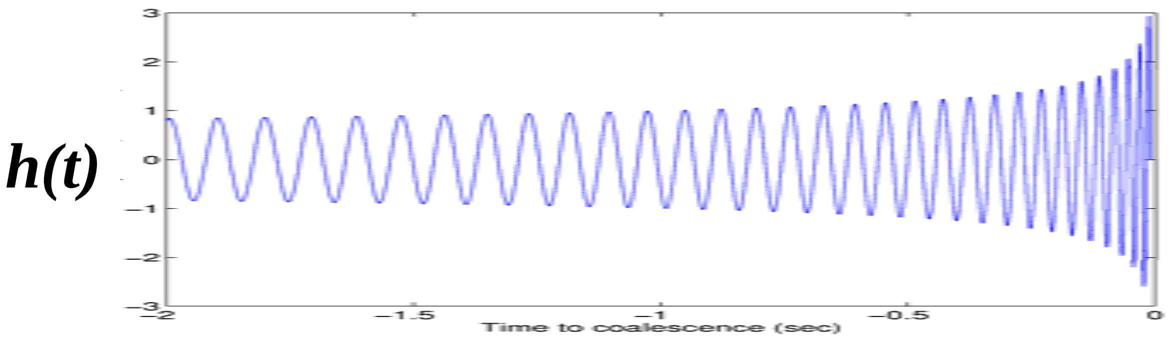
The Gravitational Wave Signal Tableau



Signal Classes



Short-duration, Known Waveforms: Inspirals, etc.



Known well, or fairly well, in some parametrized space

e.g. post-Newtonian expansion (assumes negligible spins)

$$\Psi(f) = 2\pi f t_c + \frac{3}{128\eta}(\pi m f)^{-5/3}$$

$$+\frac{5}{96\eta}\left(\frac{743}{336}+\frac{11}{4}\eta\right)(\pi m f)^{-1} \quad \text{1PN}$$

$$-\frac{3\pi}{8\eta}(\pi m f)^{-2/3} \quad \text{1.5PN}$$

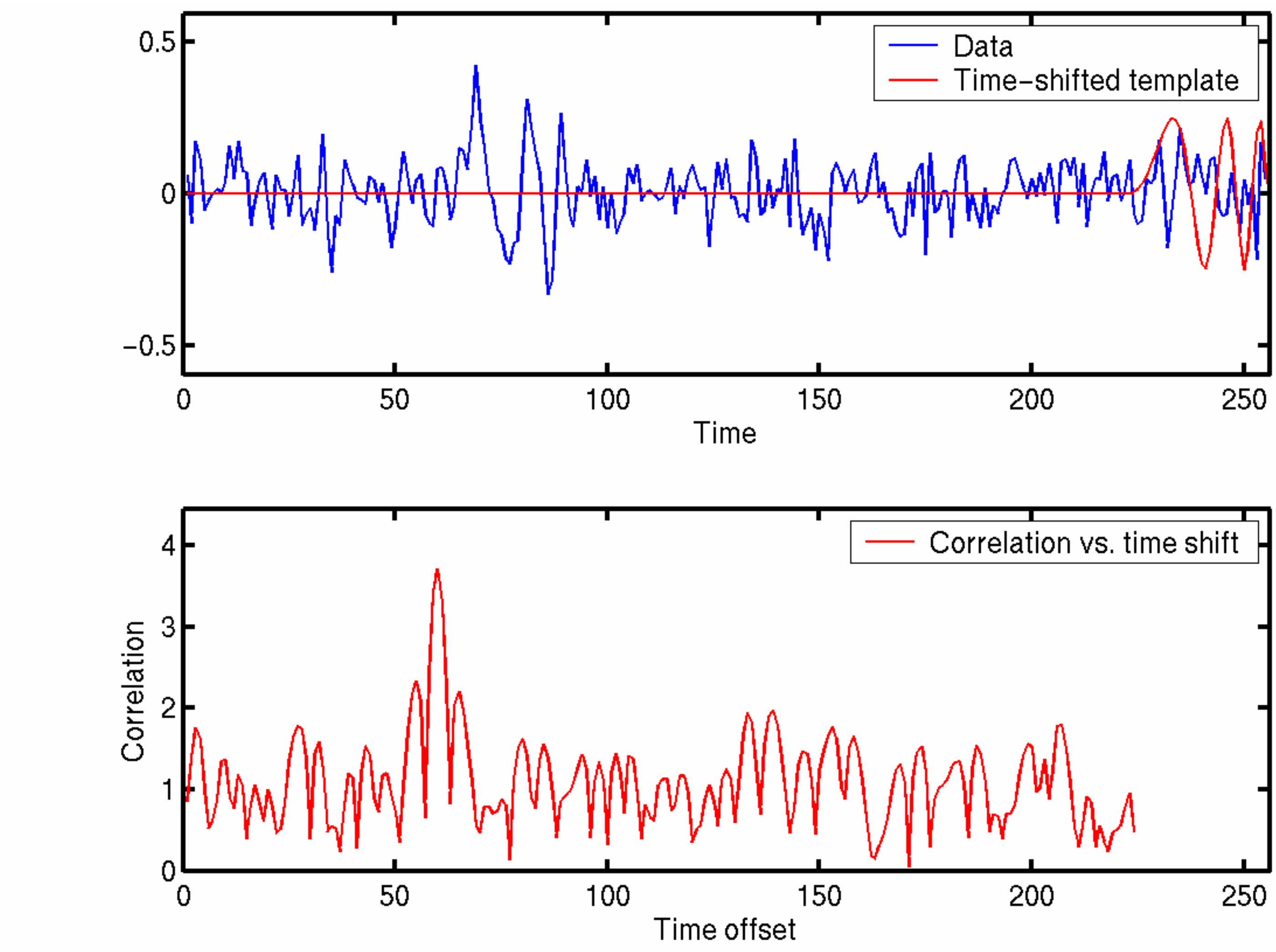
$$+\frac{15}{64\eta}\left(\frac{3058673}{1016064}+\frac{5429}{1008}\eta+\frac{617}{144}\eta^2\right)(\pi m f)^{-1/3} \quad \text{2PN}$$

$$+ \cdots$$

where $m = (m_1 + m_2)$, $\eta = \frac{m_1 m_2}{m^2}$

Known waveform ⇒ Use matched filtering

Basic Matched Filtering



Source Parameters vs. Signal Parameters

Inspiral source parameters

Masses (m1, m2)

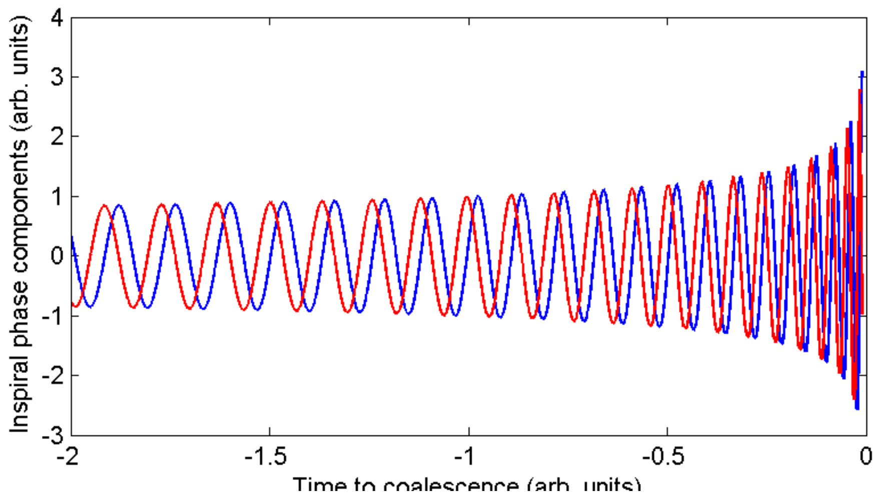
Spins → Assume negligible for now

Orbital phase at coalescence → Maximize analytically when filtering

Inclination of orbital plane
Sky location
→ Simply multiplicative for a given detector

Distance → Simply multiplicative

Filter with orthogonal templates, take quadrature sum



Optimal Matched Filtering in Frequency Domain

Data after FFT

Template, generated in freq. domain using stationary phase approx.

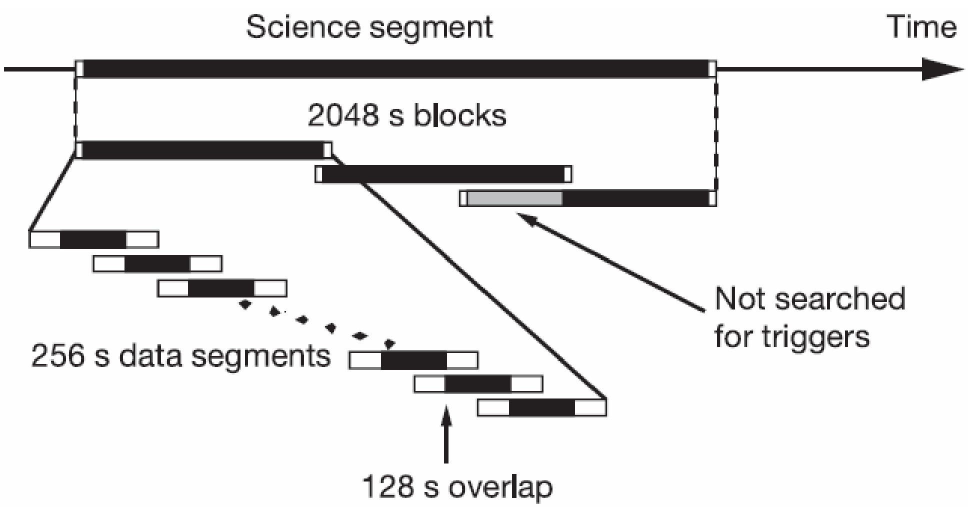
$$z(t) = 4 \int_0^\infty \frac{\tilde{s}(f)\ \tilde{h}^*(f)}{S_n(f)}\, e^{2\pi i f t}\, df$$

Noise power spectral density

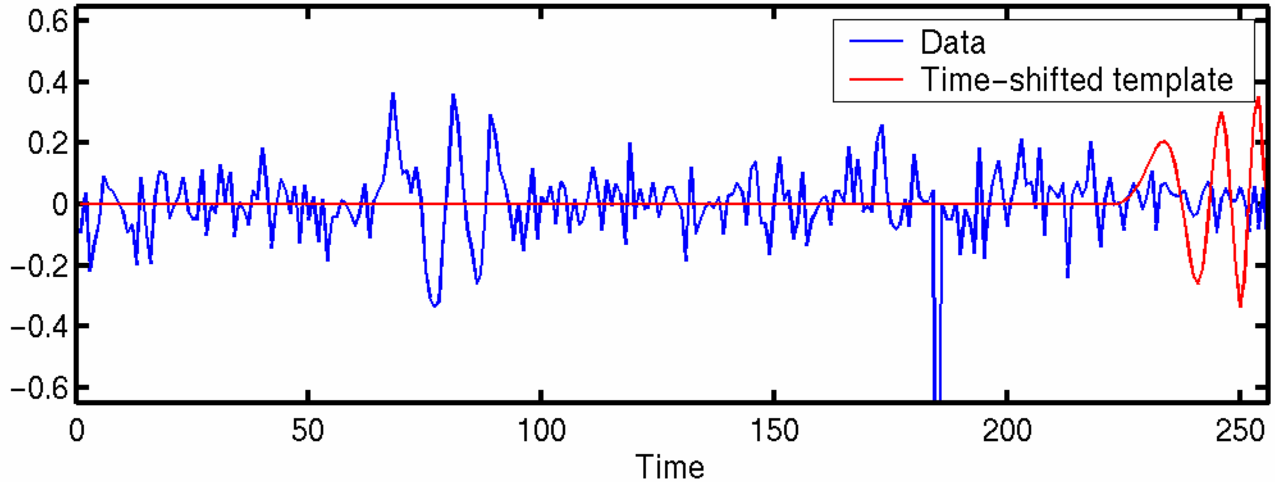
Look for maximum of $|z(t)|$ above some threshold → trigger

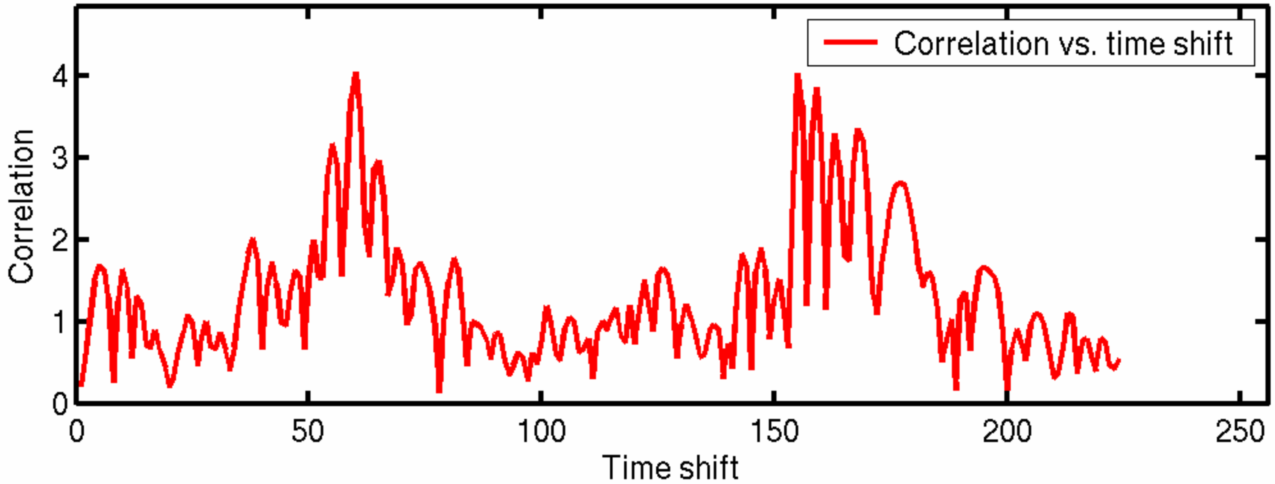
Search overlapping intervals to cover science segment, avoid wrap-around effects

Estimate power spectrum from bin-by-bin median of fifteen 256-sec data segments



Matched Filtering
Susceptibility to Glitches



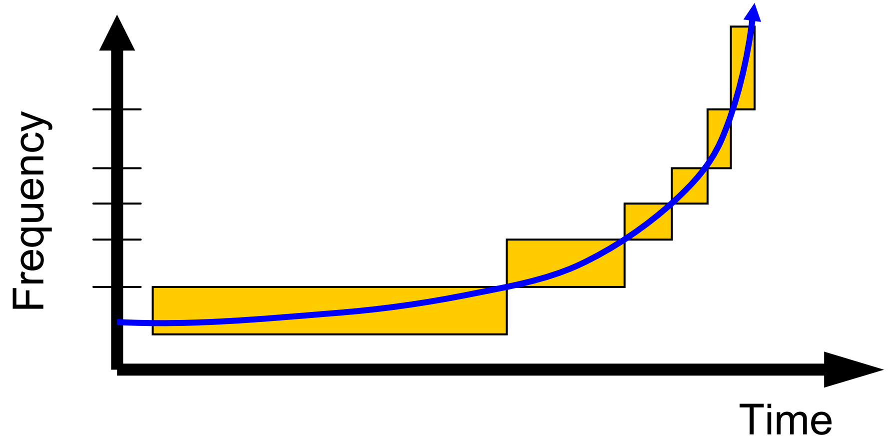


Waveform Consistency Tests

Chi-squared test

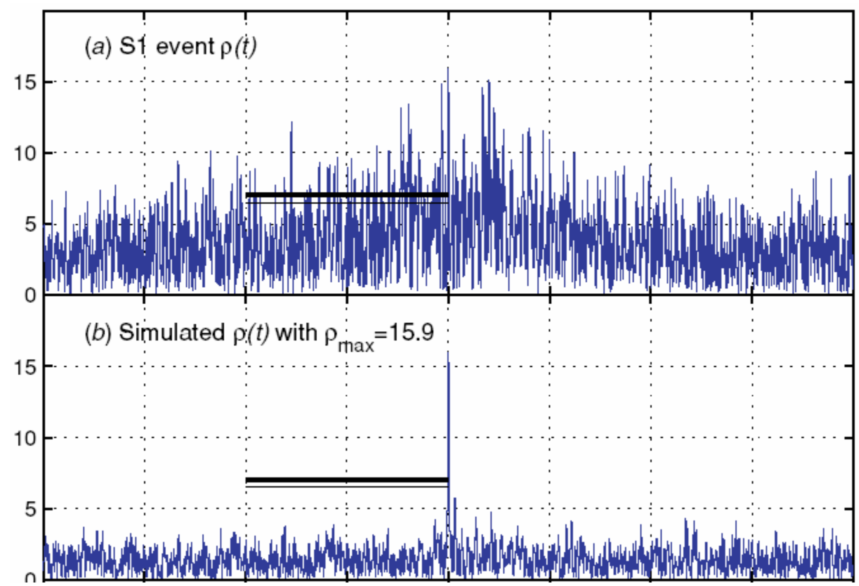
Divide template into p parts, calculate

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

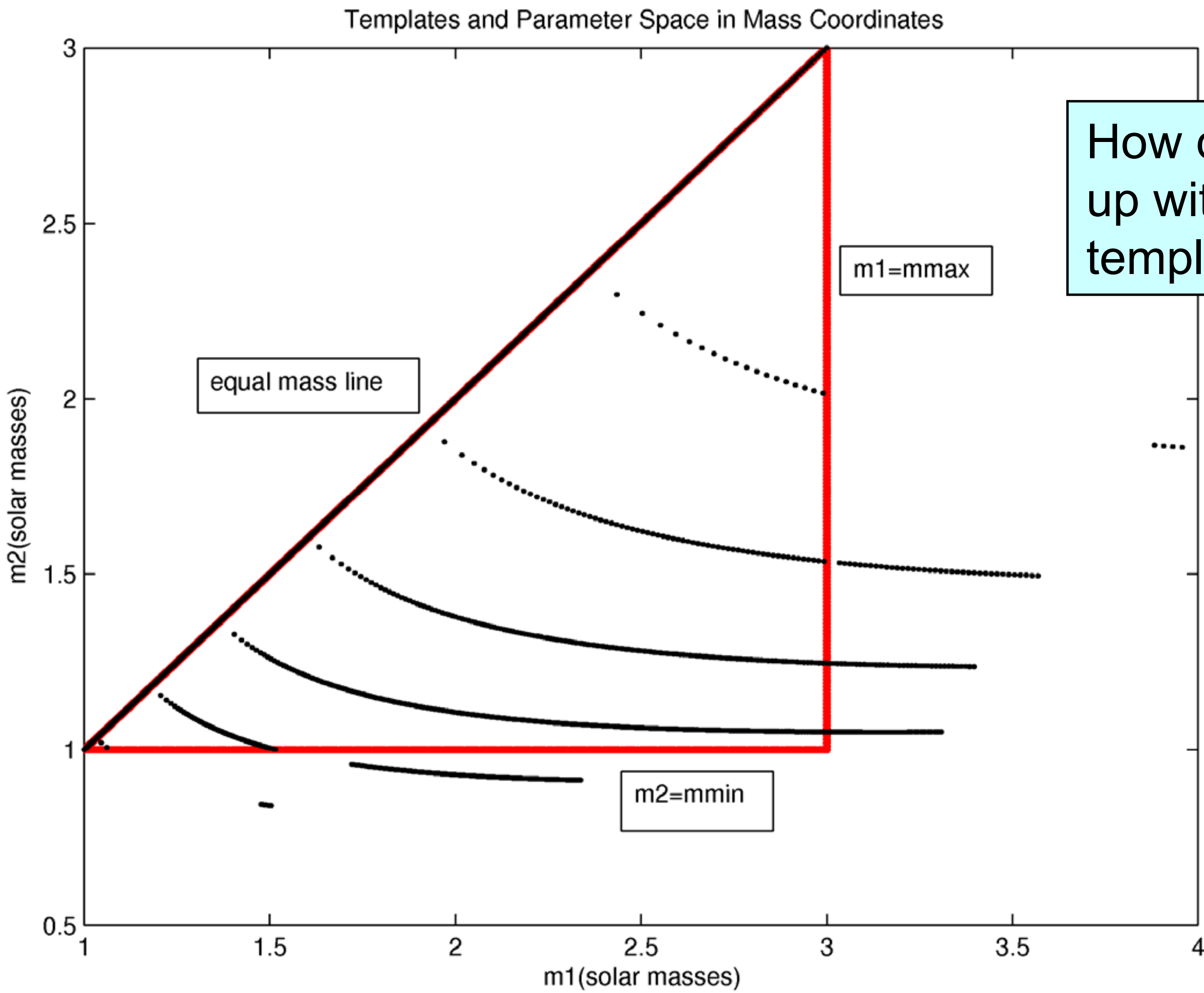


Tests using filter output

e.g. time above threshold

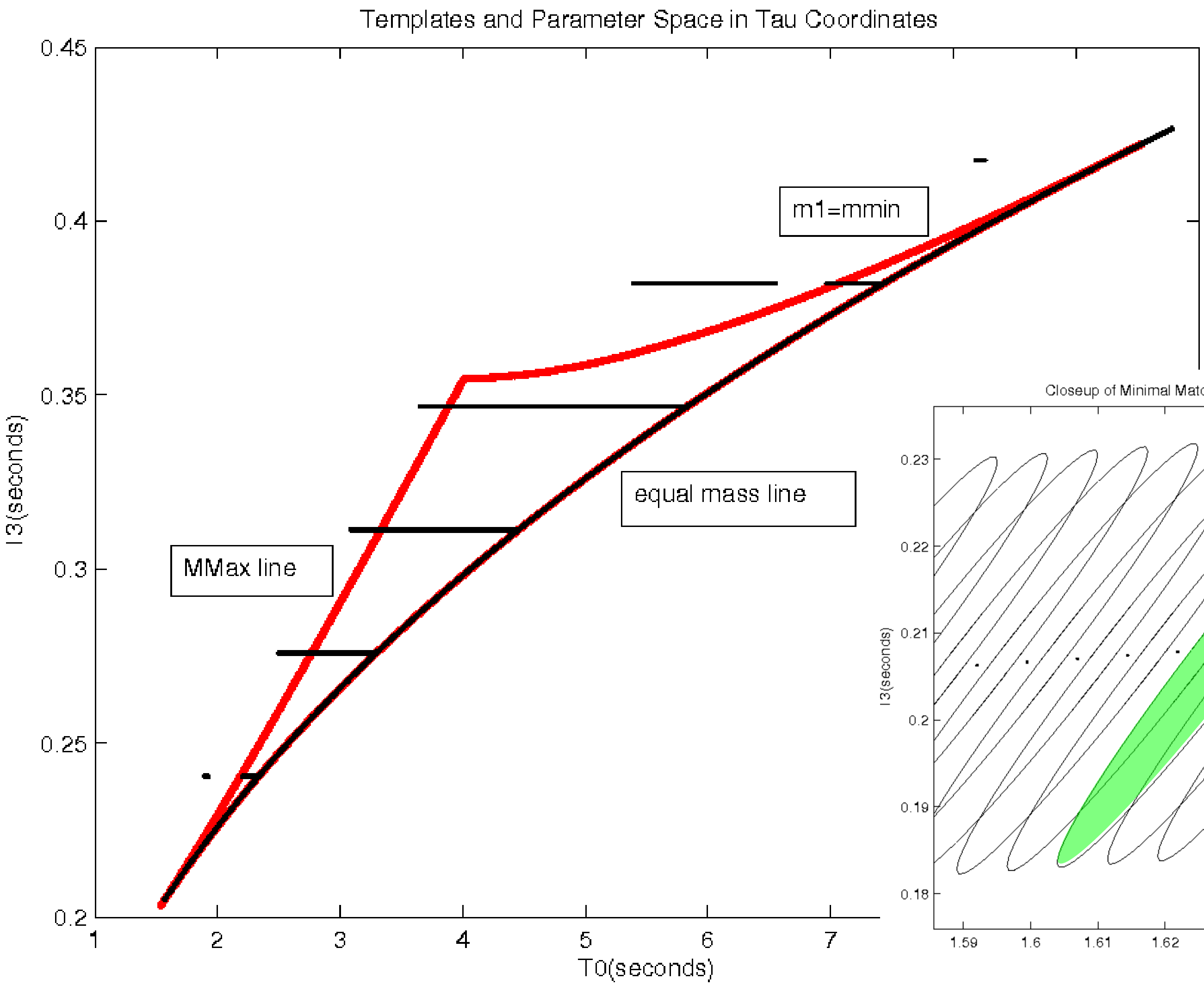


Template Bank Construction



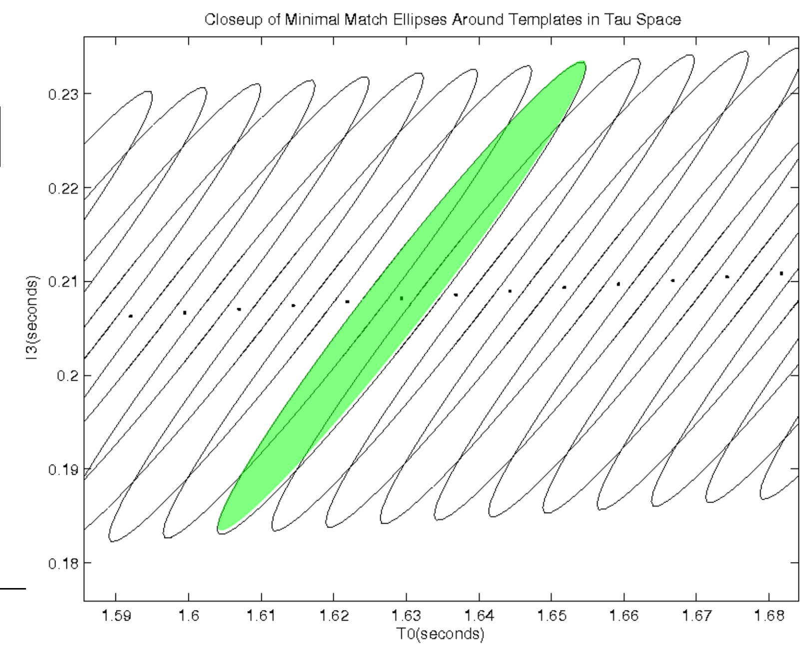
How did we come up with this set of templates???

Template Bank Construction in (τ_0,τ_3) space

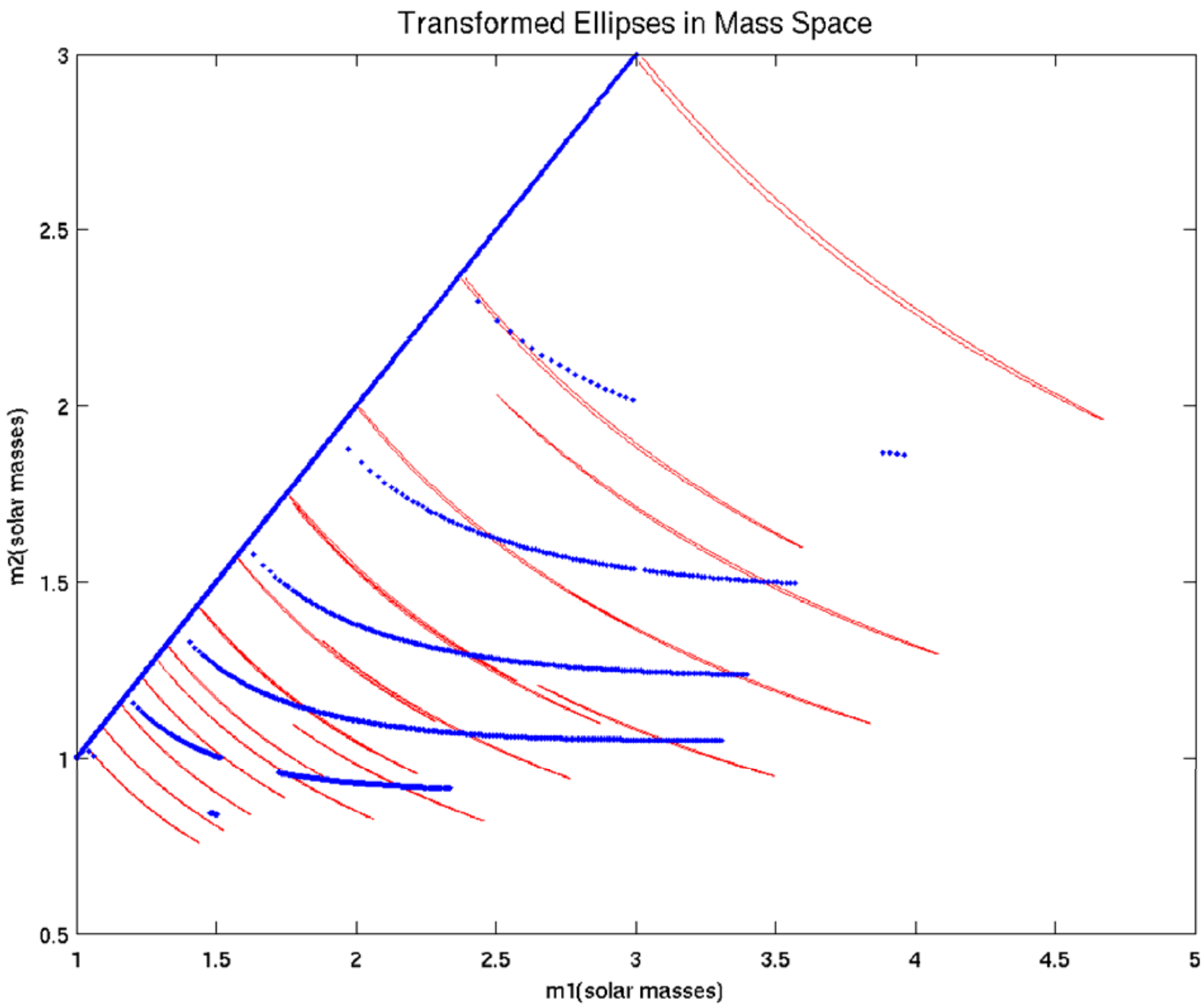


$$\tau_0 \propto m^{-5/3} / \eta$$

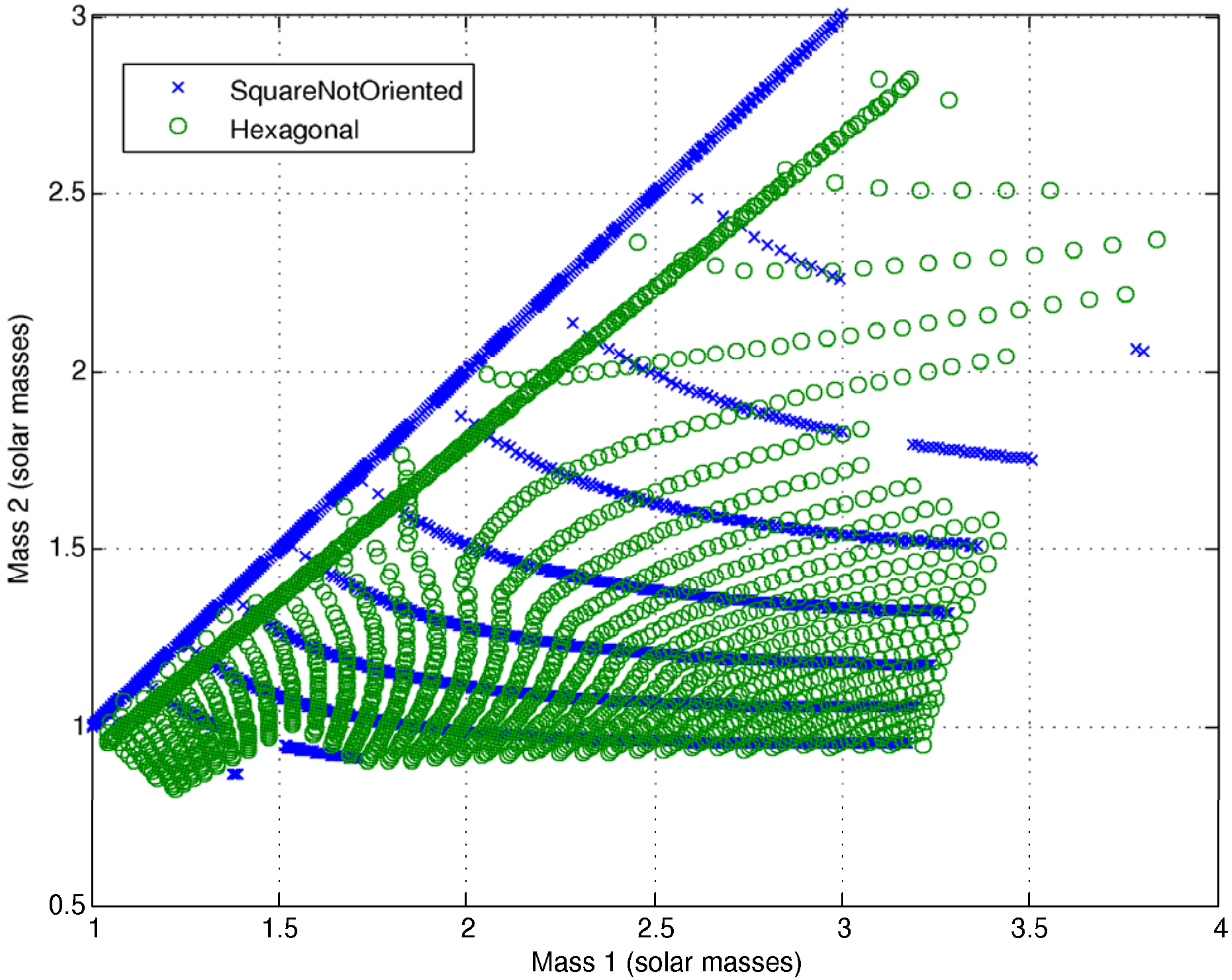
$$\tau_3 \propto m^{-2/3} / \eta$$



Ellipses in Mass Space

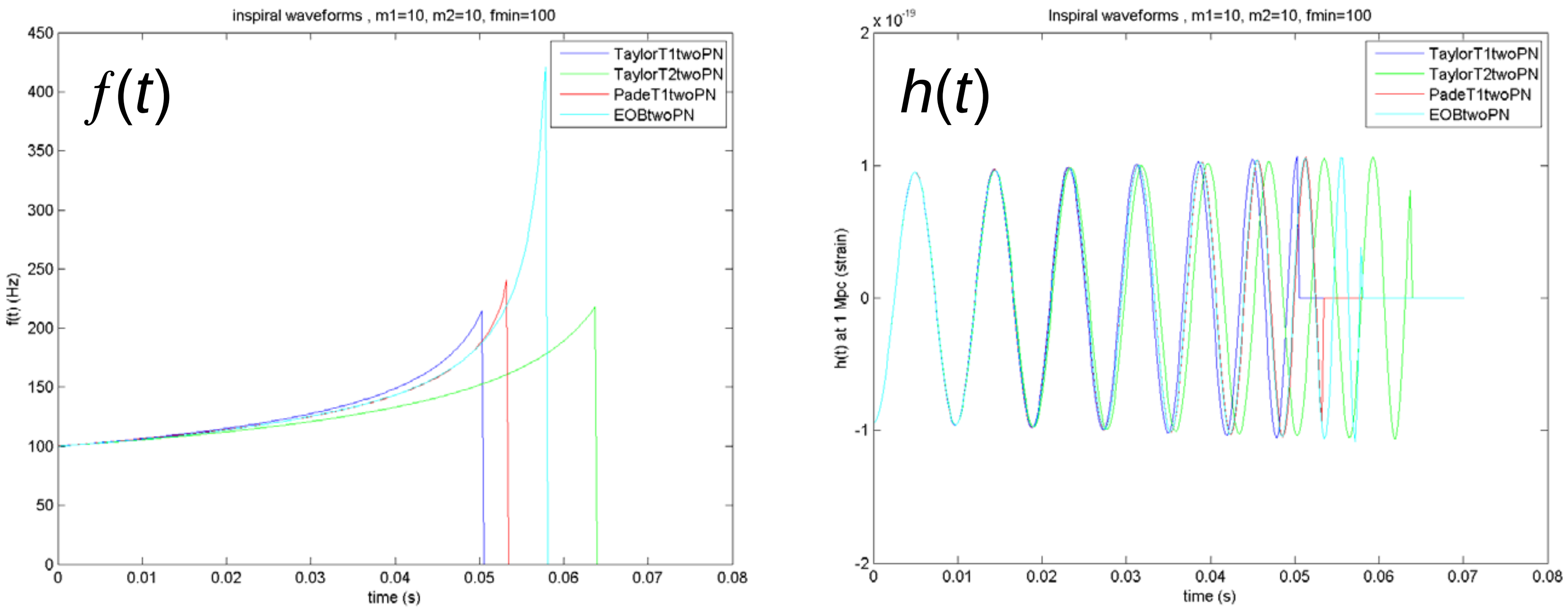


Different Bank Layout Methods



Uncertain Waveforms for High-Mass Inspirals

Different models for 10+10 M_{sun} black hole binary inspiral



Templates for Detection vs. Parameter Estimation

Can use a parametrized space of templates

e.g. Buonanno, Chen, and Vallisneri, Phys. Rev. D 67, 104025 (2003)

$$h(f) = f^{-7/6}(1-\alpha f^{2/3})\theta(f_{cut}-f)\exp[i(\phi_0+2\pi t_0 f+\psi_0 f^{-5/3}+\psi_3 f^{-2/3})]$$

Analytically calculate α to maximize SNR

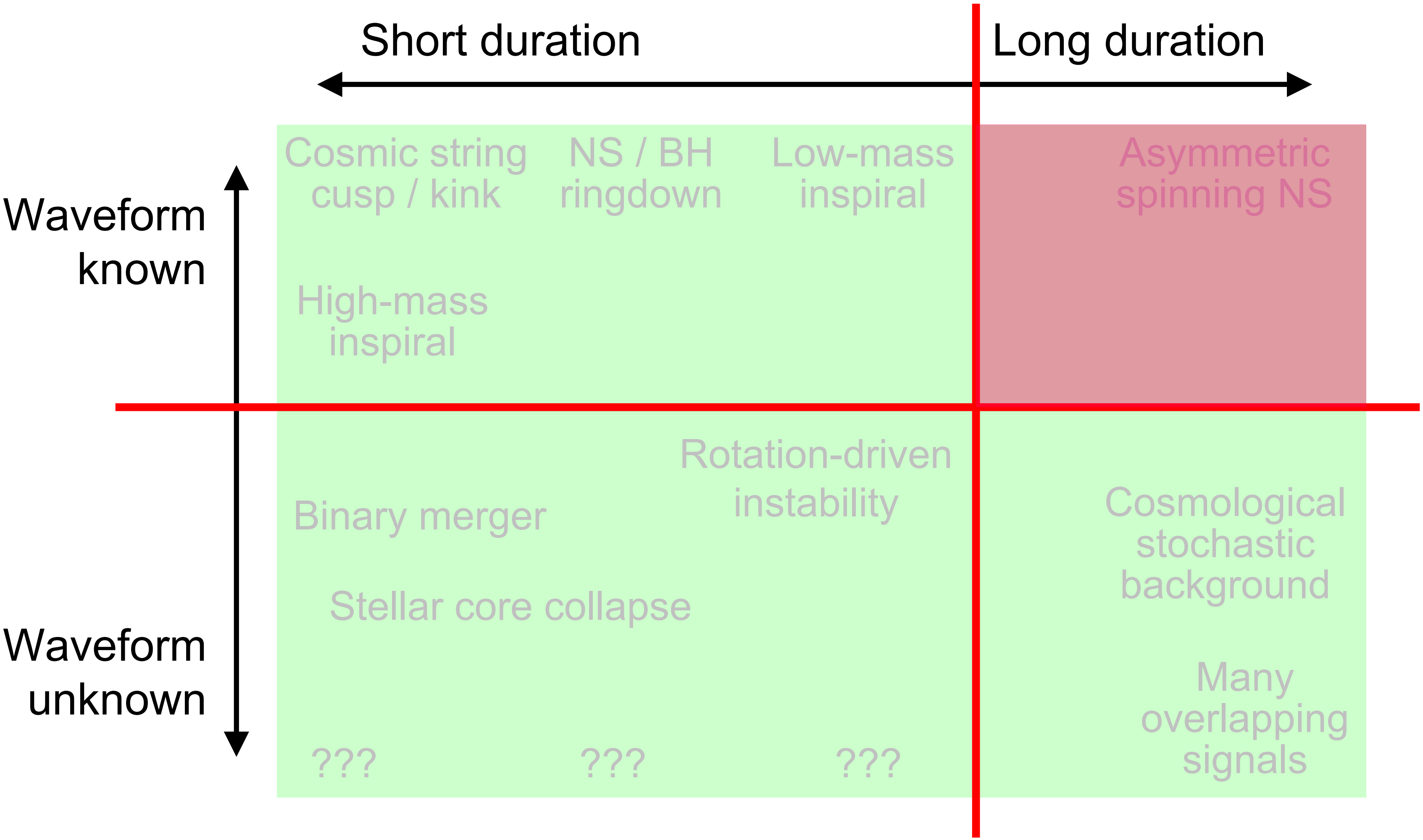
Parameters of the search

This can match the various waveform models rather well

Intended for binary components with negligible spin

Once a signal is detected, re-filter with physical templates to extract physical parameters

Signal Classes

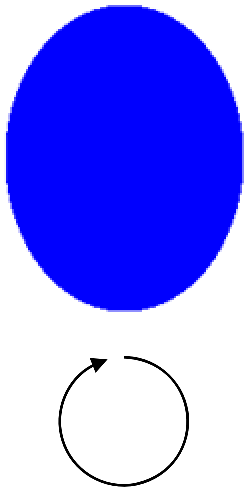


Continuous, Known Waveform: GW from Spinning Neutron Stars

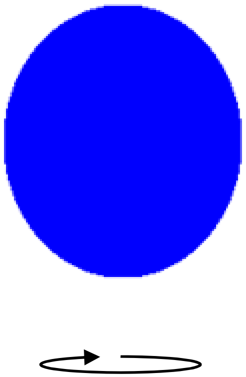
If not axisymmetric, will emit gravitational waves

Example: ellipsoid with distinct transverse axes

Along spin axis:



From side:



Continuous GW Signals at Earth

Start with a sinusoidal signal with spin-down term(s)

Polarization content depends on orientation/inclination of spin axis

Amplitude modulation

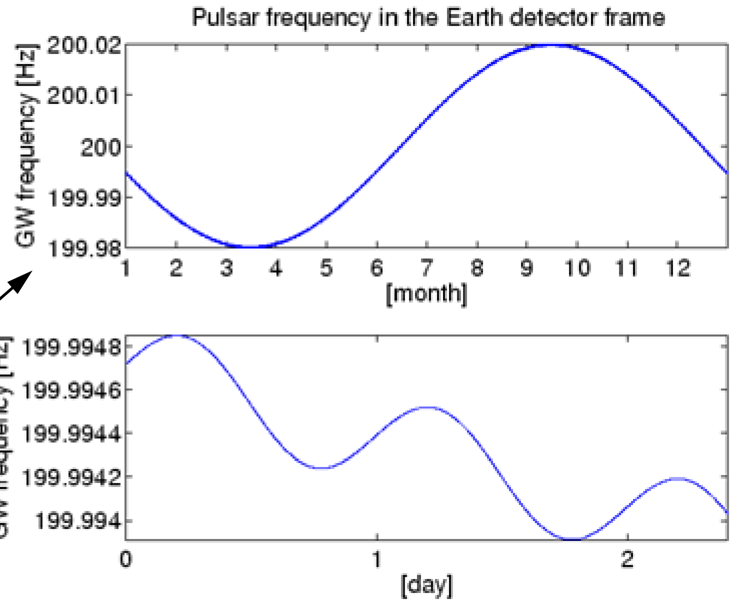
Polarization projection changes over a sidereal day

Doppler shift

$$\frac{\Delta f}{f} = \frac{\mathbf{v} \cdot \mathbf{n}}{c}$$

Annual variation: up to ~10^{-4}

Daily variation: up to ~10^{-6}



GW signals from binary systems are more complicated !

Additional Doppler shift due to orbital motion of neutron star

Varying gravitational redshift if orbit is elliptical

Shapiro time delay if GW passes near companion

Search Methods for CW signals

Several cases to consider:

- Sky position and spin frequency known accurately
- Sky position and spin frequency known fairly well
- Sky position known, but frequency and/or binary orbit parameters unknown
- Search for unknown sources in favored sky regions
- Search for unknown sources over the whole sky

Candidates

Radio pulsars
X-ray pulsars

LMXBs

Globular clusters
Galactic center
Supernova remnants

Unseen isolated neutron stars

Different computational challenges $\Rightarrow$ Different approaches

Search for Gravitational Waves from Known Pulsars

Method: **heterodyne time-domain data using the known spin phase of the pulsar**

- Requires precise timing data from radio or X-ray observations
- Include binary systems in search when orbits known accurately
- Exclude pulsars with significant timing uncertainties
- Special treatment for the Crab and other pulsars with glitches, timing noise

Wide Parameter Space Searches

Method: **matched filtering with a bank of templates**

Parameters:

- Sky position
- Spin axis inclination and azimuthal angle
- Frequency, spindown, initial phase
- Binary orbit parameters (if in a binary system)

Use a detection statistic, $\mathcal{F}$, which analytically maximizes over spin axis inclination & azimuthal angle and initial phase

- Even so, computing cost scales as ~T^6
- Detection threshold also must increase with number of templates

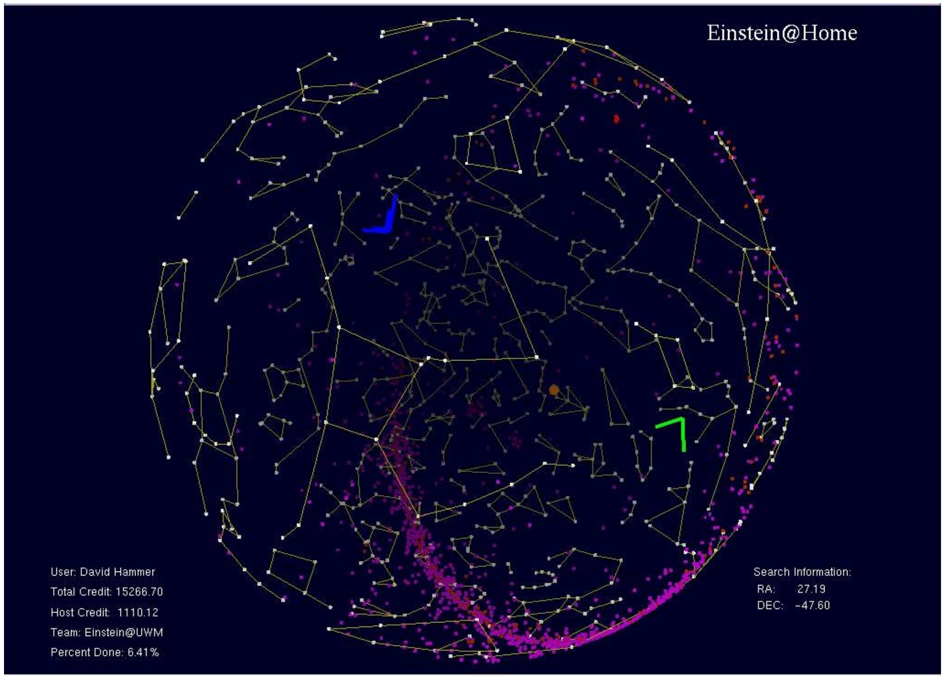
Check for signal consistency in multiple detectors

Problem: huge number of templates needed

Getting by with a Little Help from Our Friends

Public distributed computing project: Einstein@Home

Small bits of data distributed for processing; results collected, verified, and post-processed



Screen saver graphics

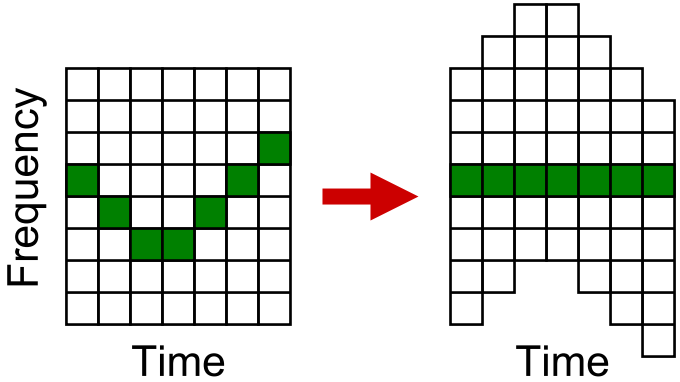
So far 156,000 users, currently providing ~77 Tflops

Semi-Coherent Search Methods

Can't do an all-sky coherent search using all of the data

Divide data into time intervals, calculate power, sum it

- Less sensitive for a given observation time, but computationally more efficient, so can use **all** the data
- Generally use 30-minute "short Fourier transforms" (SFTs)



Different methods of adding SFTs

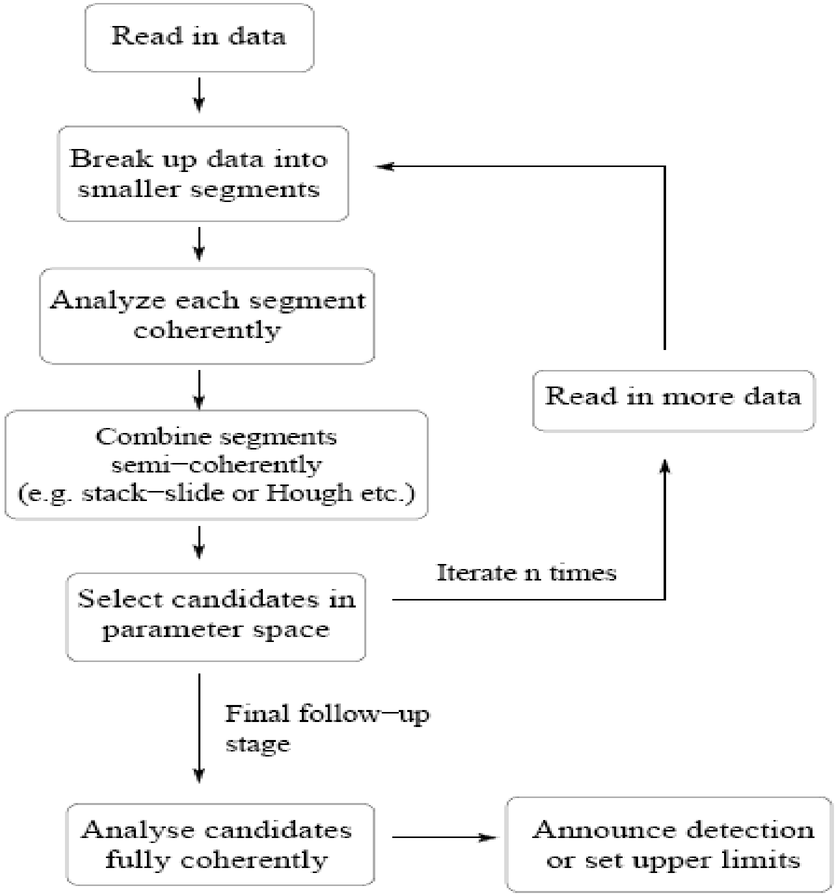
- "StackSlide" : sums normalized power
- "PowerFlux" : sums normalized power with weights for sky position, noise
- "Hough" : sums binary counts with weights for sky position, noise

Hierarchical Search

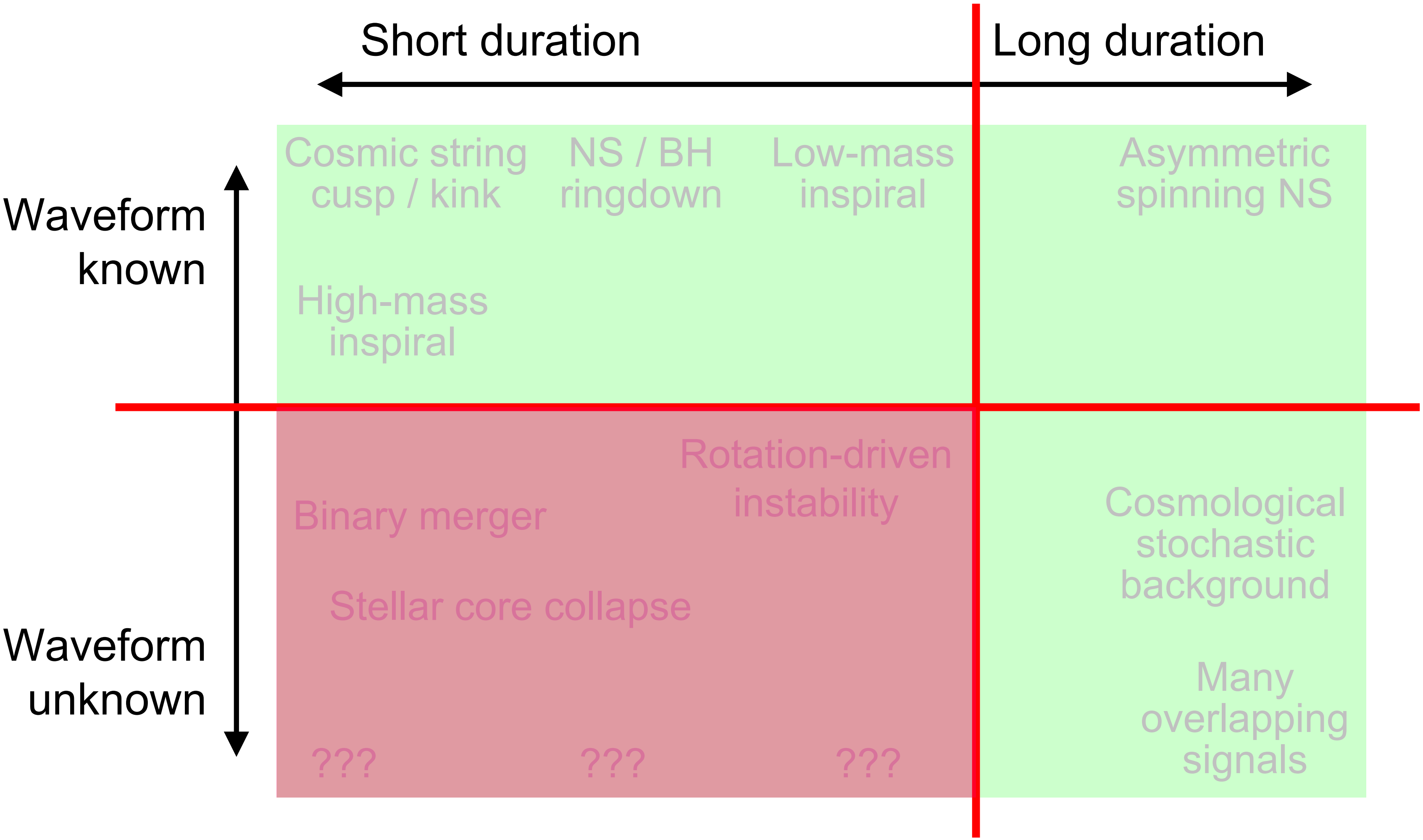
Alternate semi-coherent and fully coherent stages

Gets closer to optimal sensitivity, at a manageable CPU cost

Example:



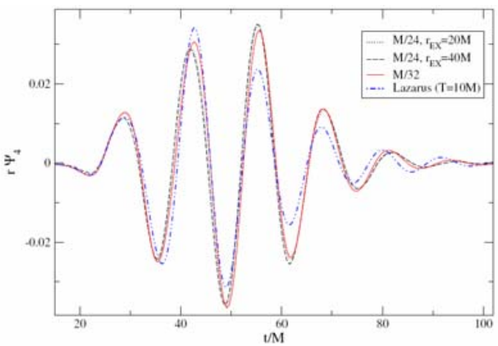
Signal Classes

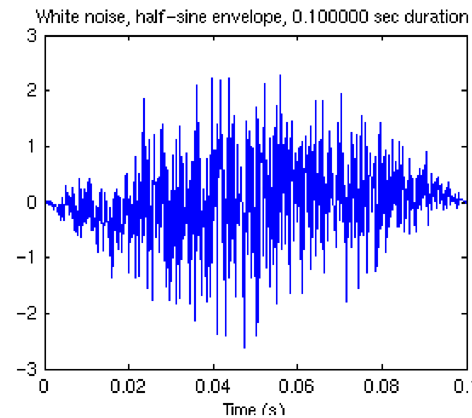


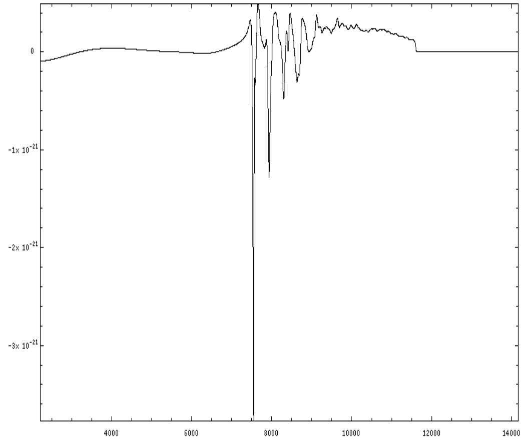
Short-duration, Unknown Waveform: Gravitational-Wave Bursts

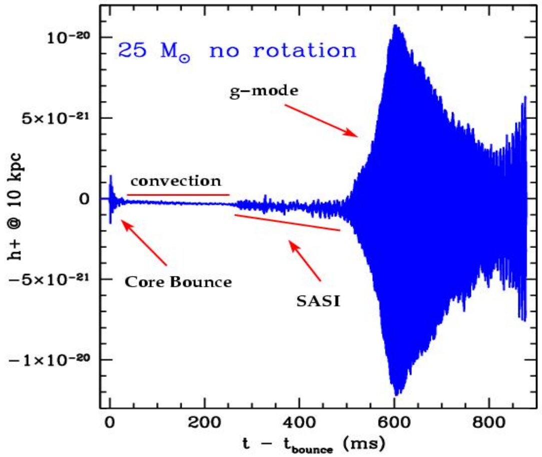
We're exploring the sky – Who knows what is out there to find?

Want to be able to detect any arbitrary signal









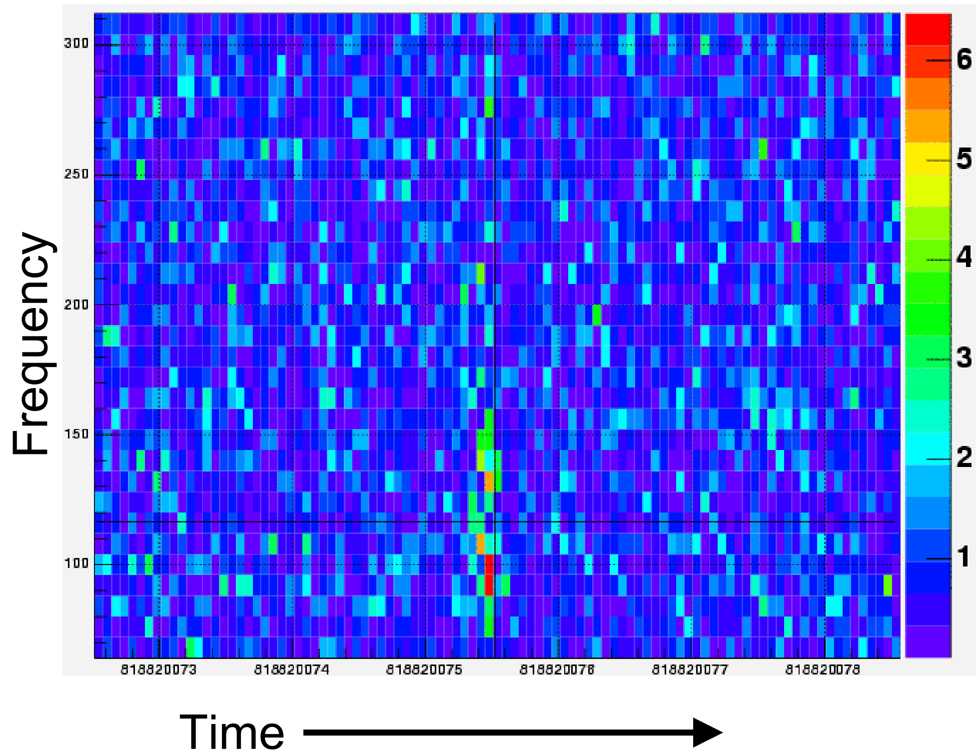
"Excess Power" Search Methods

Decompose data stream into time-frequency pixels

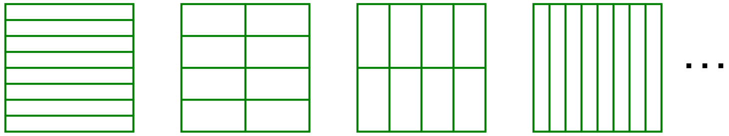
Fourier components, wavelets, "Q transform", etc.

Normalize relative to noise as a function of frequency

Look for "hot" pixels or clusters of pixels

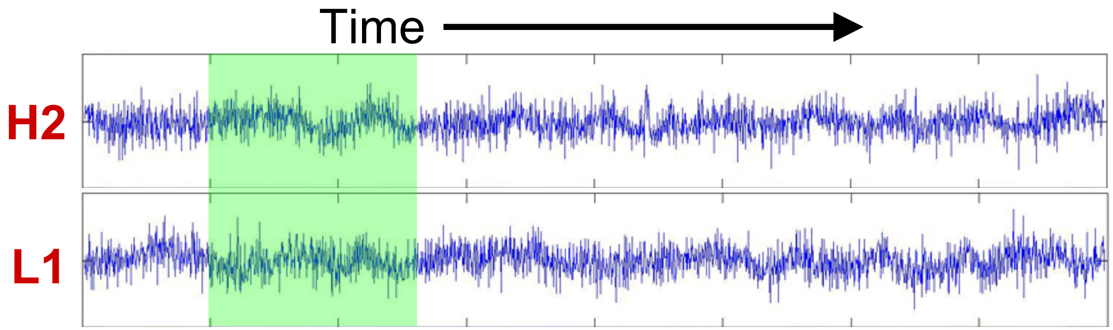


Can use multiple (Δt,Δf) pixel resolutions



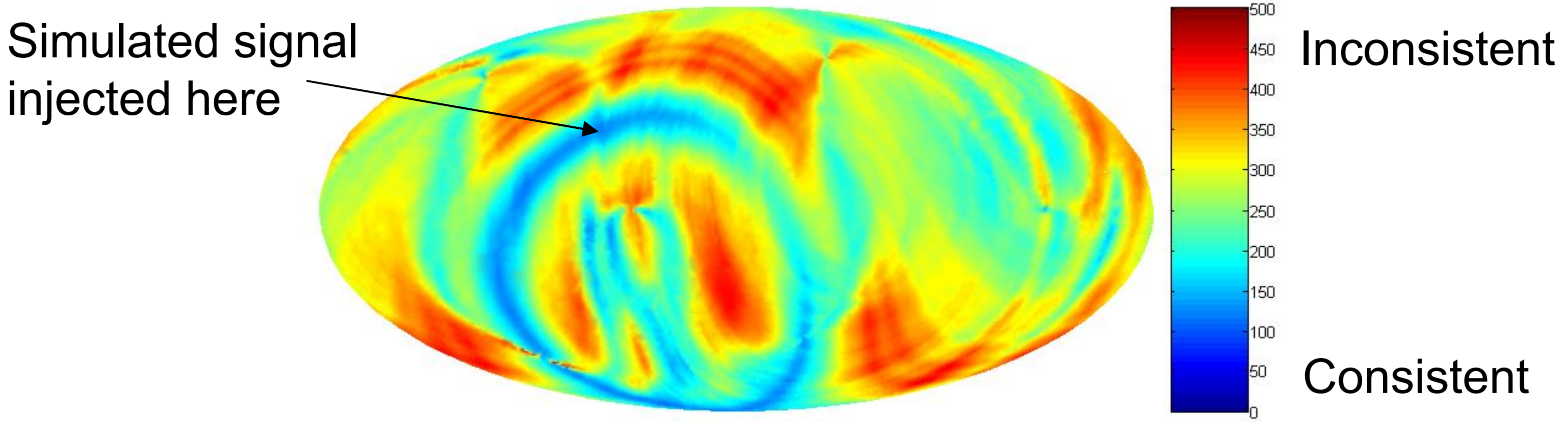
Cross-Correlation Methods

Look for same signal buried in two data streams

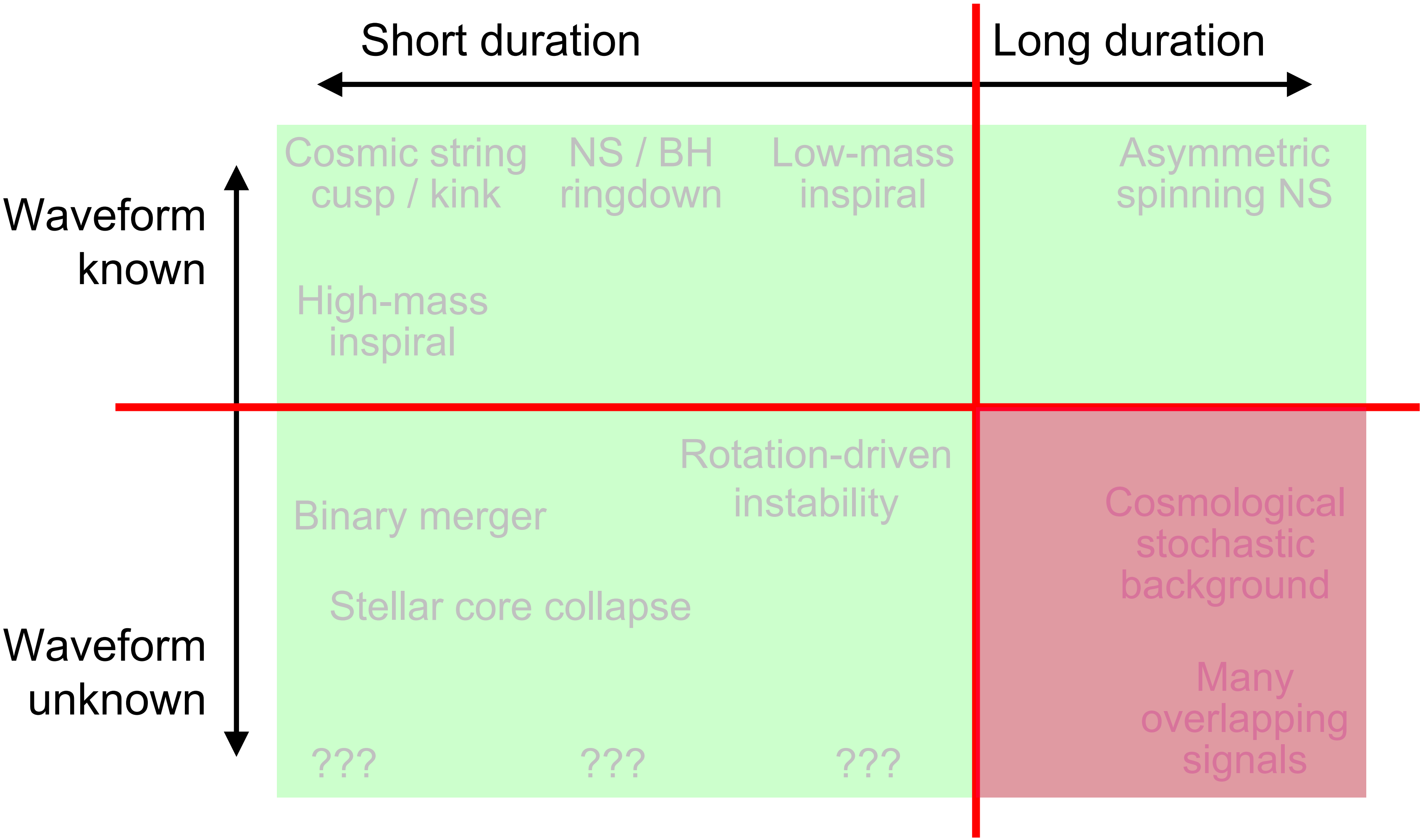


Integrate over a time interval comparable to the target signal

Extensions to three or more detector sites being worked on



Signal Classes



Continuous, Unknown Waveform: Stochastic Gravitational Waves

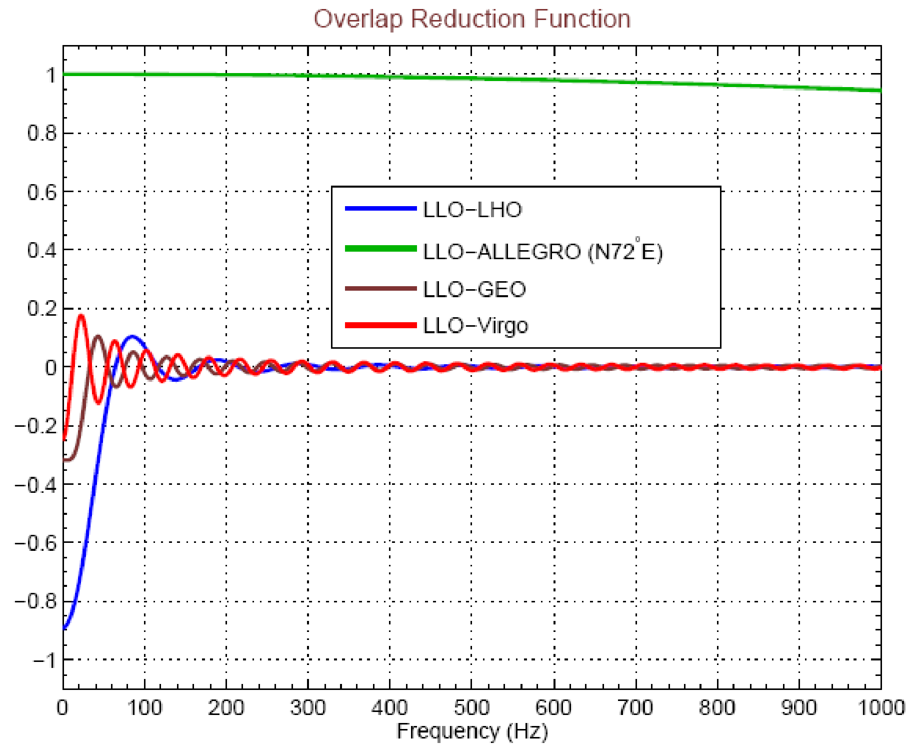
Use cross-correlation to search for signal smaller than detector noise

For isotropic stochastic GWs, know what correlation to expect between any given pair of detectors

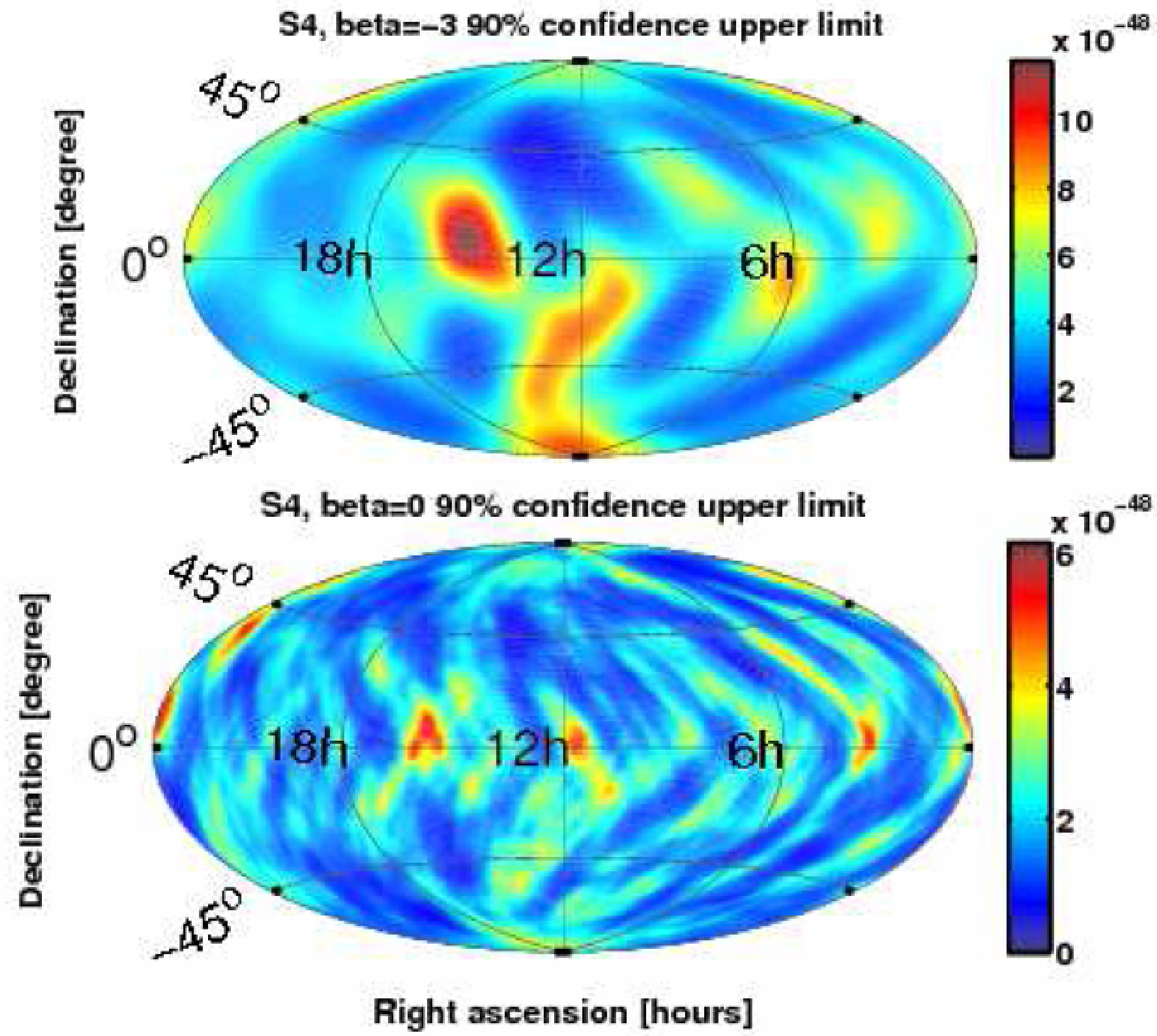
Optimal filter:

$$Y = \int df\, \tilde{s}_1^*(f)\, \tilde{Q}(f)\, \tilde{s}_2(f)$$

$$\tilde{Q}(f) \propto \frac{f^{-3}\Omega_{GW}(f)\gamma_{12}(f)}{P_1(f)P_2(f)}$$



Sky Map of Stochastic Gravitational Waves: "Radiometer"



Outline

- Gravitational-wave data
- General data analysis principles
- Specific data analysis methods
  - Classification of signals
  - Methods for each class of signals
- Idiosyncracies of real detectors
- The gravitational-wave community

Variable Data Quality

Various environmental and instrumental conditions catalogued; can study relevance using *time-shifted* coincident triggers

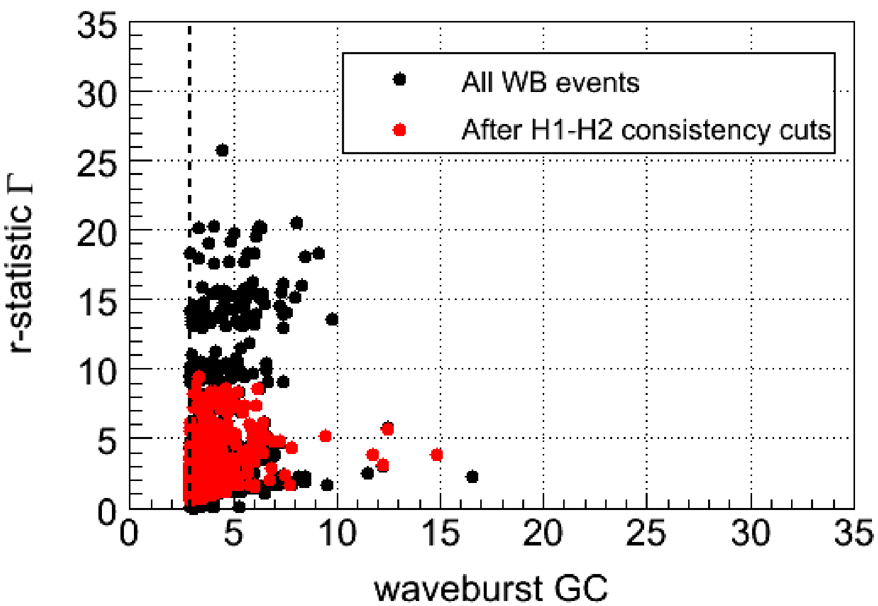
Example from S4 all-sky burst search:

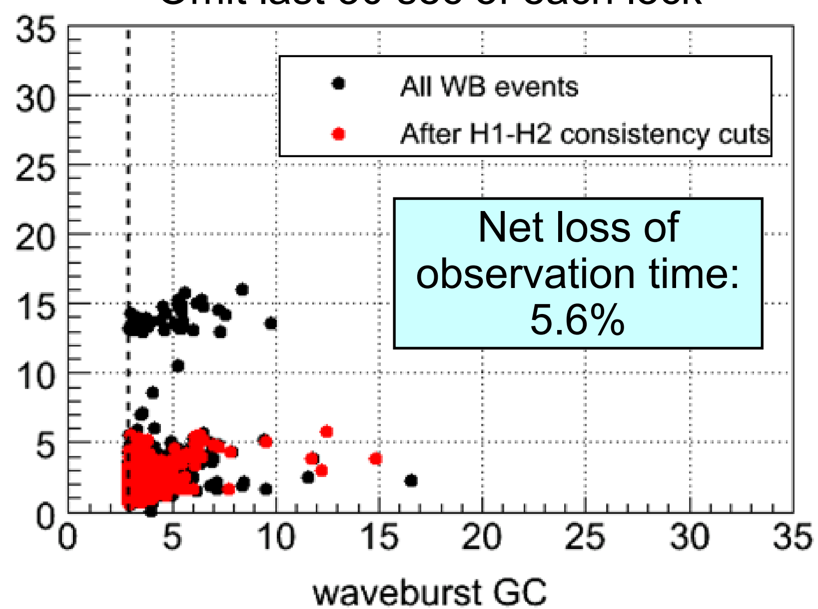
Minimal data quality cuts

Require locked interferometers
Omit hardware injections
Avoid times of ADC overflows

Additional data quality cuts

Avoid high seismic noise, wind, jet
Avoid calibration line drop-outs
Avoid times of “dips” in stored light
Omit last 30 sec of each lock



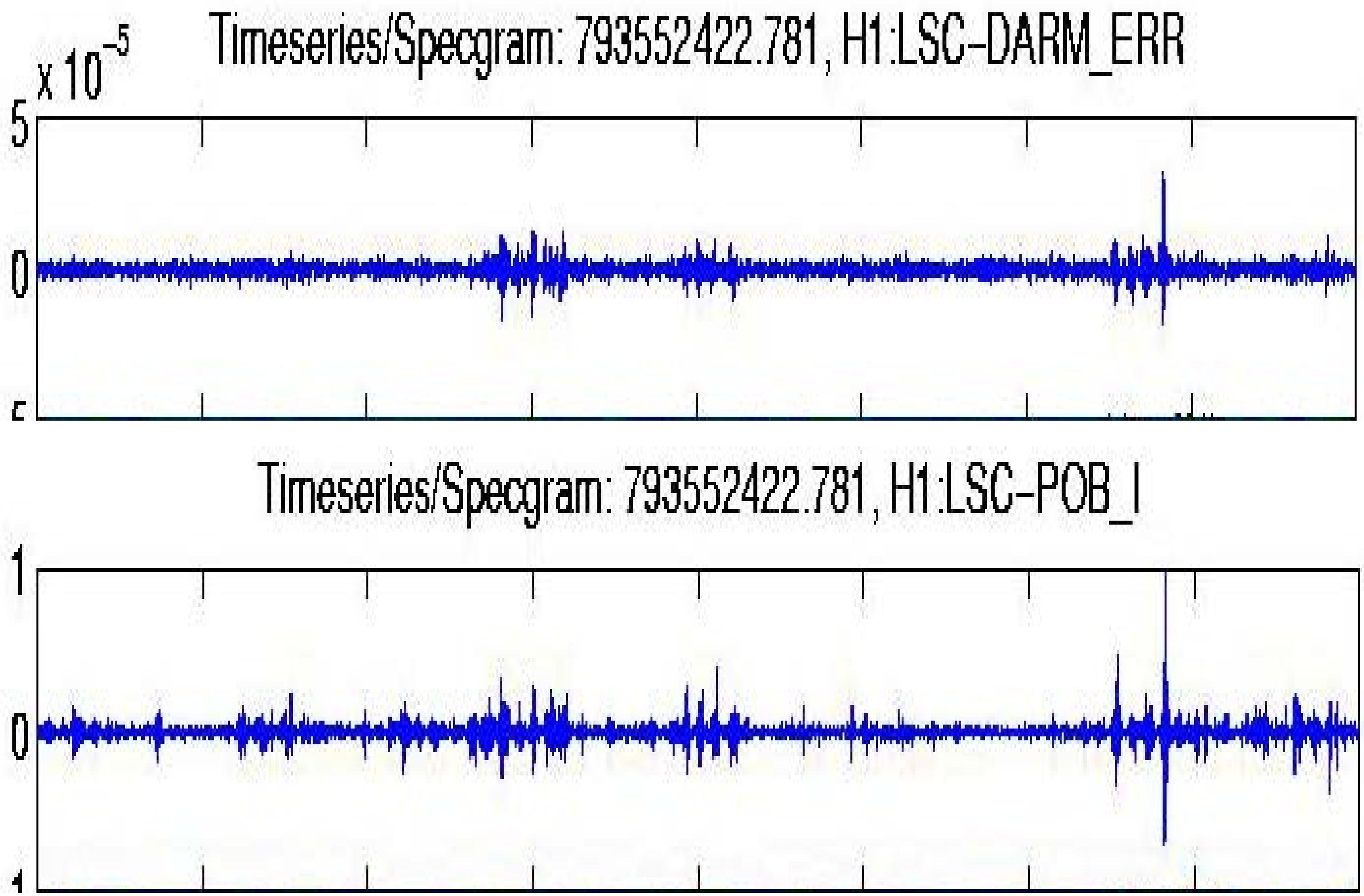


Non-Stationary Noise / Glitches

Auxiliary-channel vetoes

GW channel

Beam splitter pick-off



Outline

- **Gravitational-wave data**
- **General data analysis principles**
- **Specific data analysis methods**
  - Classification of signals
  - Methods for each class of signals
- **Idiosyncracies of real detectors**
- **The gravitational-wave community**

Big Science!

The LIGO Scientific Collaboration

- A few hundred people from ~50 institutions
- Includes everyone from GEO
- LIGO and GEO data analyzed together

Virgo

TAMA 300

Bar detectors (ALLEGRO, AURIGA, EXPLORER, NAUTILUS)

Cooperative observing and joint data analysis

- LIGO and TAMA 300, LIGO and ALLEGRO, LIGO and AURIGA, VIRGO and Explorer+Nautilus
- LIGO-VIRGO data exchange and joint analysis begins May 18