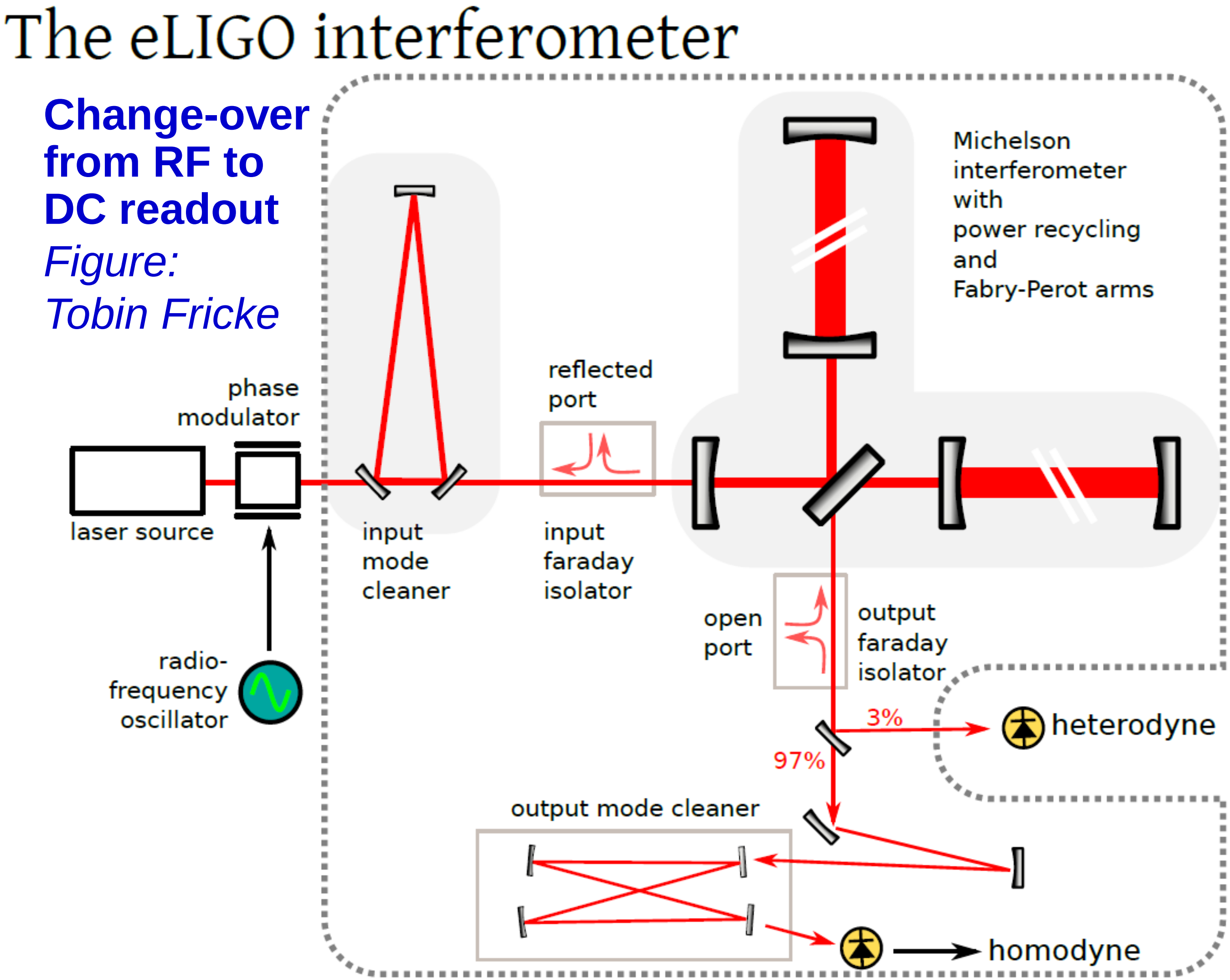
Gravitational Wave Data Analysis

Peter Shawhan

PHYS 879
May 6, 2014

GW Detector Readout – Overview



Change-over from RF to DC readout

Figure: Tobin Fricke

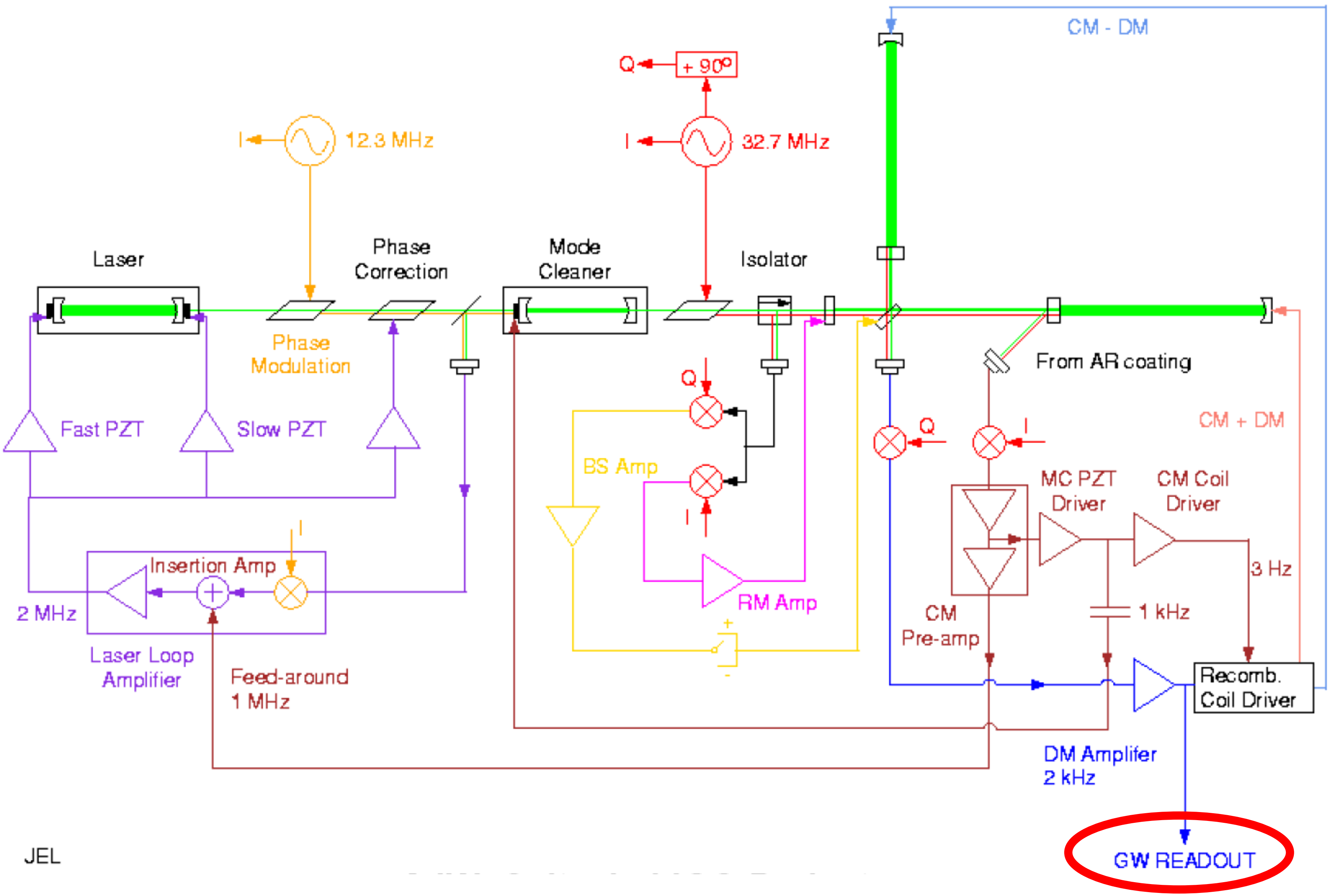
Heterodyne (RF) readout used for initial LIGO/Virgo

- Modulate phase of input light (33 MHz), demodulate signal measured by photodiode
- Perfect destructive interference on avg

Homodyne (DC) readout used for Adv. LIGO/Virgo

- Measure intensity variations
- Arm lengths offset

iLIGO Length Sensing and Control
(RF Readout)



JEL

Gravitational-Wave Data

Data = Instantaneous estimate of strain for each moment in time

i.e. demodulated channel sensitive to arm length difference

That's not the whole story – we'll come back to calibration later

Digitized discrete time series recorded in computer files

$$(t_j , x_j)$$

LIGO and GEO sampling rate: 16384 Hz ≡ f_s

VIRGO sampling rate: 20000 Hz

Synchronized with GPS time signal

Common "frame" file format (*.gwf)

***Many* auxiliary channels recorded too**

Total data volume: a few megabytes per second per interferometer

Relevance of the Sampling Rate

Is 16384 Hz a high enough sampling rate ?

The Sampling Theorem:

Discretely sampled data with sampling rate f_s can completely represent a continuous signal which only has frequency content below the Nyquist frequency, $f_s / 2$

GW signals of interest to ground-based detectors typically stay below a few kHz

e.g. binary neutron star inspiral reaches ISCO at ~1 to 1.5 kHz

Neutron star f-modes: ~3 kHz

Black hole quasinormal modes: ~1 kHz for 10 $M_\odot$

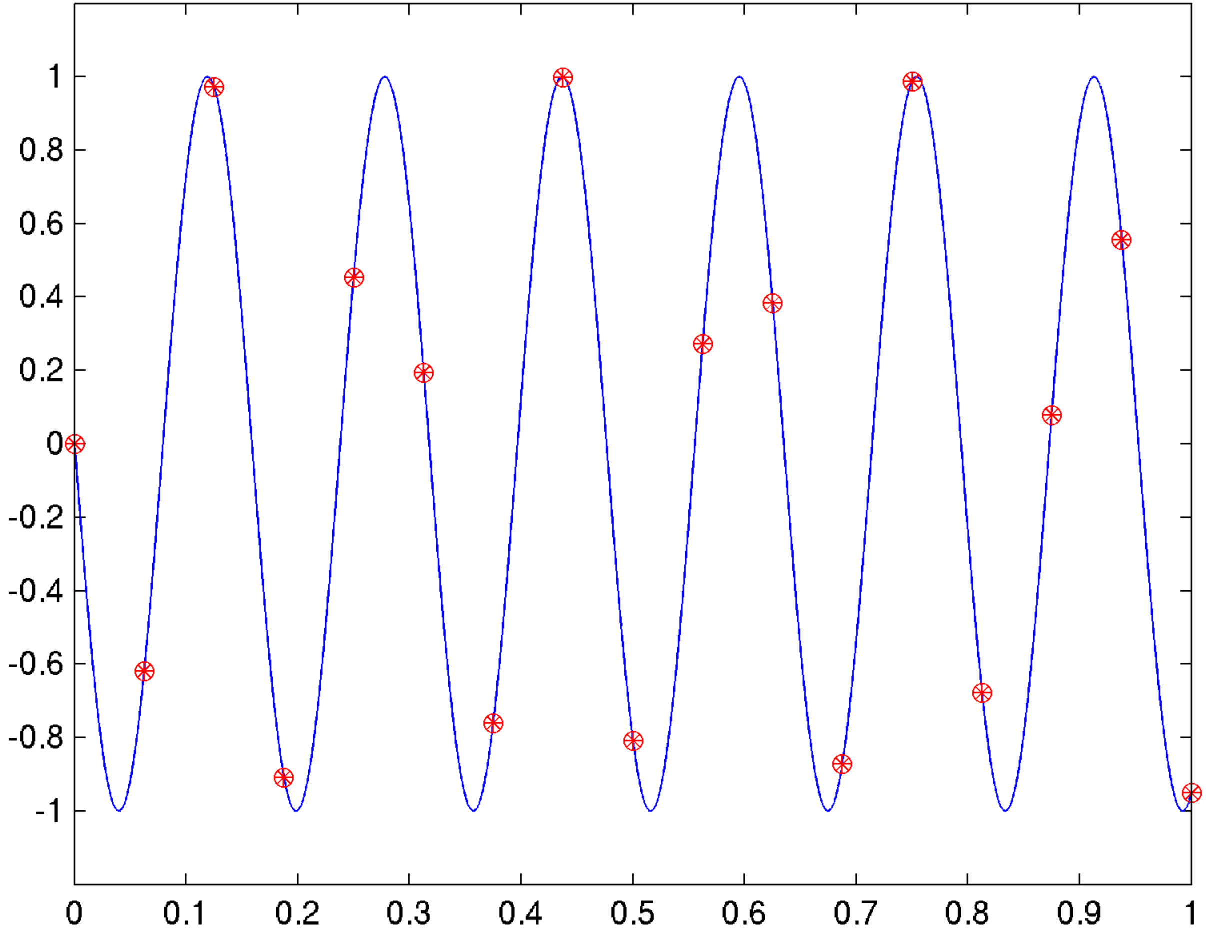
Some core collapse supernova signals could go up to several kHz

What if the signal extends above Nyquist frequency?

Higher frequencies are "aliased" down to lower frequencies

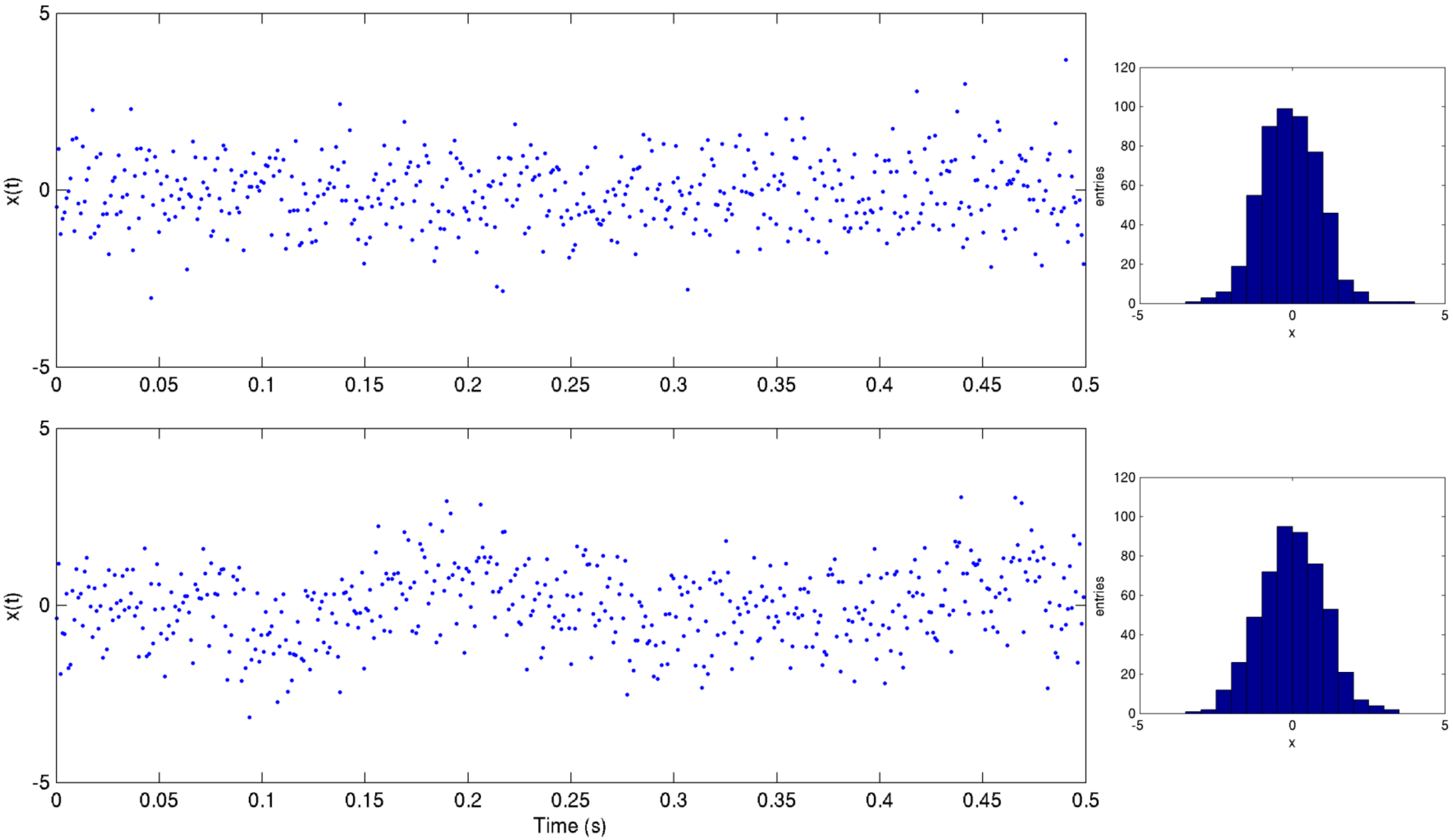
Aliasing

f_s = 16 Hz ; signal frequency = 9.7 Hz



Characterizing Noise

Noise is random, but its *properties* can be characterized



Possible Properties of Noise

Stationary : statistical properties are independent of time

Ergodic process: time averages are equivalent to ensemble averages

Gaussian : A random variable follows Gaussian distribution

For a single random variable, $$p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2}\frac{(x-\mu_x)^2}{\sigma_x^2}\right]$$

More generally, a *set* of random variables (e.g. a time series) is Gaussian if the joint probability distribution is governed by a covariance matrix

$$C_{xij} := \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle$$

such that

$$p(x_1, x_2, \ldots, x_N) = \frac{1}{(2\pi)^{N/2}\sqrt{\det C_x}} \exp\left[-\frac{1}{2}\sum_{i,j=0}^{N-1} C_{xij}^{-1}(x_i-\mu_{xi})(x_j-\mu_{xj})\right]$$

White : Signal power is uniformly distributed over frequency

$\Rightarrow$ Data samples are uncorrelated

Frequency-Domain Representation of a Time Series

Fourier transform

$$\widetilde{x}(f) = \int_{-\infty}^{\infty} dt\, x(t) e^{-i2\pi f t}$$

$$\Rightarrow \qquad x(t) = \int_{-\infty}^{\infty} df\, \widetilde{x}(f) e^{i2\pi f t}$$

A linear function, complex in general

Defined for all positive *and negative* frequencies

Frequency-Domain Representation of a *Discrete, Finite* Time Series

Time series x_j with N samples at times $t_j = t_0 + j\,\Delta t$

Discrete Fourier transform

$$\widetilde{x}_k := \sum_{j=0}^{N-1} x_j\, e^{-i2\pi jk/N}$$

$$\Rightarrow \qquad x_j = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \widetilde{x}_k\, e^{i2\pi jk/N}$$

Frequency spacing is inversely proportional to N

Efficient way to calculate complete discrete Fourier Transform: Fast Fourier Transform (FFT)

Power Spectral Density

Parseval's theorem:

$$\int_{-\infty}^{\infty} dt\, |x(t)|^2 = \int_{-\infty}^{\infty} df\, |\widetilde{x}(f)|^2$$

$\Rightarrow$ Total energy in the data can be calculated in either time domain or frequency domain

$|\widetilde{x}(f)|^2$ can be interpreted as *energy spectral density*

When noise (or signal) has infinite extent in time domain, can still define the power spectral density (PSD)

$$\lim_{T\to\infty} \frac{1}{T} \left|\widetilde{x}_T(f)\right|^2$$

Watch out for one-sided vs. two-sided PSDs

Estimating the PSD

Generally we need to determine the PSD empirically, using a finite amount of data

Simplest approach: FFT the data, calculate square of magnitude of each frequency component – this is a periodogram

For stationary noise, one can show that the frequency components are statistically independent

This estimate is unbiased , but has a large variance – equal to the square of the value itself!

Get better estimate by averaging several periodograms

Alternately, smooth periodogram; give up frequency resolution either way

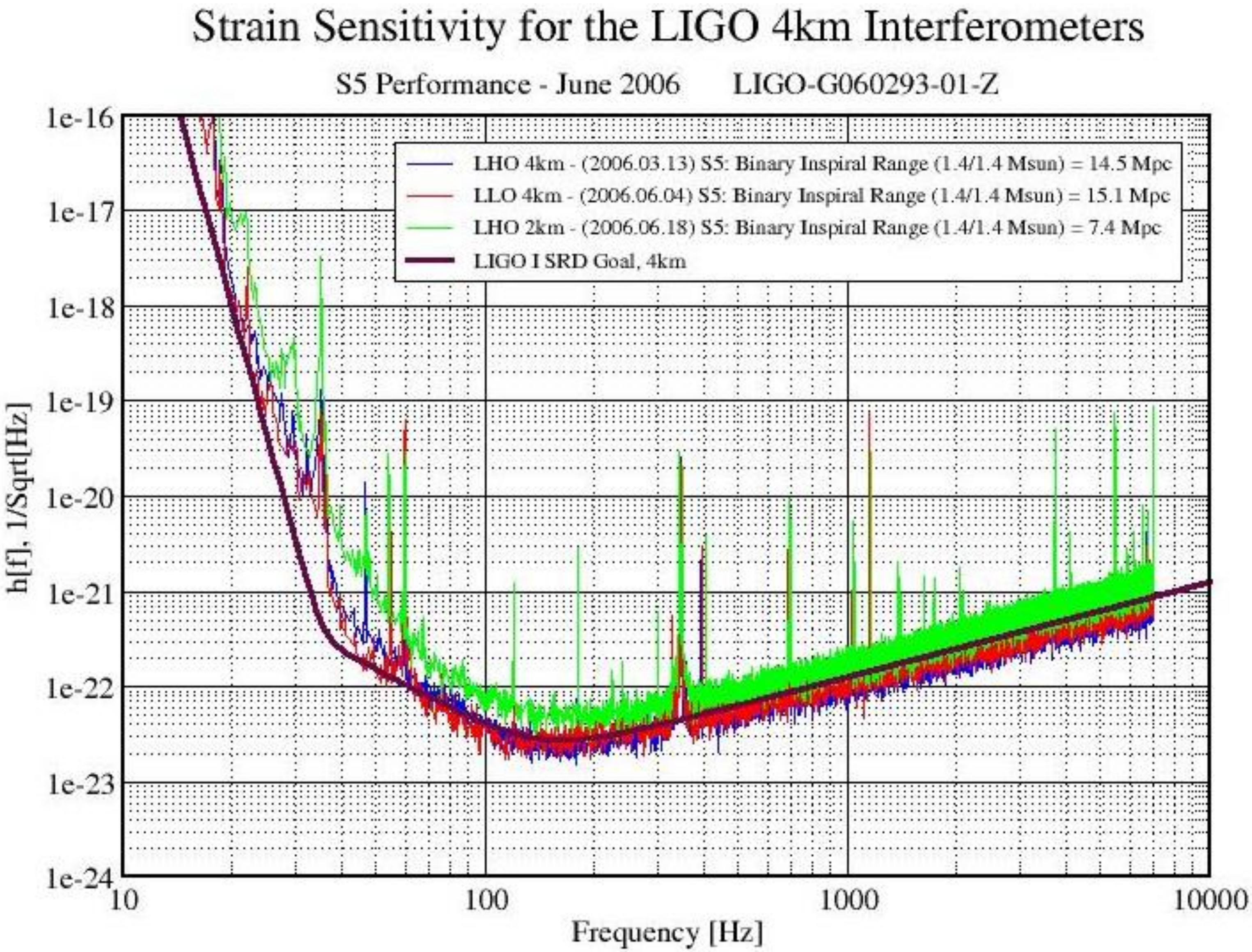
Can apply a "window" to the data to avoid spectral leakage

Leakage arises from the assumption that the data is periodic!

Tapered window forces data to go to zero at ends of time interval

Welch's method of estimating a PSD averages periodograms calculated from windowed data

Amplitude Spectral Density of LIGO Noise



Interpretation of Time Series Data

Recorded data values are *not* simply proportional to GW strain

A linear system, but that does not guarantee proportionality !

Frequency-dependent amplitude and phase relation (i.e. transfer function)

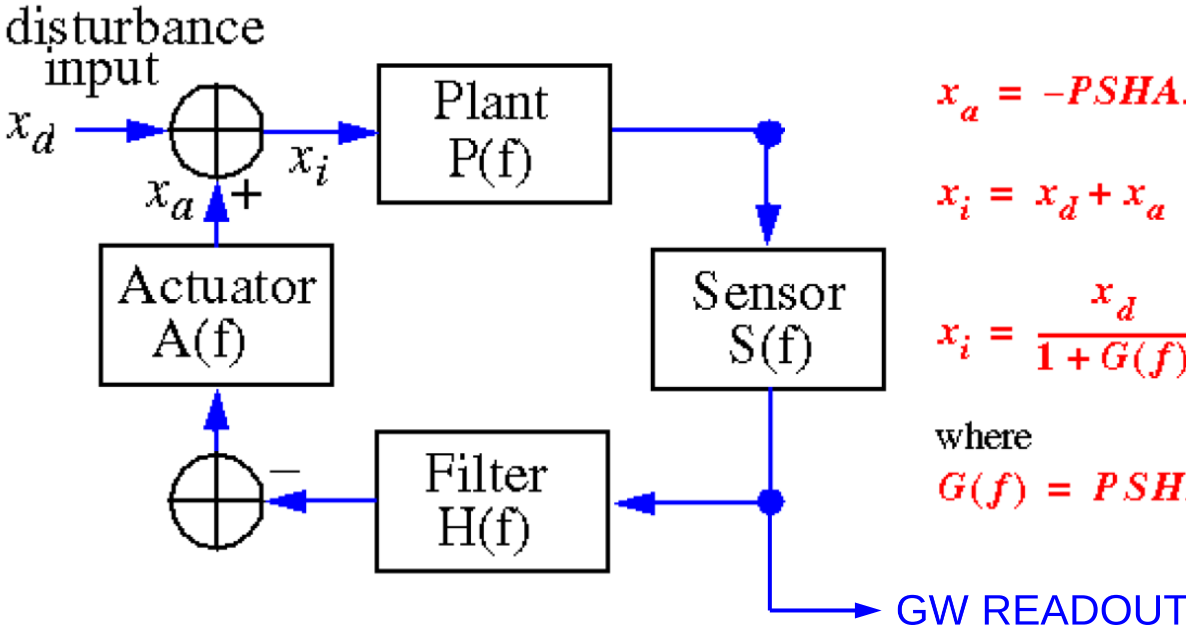
Instrumental and practical reasons

$\Rightarrow$ **Raw time series is a distorted version of GW strain signal**

e.g. a delta-function GW signal produces an output with a characteristic shape and duration ("impulse response")

Want to recover actual GW strain for analysis

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

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

Basics of Digital Filtering

A filter calculates an output time series from a linear combination of the elements of an input time series

Finite Impulse Response (FIR) filter

Calculated *only* from the input time series

Typical form: $y_i = b_0 x_i + b_1 x_{i-1} + b_2 x_{i-2} + \ldots + b_{N-1} x_{i-N}$

Infinite Impulse Response (IIR) filter

Also uses prior elements of the output time series

e.g. $y_i = b_0 x_i + b_1 x_{i-1} + b_2 x_{i-2} + \ldots + b_{N-1} x_{i-N} + a_1 y_{i-1} + a_2 y_{i-2} + \ldots$

Choice of coefficients determines transfer function

Many filter design methods, depending on goals

Applications of filtering

High-pass, low-pass, band-pass, band-stop, etc.

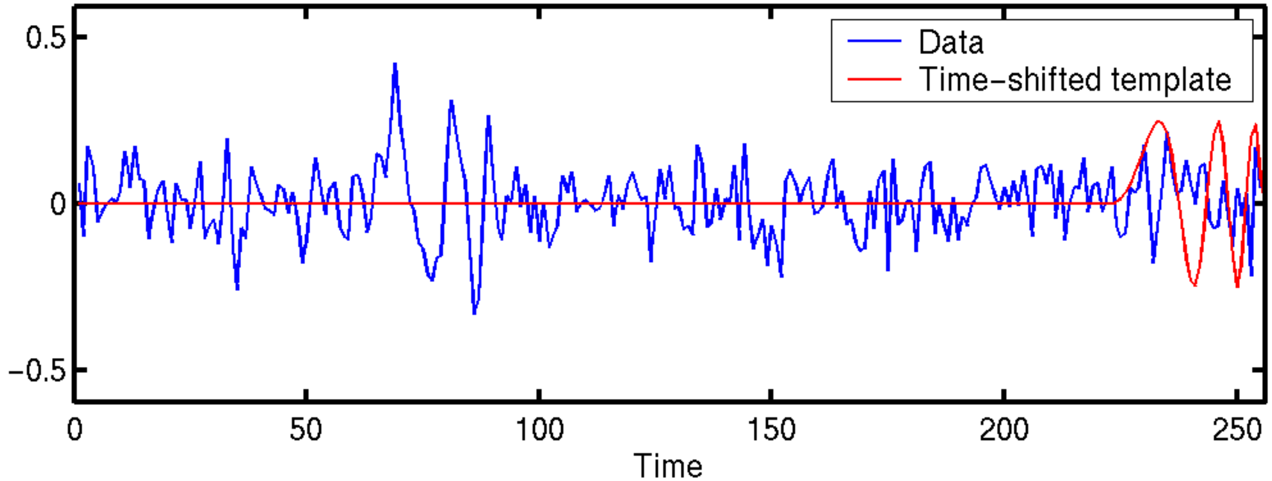
Anti-aliasing for down-sampling

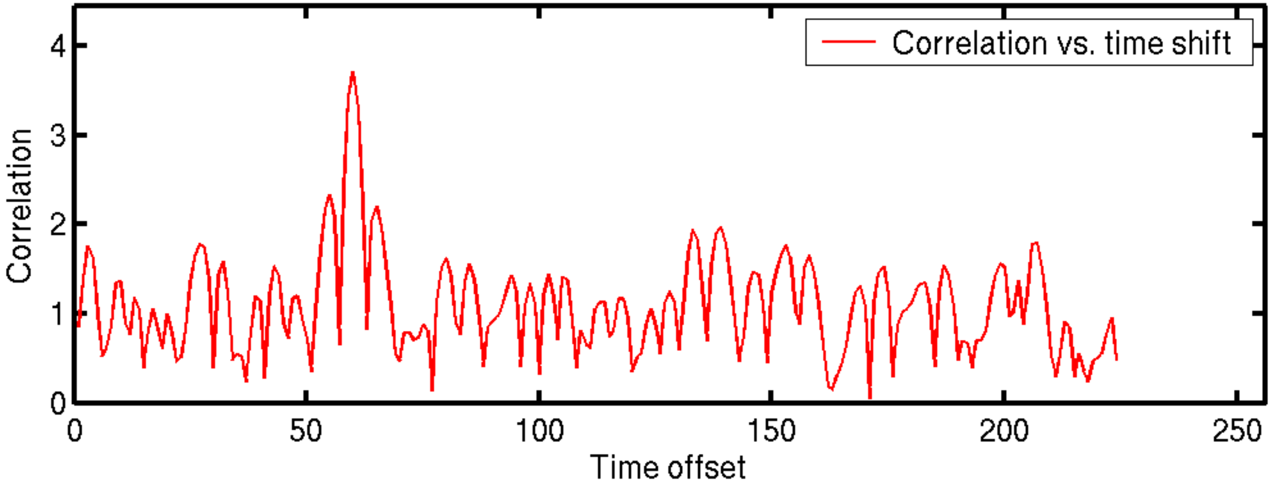
Low-pass filter to cut away signal content above new Nyquist frequency

Whitening / Dewhitening

Basic Illustration of Matched Filtering

<u>Correlate</u> data with expected signal (Here, plotting absolute value)





Other Ways to Think About Matched Filtering (1)

Correlating data with template can be thought of as taking an *inner product*

$$C = \langle s \mid h \rangle = \int_{-\infty}^{\infty} dt' \, s(t') \, h(t') \quad or \quad \sum_j s_j h_j$$

Data

Template

For GW data analysis, there is usually a free param: time offset

$$C(t) = \int_{-\infty}^{\infty} dt' \, s(t') \, h(t'-t) \qquad or \qquad C_i = \sum_j s_j h_{j-i}$$

Time offset

Data

Template with time offset

Yields a time series of correlation values

Other Ways to Think About Matched Filtering (2)

Correlating data with template is equivalent to an *FIR filter* with coefficients following the template

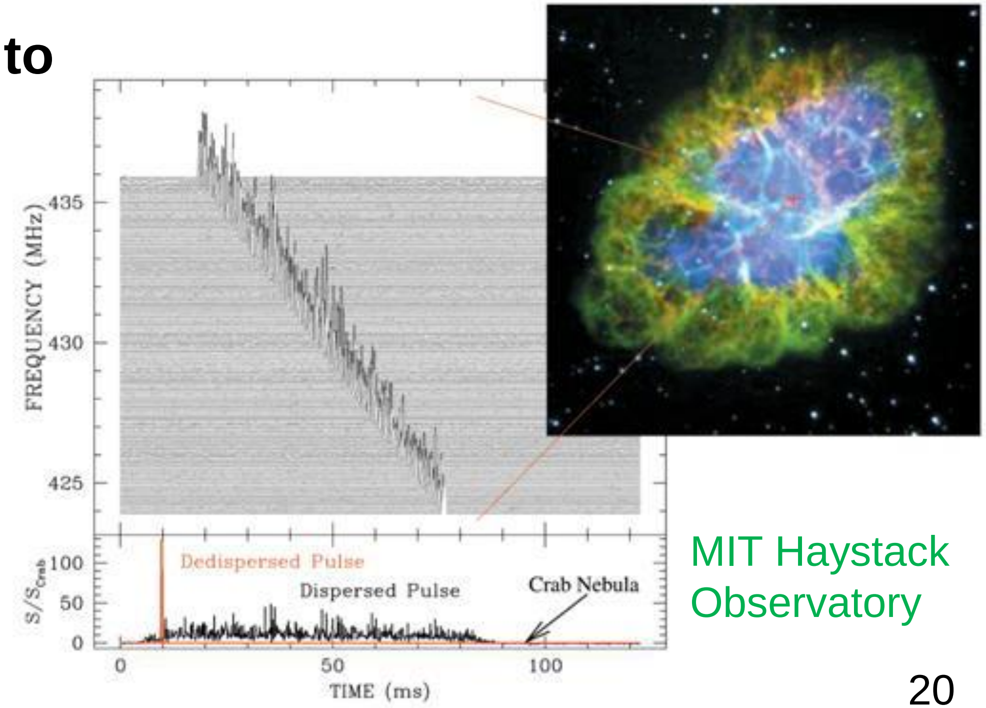
$$y_n = b_0 x_n + b_1 x_{n-1} + b_2 x_{n-2} + \cdots + b_N x_{n-N}$$

The impulse response of that FIR filter looks like the template, but *time-reversed*

The goal of this kind of filter is to "compress" an extended signal into a delta function

Shift all parts of the signal in the data to a common time, and add them together with the same sign

Similar to "de-dispersion" in radio telescope pulse search

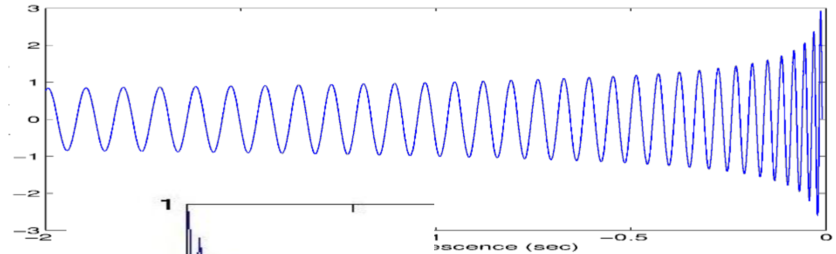


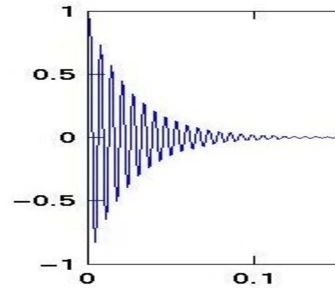
MIT Haystack Observatory

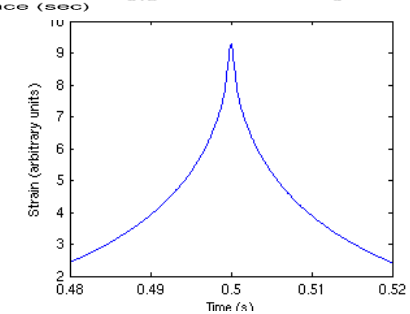
Usefulness

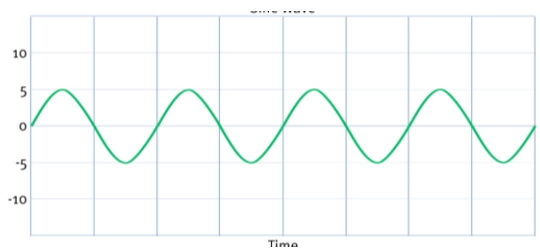
Useful for GW data analysis whenever target waveform is known

- Binary coalescence
- Ringdown
- Cosmic string cusp
- Continuous-wave signal









Phase coherence is more important than amplitude matching

Also known as "Wiener optimal filter"

- Optimal detection statistic if noise is Gaussian

Books:
Creighton & Anderson,
"GW Physics and Astronomy";
Wainstein and Zubakov,
"Extraction of Signals from Noise"

Matched Filtering in Frequency Domain

$$C(t) = \int_{-\infty}^{\infty} dt' \, s(t') \, h(t' - t)$$

Time offset

Data

Template with time offset

Rewrite correlation integral using Fourier transforms…

$$\Rightarrow \quad C(t) = 4 \int_{0}^{\infty} \tilde{s}(f) \, \tilde{h}^{*}(f) \, e^{2\pi i f t} \, df$$

This is simply the inverse FFT of $\tilde{s}(f) \, \tilde{h}^{*}(f)$ **!**

(Correlation in time domain is product, with complex conj, in freq domain)

Computationally efficient way to calculate filter output for <u>all</u> time offsets!

Optimal Matched Filtering with Frequency Weighting

FFT of data

Template can maybe be generated in frequency domain using stationary phase approximation

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

Noise power spectral density

This de-weights frequencies with larger noise power

Or, can apply a time-domain filter which is essentially the inverse Fourier transform of $\tilde{h}(f)/S_n(f)$

Look for maximum of $|C(t)|$ above some threshold ➔ trigger

See, for instance, Allen et al., PRD 85, 122006 (2012)

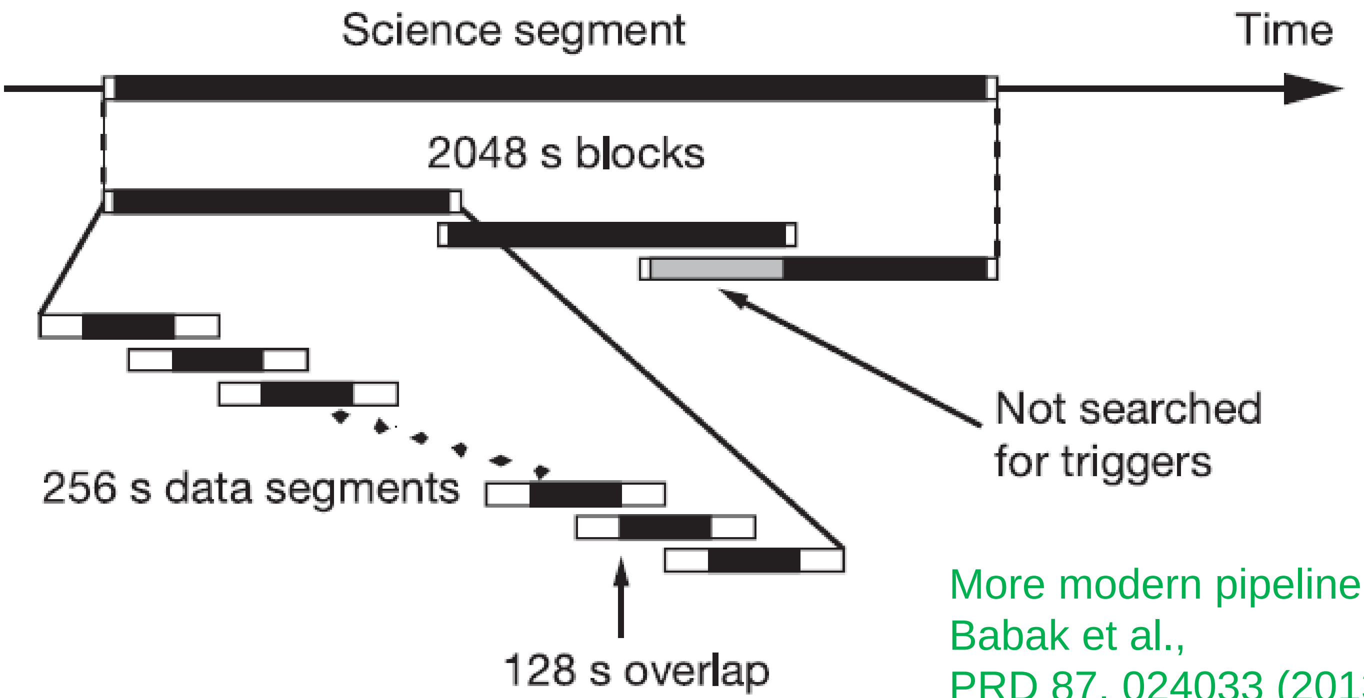
Searching a Full Data Set

Search *overlapping* intervals to avoid wrap-around effects when filtering in frequency domain

Do inverse FFTs on, say, 256 s of data at a time

Estimate power spectrum from longer stretches of data, e.g. using median

Original iLIGO scheme: D. A. Brown for the LSC, CQG 22, S1097 (2005)



More modern pipeline:
Babak et al.,
PRD 87, 024033 (2013)

Binary Coalescence Source Parameters vs. Signal Parameters

Inspiral source parameters

Masses (m1, m2)

Spins → Negligible for neutron stars

Orbital phase at coalescence → Maximize analytically when filtering

Inclination of orbital plane

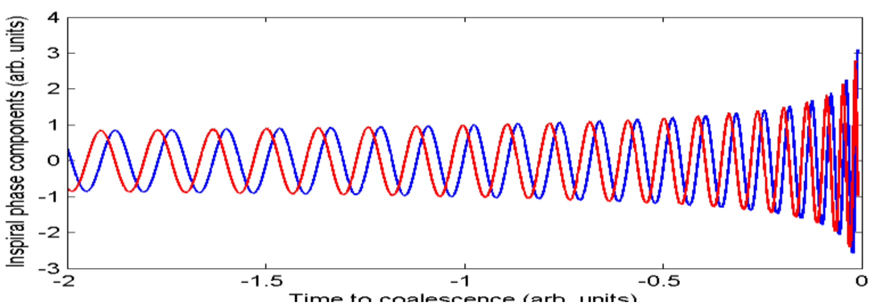
Sky location

→ Simply multiplicative for a given detector (long-wavelength limit)

Distance → Simply multiplicative

Coalescence time

Filter with orthogonal templates, take quadrature sum



➔ Only have to explicitly search over masses and coalescence time (“intrinsic parameters”)

Template Matching

Want to be able to detect any signal in a *space* of possible signals

All with different phase evolution

… but do it with a finite set of templates!   i.e., a "template bank"

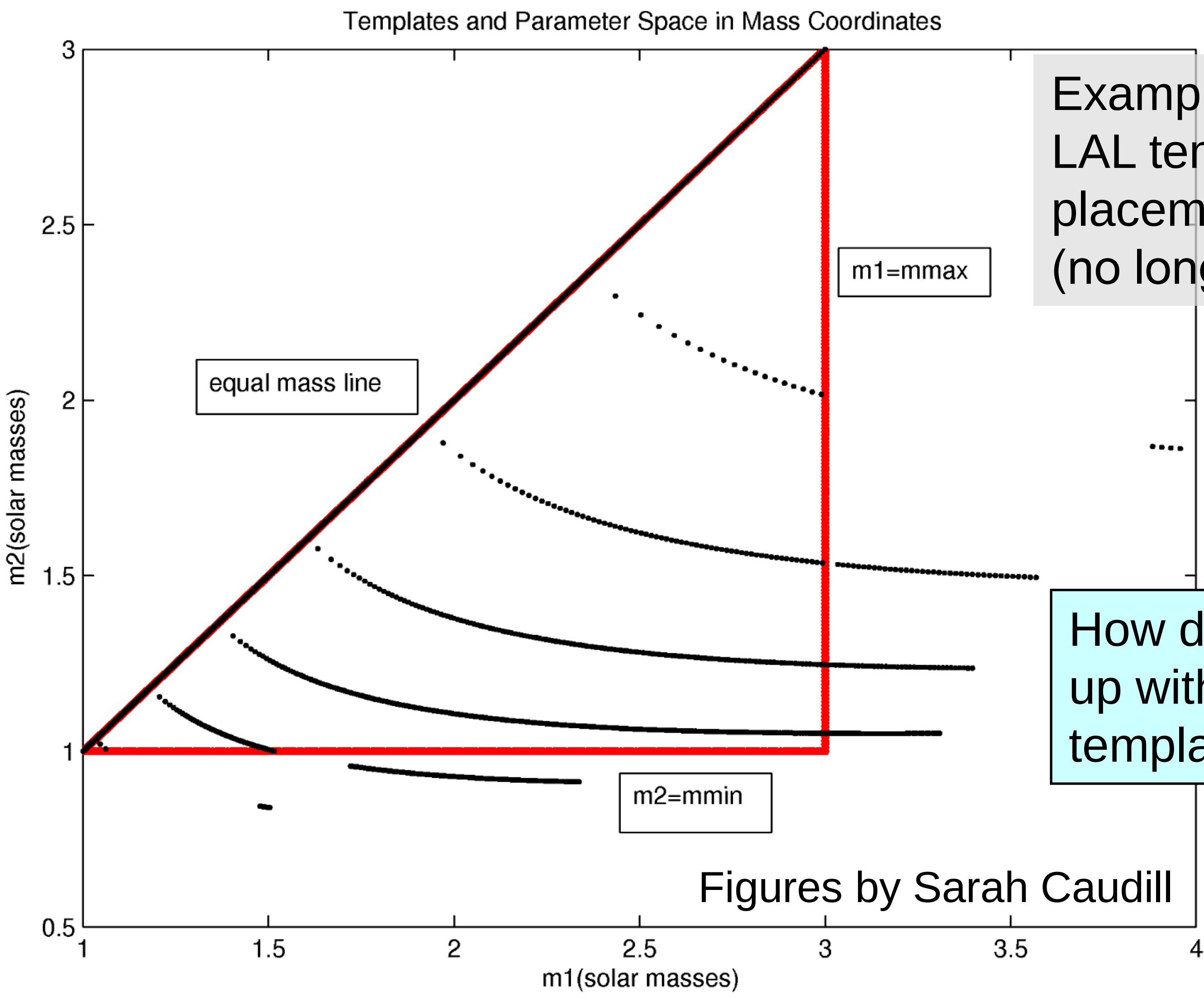
So make sure there is a "close enough" template for every part of the signal space

Require a minimum overlap between signal and template, e.g. 0.97

Often can calculate a "metric" which parametrizes the mismatch for small mismatches

See, for instance, Sathyaprakash and Dhurandhar, PRD 44, 3819 (1991);
Balasubramanian, Sathyaprakash and Dhurandhar, PRD 53, 3033;
Owen and Sathyaprakash, PRD 60, 022002 (1999)

Template Bank Construction

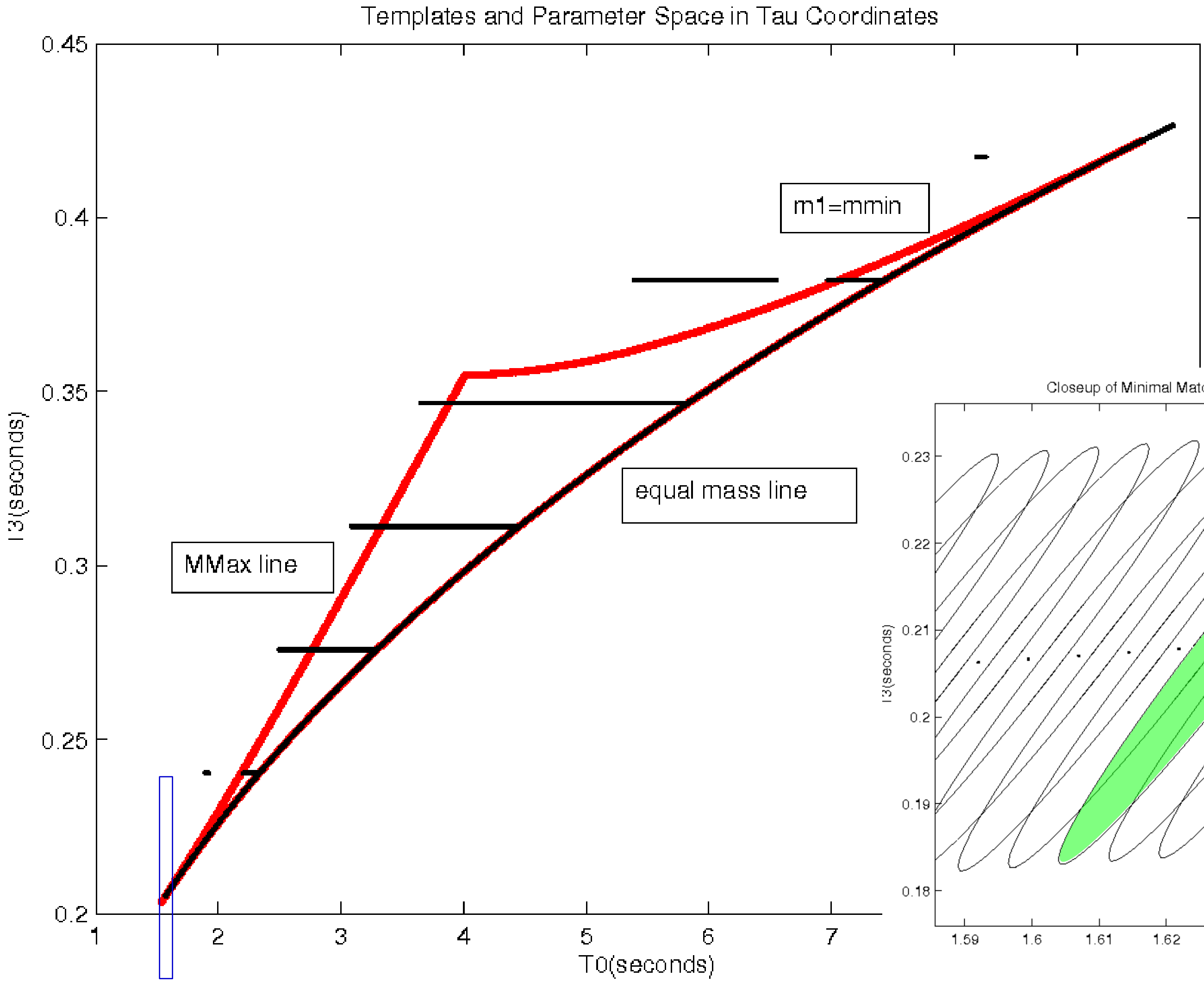


Example from early LAL template bank placement algorithm (no longer in use…)

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

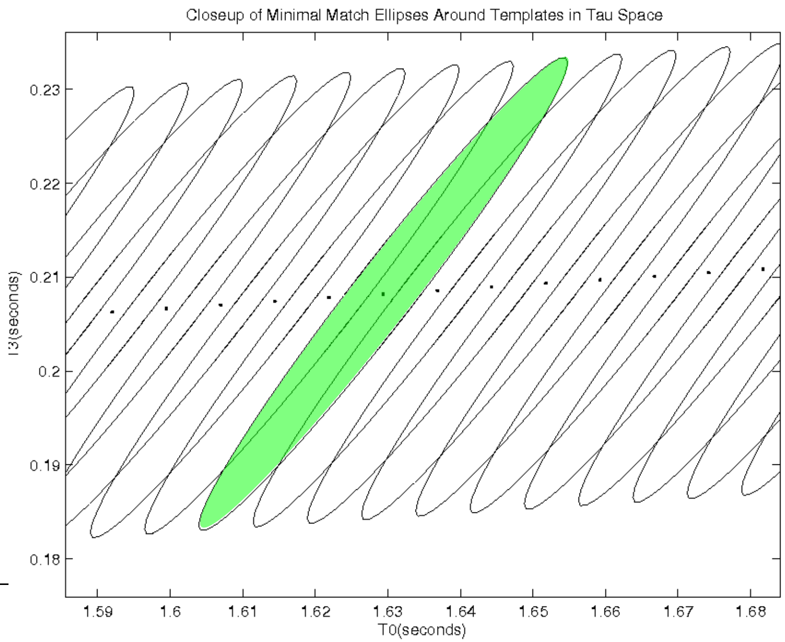
Figures by Sarah Caudill

Template Bank Construction in (τ_0,τ_3) space

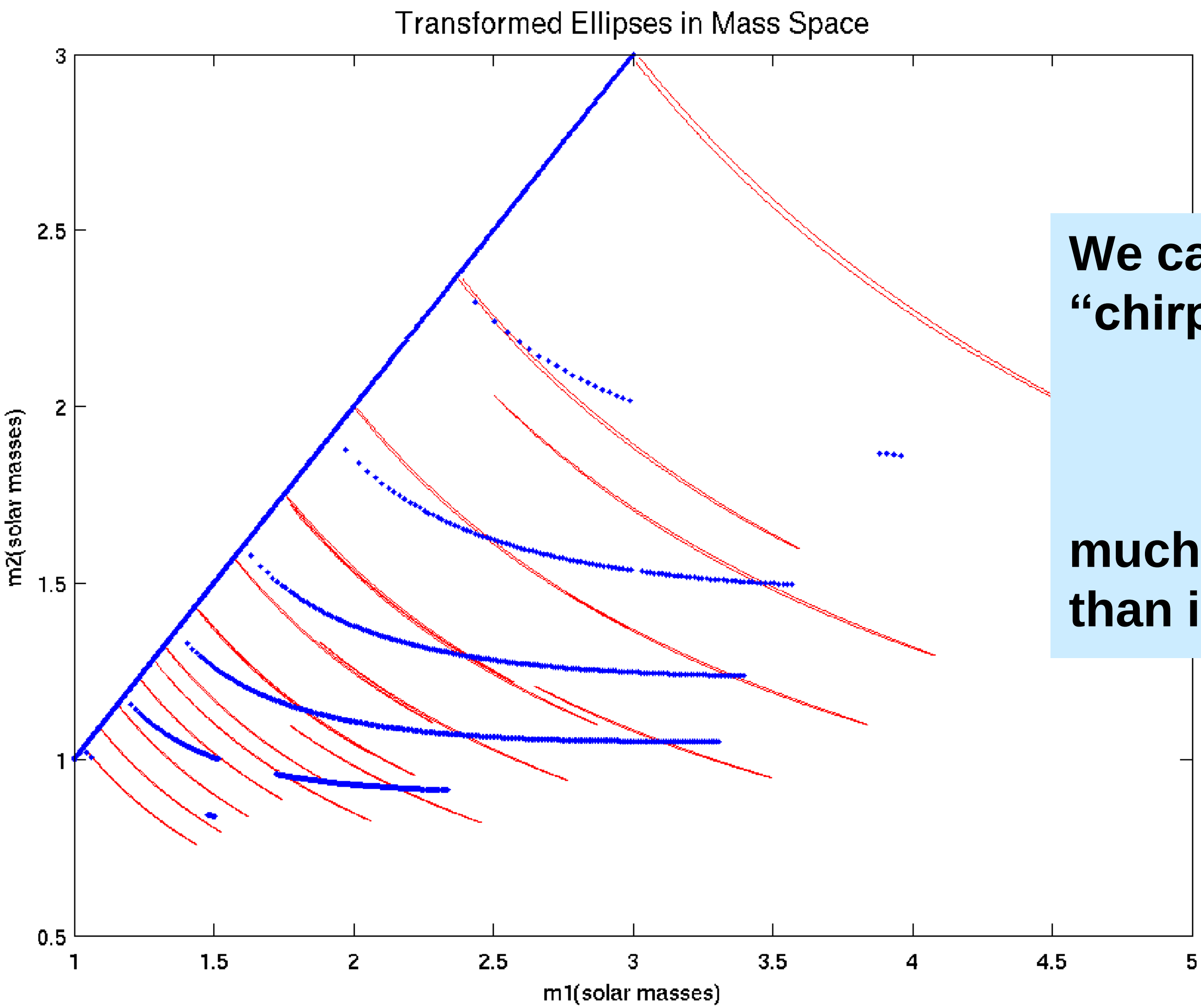


$$\tau_0 \propto m^{-5/3}/\eta$$

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



Ellipses in Mass Space



We can determine the "chirp mass",

$$\frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$

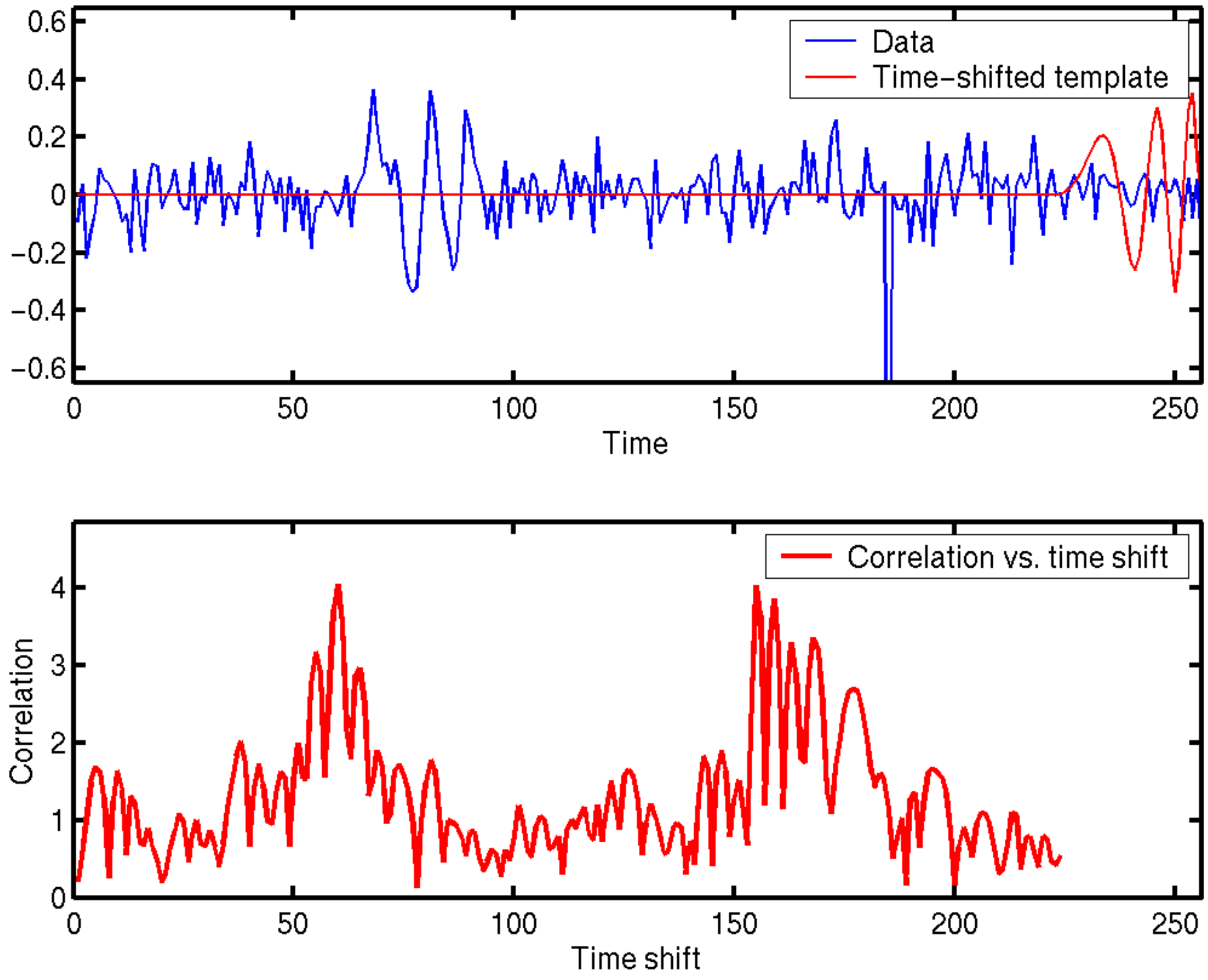
much more precisely than individual masses

Peters and Mathews, Phys. Rev. 131, 435 (1963)

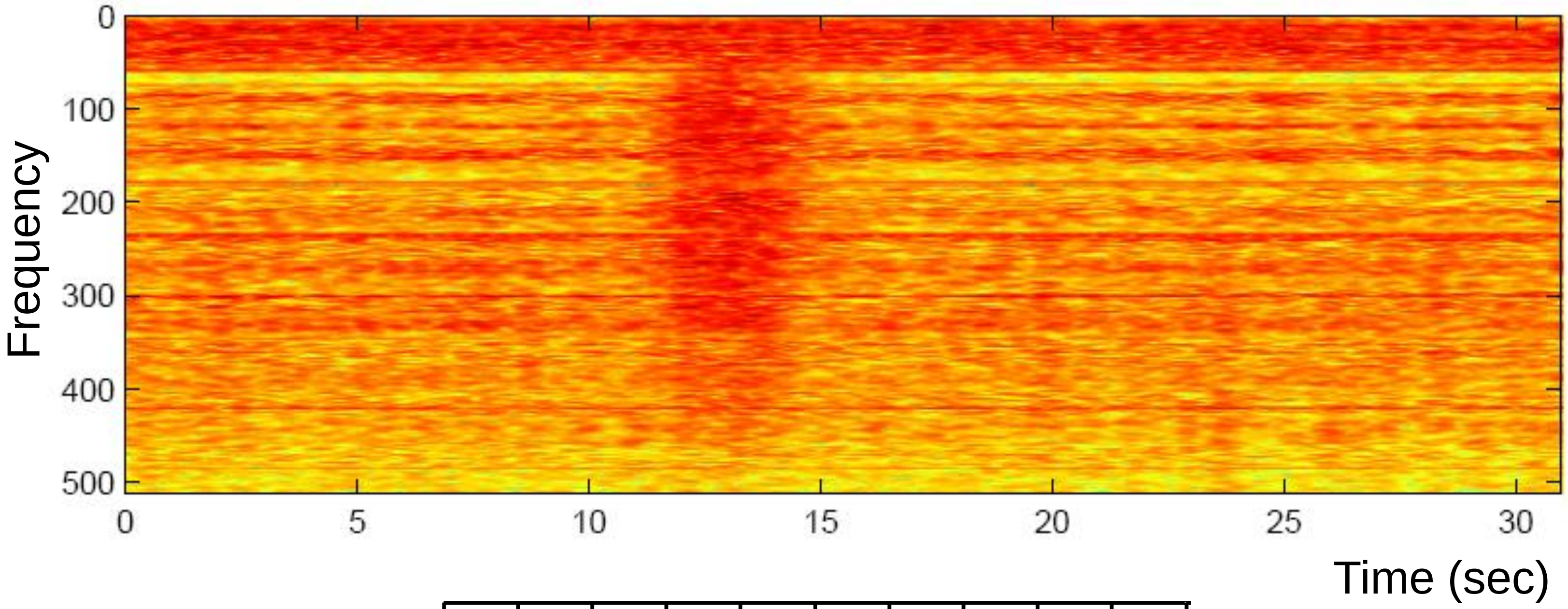
How Sensitive is Matched Filtering?

- **An extended signal, searched for with matched filtering, can be detected even if the instantaneous S/N amplitude ratio is small**
  - Signal adds coherently
  - Noise adds incoherently

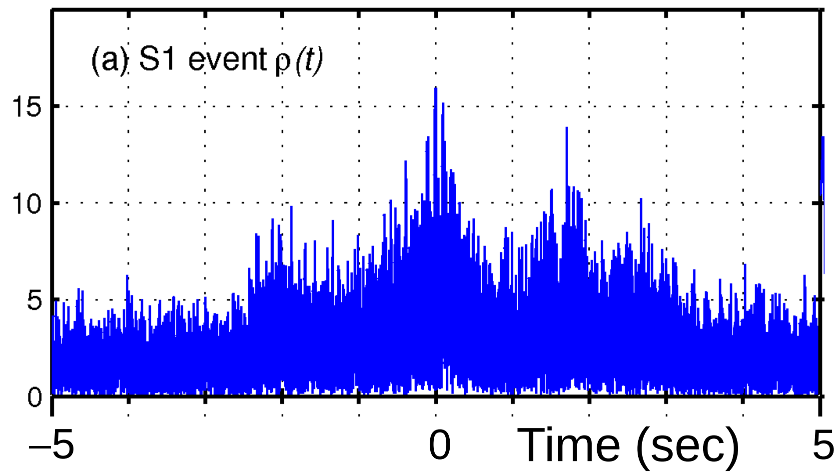
Matched Filtering is *Sensitive*, But Not *Selective*



Dealing with Non-Stationary Noise



Inspiral filter output:



Shawhan and Ochsner, CQG 21, S1757 (2004)

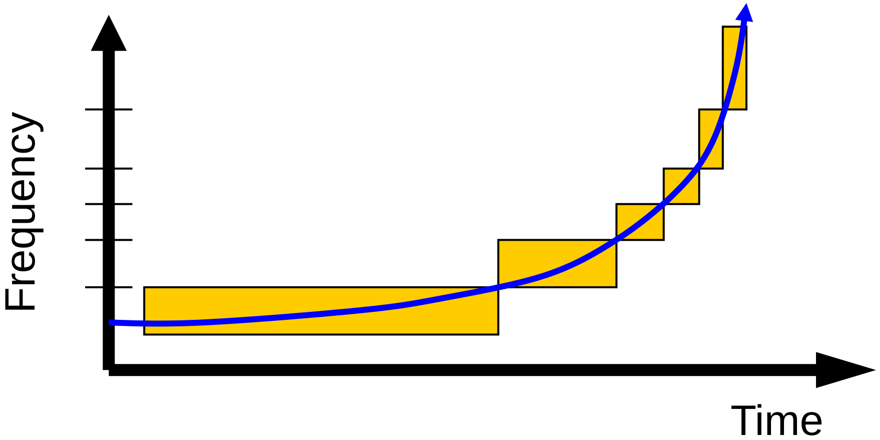
Waveform Consistency Tests

Chi-squared test

Allen, PRD 71, 062001 (2005)

Divide template into p parts, calculate

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



Can use χ^2 with ρ to form some sort of "effective SNR", e.g.:

$$\rho_{\text{eff}}^2 = \frac{\rho^2}{\sqrt{\left(\frac{\chi^2}{2p-2}\right)\left(1+\frac{\rho^2}{250}\right)}}$$

$$\rho_{\text{new}} = \begin{cases} \rho, & \chi^2 \le n_{\text{dof}} \\ \dfrac{\rho}{\left[\left(1+\frac{\chi^2}{n_{dof}}{}^{4/3}\right)/2\right]^{1/4}}, & \chi^2 > n_{\text{dof}} \end{cases}$$

$$\hat{\rho} = \begin{cases} \dfrac{\rho}{[(1+(\chi_r^2)^3)/2]^{1/6}} & \text{for } \chi_r^2 > 1, \\ \rho & \text{for } \chi_r^2 \le 1. \end{cases}$$

Empirical – to separate signals from background as cleanly as possible