# *Data Analysis and the Detection of Gravitational Waves with LIGO*

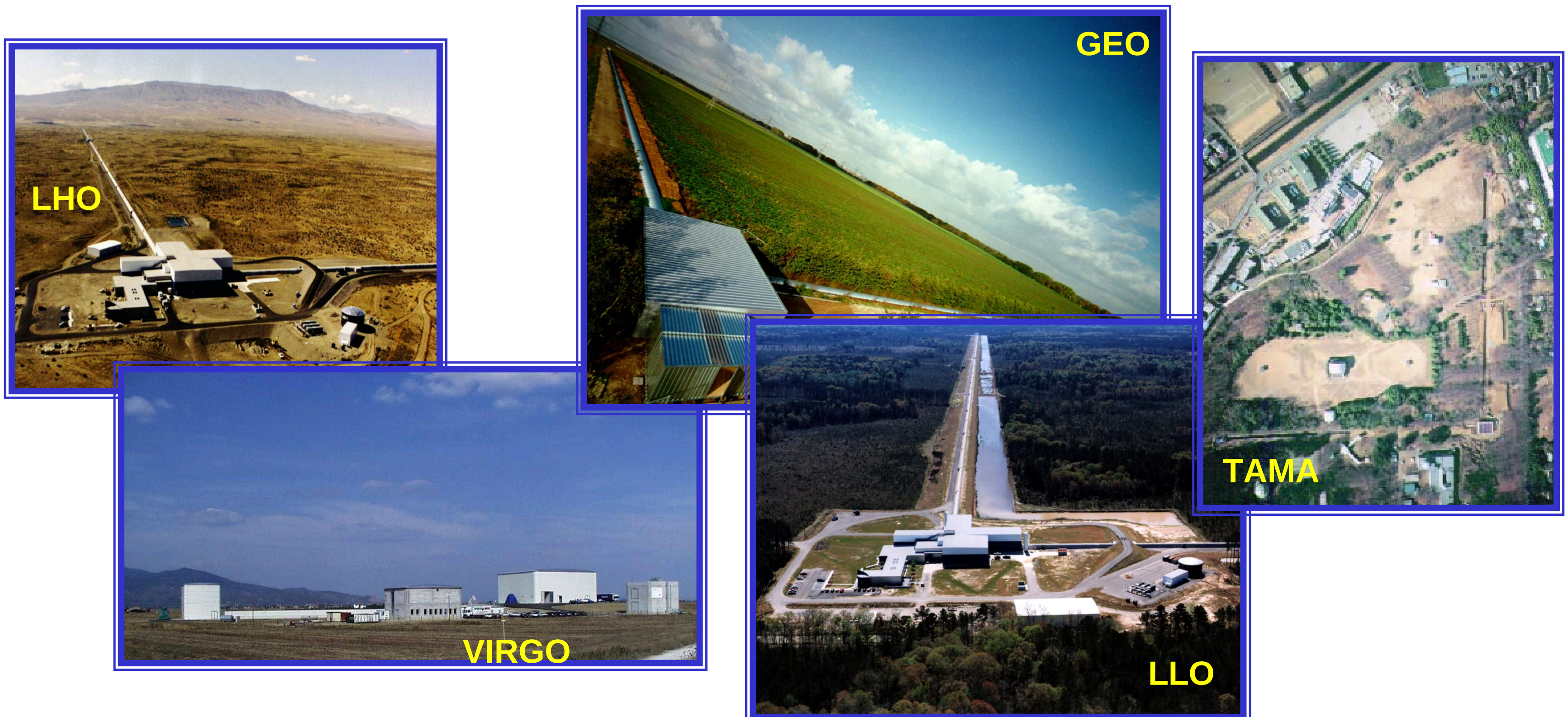
## LIGO Summer SURF Seminar Series


*Albert Lazzarini*
*LIGO Laboratory*
*California Institute of Technology*
*Pasadena, California 91125*


*LIGO-G010259-00-E*

# *The Opening of a New Observational Window on the Universe*

- LIGO, VIRGO, GEO, TAMA …
  - *4000m, 3000m, 2000m, 600m, 300m interferometers built to detect gravitational waves from compact objects*



# *Ultimate Goals for the Observation of Gravitational Waves*

- Tests of Relativity
  - *Black holes & strong-field gravity (ringdown of excited BH)*
  - *Spin character of the radiation field (polarization of radiation from CW sources)*
  - *Wave propagation speed (delays in arrival time of bursts)*
- Gravitational Wave Astronomy
  - *Compact binary inspirals*
  - *Gravitational waves and gamma ray burst associations*
  - *Black hole formation*
  - *Supernovae in our galaxy*
  - *Newly formed neutron stars -  spin down in the first year*
  - *Pulsars and rapidly rotating neutron stars*
  - *LMXBs*
  - *Stochastic background*

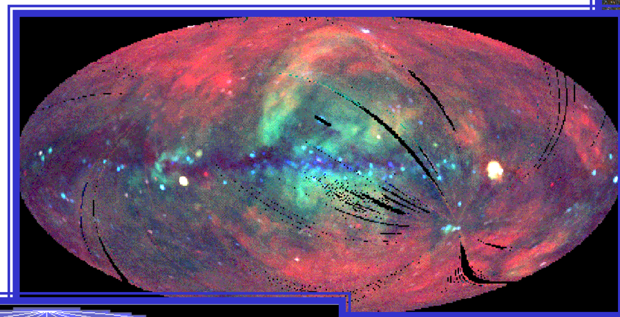
# *Observing the Galaxy with Different Electromagnetic Wavelengths*

*http://antwrp.gsfc.nasa.gov/apod/image/SagSumMW_dp_big.gif*

***λ = 5 x 10 $^{-7}$ m***

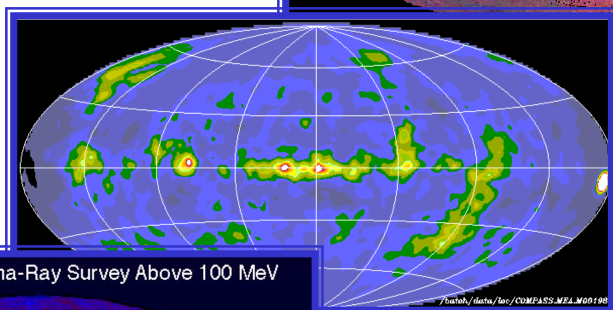
*http://antwrp.gsfc.nasa.gov/apod/image/xallsky_rosat_big.gif*

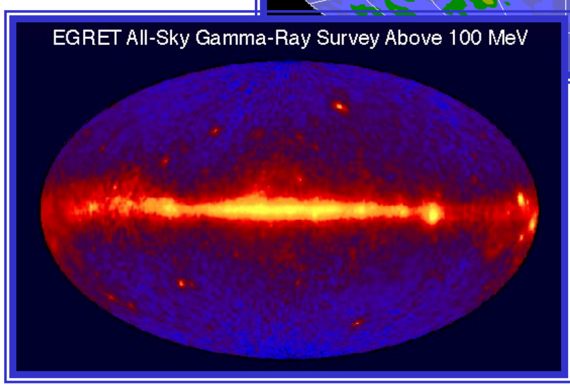
***λ = 5 x 10 $^{-10}$ m***



*http://antwrp.gsfc.nasa.gov/apod/image/comptel_allsky_1to3_big.gif*

***λ = 6 x 10 $^{-13}$ m***



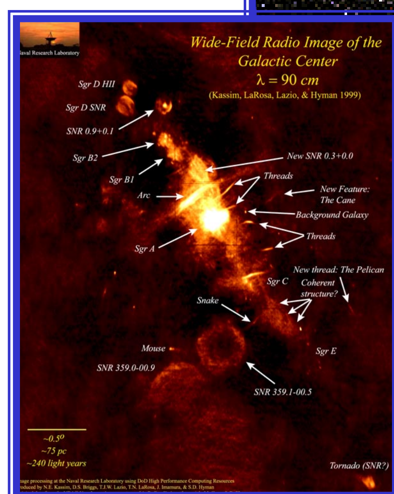



*http://cossc.gsfc.nasa.gov/cossc/egret/*

***λ = 1 x 10 $^{-14}$ m***


*http://www.gsfc.nasa.gov/astro/cobe*

***λ = 2 x 10 $^{-6}$ m***

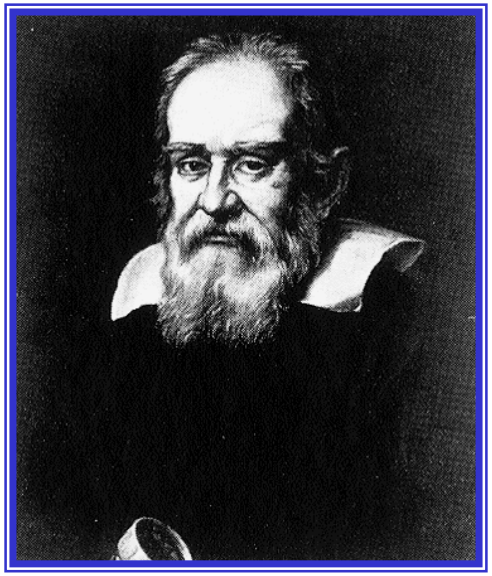


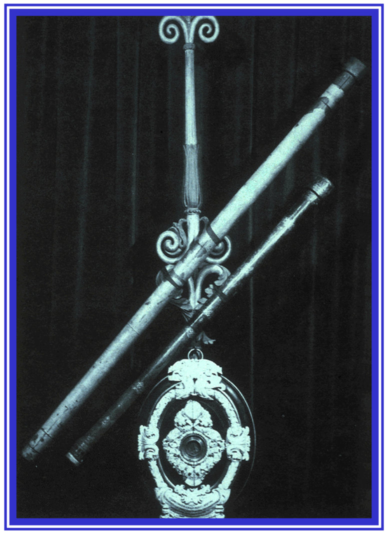

*//rsd-www.nrl.navy.mil/7213/lazio/GC/*

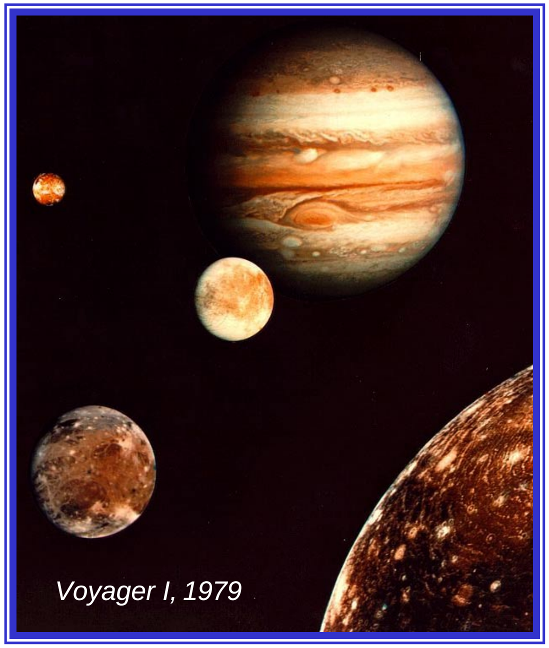
***λ = 9 x 10 $^{-1}$ m***

# *The Opening of a New Observational Window on the Universe*

- Galileo Galilei, 1610
  - *Improves on an invention by Hans Lipperhey to build a 9X telescope*
    - ***Discovers the "Gallilean" moons of Jupiter***





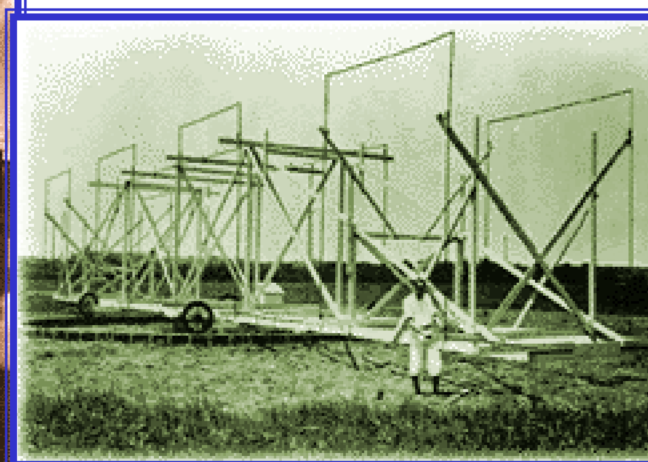


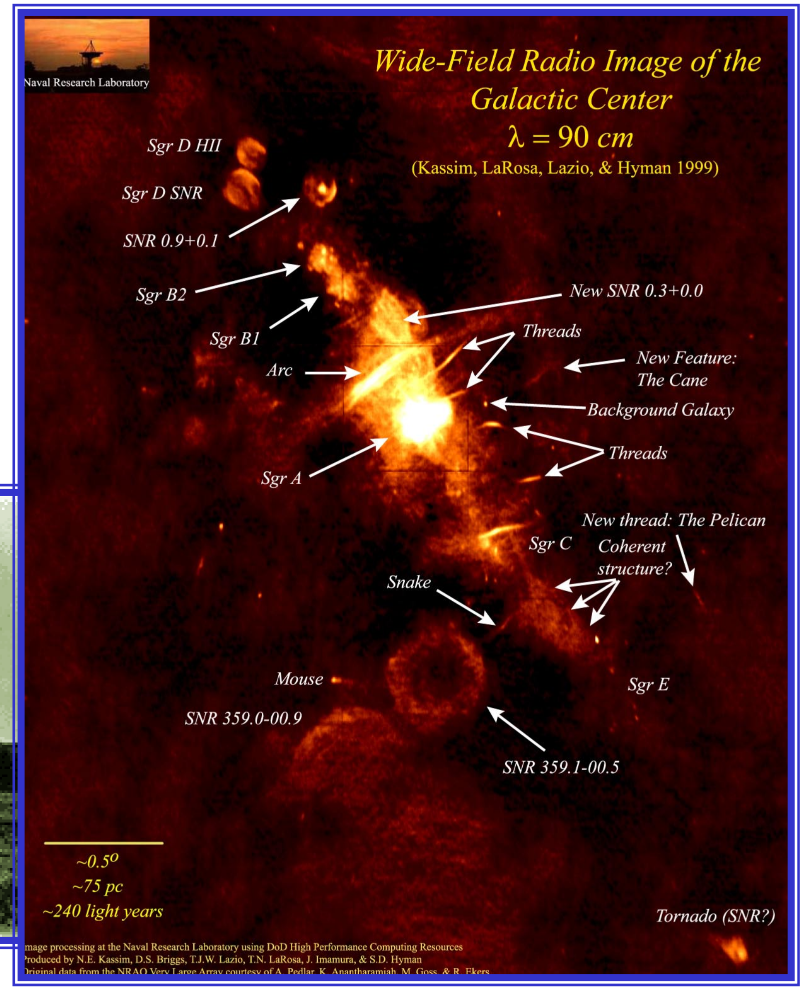

*http://es.rice.edu:80/ES/humsoc/Galileo//*

*http://photojournal.jpl.nasa.gov*

# *The Opening of a New Observational Window on the Universe*

- Karl Janksy, 1933
  - *Builds a radio antenna array to study interference in transatlantic telecommunications*
    - ! ***Discovers radio emissions from the galactic center***



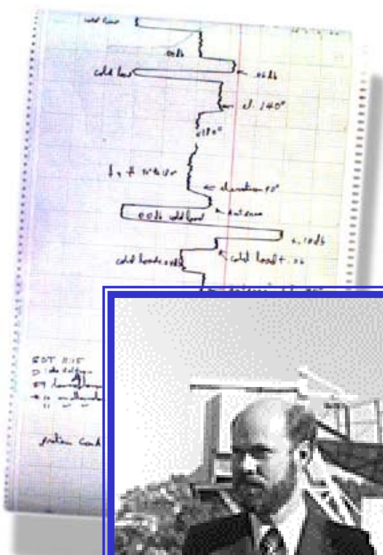



*http://www.lucent.com/museum/1933rt.html*

*//http://rsd-www.nrl.navy.mil/7213/lazio/GC/*

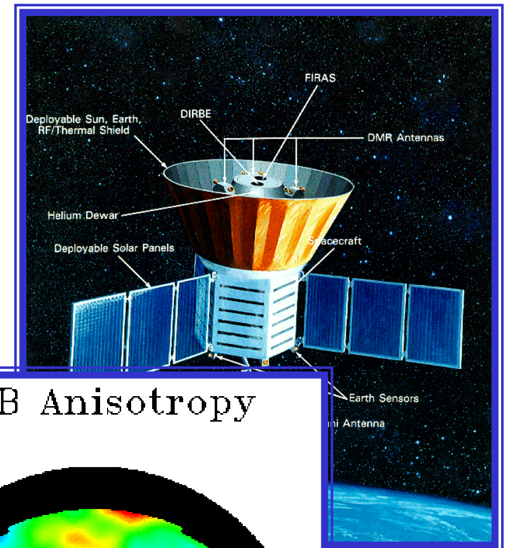
# *The Opening of a New Observational Window on the Universe*

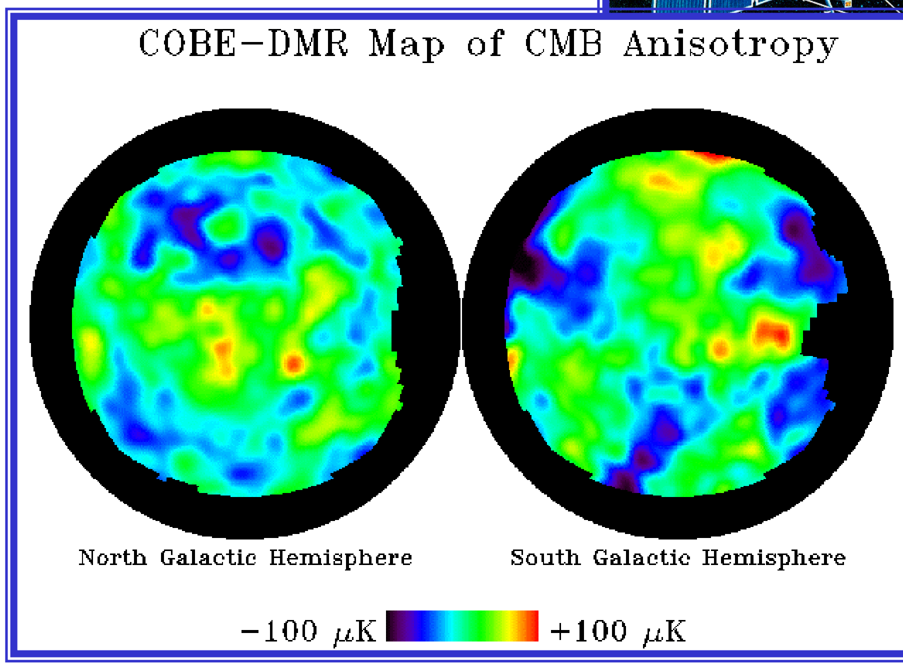
- Penzias & Wilson,1963
  - *Track down excess antenna noise*
    - ! ***Observe the cosmic microwave background radiation (CMBR)***



*http://www.lucent.com/museum/1964bang.html*

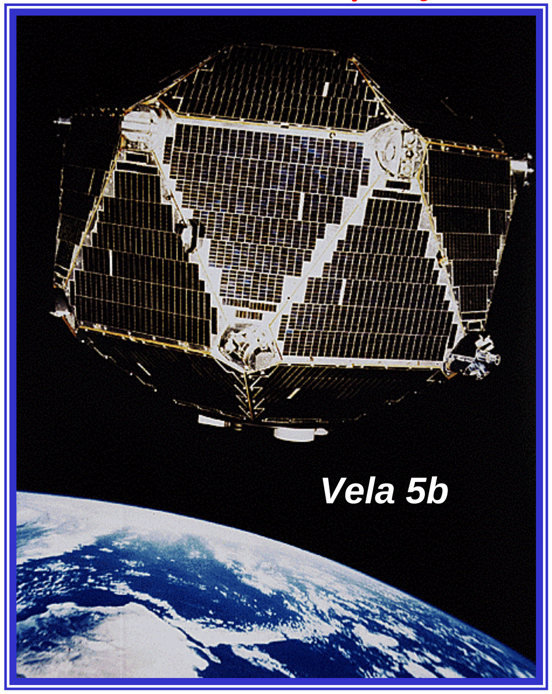
*http://www.gsfc.nasa.gov/astro/cobe/cobe_home.html*



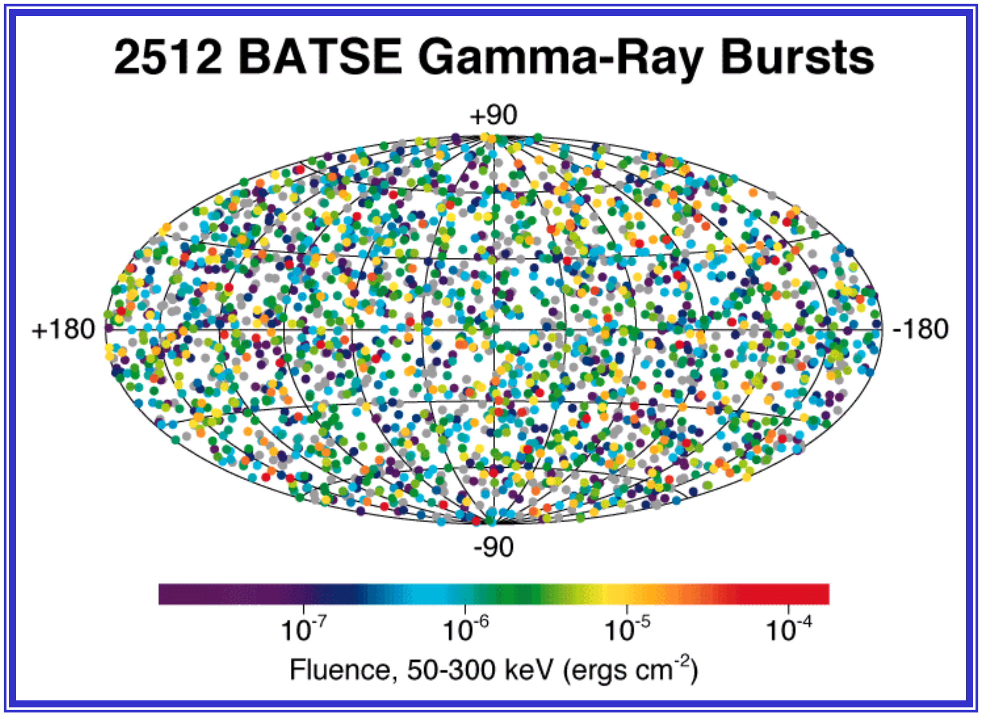



# *The Opening of a New Observational Window on the Universe*

- Klebesadel, Strong & Olsen (LANL), 1969
  - *Review of Vela 5 satellite data from 1967.07.02 shows a $\gamma$ event of non-terrestrial origin*
    - ! ***Discover $\gamma$-ray bursts (GRBs), X-ray sources***

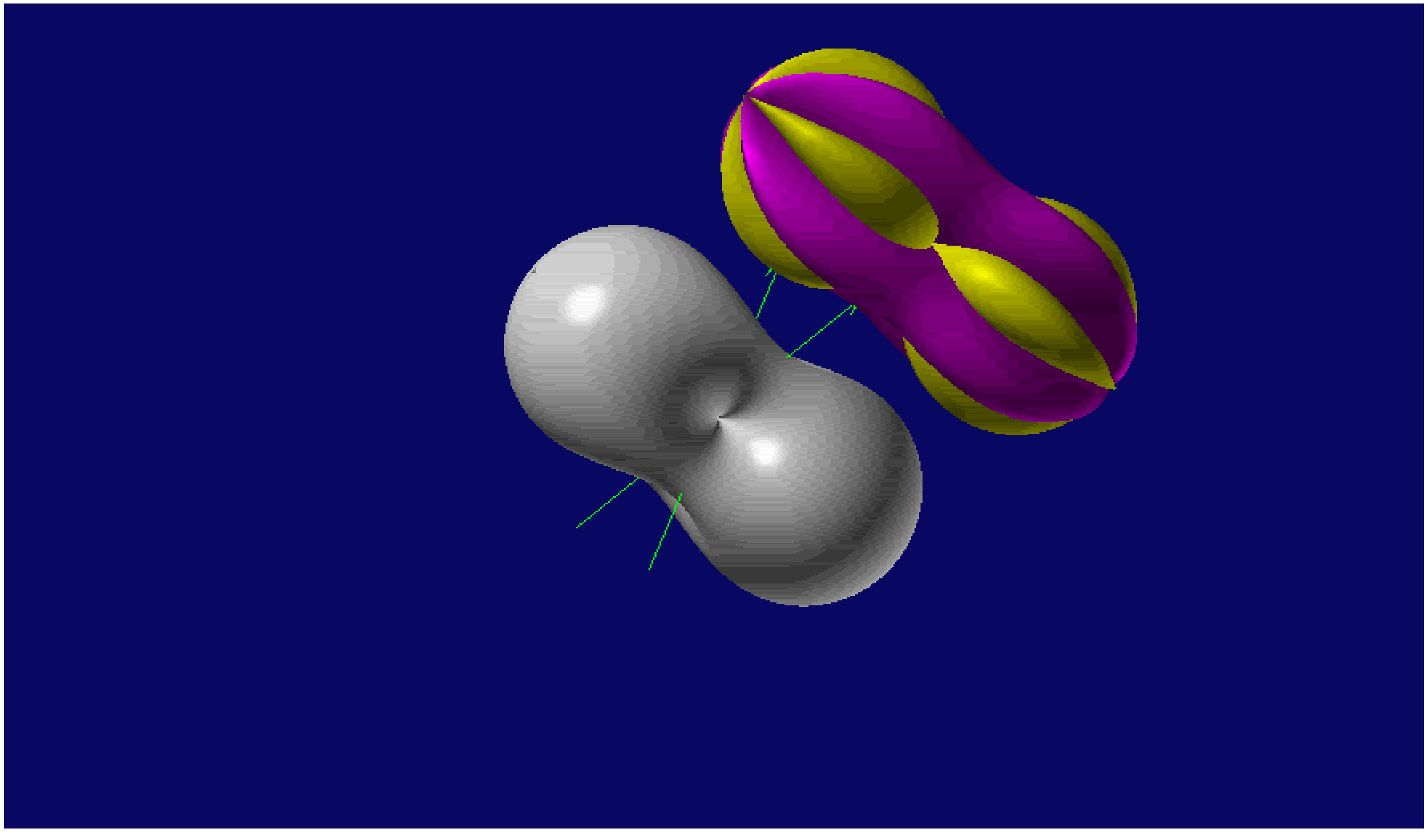



*http://science.msfc.nasa.gov/newhome/headlines/ast19sep97_2.htm*



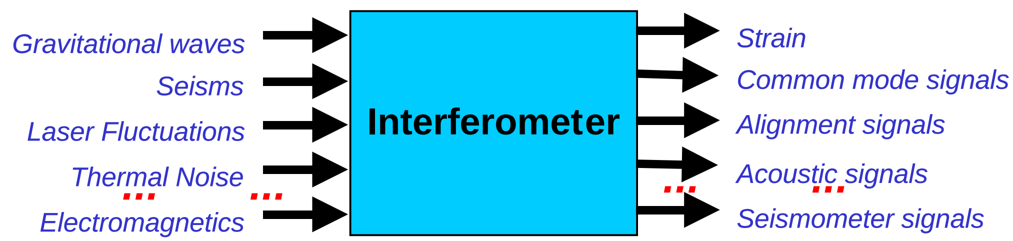

*http://www.batse.com/*

# Interferometer GW Antenna Pattern

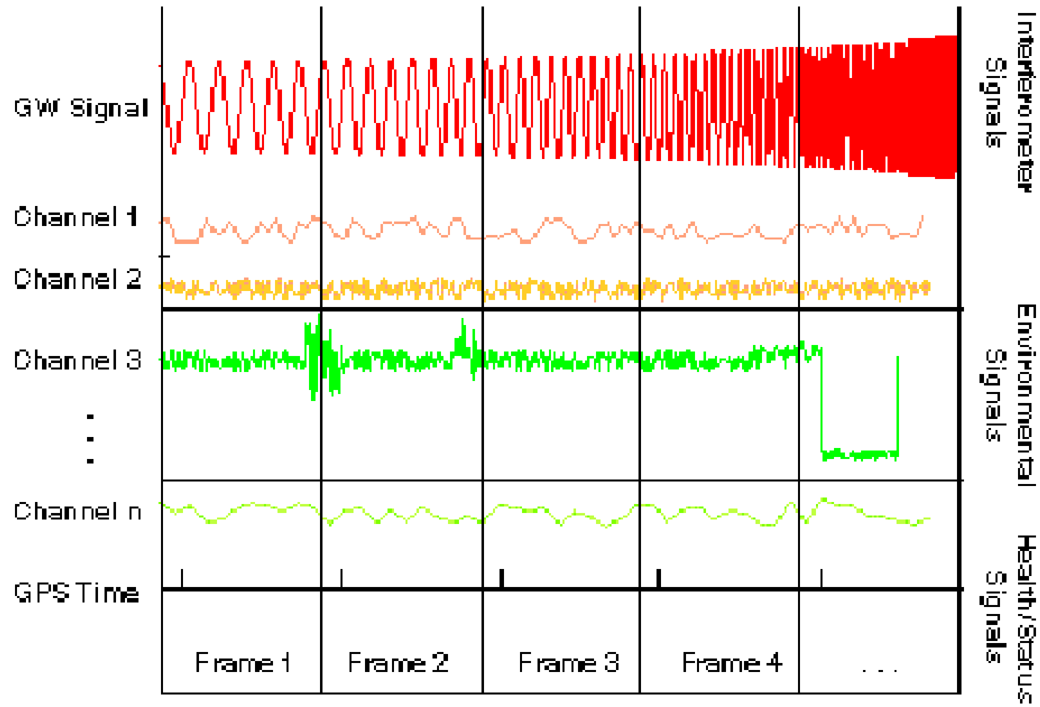


- ***Omni-directional*** **- multiple detectors, lots of signal processing to determine direction**
- ***Extremely weak signal*** **- lots of environmental, instrumental monitoring to determine validity**

# Interferometer Data Channels




- **All interferometric detector projects have agreed on a standard data format**
- **Anticipates joint data analysis**
- **LIGO Frames for 1 interferometer are ~3MB/s**
  - **32 kB/s strain**
  - **~2 MB/s other interferometer signals**
  - **~1MB/s environmental sensors**
  - ***Strain is ~1% of all data***



Mathematical Interlude …
Random Processes & Signal Noise

- n(t) is a randomly varying signal
  - *Gaussian process:*

$$P(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(n-\mu)^2}{2\sigma^2}}$$

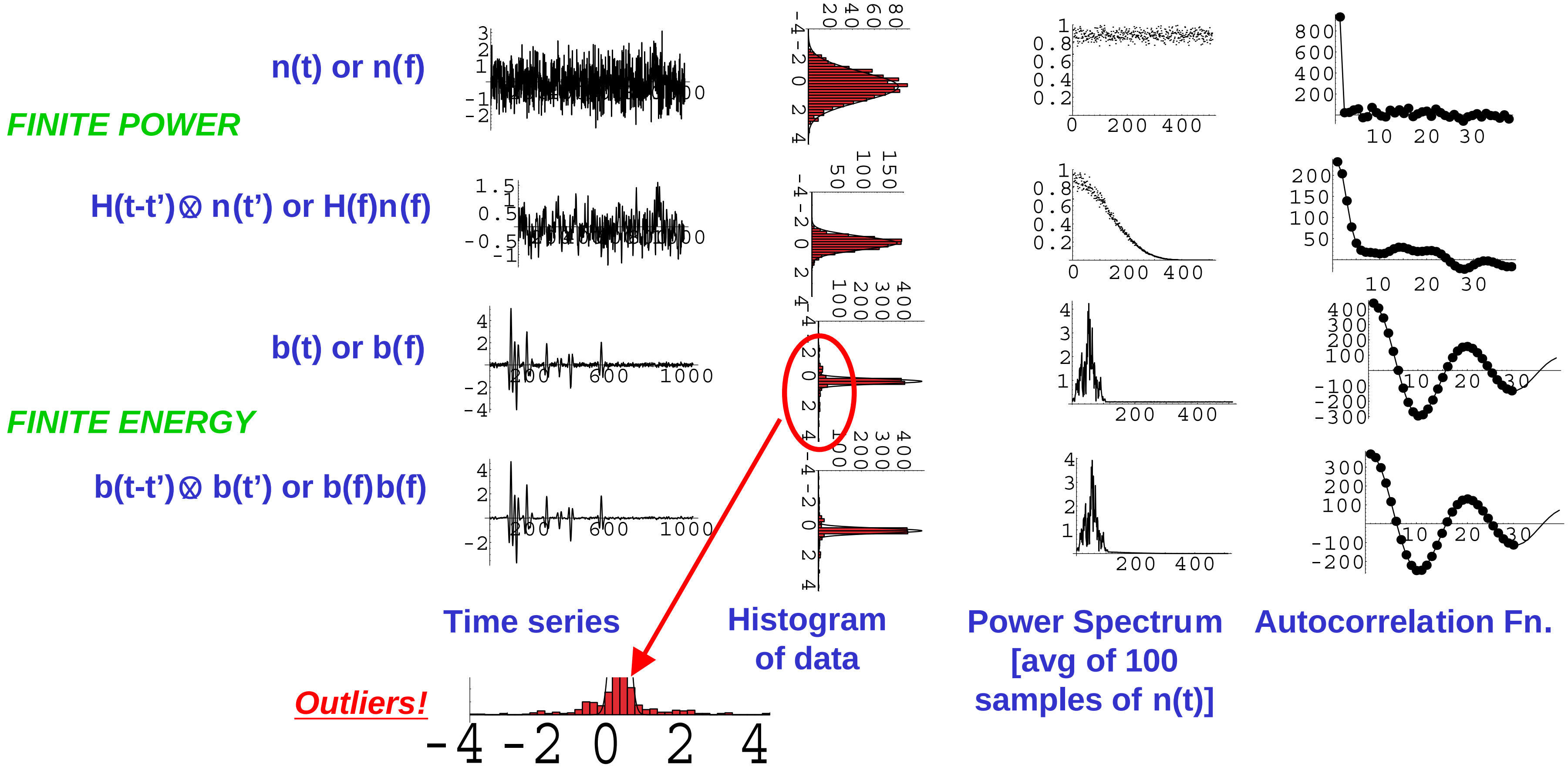
$$< n >= \mu = \int_{-\infty}^{\infty} nP(n)dn$$

$$< n^2 >= \sigma^2 + \mu^2 = \int_{-\infty}^{\infty} n^2 P(n)dn$$

  - *Can assume $\mu$=0 without loss of generality*
- **Ensemble averages, time averages**
  - *Time averge:* $$< n(t) >_t = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} n(t)dt$$
  - *Ensemble average:* $$< n >_P = \int_{-\infty}^{\infty} nP(n)dn$$
  - *Stationarity: $\mu$,$\sigma$ , etc. do not vary with time*
  - *Ergodicity: Probability distribution of n(t) over a long period T is the same as P(n) at one instant, t'*

# *Characteristics of Random Processes*

- n(t) is a (white) random Gaussian process; b(t) is a transient or burst process
- H(f)n(f) is a (colored) random Gaussian process; H(f)b(f) is a filtered burst process



# *Mathematical Interlude …*
# *Fourier Transforms of Time Dependent Signals*

- Continuous-time
  (infinite time duration)

$$\hat{R}(f) = \int_{-\infty}^{\infty} R(t) e^{-2\pi i f t} dt \equiv \boldsymbol{F}[R(t)]$$

$$R(t) = \int_{-\infty}^{\infty} \hat{R}(f) e^{2\pi i f t} df \equiv \boldsymbol{F}^{-1}[\hat{R}(f)]$$

$$\boldsymbol{F}^{-1}\boldsymbol{F}[R(t)] = R(t) \Rightarrow \delta(t-t') = \int_{-\infty}^{\infty} e^{-2\pi i f (t-t')} df$$

$$\int_{-\infty}^{\infty} \delta(t-t')dt = 1$$

- Discrete-time
  (finite time stretch T)

$$\hat{R}(f_i) = \sum_{n=0}^{N-1} R(t_n) e^{-2\pi i f_i t_n} \Delta t \equiv \boldsymbol{F}[R(t_n)] \; ; \; t_n = n\Delta t = \frac{nT}{N}$$

$$R(t_n) = \sum_{i=0}^{N-1} \hat{R}(f_k) e^{-2\pi i f_k t_n} \Delta f \equiv \boldsymbol{F}^{-1}[\hat{R}(f_k)] \; ; \; f_k = k\Delta f = \frac{k}{T}$$

$$\boldsymbol{F}^{-1}\boldsymbol{F}[R(t_j)] = R(t_i) \Rightarrow \delta_{ij}$$

- Computational cost to perform transform (FFT): *5 N $\log_2$ [N]*
  - ***100 s of 1024 Sample/s data: $8.5\times10^6$ floating point operations (FLOP)***
  - ***To keep up with data: 85 kFLOP/s (kFLOPS)***

## *Mathematical Interlude …*
## *Useful Theorems and Formulae*

- Parseval's Theorem

$$\int_{-\infty}^{\infty} R(t)^2\,dt = \text{"Integrated Signal Energy"}$$

$$= \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty}\hat{R}(f)e^{2\pi ift}df\Big]\Big[\int_{-\infty}^{\infty}\hat{R}(f'')e^{2\pi if''t}df''\Big]dt$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}df''\,df\{\hat{R}(f)\hat{R}(f'')\}\int_{-\infty}^{\infty}e^{2\pi i(f+f'')t}dt$$

$$= \int_{-\infty}^{\infty}\left|\hat{R}(f)\right|^2 df$$

- Translation:

$$\boldsymbol{F}\,[R(t+t')] = \hat{R}(f)e^{-2\pi ift'}$$

- Useful identity:

$$\hat{n}(f) = \hat{n}^*(-f) \quad for \quad real \quad n(t)$$

- Convolution

$$c(\tau) = \int_{-\infty}^{\infty} a(t)b(\tau - t)dt$$

$$\hat{c}(f) = \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} dt \;\; a(t)b(\tau - t)e^{-2\pi if\tau}; \;\; t' = \tau - t$$

$$\hat{c}(f) = \int_{-\infty}^{\infty} dt'\, b(t')e^{-2\pi ift'} \int_{-\infty}^{\infty} dt \;\; a(t)e^{-2\pi ift}$$

$$\hat{c}(f) = \hat{a}(f)\hat{b}(f)$$

- Autocorrelation

$$R_a(\tau) = \int_{-\infty}^{\infty} a(t)a(\tau + t)dt$$

$$\frac{1}{2}S_a(f) \equiv \int_{-\infty}^{\infty} R_a(\tau)e^{-2\pi if\tau}d\tau$$

$$= \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} dt \;\; a(t)a(\tau + t)e^{-2\pi if\tau}$$

$$= \int_{-\infty}^{\infty} dt'\, a(t')e^{-2\pi ift'} \int_{-\infty}^{\infty} dt \;\; a(t)e^{-2\pi ift}$$

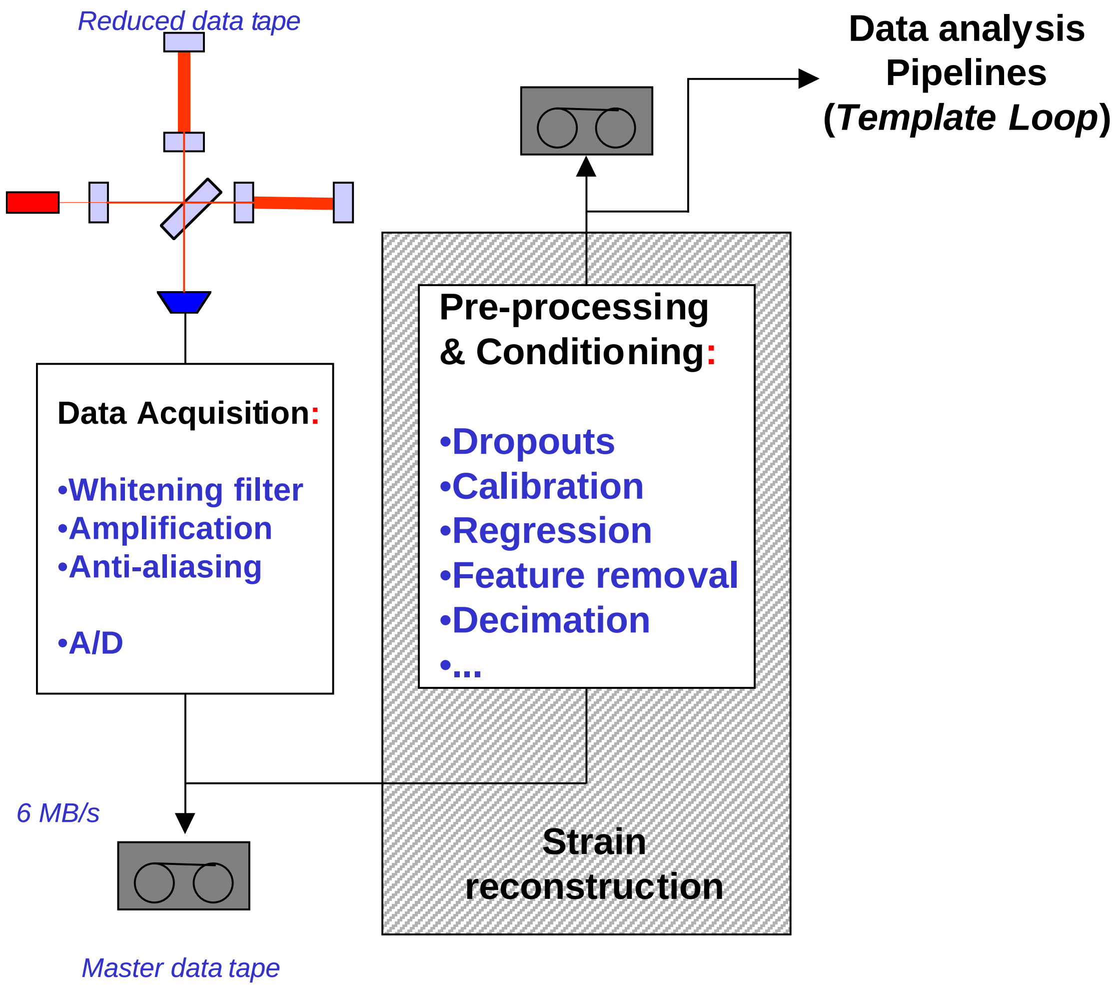
$$= \left|a(f)\right|^2$$

## *Mathematical Interlude …*
## *Property of an Ergodic Gaussian Process*

- **Different Fourier components of n(t) are independent of each other:**

$$
\begin{aligned}
\left\langle n(f)n^{*}(f')\right\rangle_{P} &= \left\langle \int n(t)e^{-2\pi ift}dt \int n(t+\tau)e^{2\pi if'(t+\tau)}d\tau \right\rangle_{t} \\
&= \int dt' \int \left\langle n(t+\tau)n(t)\right\rangle_{t} e^{-2\pi ift}e^{2\pi if'(t+\tau)}d\tau \\
&= \int dt' e^{-2\pi i(f-f')t} \int R_{n}(\tau)e^{2\pi if'\tau}d\tau \\
&= S_{n}(f)\delta(f-f')
\end{aligned}
$$

# Data Flow: Pre-processing

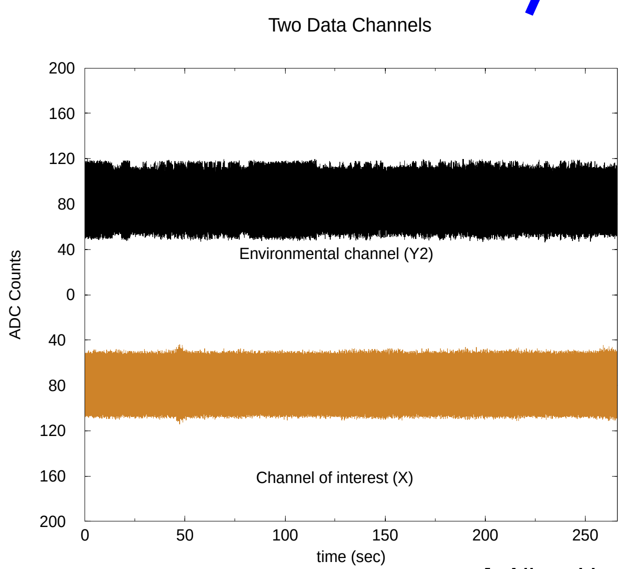


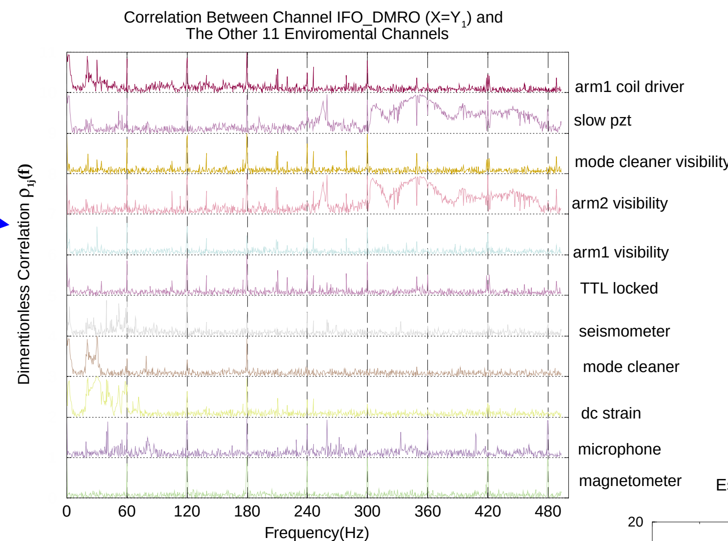
# *Data Pre-processing: removing instrumental effects*

- Cross channel regression will be used to improve signal to noise ratios when possible (need adequate SNR)

**Raw channel data (40m prototype)**

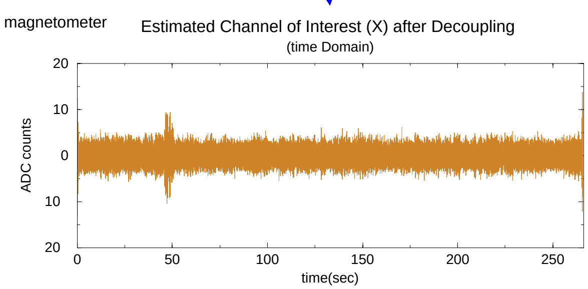
$$s_a(t) \Rightarrow \hat{s}_a(f)$$
$$s_b(t) \Rightarrow \hat{s}_b(f)$$
$$s_a(t)s_b(t+\tau) \Rightarrow R_{ab}(\tau) \Leftrightarrow \hat{S}_{ab}(f)$$




**Cross channel spectral correlation**

$$\Gamma_{ab}(f) \equiv \frac{\left|\hat{S}_{ab}(f)\right|^2}{\hat{s}_a(f)\;\hat{s}_b(f)}$$




$$\begin{pmatrix}\hat{s}'_a(t)\\ \hat{s}'_b(t)\end{pmatrix} \Leftarrow \begin{pmatrix}\hat{s}'_a(f)\\ \hat{s}'_b(f)\end{pmatrix} = M(f)\cdot\begin{pmatrix}\hat{s}_a(f)\\ \hat{s}_b(f)\end{pmatrix}$$

**Reduced data channel**

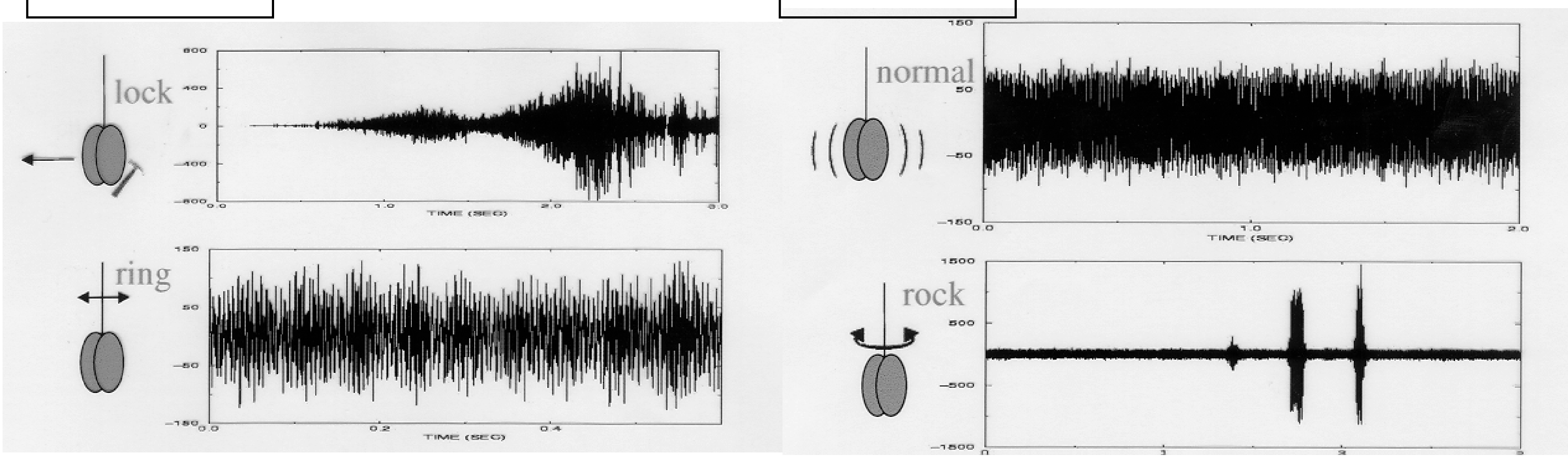
*ref: Allen, Hua, Ottewill (gr-qc/9909083)*

# *Interferometer Transients -- Examples from 40m Data*

## Real interferometer data is UGLY!!!

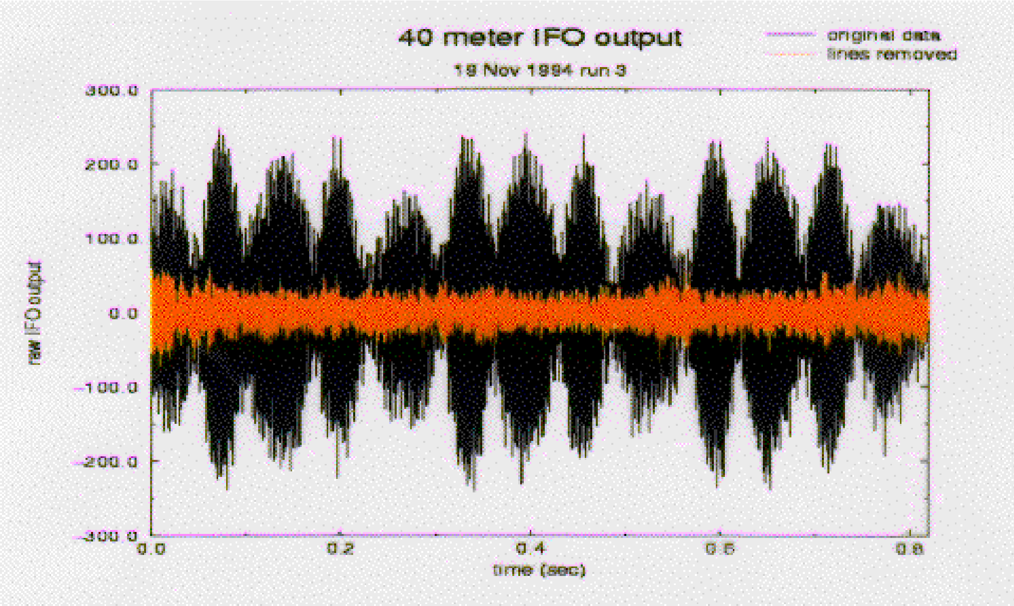
**(Glitches - known and unknown)**

**LOCKING** **NORMAL**

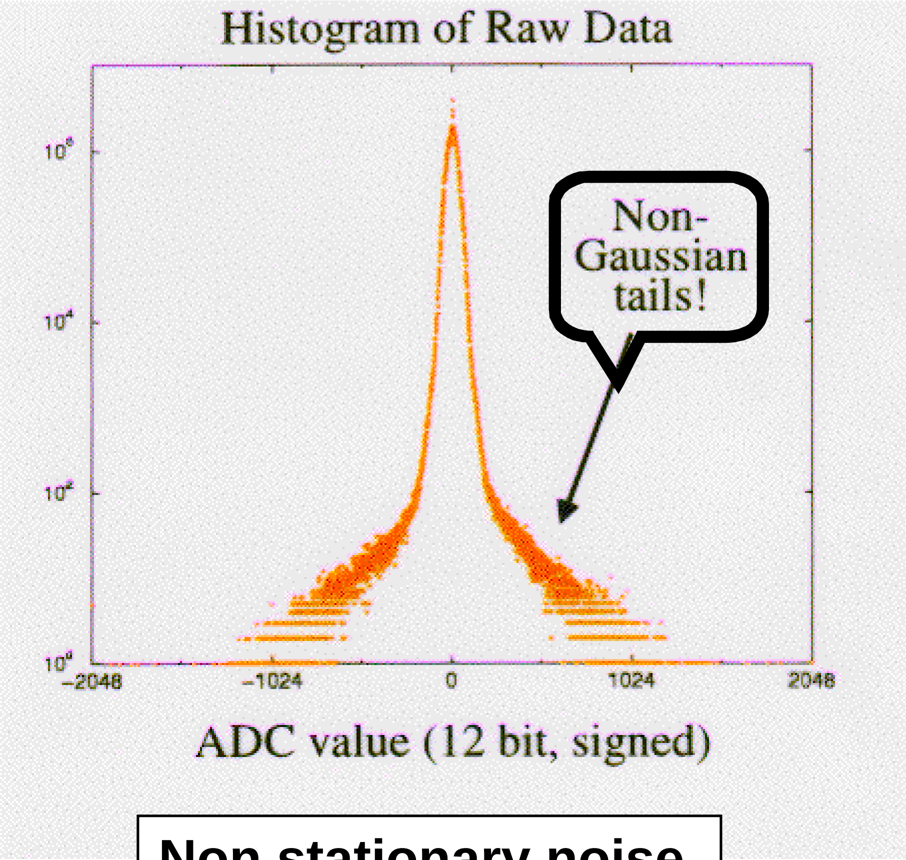



**RINGING** **ROCKING**

# *"Clean up" data stream*

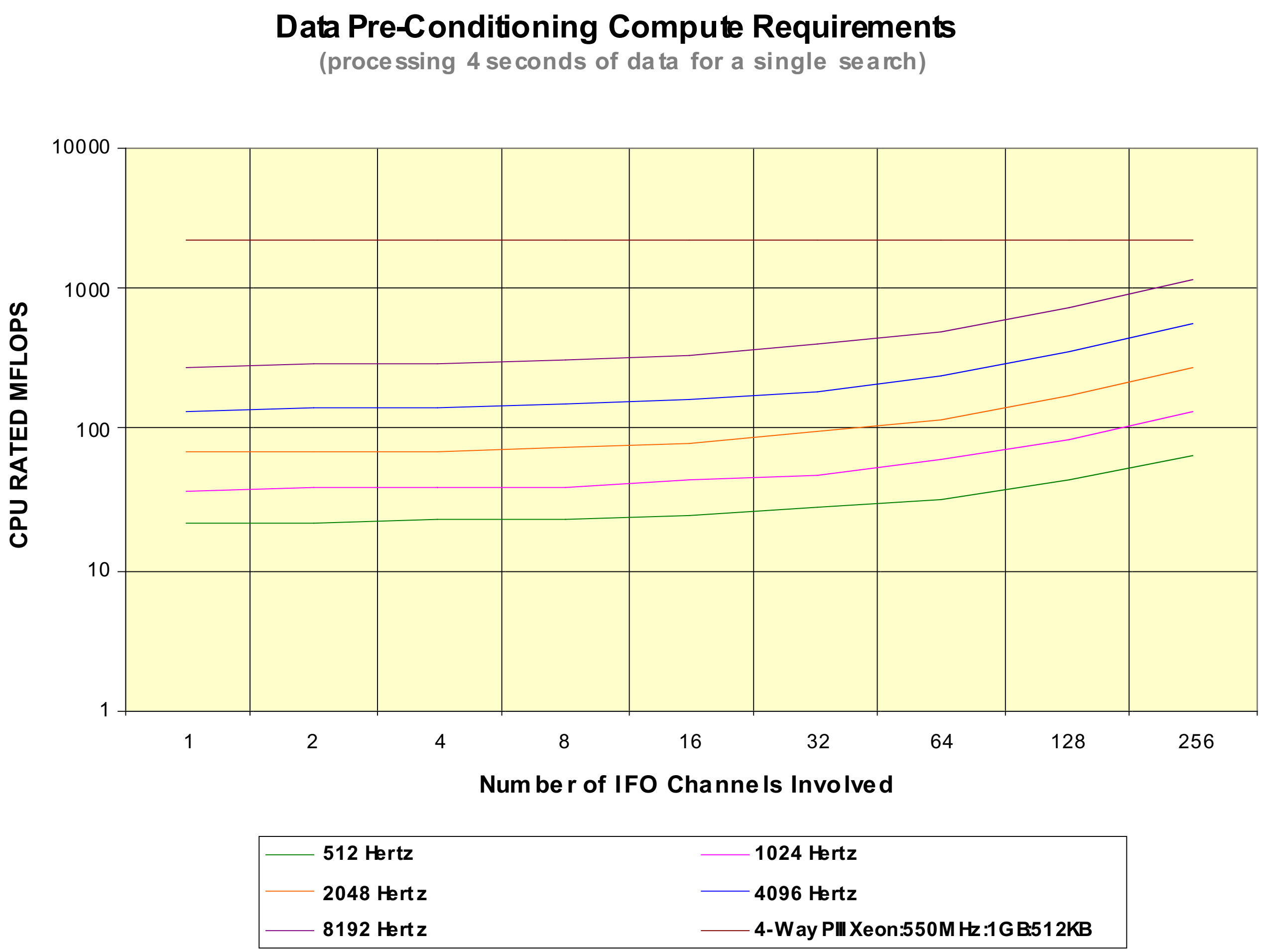



**Effect of removing sinusoidal artifacts using multi-taper methods**




**Non stationary noise**
**Non gaussian tails**

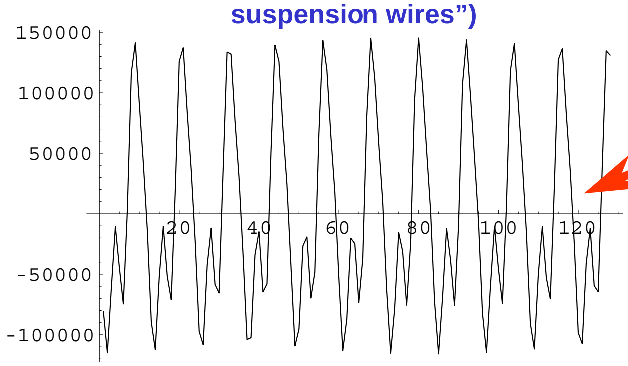
# LIGO Data(Pre)-Conditioning Computational Costs



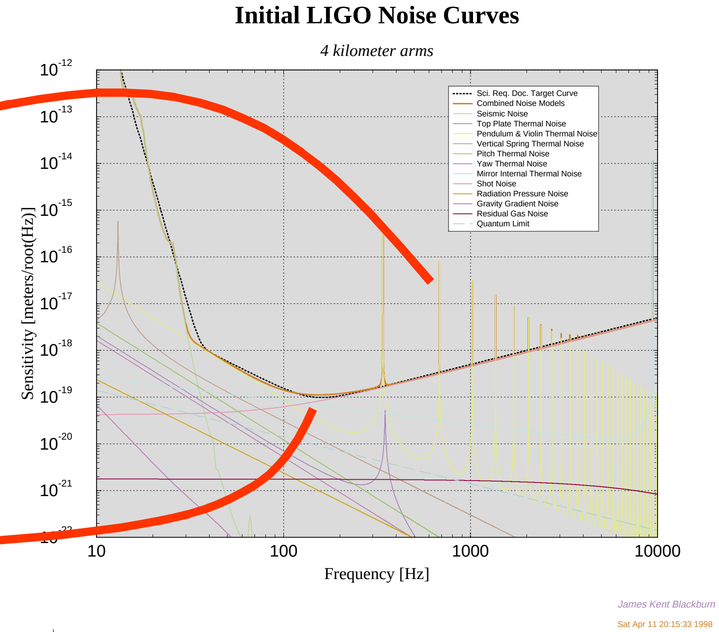

- Pre-Conditioning steps involved:
  - *64K samples of input data per channel*
  - *Data Drop-Out Correction on 10% of data.*
  - *Line Removal of 64 lines.*
  - *Calibration of GW channel.*
  - *Resampling (see legend).*
  - *Linear Regression using all input signals.*
- Roughly 4-8 unique searches expected to be active in LDAS.

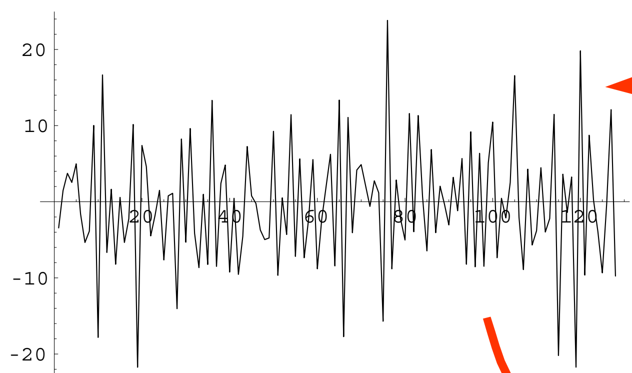
# *Interferometer Strain Signal (Simulated)*

**Dominated by narrowband features in spectrum (“violin resonances” of suspension wires”)**

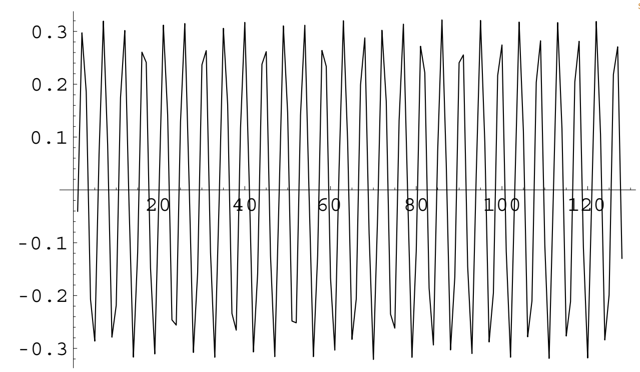



**Design LIGO I limiting strain sensitivity**




**Broadband noise spectrum**

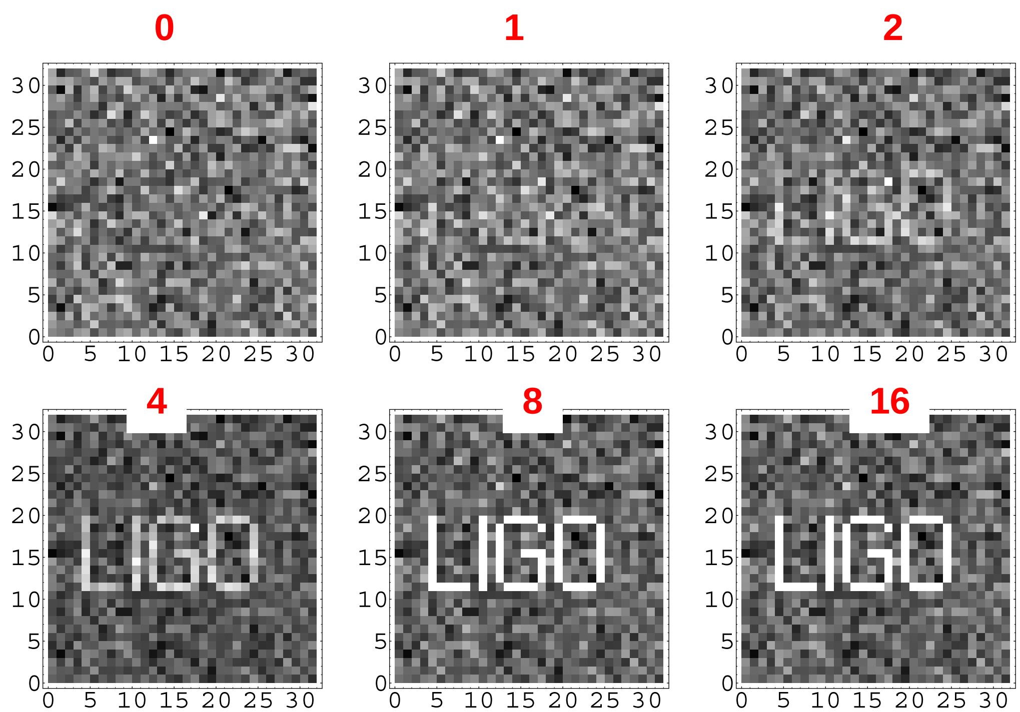



**Chirped waveform**

# *Detections vs. Observations*

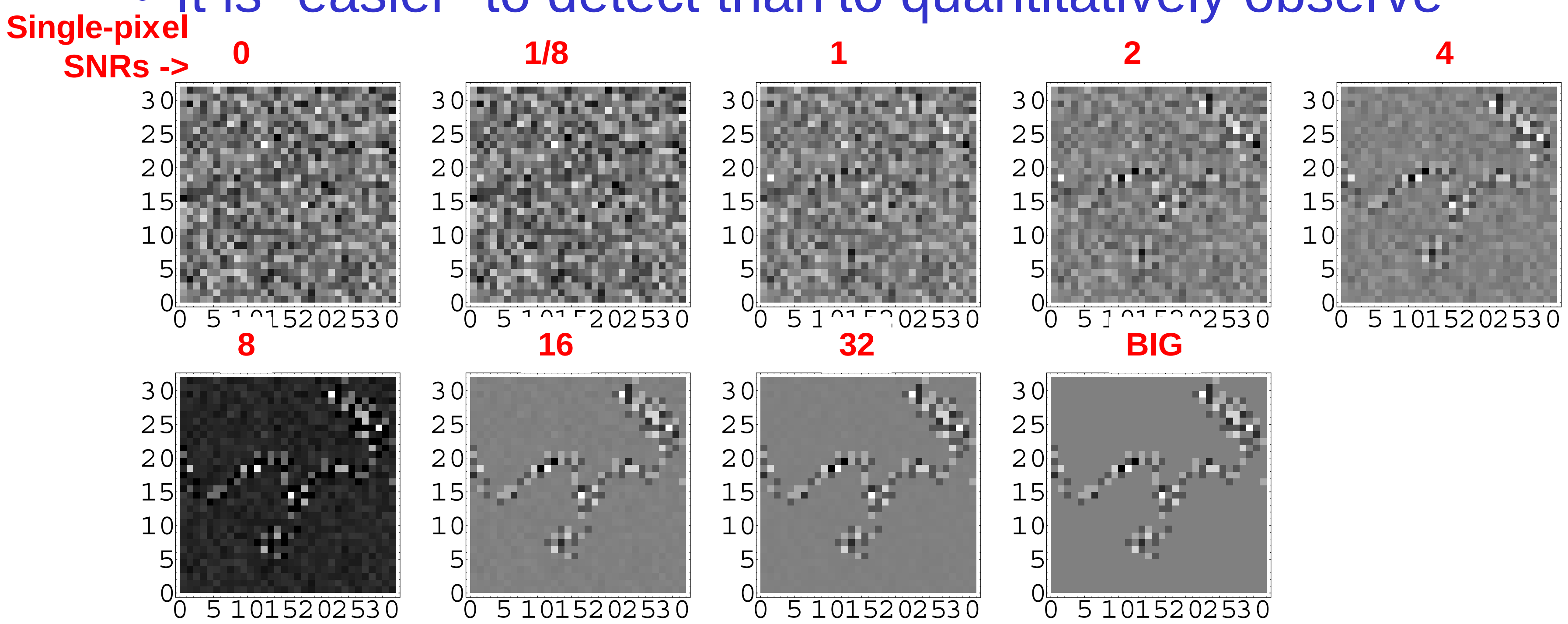
- It is "easier" to detect than to quantitatively observe

**Single-pixel SNRs ->**



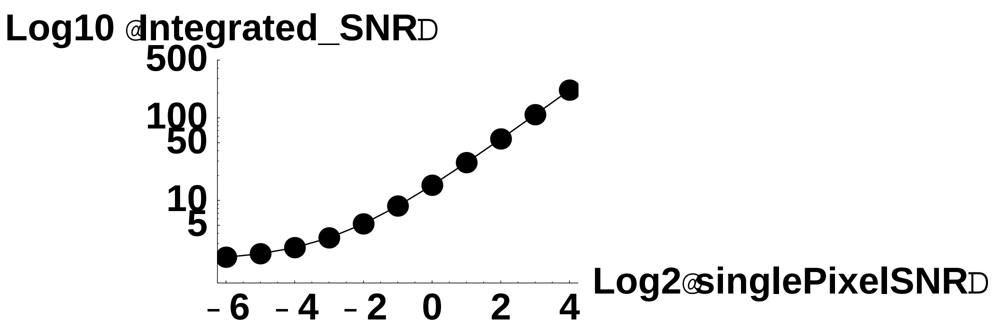
# *Detections vs. Parameter Estimation*

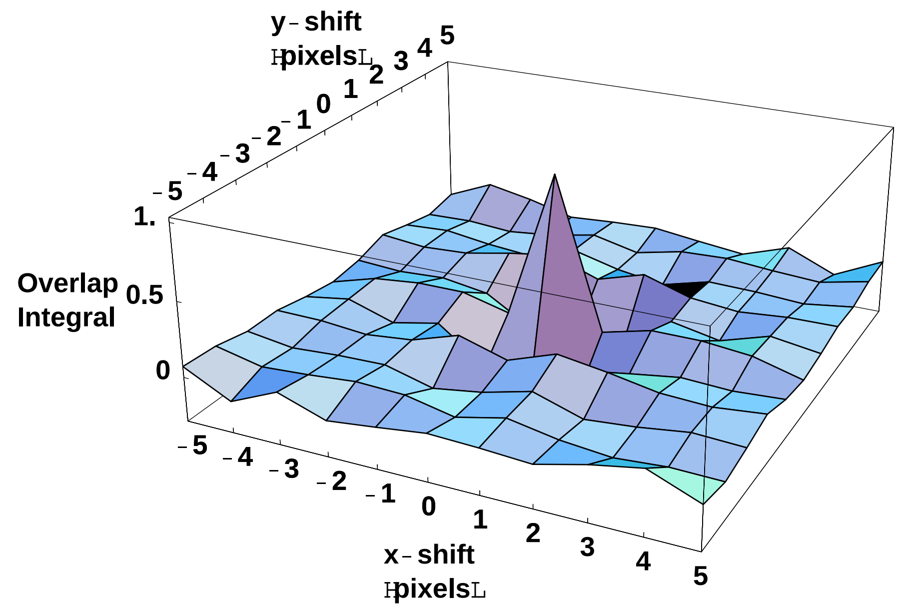
- It is "easier" to detect than to quantitatively observe



# Detections vs. Parameter Estimation

- It is "easier" to detect than to quantitatively observe




$$n = noise;$$
$$d = noisy\ \ data;$$
$$T = noiseless\ \ template$$

$$SNR = \frac{\sum_{i,j} d(i,j)T(i,j)}{\sqrt{\sum_{i,j} n^2(i,j)T^2(i,j)}}$$

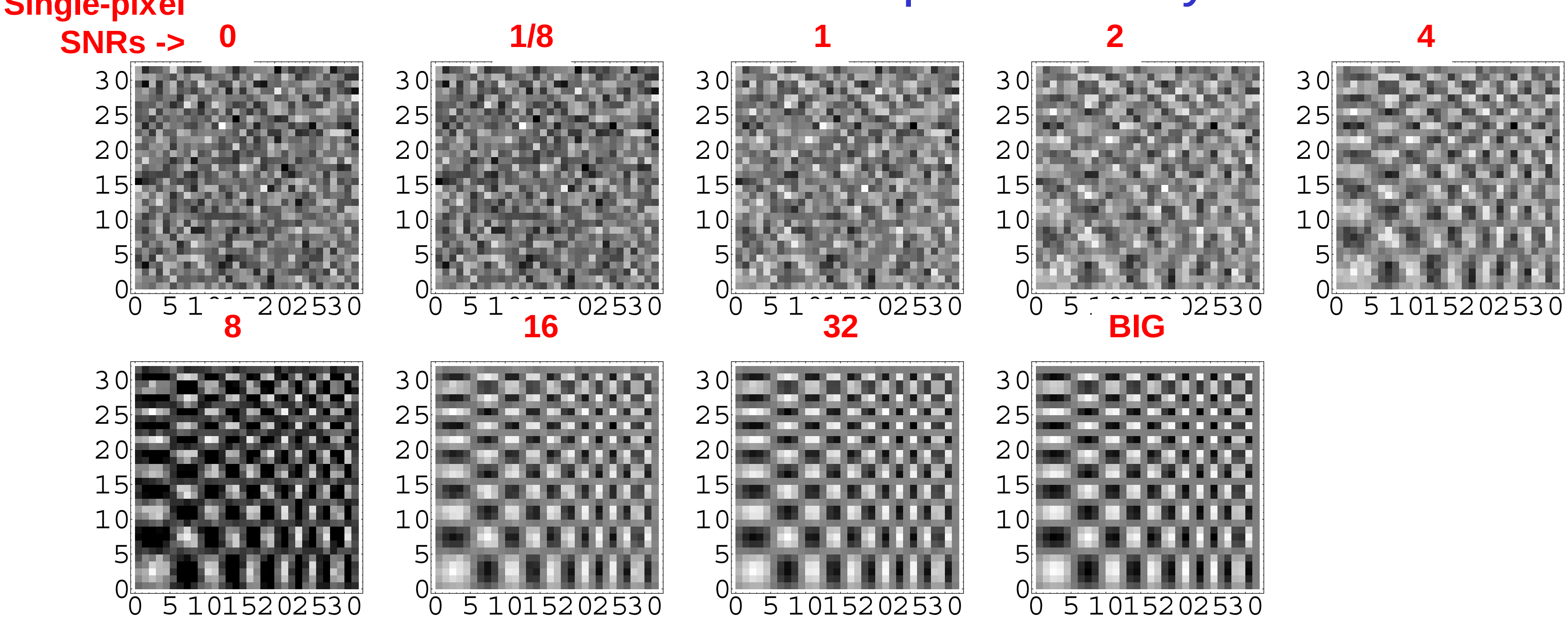
$$A(\Delta i, \Delta j) = \sum_{i,j} T(i,j)T(i+\Delta i, j+\Delta j)$$

- **Integrated SNR == Sum over entire image after it is multiplied by noiseless signal (template)**

- **Effect on Integrated SNR of (unknown) shift between *noiseless* signal and template**
- **Shift +/- 1 pixel kills correlations => need $32^2$ templates to cover a 32x32 pixel image...**

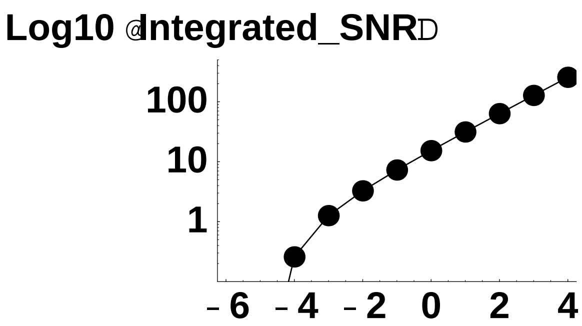
# *Detections vs. Parameter Estimation*

- It is “easier” to detect than to quantitatively observe



# *Detections vs. Parameter Estimation*

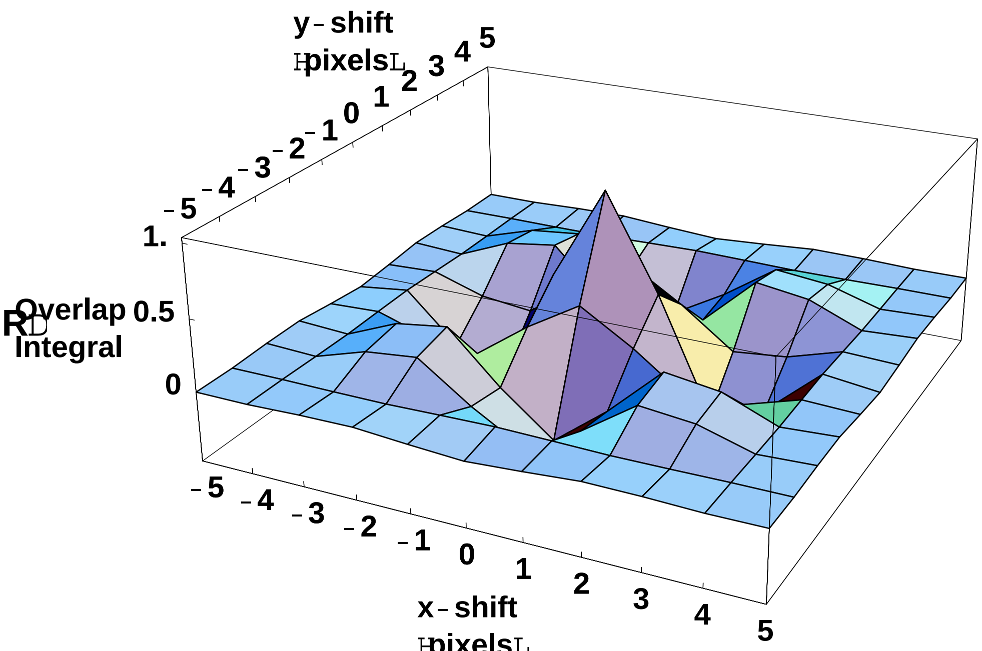
- It is "easier" to detect than to quantitatively observe




$$n = noise;$$
$$d = noisy\ \ data;$$
$$T = noiseless\ \ template$$

$$SNR = \frac{\sum_{i,j} d(i,j)T(i,j)}{\sqrt{\sum_{i,j} n^2(i,j)T^2(i,j)}}$$

- **Integrated SNR == Sum over entire image after it is multiplied by noiseless signal (template)**




$$A(\Delta i, \Delta j) = \sum_{i,j} T(i,j)T(i+\Delta i, j+\Delta j)$$

- **Effect on Integrated SNR of (unknown) shift between *noiseless* signal and template**
- **Shift +/- 1 pixel reduces correlations => need $32^2$ templates to cover a 32x32 pixel image...**

# Optimal Filtering and Signal Processing

- **Measurement s, (may contain) signal h, contains noise n:**

$$s(t) = h(t) + n(t)$$
$$\langle n(t) \rangle = 0 \;\; ; \; \langle h \rangle = \alpha T_i(t - t_0);$$

- **$\alpha T_i$(t-$t_0$) is one of a family of fiducial (expected) waveform or *templates* of type *i* with unknown time origin $t_0$ and amplitude $\alpha$ .**

- **Design optimal correlation filter,Q, that maximizes chance of detecting *T* in s(t) :**

$$C(t_0) = \int_{-\infty}^{\infty} s(t - t_0) Q(t) dt;$$
$$\langle C \rangle = \int_{-\infty}^{\infty} \alpha T(t - t_0) Q(t) dt$$
$$= \alpha \int_{-\infty}^{\infty} \hat{T}(f) e^{-2\pi i f t_0} \hat{Q}^*(f) df$$

$$N = C - \langle C \rangle; \; \langle N \rangle \equiv 0$$
$$\langle N^2 \rangle = \langle C^2 \rangle - \langle C \rangle^2$$
$$= \left\langle \int \hat{n}(f) \hat{Q}^*(f) df \int \hat{n}(f') \hat{Q}^*(f') df' \right\rangle$$
$$= \int df \int df' \, \hat{Q}^*(f) \hat{Q}^*(f') \langle \hat{n}(f) \hat{n}(f') \rangle$$
$$= \frac{1}{2} \int S_n(f) |Q(f)|^2 df$$

## *Optimal Filtering and Signal Processing*

- **Maximize P(detection) => Maximize signal-to-noise ratio (SNR):**

$$Max[SNR = \frac{\langle C\rangle}{\langle N^2\rangle^{1/2}}] => \quad Max[(SNR)^2 = \frac{\langle C\rangle^2}{\langle N^2\rangle}]$$

$$\delta(SNR)^2 = \frac{\delta\langle C\rangle^2}{\langle N^2\rangle} - \frac{\langle C\rangle^2 \delta\langle N^2\rangle}{\langle N^2\rangle^2} = 0$$

$$= \frac{2\langle C\rangle}{\langle N^2\rangle^2}\left(\langle\delta C\rangle\langle N^2\rangle - \langle C\rangle\langle N\delta N\rangle\right) = 0$$

$$\langle\delta C\rangle\langle N^2\rangle = \langle C\rangle\langle N\delta N\rangle$$

- **Variation and maximization is with respect to optimal filter, δQ:**

$$\langle C\rangle = \alpha\int_{-\infty}^{\infty}\hat{T}(f)e^{-2\pi i f t_0}\hat{Q}^*(f)df$$

$$\langle\delta C\rangle = \alpha\int_{-\infty}^{\infty}\hat{T}(f)e^{-2\pi i f t_0}\delta\hat{Q}^*(f)df$$

$$\langle N^2\rangle = \frac{1}{2}\int S_n(f)|Q(f)|^2 df$$

$$\langle N\delta N\rangle = \frac{1}{2}\int S_n(f)Q(f)\delta Q^*(f)df$$

## Optimal Filtering and Signal Processing

- ***Equate LHS, RHS:***

$$\langle \delta C \rangle \langle N^2 \rangle = \langle C \rangle \langle N \delta N \rangle$$

$$\hat{T}(f)e^{-2\pi i f t_0}\left[\int S_n(f')|Q(f')|^2 df'\right] = S_n(f)Q(f)\left[\int \hat{T}(f')e^{-2\pi i f t_0}\hat{Q}^*(f')df'\right]$$

- ***Require coefficients of […] to be equal:***

$$\hat{T}(f)e^{-2\pi i f t_0} = S_n(f)Q(f) \Rightarrow \boxed{Q(f) = \frac{\hat{T}(f)e^{-2\pi i f t_0}}{S_n(f)}}$$

***Optimal filter for this problem!***

- ***Check that […] == […] with Q(f) from above:***

$$\left[\int S_n(f')|Q(f')|^2 df'\right] ?=? \left[\int \hat{T}(f')e^{-2\pi i f t_0}\hat{Q}^*(f')df'\right]$$ ***Yes!!!!***

# *Matched Filtering with Templates*

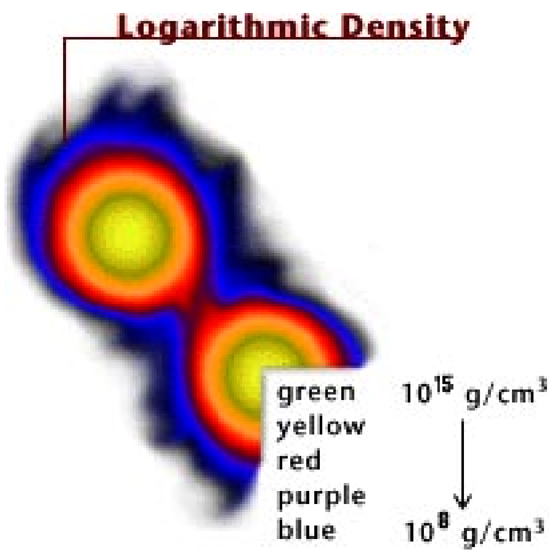
$$C(t_0 : \{p\}) = \int_{-\infty}^{\infty} \frac{\hat{s}(f)\hat{T}^*_{\{p\}}(f)}{S_n(f)} e^{2\pi i f t_0} df$$

- $C(t_0{:}\{p\})$ is a *family* of derived correlation time series
  - *{p} is a set of parameters used to characterize the templates T*
    - *Intrinsic parameters: masses, other GR parameters*
    - *Extrinsic parameters: distance, orbital inclination, phase, position in the sky, …*
    - *Dimensions of {p} can be HUGE (e.g., $10^4$ - $10^5$) if one wants a reasonable certainty of detecting a weak signal with unknown parameters*
  - *If statistics of noise n(t) are sufficiently well understood , it is possible to make a confidence statement of the likelihood that a certain value for $C(t_0)$ is due to a signal characterized by {p}, and not noise alone.*

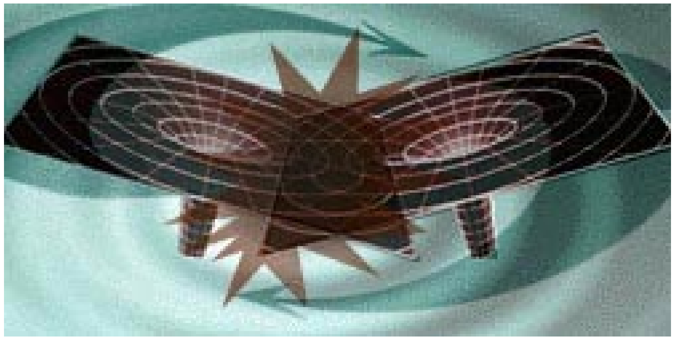
# *Science in LIGO I*
## Data analysis plan

- Compact binary inspiral: *"chirps"*
  - *NS-NS waveforms are well described*
  - *BH-BH need better waveforms*
  - *search technique: matched templates*
- Supernovae / GRBs: *"bursts"*
  - *burst search algorithms – excess power; time-freq patterns*
  - *burst signals - coincidence with signals in E&M radiation*
  - *prompt alarm (~ 1 hr) with $\nu$ detectors [SNEWS]*
- Cosmological Signals *"stochastic background"*
  - *Search technique: optimal Wiener filter for different models*
- Pulsars in our galaxy: *"periodic"*
  - *search for observed neutron stars (freq., doppler shift) - matched filters*
  - *all sky search (computing challenge)*
  - *r-modes*

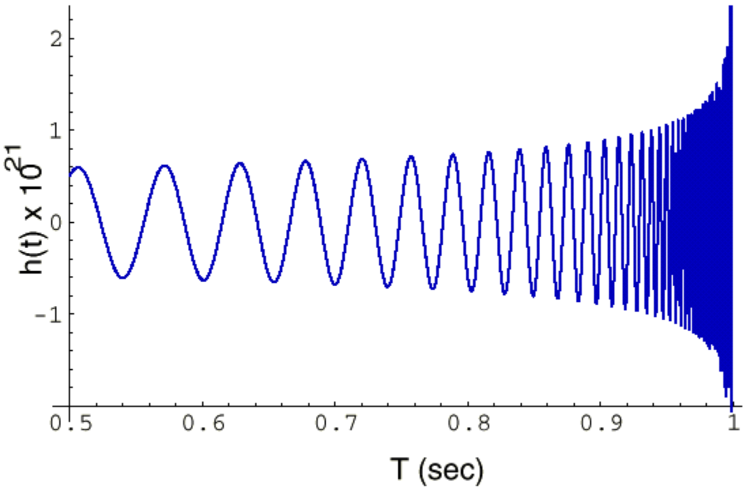
# *Sources of Gravitational Waves*
# *Closely orbiting compact stars*



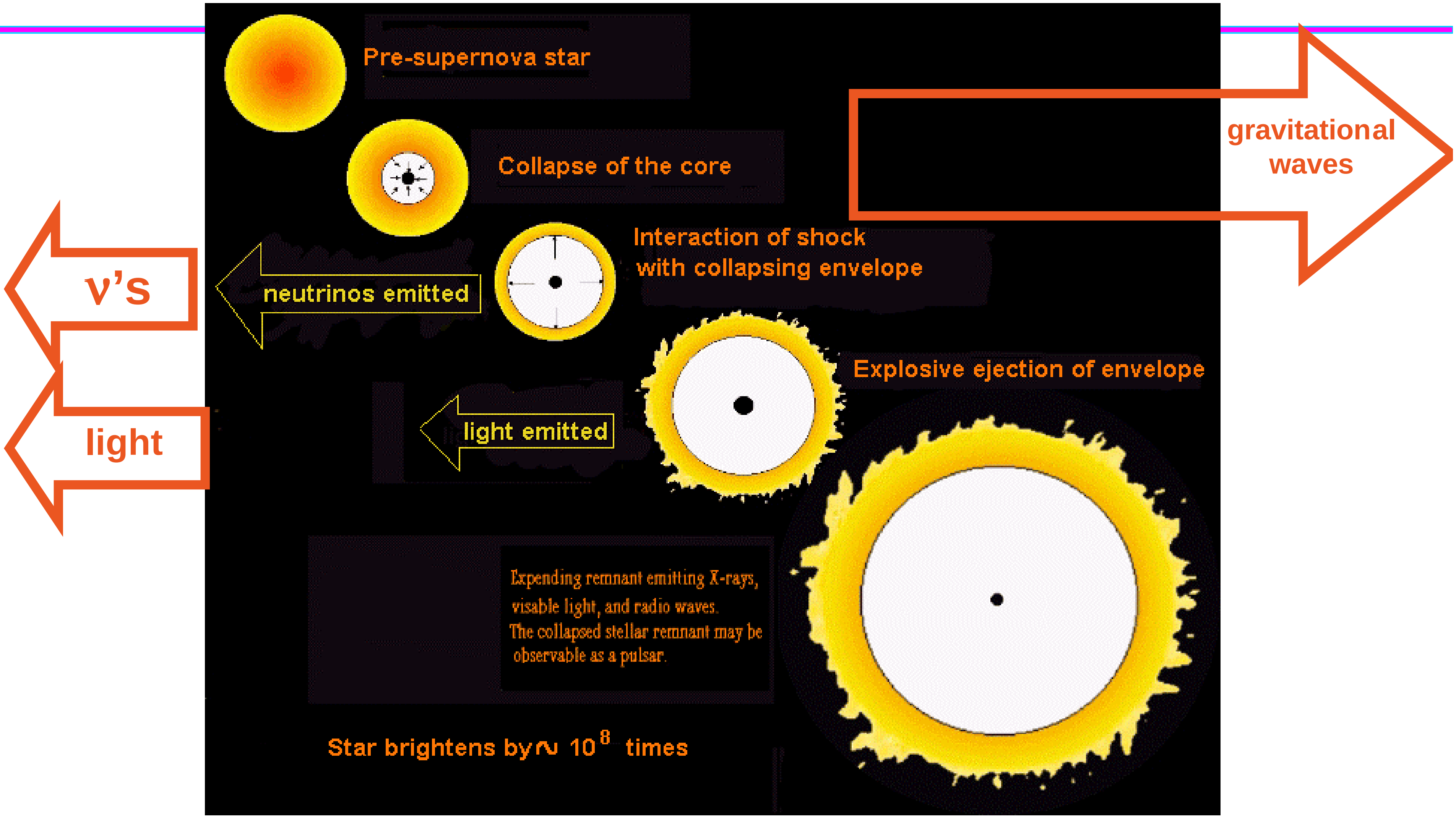

## Inspiral of Neutron Stars



## "Chirp Signal"

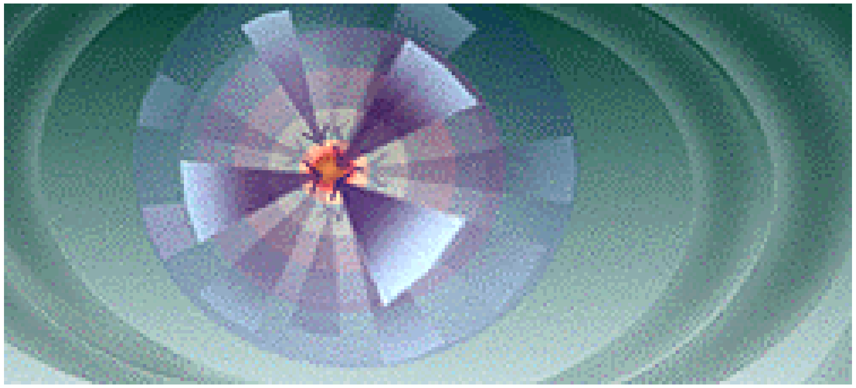


# *Sources of Gravitational Waves*
# *Supernovae*

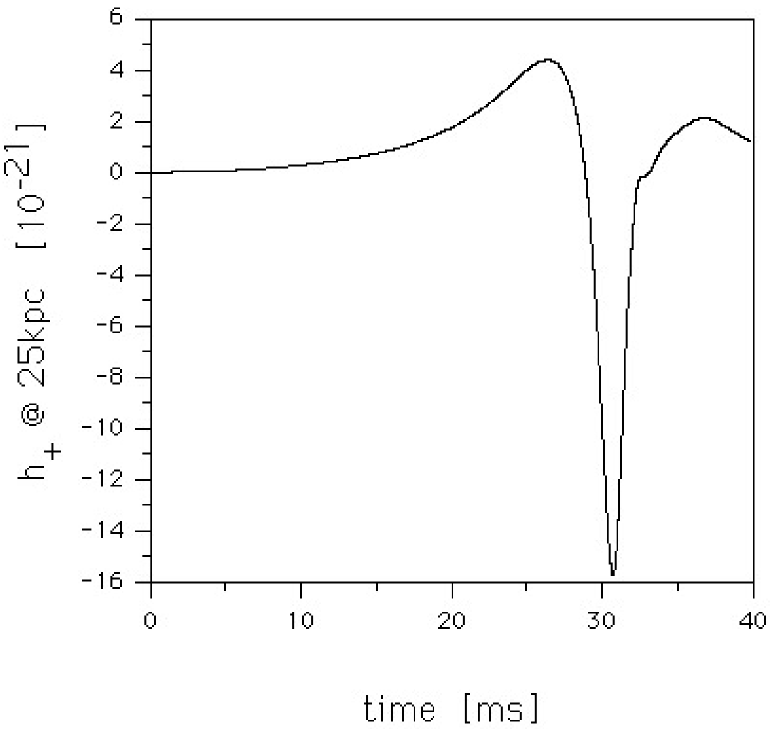


# *Supernovae Gravitational Waves*

**Non axisymmetric collapse**



**'burst' signal**




**Rate**

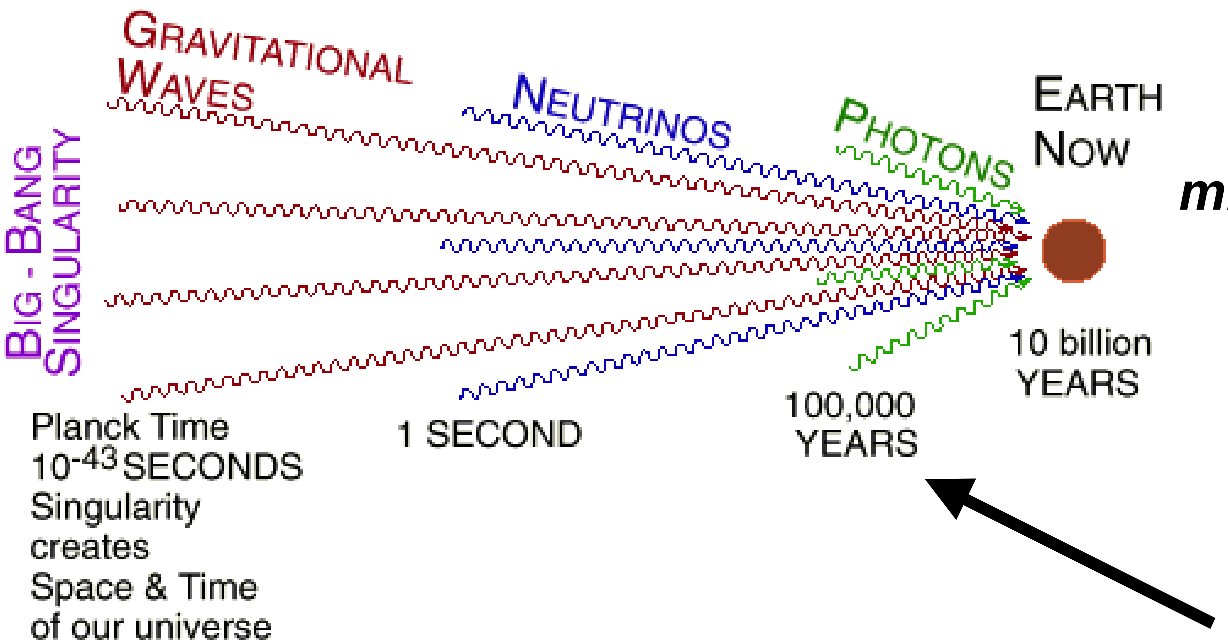
***1/50 yr - our galaxy***

***3/yr - Virgo cluster***

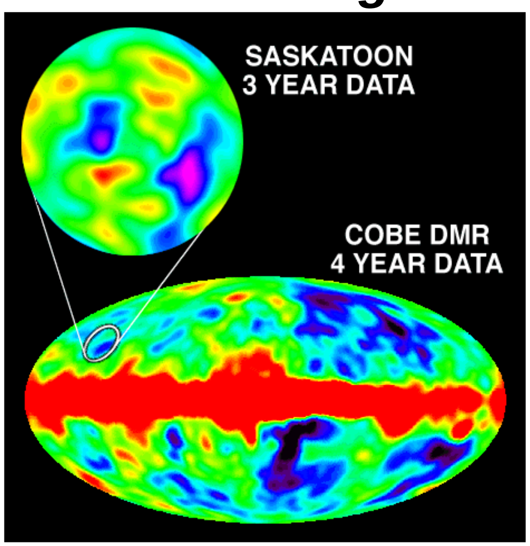
# *Sources of Gravitational Waves*
# *The Big Bang*

## 'Murmurs' from the Big Bang

***Signals from the early universe:*** ***Gaussian*** **white noise (hiss) that is** ***correlated*** **among** ***multiple detectors*** **in predictable manner (frequency, orientations, etc.)**



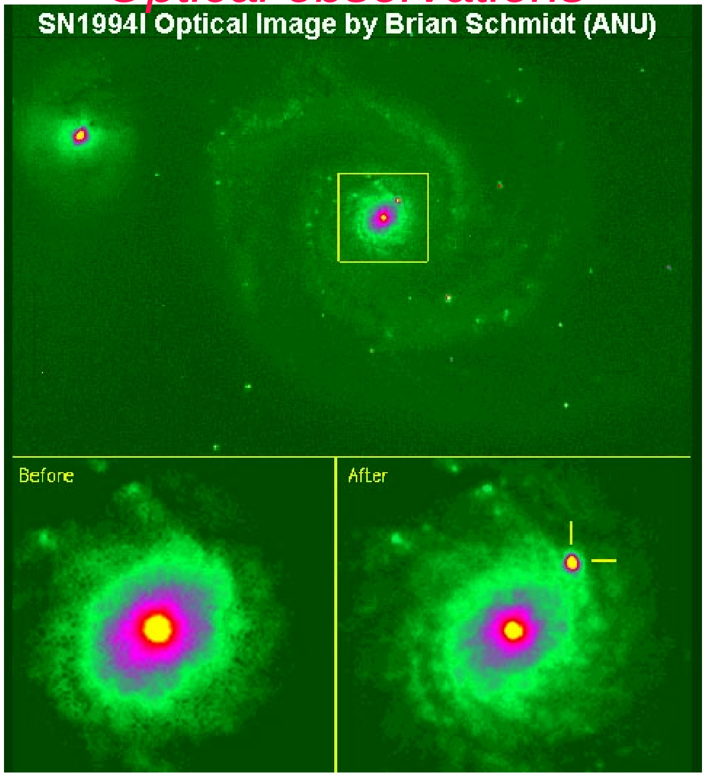

***Cosmic microwave background***



# *Sources of Gravitational Waves*
# *The Rotating Deformed Neutron Stars (Pulsars)*

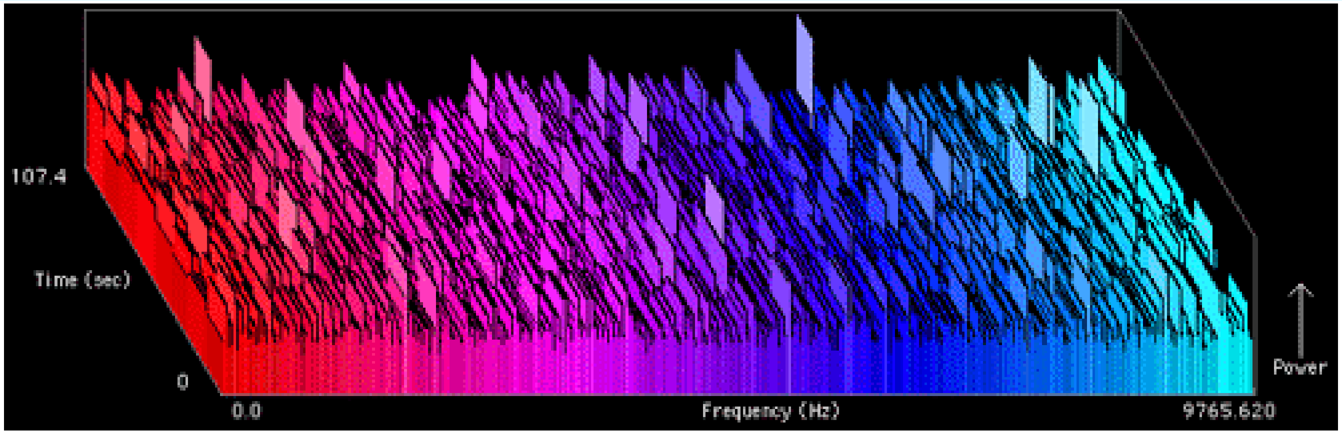
**Pulsars (e.g., Crab Nebulaa 1054 AD)**
*Continuous waves*

**Supernovae (e.g., SN 1994I)**
*Optical observations*



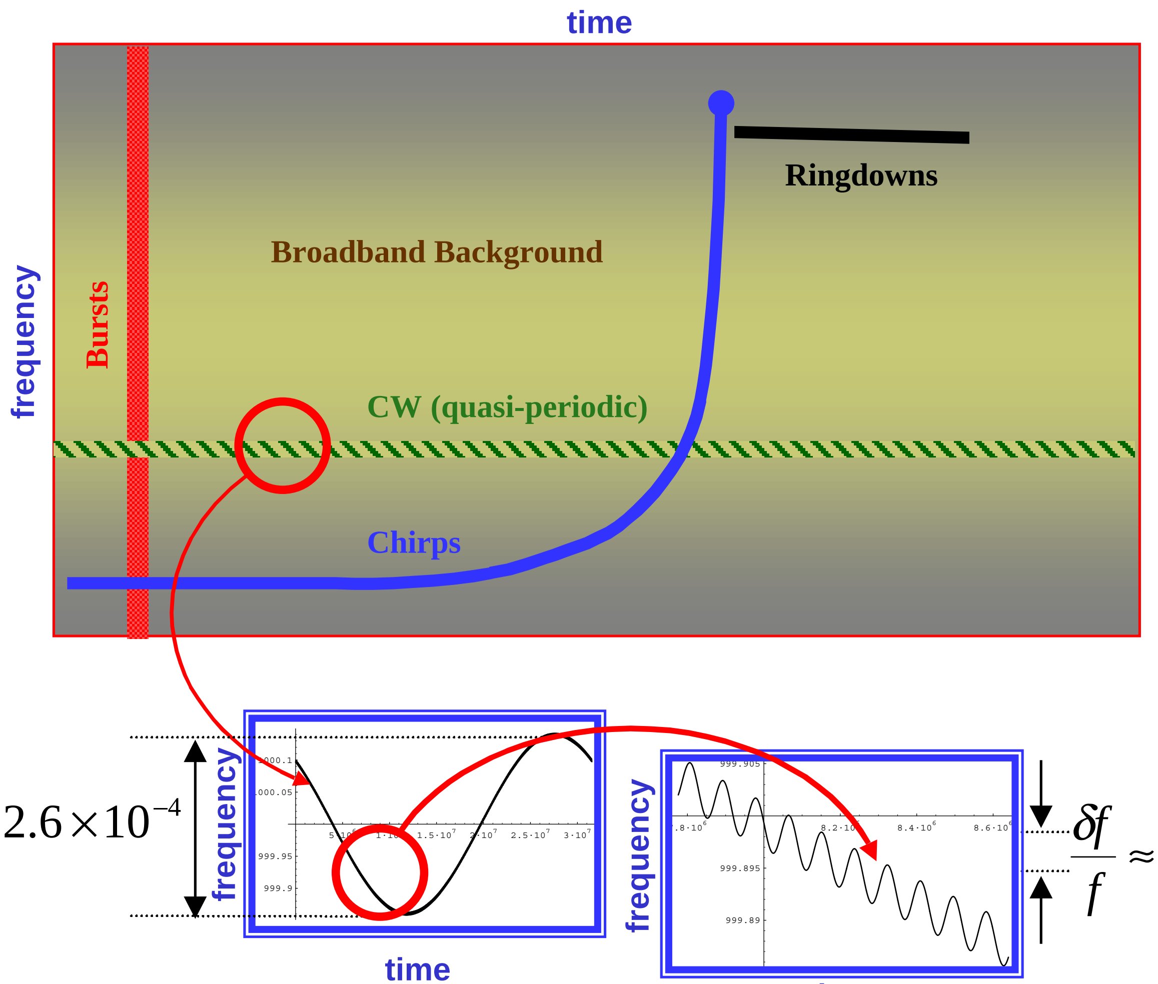

- **Periodic sources last "forever" have AM and FM due to motion the earth**
  - **Look for these signatures**
  - **Depend on sky position of source**
  - **Sources will have intrinsic variations ( spindown)**

# *Frequency-Time Maps ("Images")*




***SETI@home* uses frequency-time analysis methods to detect *unexpected* or *novel* features in otherwise *featureless "hiss"***

# Frequency-Time Characteristics of GW Sources




$$\frac{\delta f}{f} \approx 2.6\times10^{-4}$$

$$\frac{\delta f}{f} \approx 4\times10^{-6}$$

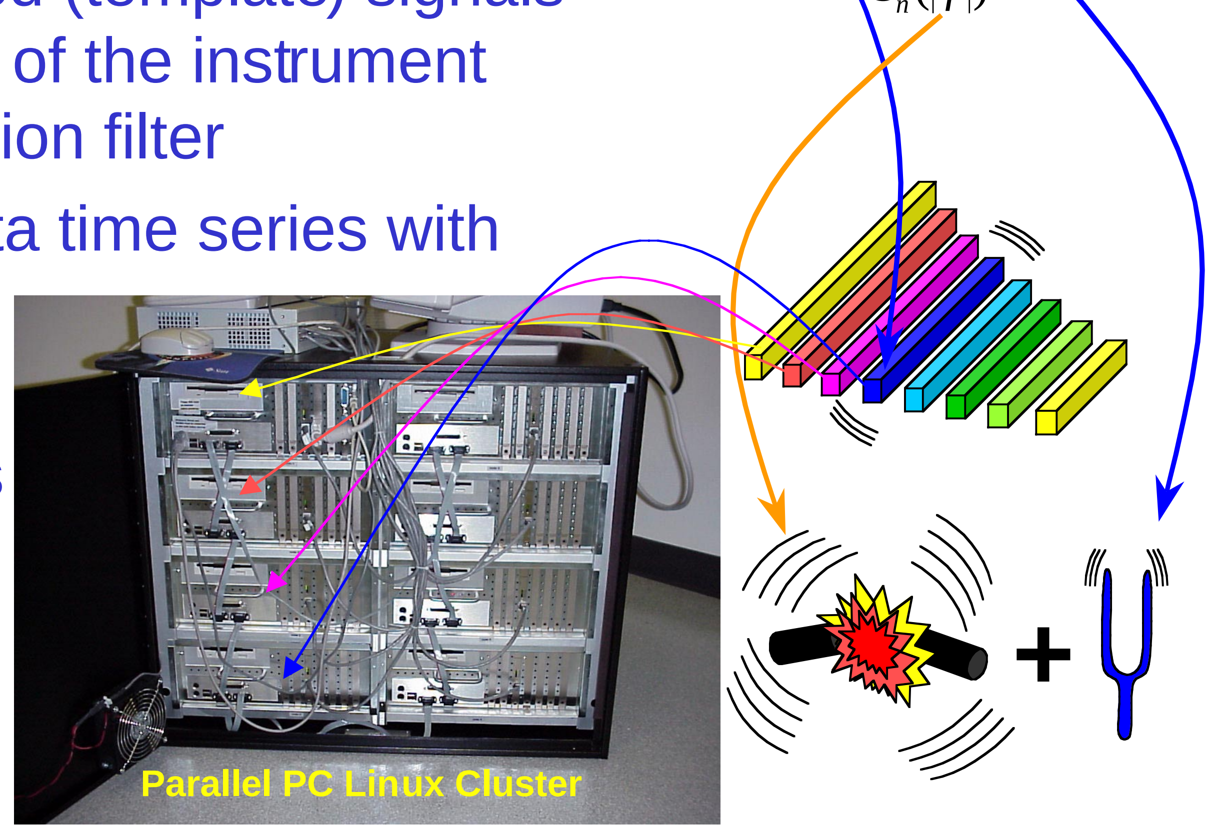
- **Bursts are short duration, broadband events**
- **Chirps explore the greatest time-frequency area**
- **BH Ringdowns expected to be associated with chirps**
- **CW sources have FM characteristics which depend on position on the sky (*and source parameters*)**
- **Stochastic background is stationary and broadband**

- **For each source, the optimal signal to noise ratio is obtained by integrating signal *along* the trajectory**
  - **If SNR >> 1, kernel ∝ |signal|^2**
  - **If SNR ≤ 1, kernel ∝ |template* signal| or $|signal_j^* \, signal_k|$**
  - **Optimal filter: kernel ∝ 1/(noise power)**

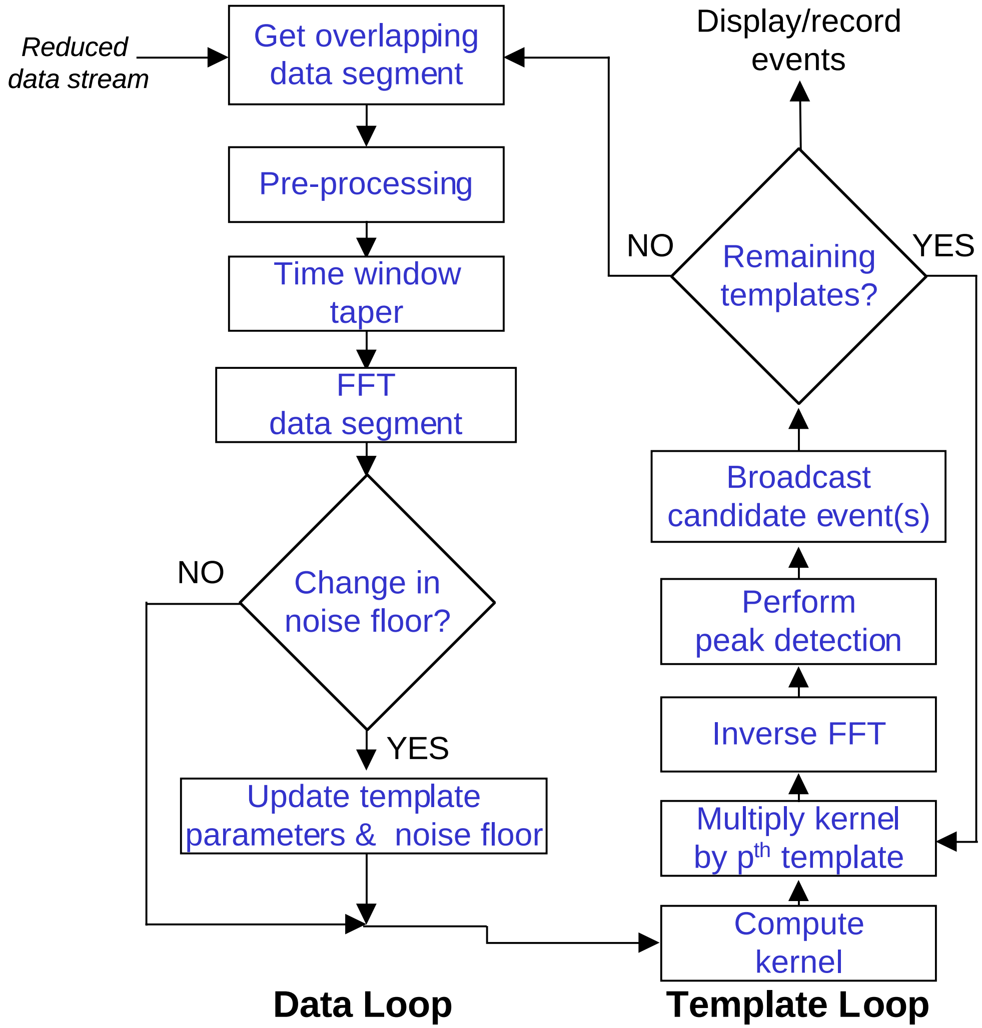
# Optimal Wiener Filtering

- Matched filtering (optimal) looks for best overlap between a signal and a set of expected (template) signals in the presence of the instrument noise -- correlation filter
- Replace the data time series with an SNR time series
- Look for excess SNR to flag possible detection

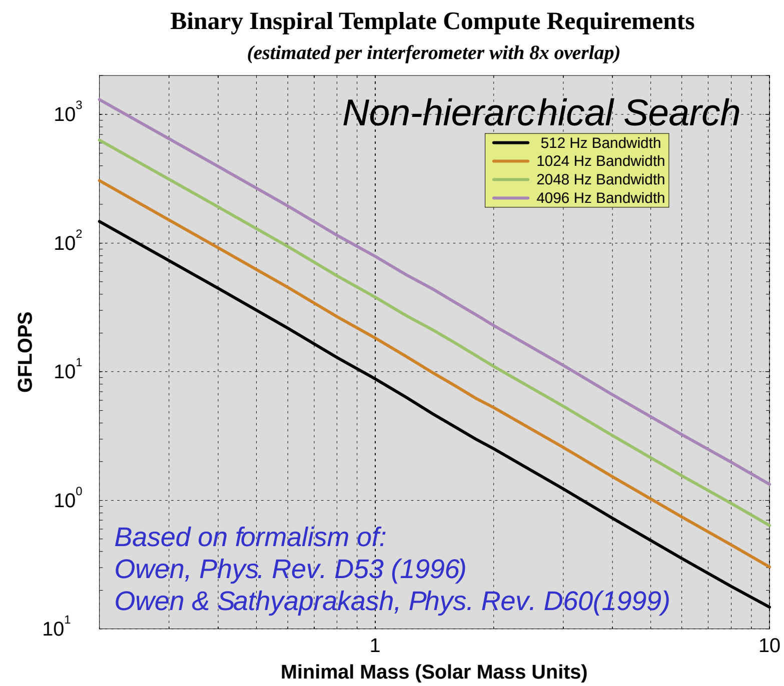
$$\xi_p[t_c] = \int \frac{\hat{T}_p^*(f)\ \hat{s}(f)}{\hat{S}_n(|f|)} e^{-2\pi i f t_c} df$$



# Compact Binary Inspirals
# Data Analysis Flow

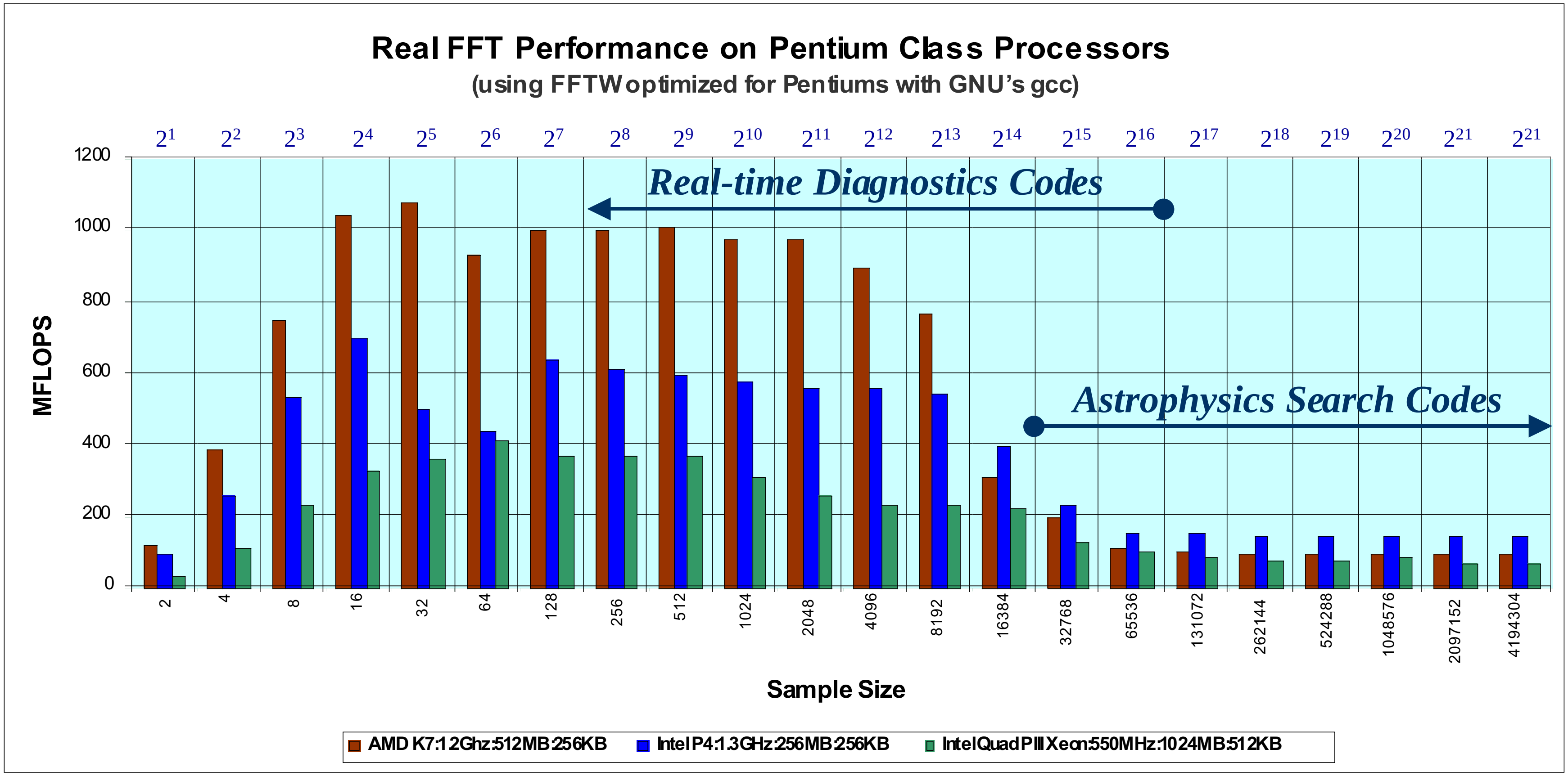



$$\xi_p[t_c] = 2\int \frac{\hat{T}_p^*(f)\ \hat{s}(f)}{\hat{S}_n(|f|)} e^{-2\pi i f t_c} df$$




- **Process data at real time rate**
- **Improvements:**
  - ***Hierarchical searches developed***
  - ***Phase coherent analysis of multiple detectors (Finn, in progress)***

# Recent CPU Node Performance

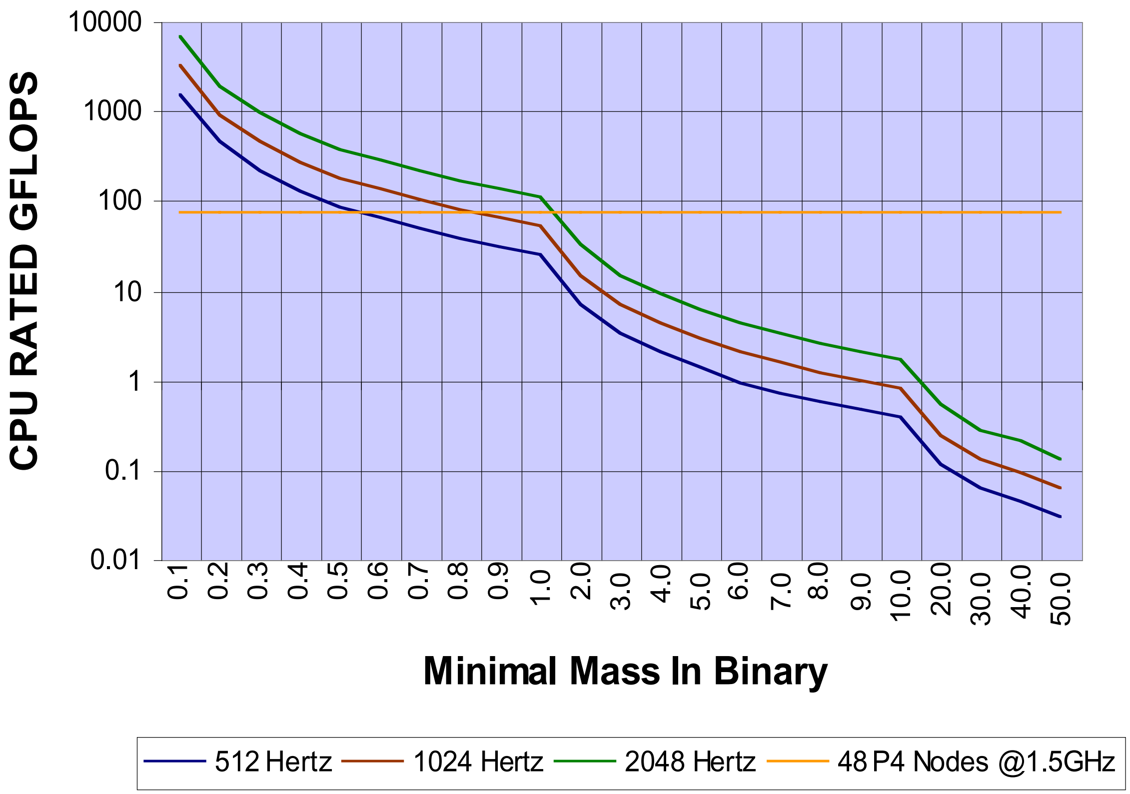



- **Pipeline analysis of LIGO data computationally dominated by cost of Fast Fourier Transforms (FFT).**
  - **Non-Hierarchical Binary Inspiral Search spends an average of ~90% of CPU cycles performing FFT.**
- **Most practical/efficient data segment size as much as $2^{20}$ points for Binary Inspiral Search.**

# *Computation of Templated Binary Inspiral Search*

**Computation for Binary Inspiral Search Using 1.5GHz P4's**

(8x overlap, no hierarchy, 10% detection loss, 10% efficiency for FFTs)



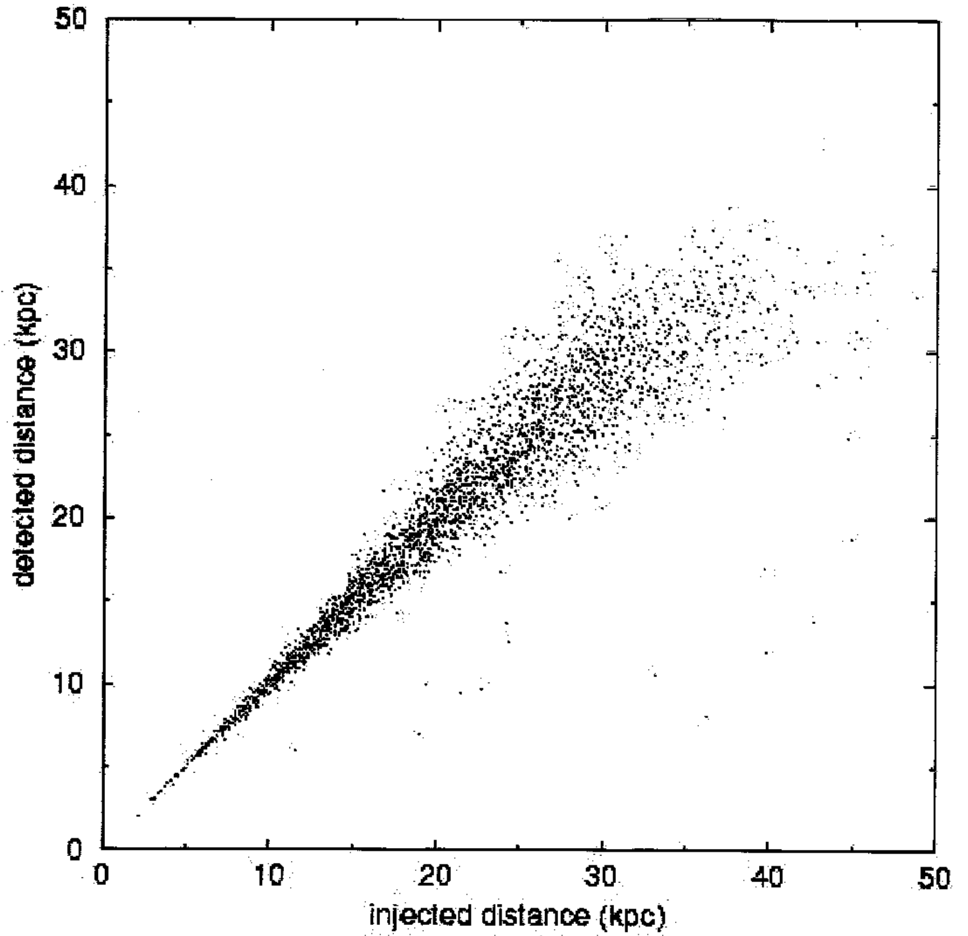

- **Non-Hierarchical Search for NS-NS Inspiral (1.4$M_S$, 1.4$M_S$) ~ {15, 32, 67} GFLOPS.**
- **512 Hertz band is adequate for detections (blue curve).**
- **Hierarchical strategies expected to decrease cost by 5x to 30x.**
- **Ringdown Search estimated to be roughly 10% as costly.**
- **Other searches (excluding all-sky pulsar search) are single node compute problems.**

# *Detection Efficiency*

***Monte Carlo (Statistical) techniques are needed to characterize complex detection probabilities***

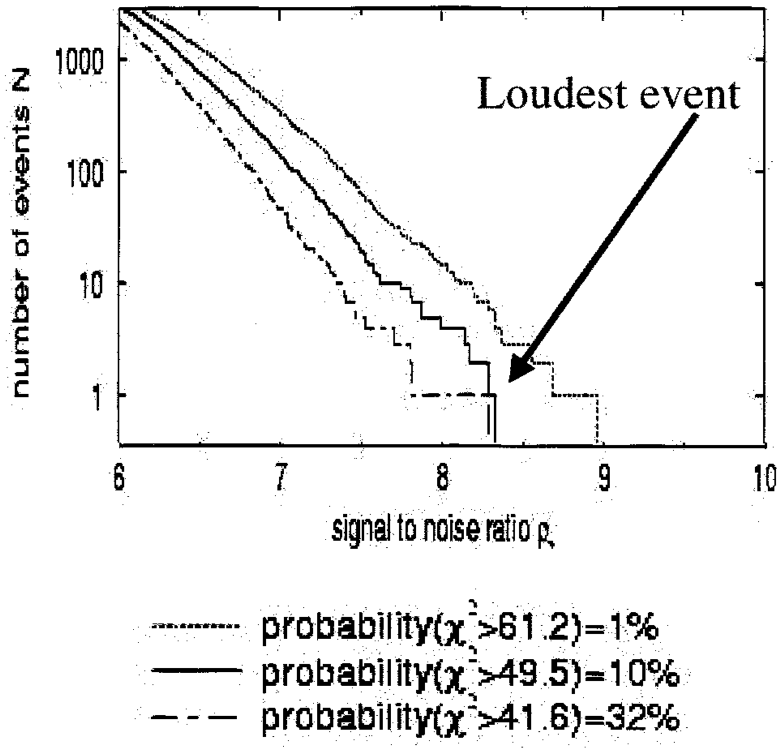
- ***Simulated inspiral events provide end to end test of analysis and simulation code for reconstruction efficiency***

- ***Errors in distance measurements from presence of noise are consistent with SNR fluctuations***



# *Setting a limit*

***Quantitative Science: making a probabilistics statement about the likelihood of an observation (or lack thereof)***




**Upper limit on event rate can be determined from SNR of 'loudest' event**

**Limit on rate:**

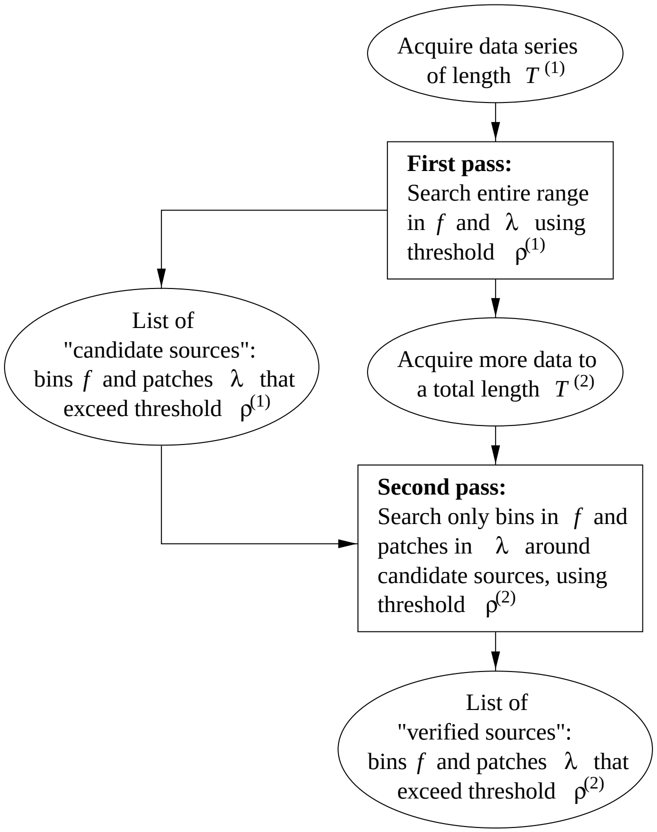
**R < 0.5/hour with 90% CL**

**ε = 0.33 = detection efficiency**

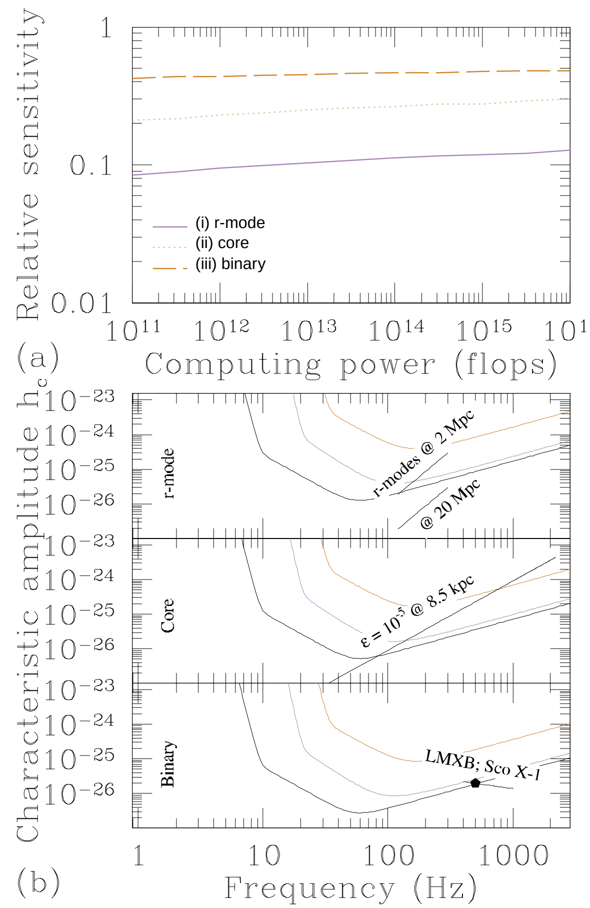
**An ideal detector would set a limit:**

**R < 0.16/hour**

# *CW Sources*
# *Hierarchical (Constrained) Search*




*ref: Brady & Creighton gr-qc/981204*



# Search Approaches for Other GW Sources

- Burst events (unmodeled)
  - *Cross correlate detector outputs over narrow time window*
  - *Look for excess power*
  - *Use environmental vetoes*
  - *Look for few parametric templates (e.g., wavelets)*
  - *CPU: Workstation(s)*

$$S(t) = \iint_0^{T_B} dt'\, dt''\, s_1(t-t')Q(t''-t')s_2(t-t'')$$

$$Q(\tau) = \int df \quad e^{-2\pi i f \tau} \frac{\hat{h}^*_{B,1}(f)\hat{h}_{B,2}(f)}{S_1(|f|)S_2(|f|)}$$

*ref: Finn, Mohanty, Romano gr-qc/9903101*

- Stochastic background search
  - *Correlate & integrate signals from pairs of interferometers*
  - *Look for excess power in band consistent with baseline separation*
  - *CPU: Workstation(s)*

$$S_{(\alpha)} = T_{\text{int}} \int_{f_{\min}}^{f_{\max}} \frac{\hat{s}_1^*(f)\hat{s}_2(f)\Omega_{GW}^{(\alpha)}(|f|)\gamma(|f|)}{f^3 S_1(|f|)S_2(|f|)} df$$

*ref:*
*Michelson, 1987*
*Christensen, 1992*
*Flannigan, 1993*
*Allen & Romano gr-qc/9710117*

# Software Block Diagram

LIGO-T990101-E
LAL Specification

LIGO-T990097-E
wrapperAPI Requirements

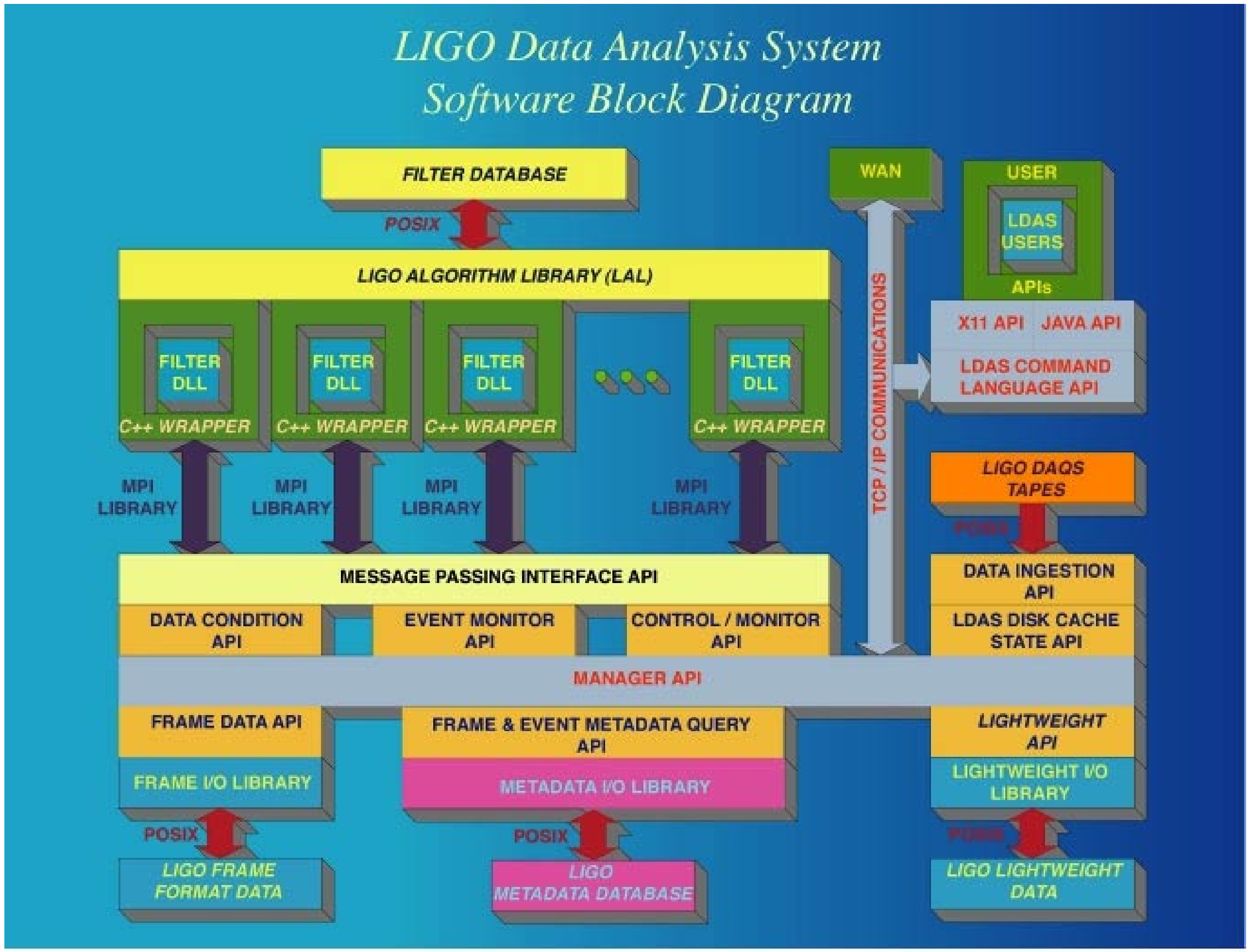
www.lam.org
LAM version 5.6.1

LIGO-T990086-E
mpiAPI Requirements

LIGO-T990002-E
dataConditionAPI Requirements
FFTW version 2.3.1
CLAPACK version

LIGO-T980115-E
managerAPI Requirements

LIGO-T980117-E
frameAPI Requirements




GUILD
TCL/TK GUI

UNIX Sockets (TELNET)
User Commands (TCL)

LIGO-T010052-E
dataIngestionAPI Requirements
"In Development"

LIGO-T010051-E
diskCacheAPI Requirements
"In Development"

LIGO-T000026-E
controlMonitorAPI Requirements

LIGO-T990037-E
lightWeightAPI Requirements

www.apache.org
C++ XML Parser Library

LIGO-T970130-E
LIGO-VIRGO Frame Specification
FrameCPP Version 4

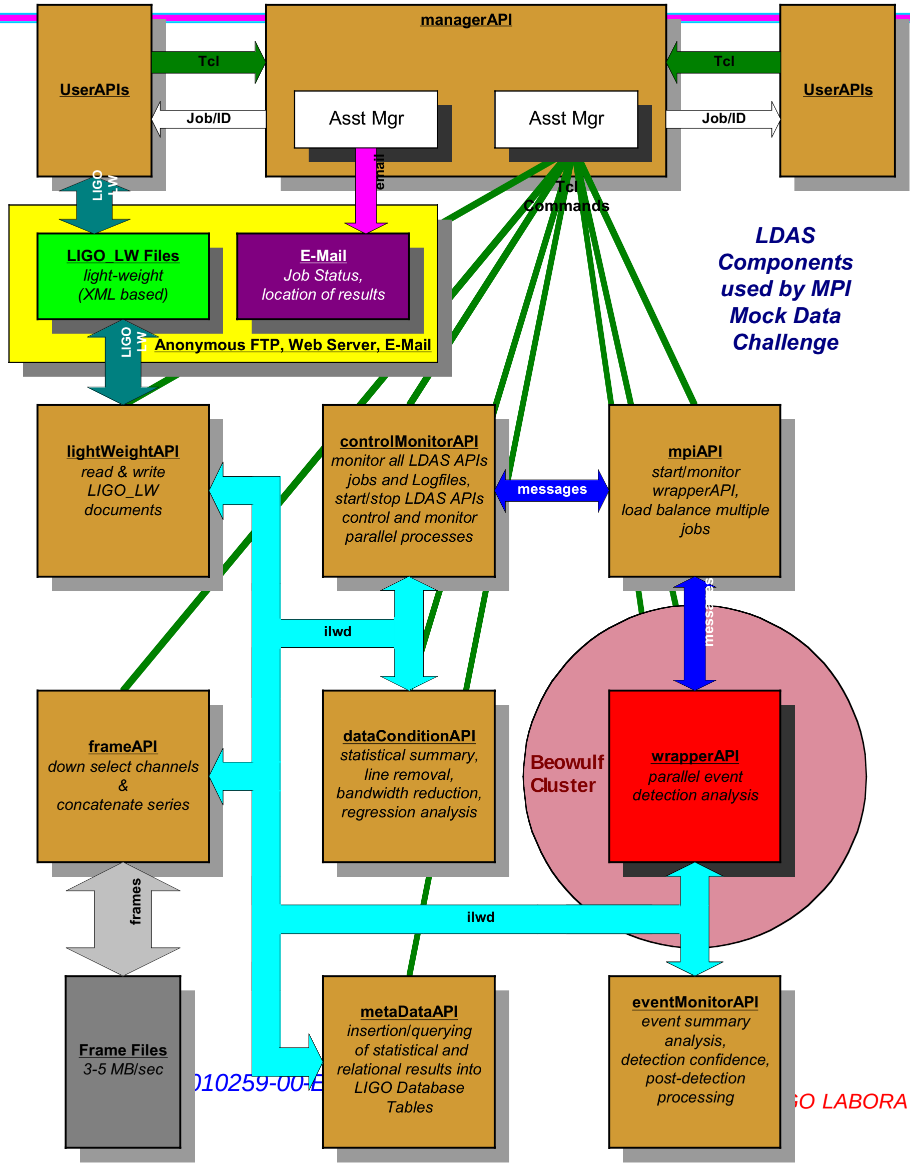
LIGO-T980119-E
metaDataAPI Requirements

dbEasy Library
ODBC Level 3

LIGO-T990101-E
LDAS Database Tables

LIGO-T990023-E
LIGO-Lightweight Format

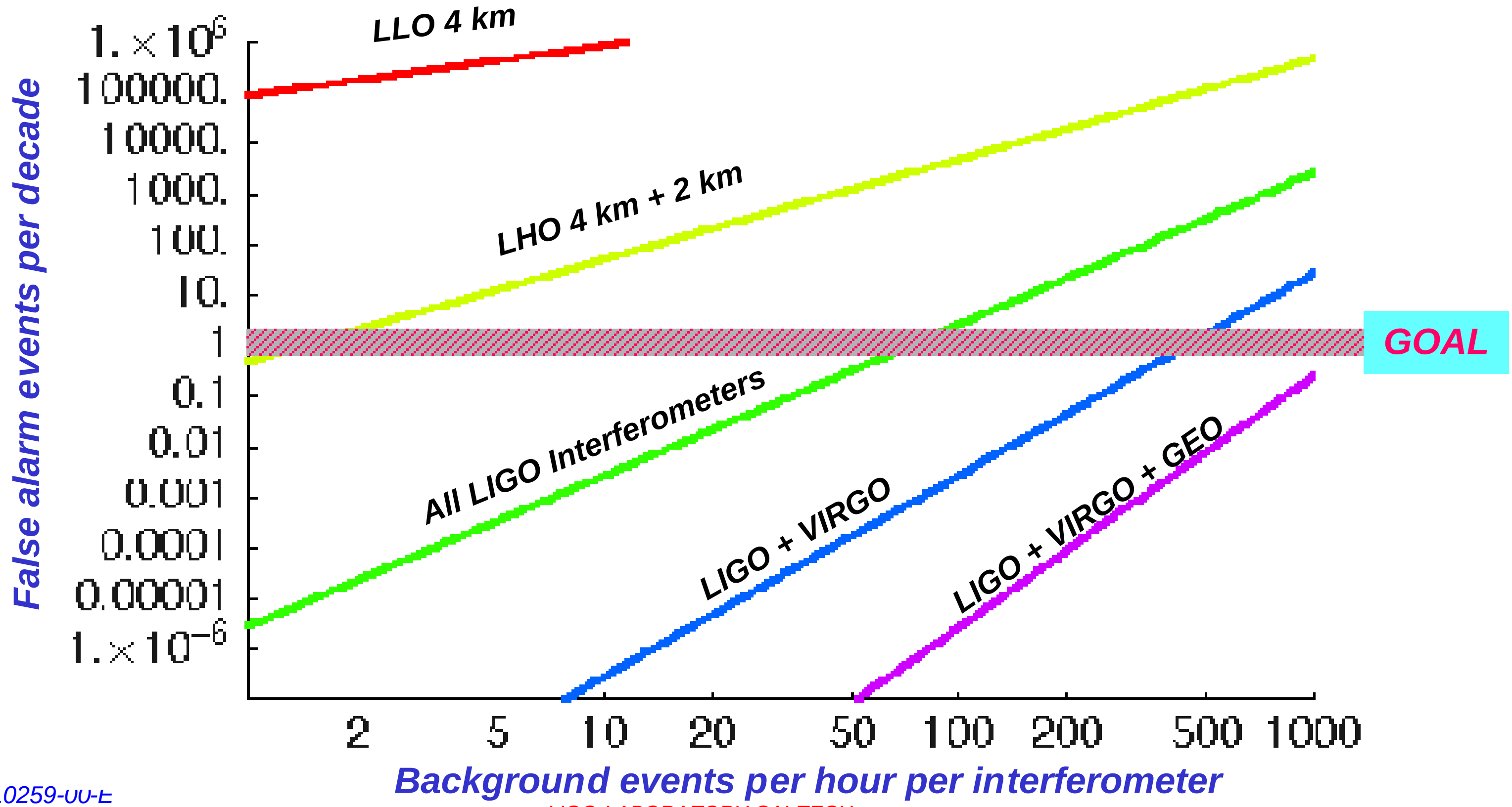
# Simple Inspiral Pipeline Example



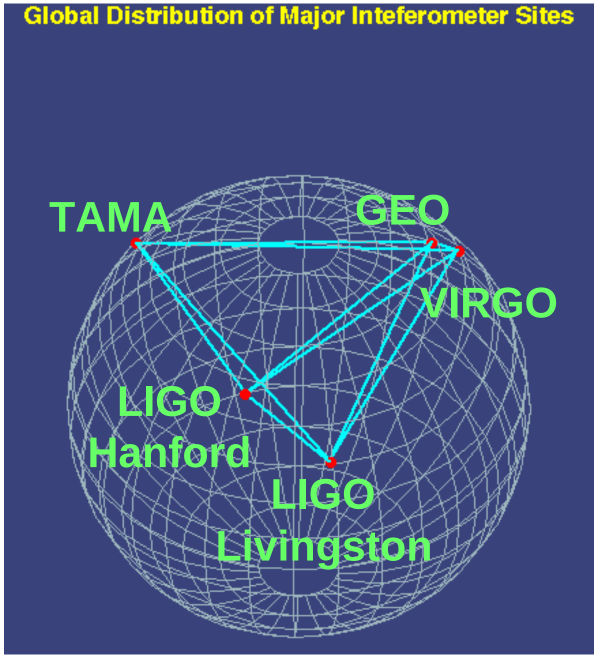

- **UserAPI** connects to **managerAPI** and starts pipeline.
- **managerAPI** assigns *assistant manager* to control job.
- **frameAPI** reads *frames* into the LDAS system, down-selects and sends data to the **dataConditionAPI**.
- **dataConditionAPI** pre-conditions data and forwards this to the **wrapperAPI** under the control of **mpiAPI**.
- **wrapperAPI** performs template based search using a dynamically loaded shared object.
- **eventMonitorAPI** receives search results and directs events to the **metaDataAPI** where they are placed in database.
- **eventMonitorAPI** sends user data products from search to **lightWeightAPI** where they are sent to user via FTP or URI commands.
- **managerAPI** notifies user via email of job completion.

# Coincidence windows among detectors

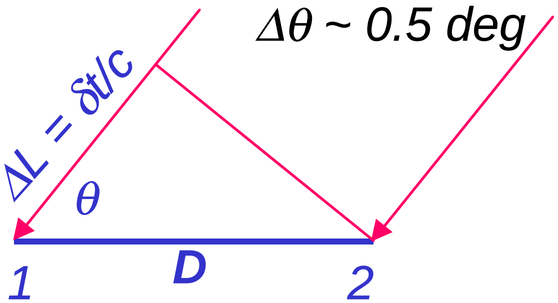
- Rejection of statistically uncorrelated random events
  - *Coincidence window duration determine by baselines*

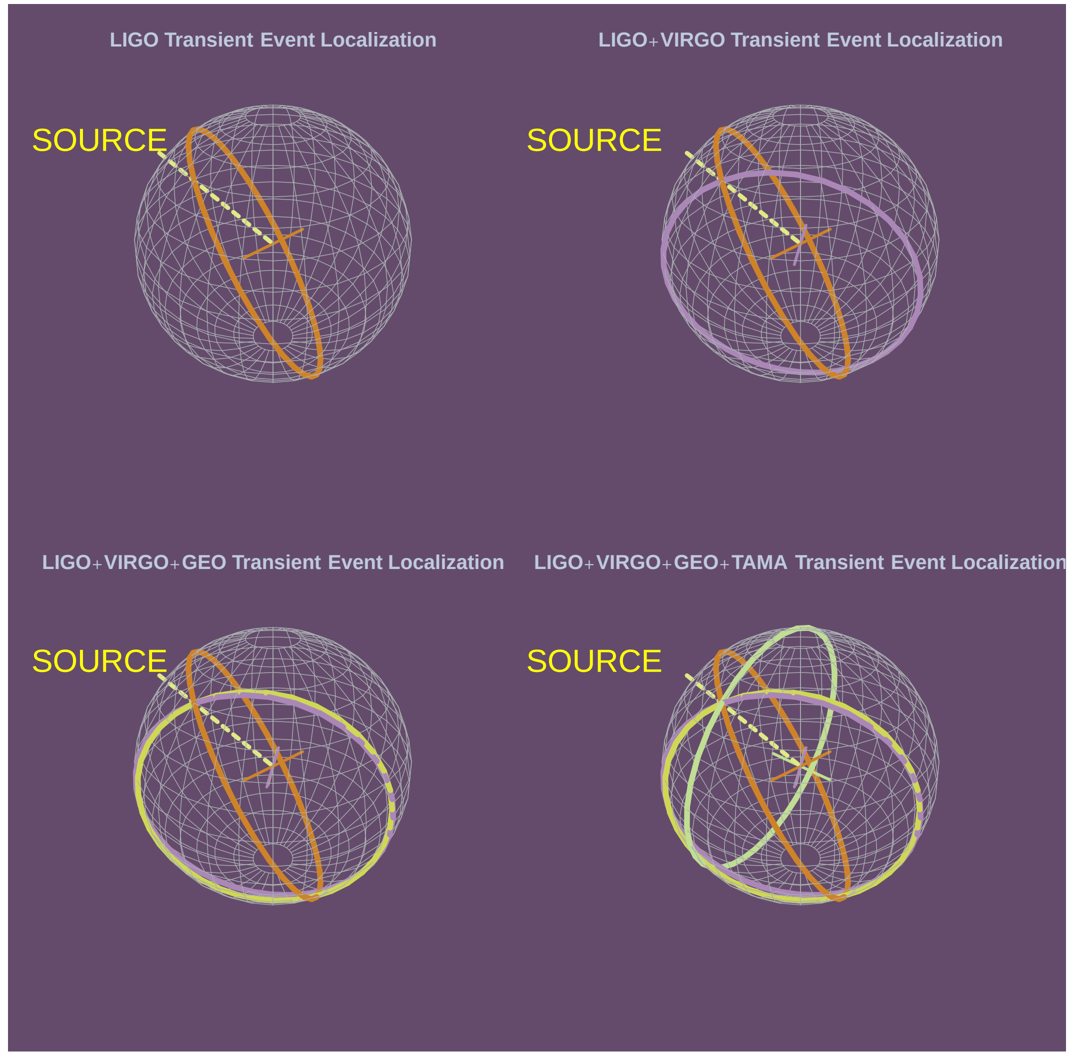


# *Event Localization With An Array of GW Interferometers*




$$\cos\theta = \delta t / (c\, D_{12})$$
$$\Delta\theta \sim 0.5 \text{ deg}$$



# Joint Data Analysis Among GW projects
# From detection to validation

- For a *putative* detection:
  - Environmental, instrumental vetoes?
  - ($\Delta t_i$, $\Delta\Omega_i$ ): Seen by all detectors within consistent (time, position) windows?
  - $\Delta h_i$: Is the amplitude of the signal consistent among detectors*?
  - $\Delta\alpha_i$: Are the deduced model parameters consistent?
- Follow up analyses
  - Independent
  - Coherent multi-detector analysis - maximum likelihood over all detectors: $\{t,\Omega,h,\alpha\}$

$$h_i \rightarrow \vec{h}$$

$$\ln\Lambda(h_i,\theta_i) \rightarrow \ln\Lambda(\vec{h},\vec{\theta})$$

$$\sigma_i^2 \rightarrow C_{kl} \equiv \langle \vec{n}_k \otimes \vec{n}_l \rangle$$

- Discrepancies should be explainable, e.g.:
  - Not on line
  - Below noise floor
  - *Different polarization sensitivity, *etc.*