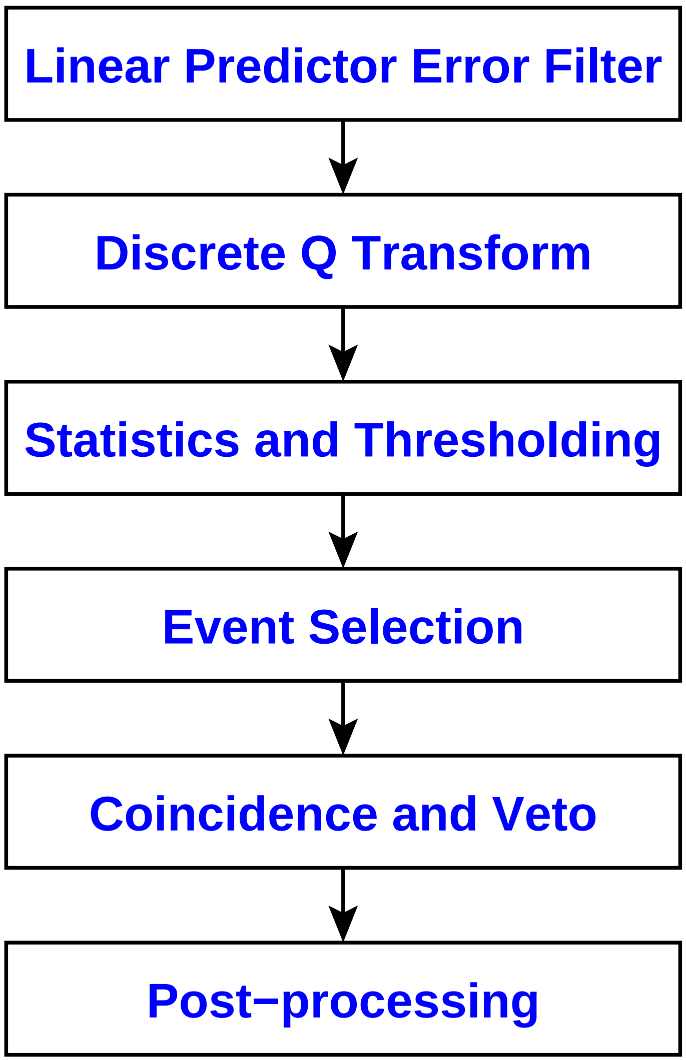
New tools for gravitational-wave burst data analysis

Shourov K. Chatterji

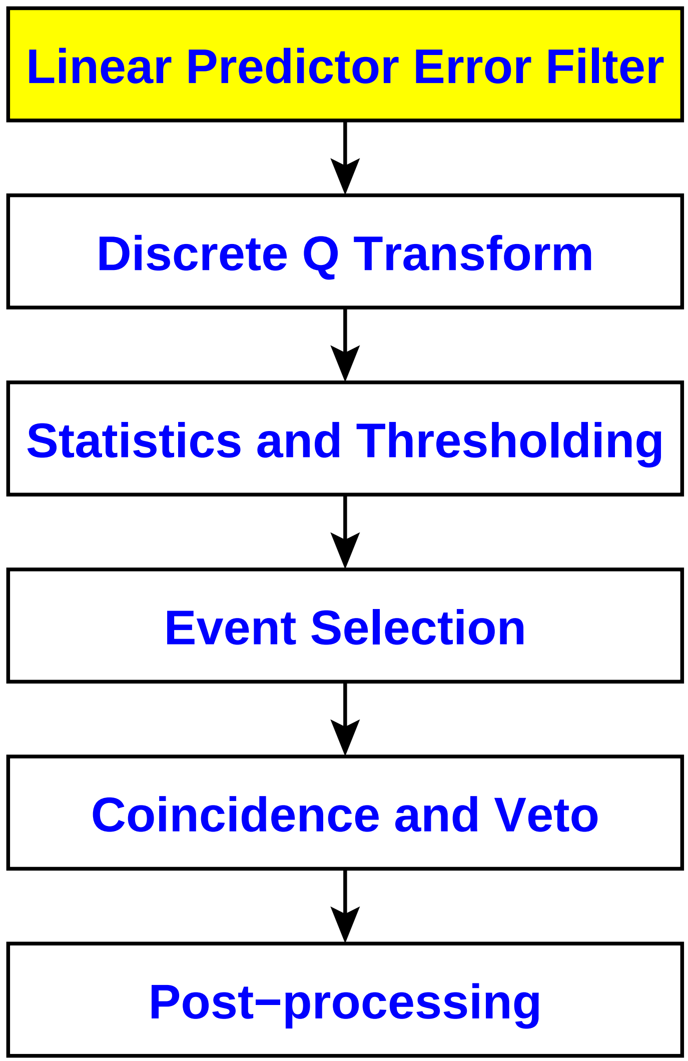
Massachusetts Institute of Technology

Presentation Overview



- Analysis Pipeline
- Pipeline Tuning
- Detection Efficiency
- Black Hole Mergers
- Current Status
- Future Plans

Linear Predictor Error Filter



- Removes predictable signal content
- Whitening
- Line removal
- Simplifies statistics
  - Time-frequency
  - Cross-correlation

LIGO

LPEF Definition

- Linear Prediction: Assume each sample is a linear combination of the previous M samples.

$$\tilde{x}[n] = \sum_{m=1}^{M} c[m]x[n-m]$$

- Prediction Error: We are interested in the unpredictable signal content.

$$e[n] = x[n] - \tilde{x}[n]$$

- Training: Choose $c[m]$ to minimize the mean squared prediction error.

$$\sigma_e^2 = \frac{1}{N}\sum_{n=1}^{N} e[n]^2$$

LIGO

LPEF Computation

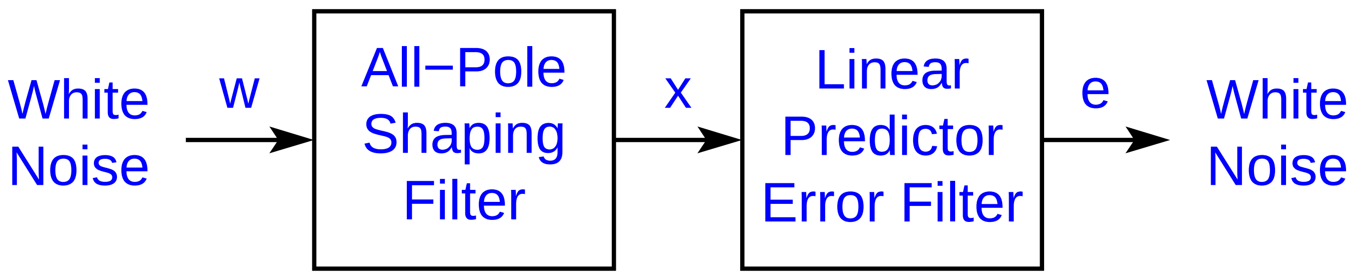
- Linear least squares optimal filter problem well known
- Yields Yule-Walker equations

$$\sum_{m=1}^{M} c[m]r[m-k] = r[k] \quad \text{for} \quad 1 \leq k \leq M$$

- Robust efficient algorithms exist to train and apply
- FFT allows computation of signal autocorrelation in order $N \log N$
- Levinson-Durbin recursion solves Yule-Walker equations in order M^2
- Block filtering using FFT allows application of the filter in order $N \log N$

LPEF Properties

- Compensates exactly for all-pole filters



- In general, compensation is not exact
- Filter order, M, can compensate for features

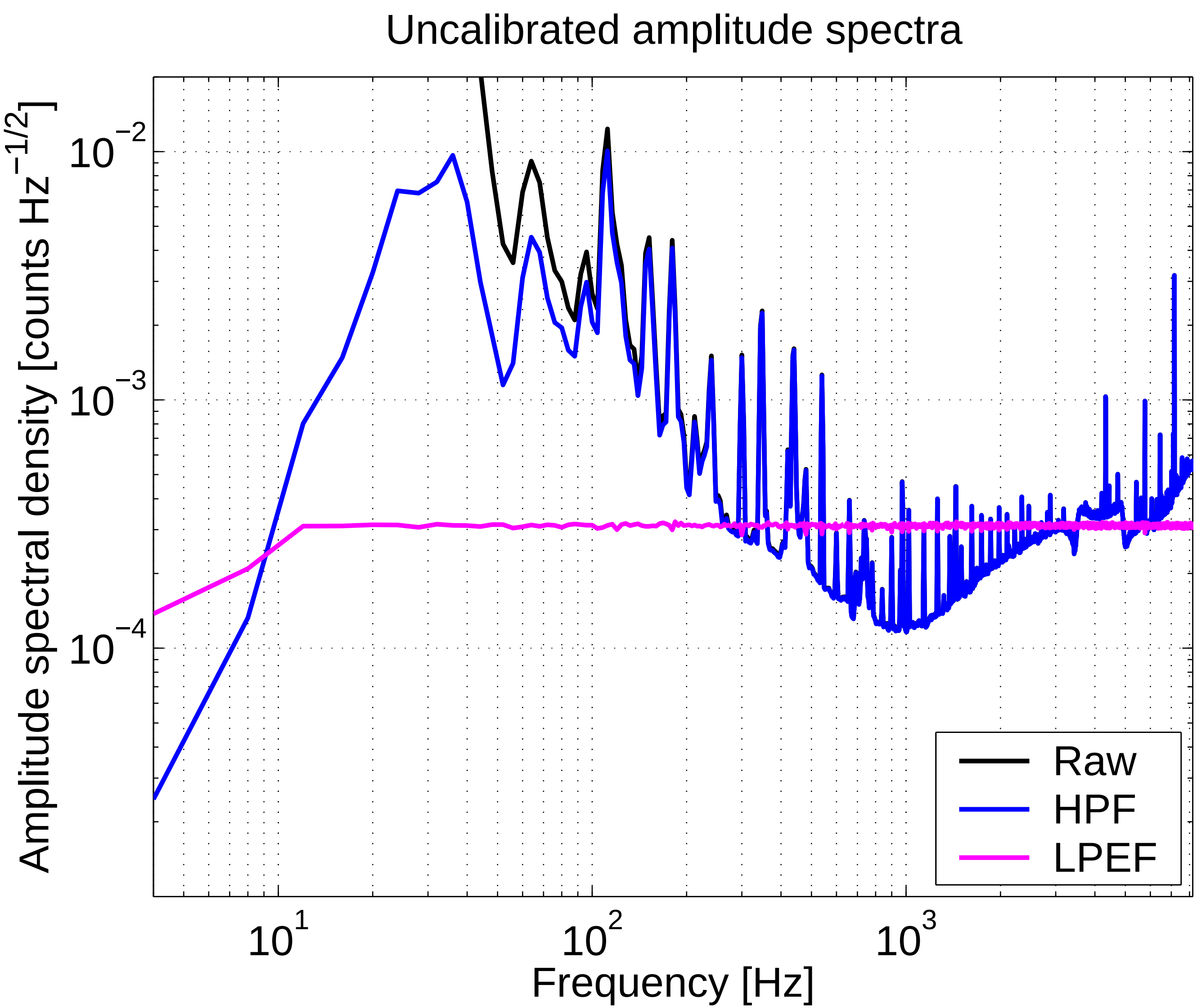
$$\Delta f \gtrsim f_s/M$$

- Training length, N, can learn about features

$$\Delta f \gtrsim f_s/N$$

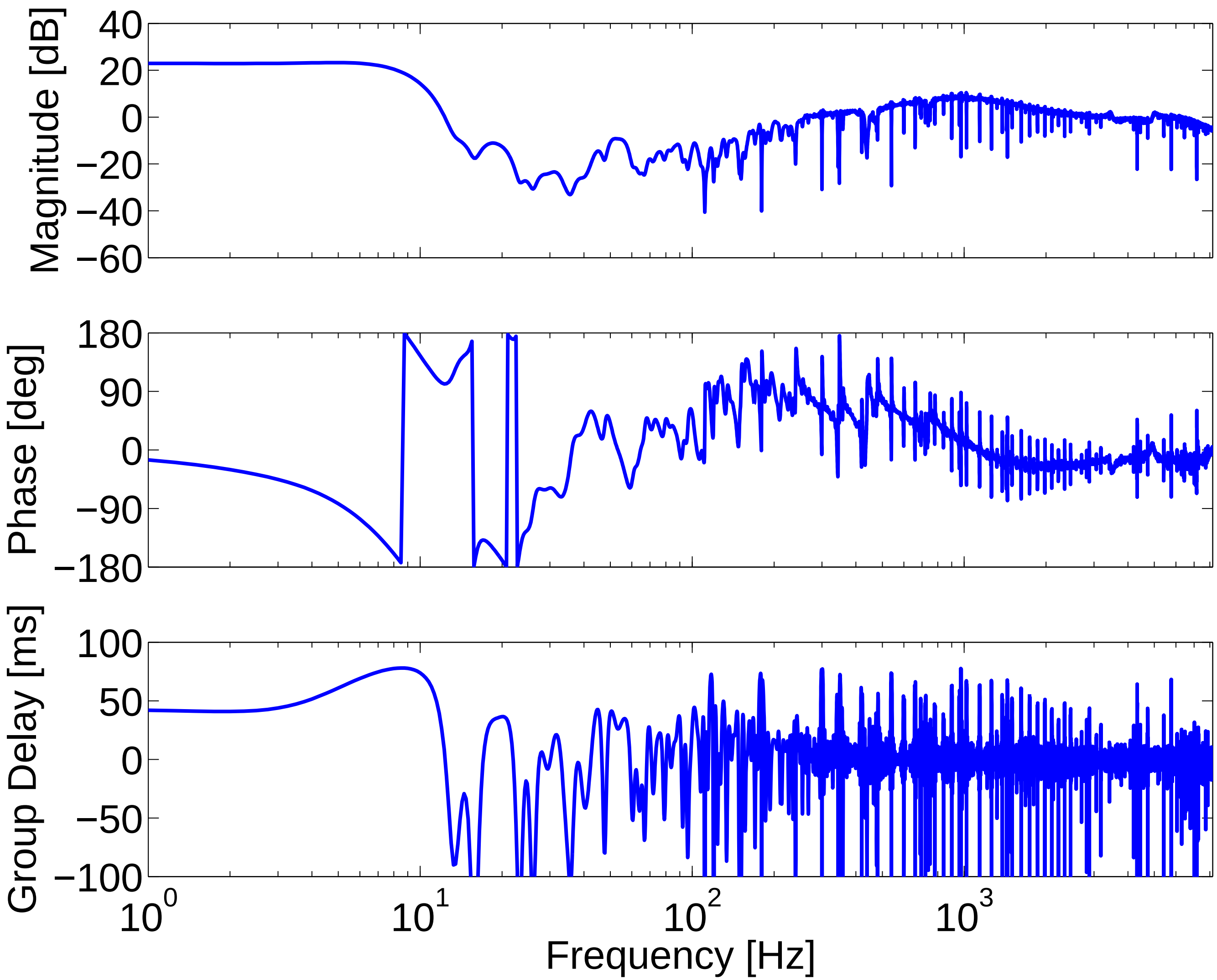
LPEF Whitened Spectra

Amplitude spectra of S2 H1 data after whitening by LPEF



LIGO LPEF Frequency Response

Frequency response of LPEF trained on S2 H1 data

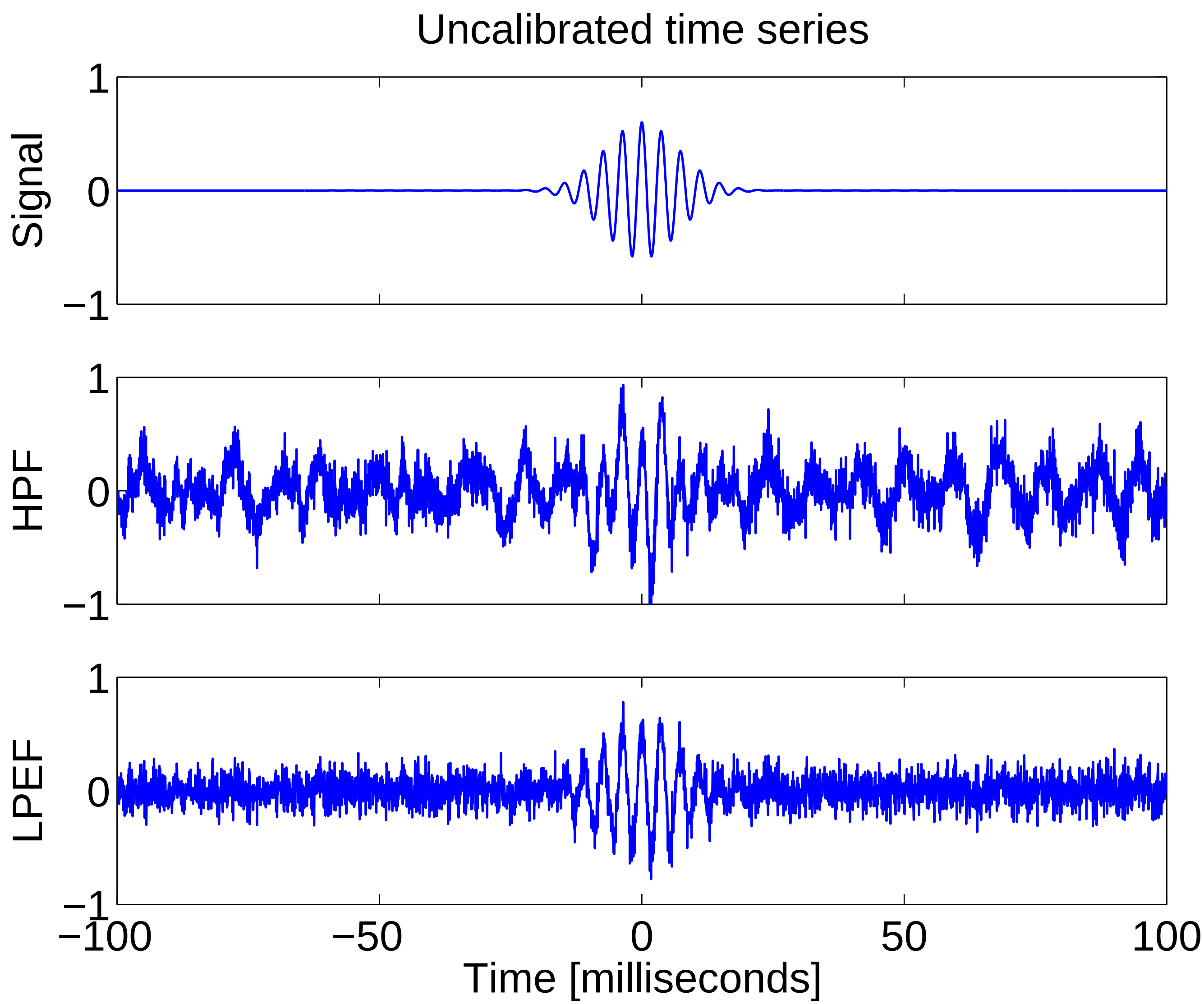


Zero-Phase LPEF

- Yule-Walker solution is minimum phase
- Problem: Unkown phase response produces error coincidence time delay determination
- Solution: Symmetric FIR filters have linear phase (causal) or zero-phase (acausal)
- Form a symmetric FIR filter by auto-correlation of LPEF coefficients (equivalent to forward and reverse filtering)
- Problem: Magnitude response of auto-correlation of LPEF coefficients is square of magnitude response of LPEF coefficients
- Solution: First, find new filter with approximate square root response
  FFT $\longrightarrow$ complex square root $\longrightarrow$ inverse FFT

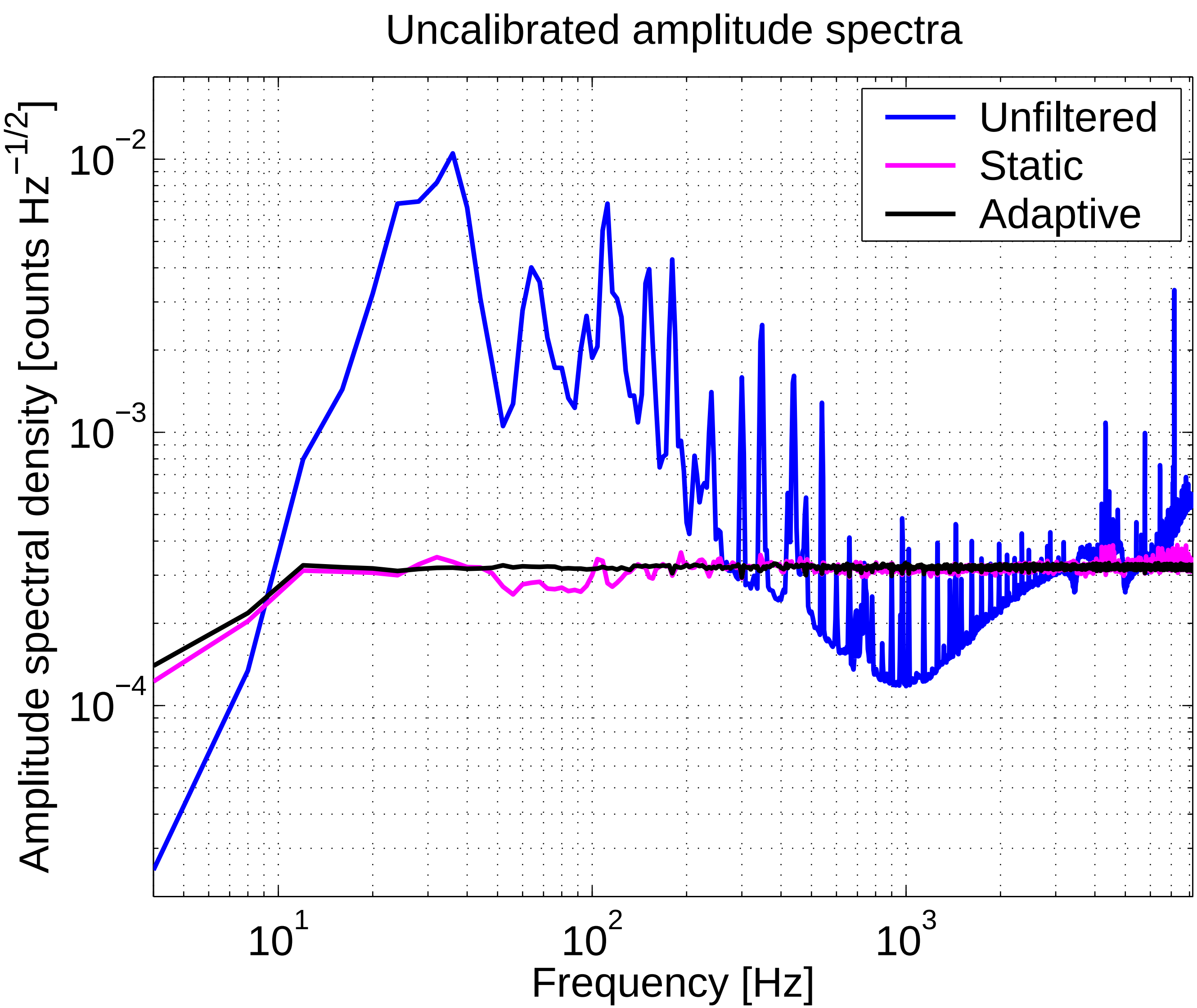
LPEF Time Series

Effect of zero-phase LPEF on simulated Sine-Gaussian burst



LPEF Stationarity

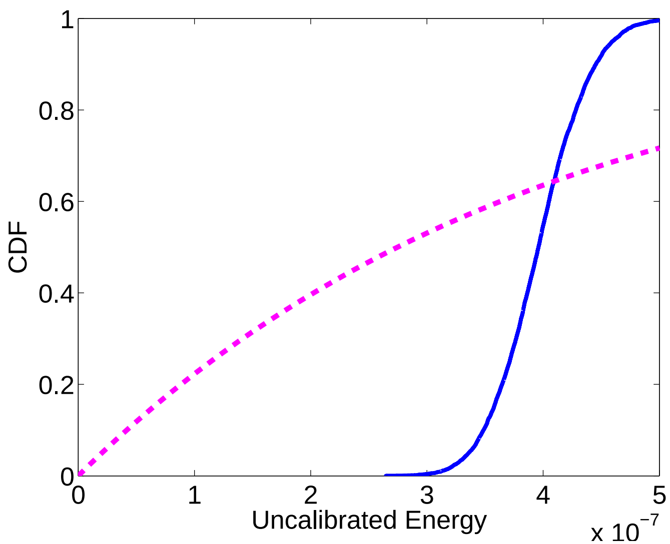
Whitening performance on S2 H1 data after 45 minutes



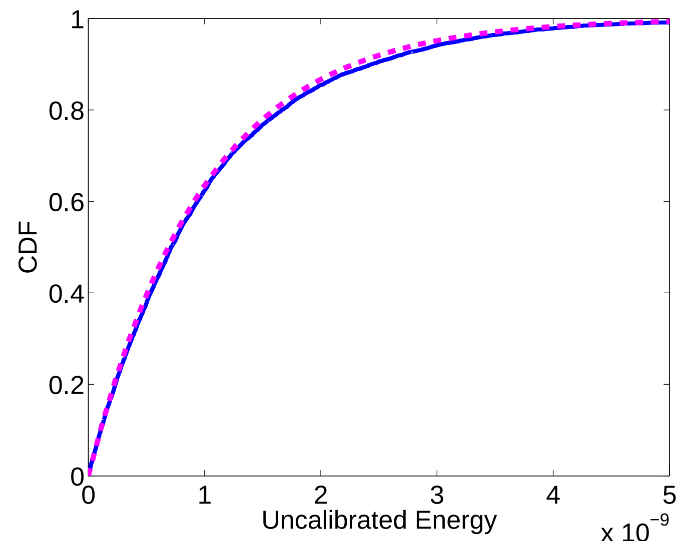
LPEF Statistics

The distribution of time-freqeuncy pixel energies is Rician in general, but exponential after linear predictor error filtering.

Before LPEF



After LPEF



Blue: Empirical CDF of pixel energies

Magenta: Exponential CDF with same mean

LPEF Cross-Correlation

What is the effect of linear predictor error filtering on cross-correlation analysis? Consider the observation of a gravitational wave burst by two interferometers:

$$
\begin{aligned}
x_1(t) &= b_1(t) * [g_1(t) * h_1(t) + n_1(t)] \\
x_2(t) &= b_2(t) * [g_2(t) * h_2(t) + n_2(t)]
\end{aligned}
$$

$h_i(t)$	-	incident gravitational wave burst
$g_i(t)$	-	interferometer impulse response
$n_i(t)$	-	additive detector noise
$b_i(t)$	-	linear predictor impulse response
$x_i(t)$	-	observed time series

LIGO

LPEF Cross-Correlation

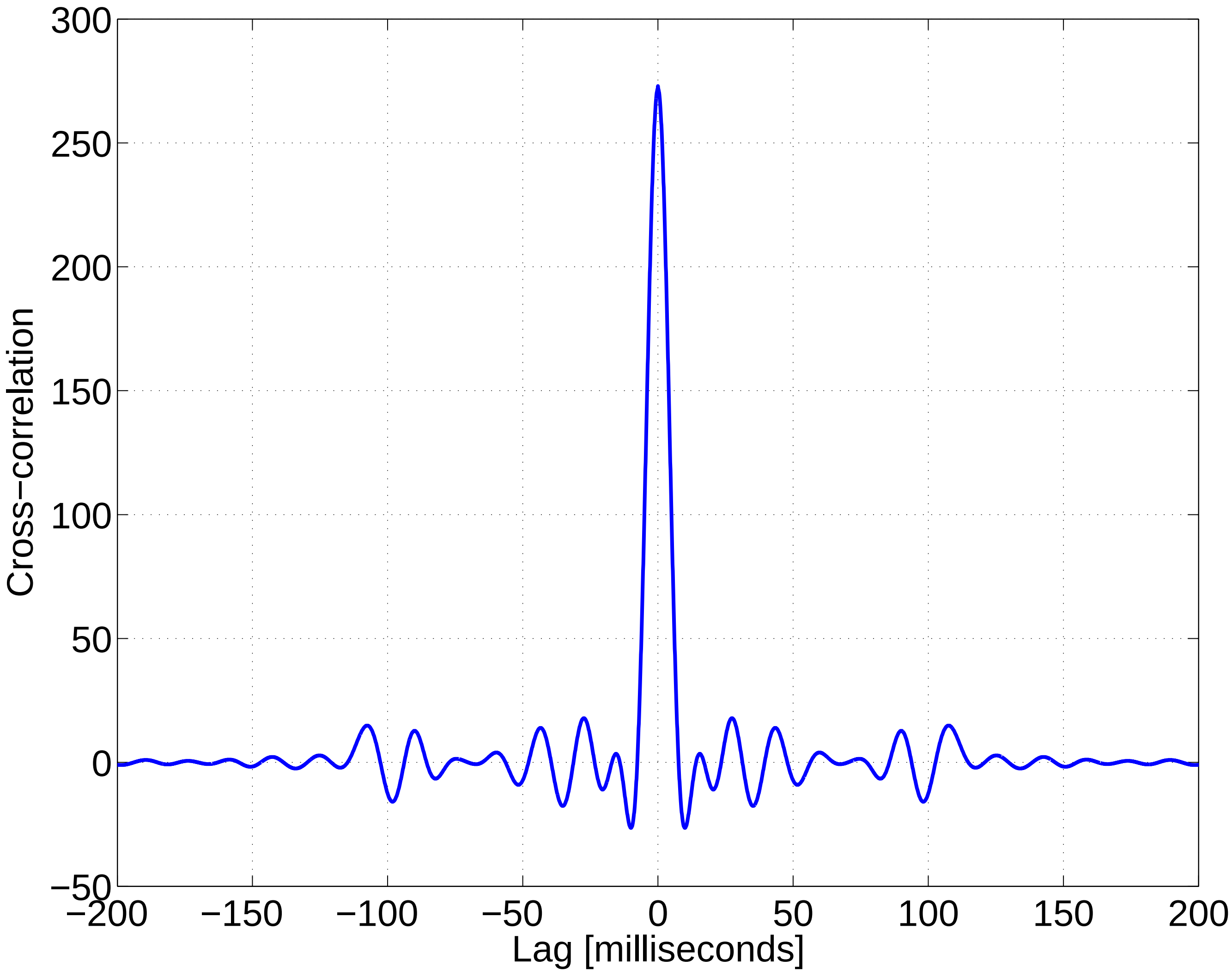
Cross-correlate the two observations:

$$r_{x_1,x_2}(\tau) = r_{b_1,b_2}(\tau) * [r_{g_1,g_2}(\tau) * r_{h_1,h_2}(\tau) + r_{n_1,n_2}(\tau) + \ldots]$$

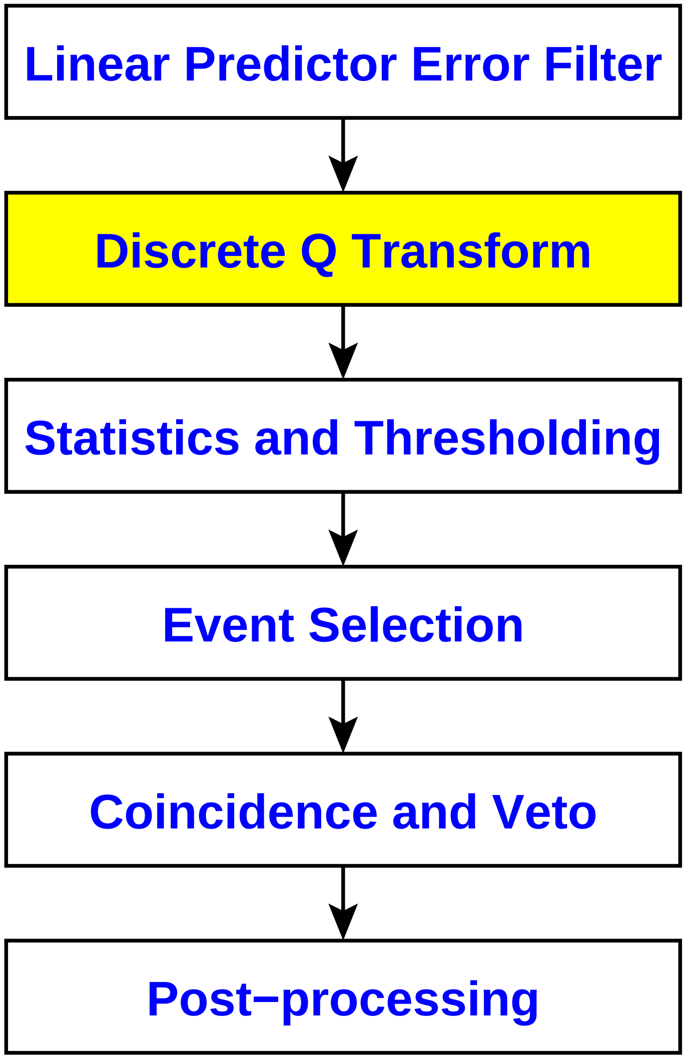
- Assume the gravitational wave burst and detector noise are uncorrelated.
- The cross-correlation of detector noise is the dominant term inside the brackets.
- The linear predictor error filter is trained to minimize the detector noise term.
- The desired result is “blurred” by convolving with the cross-correlated interferometer impulse responses and cross-correlated linaer predictor error filter coefficients.

Cross-Correlation Example

Cross-correlation of S2 H1 and L1 zero-phase LPEF coeffi cients



Discrete Q Transform



- Multi-resolution search for time-frequency excess power
- Targets a specific range of Q
- Achieves the optimal signal to noise ratio measurement

Burst Parameters

Burst “energy” and normalized burst waveforms:

$$\int_{-\infty}^{+\infty} |h(t)|^2 \, dt = \int_{-\infty}^{+\infty} |\tilde{h}(f)|^2 \, df = h_{\text{rss}}^2 \qquad h(t) = h_{\text{rss}} \psi(t)$$

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 \, dt = \int_{-\infty}^{+\infty} |\tilde{\psi}(f)|^2 \, df = 1 \qquad \tilde{h}(f) = h_{\text{rss}} \tilde{\psi}(f)$$

Central time, central frequency, duration, bandwidth, and Q:

$$t_0 = \int_{-\infty}^{+\infty} t |\psi(t)|^2 \, dt \qquad f_0 = 2 \int_{0}^{+\infty} f |\tilde{\psi}(f)|^2 \, df$$

$$\sigma_t^2 = \int_{-\infty}^{+\infty} (t - t_0)^2 |\psi(t)|^2 \, dt \qquad \sigma_f^2 = 2 \int_{0}^{+\infty} (f - f_0)^2 |\tilde{\psi}(f)|^2 \, df$$

$$\sigma_t \sigma_f \geq \frac{1}{4\pi} \qquad Q = \frac{f_0}{\sigma_f}$$

Multiresolution Analysis

Optimal time-frequency signal to noise ratio measurement:

$$\rho^2 = \int_0^\infty \frac{2|\tilde{h}(f)|^2}{S_h(f)}\, df \simeq \frac{h^2_{\text{rss}}}{S_h(f_c)}$$

This is only obtained if the measurement pixel exactly matches the signal:

- Maximal burst energy in pixel
- Minimal background energy in pixel

Characterize "bursts" as signals with $Q \lesssim 10$

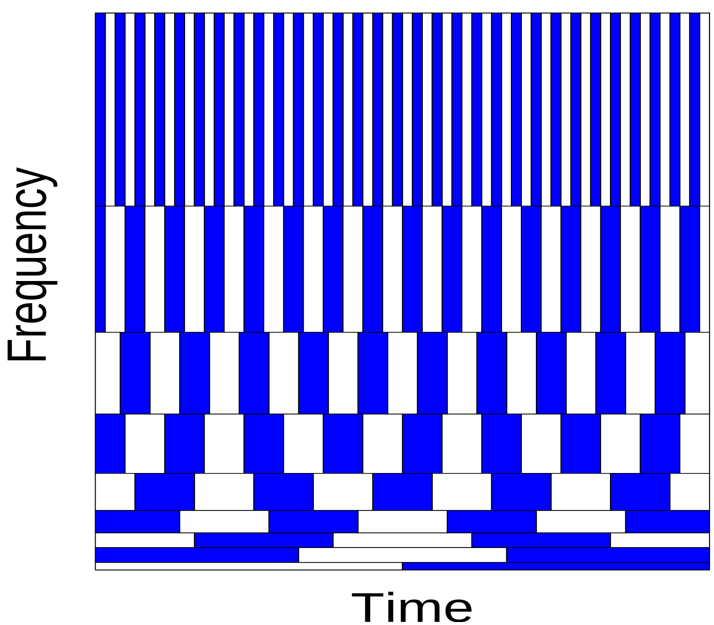
Tile the time frequency plane to maximize the detectability of bursts within a specific range of Q

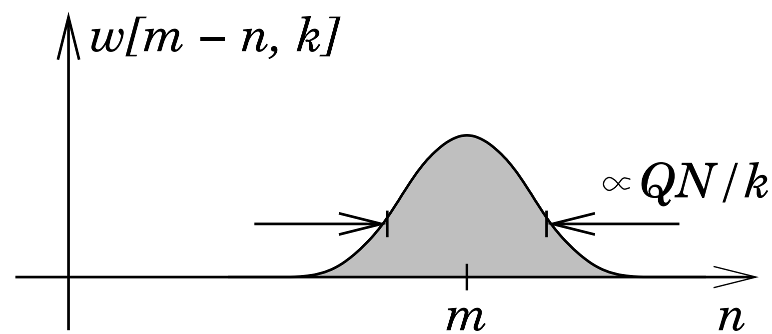
LIGO

DQT Definition

Project $x[n]$ onto time-shifted windowed sinusoids, whose widths are inversely proportional to their center frequencies.

$$X_Q[m,k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi nk/N} w[m-n,k]$$





Fast Q Transform

Efficient computation is possible in frequency domain.

$$X_Q[m,k] = \sum_{l=0}^{N-1} \tilde{X}[l+k]\tilde{W}[l,k]e^{-i2\pi ml/N}$$

- One time FFT of signal: $\tilde{X}[l]$
- Frequency domain window: $\tilde{W}[l]$
- Inverse FFT for each frequency bin
  - Only for frequency bins of interest
  - Only for samples in proximity of window
  - Length determines overlap in time
- Computational cost
  - Dominated by initial FFT
  - Varies with overlap and Q

DQT Window Normalization

A frequency domain Hanning window is chosen for simplicity

- Near optimal time-frequency localization
- Smoothly goes to zero with finite support

The window normalization is chosen to obey a generalized Parseval's theorem.

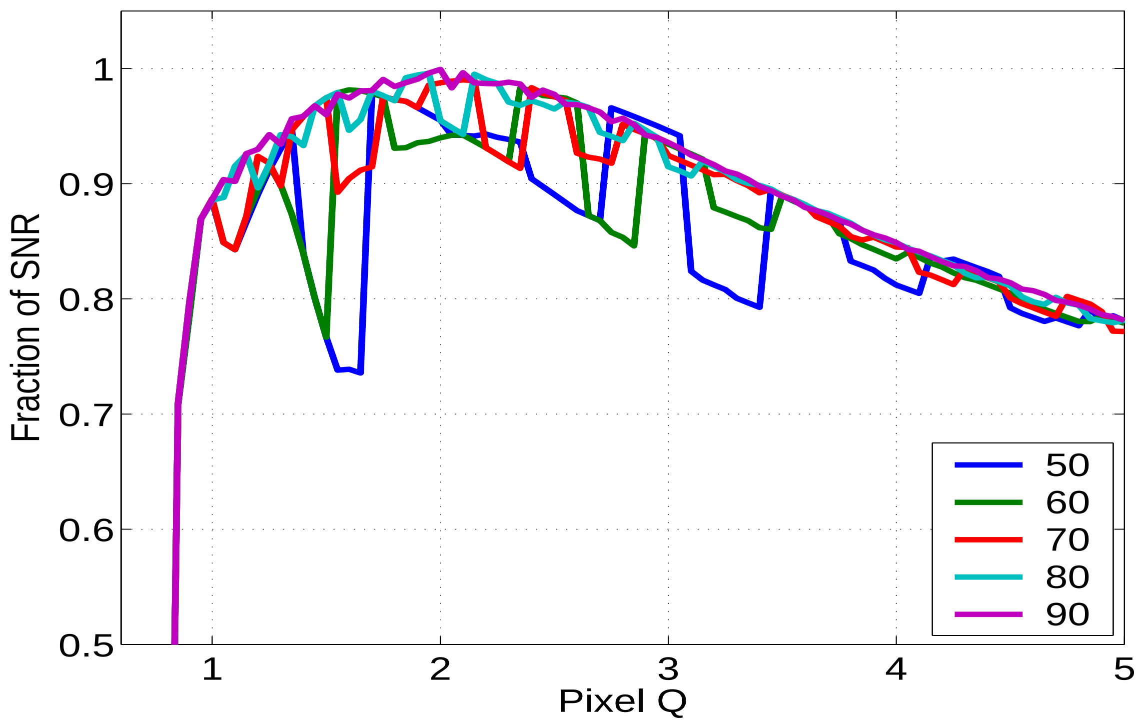
$$\frac{f_s}{N^2} \sum_{m=0}^{N-1} \sum_{k=0}^{N-1} |X_Q[m,k]|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|^2 = \sigma_x^2$$

The reported pixel amplitude is a combination of the noise amplitude spectral density and the root sum square signal amplitude in units of Hz$^{-1/2}$.

LIGO

DQT Pixel Mistmatch

- Mismatch between a signal and the nearest time frequency pixel results in a loss in measured SNR.
- Fractional loss in detected SNR for a sine-gaussian burst as a function of measurement Q and percentage overlap:



DQT Pixel Mistmatch

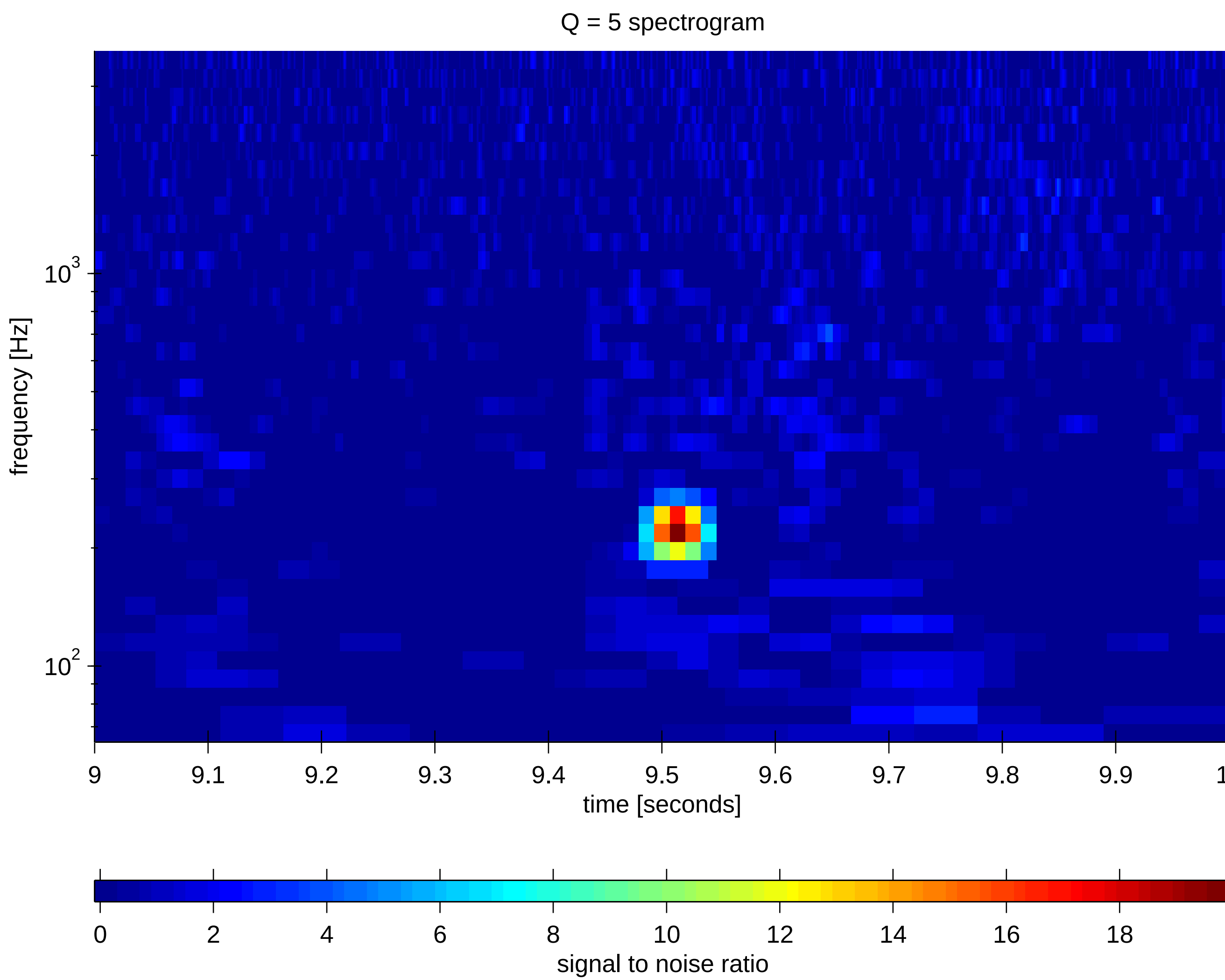
- This is similar to the problem of selecting discrete template banks in a matched filtering analysis.
- Find the maximum pixel spacing in time, frequency, and Q such that the SNR loss due to mismatch never exceeds a specified threshold.
- This is conveniently represented by a pixel space metric for fractional SNR loss.

$$ds^2 = g_{tt}\,dt^2 + g_{ff}\,df^2 + g_{QQ}\,dQ^2$$

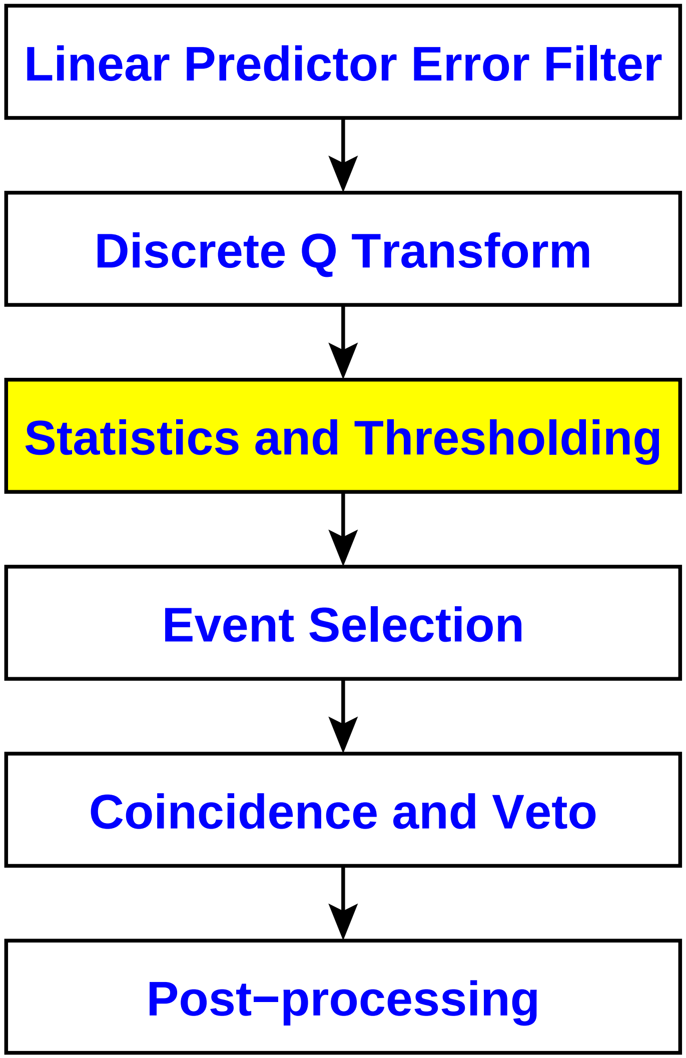
- For a given test waveform:
  Find g_{tt}, g_{ff}, and g_{QQ}.
  Find dt, df, and dQ such that ds never exceeds a specified threshold.

DQT Example

Hardware injection seen in H1:LSC-AS_Q



Statistics & Thresholding



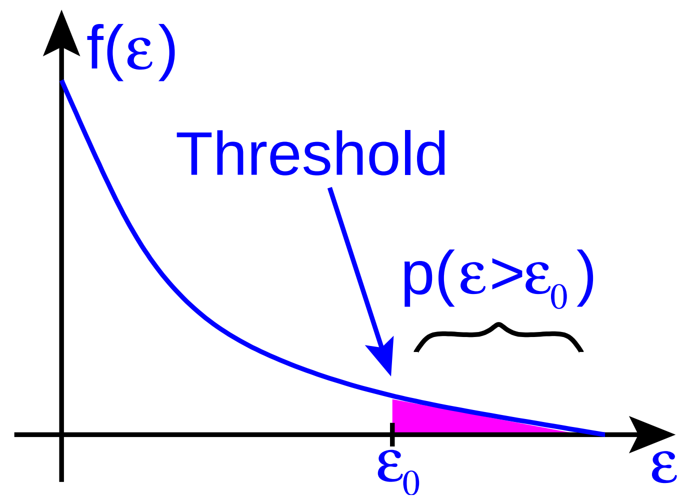
- Assume white Gaussian noise statistics
- Threshold for desired Gaussian noise false rate
- Achieves fundamental measurement accuracy

White Gaussian Noise

After linear predictor error filtering, pixel energies are exponentially distributed with mean and standard deviation ε_μ.

Probability density function:

$$f(\varepsilon)\, d\varepsilon = \frac{1}{\varepsilon_\mu} e^{-\varepsilon/\varepsilon_\mu}\, d\varepsilon$$



Significance (false rate):

$$p(\varepsilon > \varepsilon_0) = e^{-\varepsilon_0/\varepsilon_\mu}$$

Signal energy:

$$h_0^2 = \varepsilon - \varepsilon_\mu$$

Signal to noise ratio:

$$\rho^2 = \frac{\varepsilon - \varepsilon_\mu}{\varepsilon_\mu}$$

Measurement Errors

Consider the measured signal to noise ratio for a true signal energy ε_s and noise energy ε_μ.

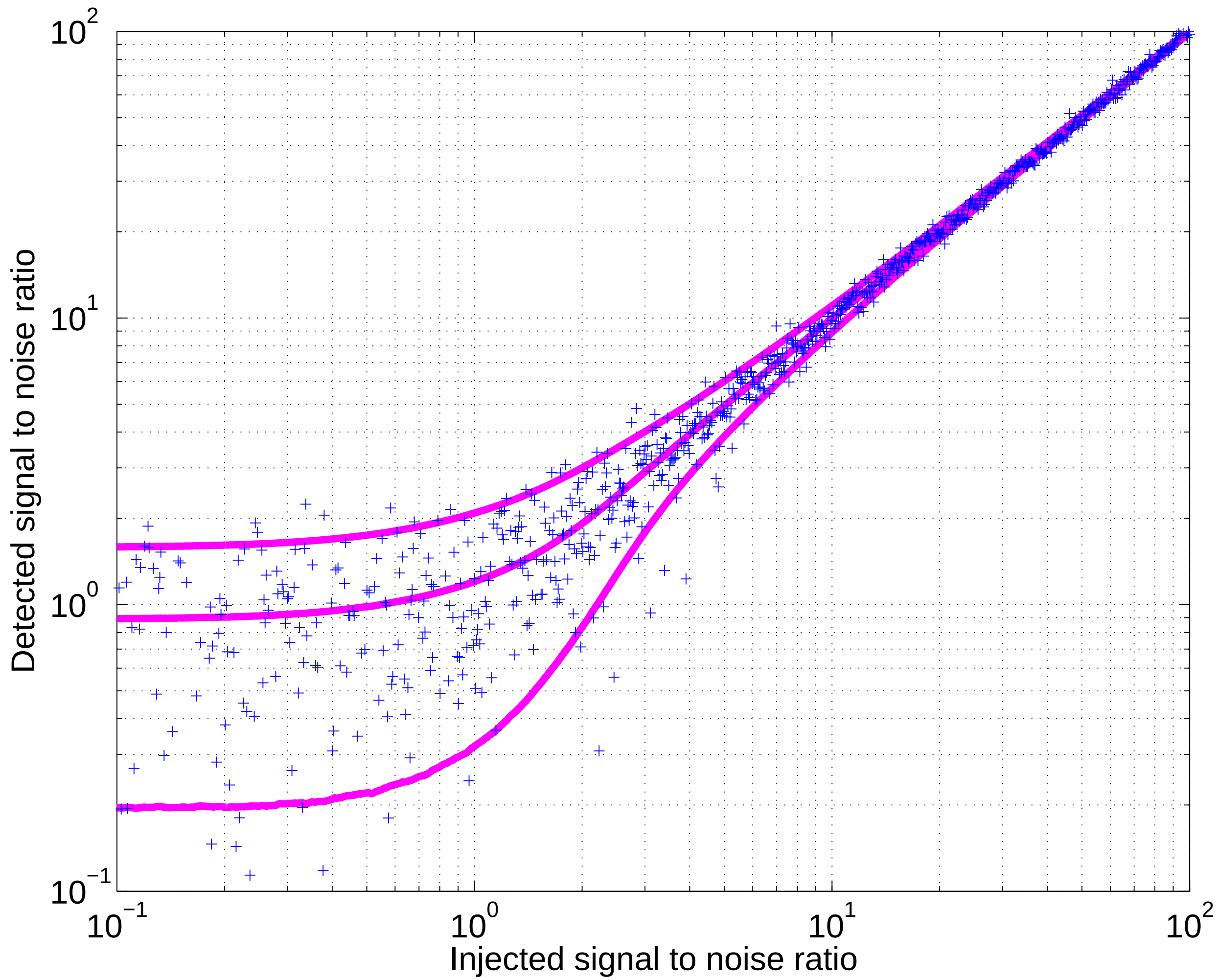
$$\rho^2 = \frac{\varepsilon_s + \varepsilon_n + 2(\varepsilon_s \varepsilon_n)^{1/2} \cos\phi - \varepsilon_\mu}{\varepsilon_\mu}$$

There are four sources of measurement error:

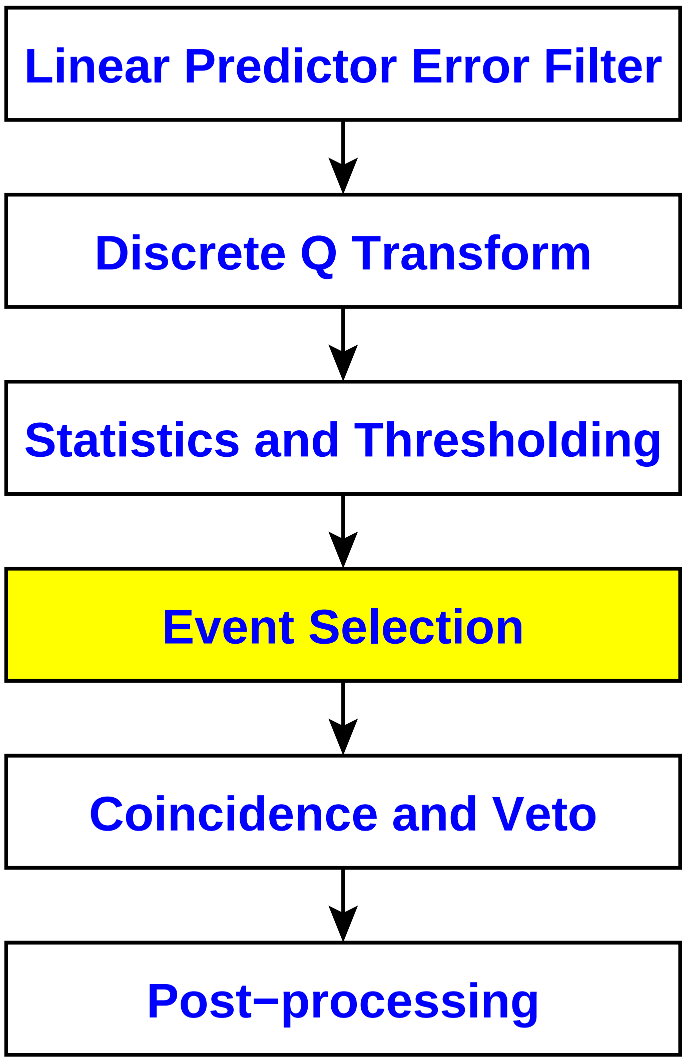
- Time-frequency pixel mismatch:
  Vanishes with increasing pixel overlap
- Guassian distribution of mean background energy ε_μ:
  Vanishes with increasing measurement time
- Exponential distribution of background energies ε_n:
  Fundamental to time-frequency measurement
- Uniform distribution of background phase ϕ:
  Fundamental to time-frequency measurement

Simulated Measurements

Optimal detection of Sine-Gaussian bursts in the presence of white Gaussian noise.



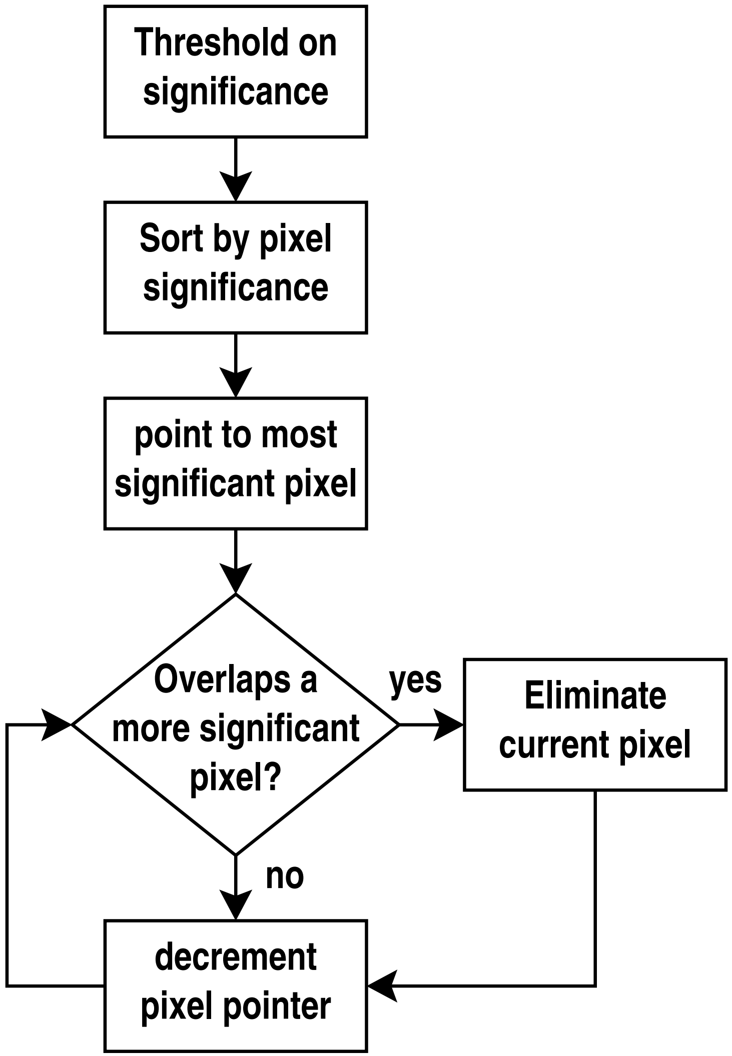
Event Selection



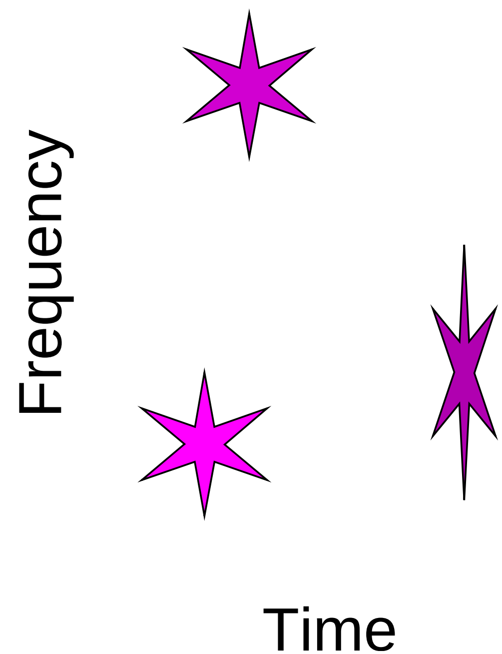
- Goal: Estimate of true signal parameters
- Problem: Thresholding yields mulitple pixels per event
- Approach: Choose most significant pixel within cluster
- Simple robust algorithm exists

Event Selection Algorithm

Selection algorithm



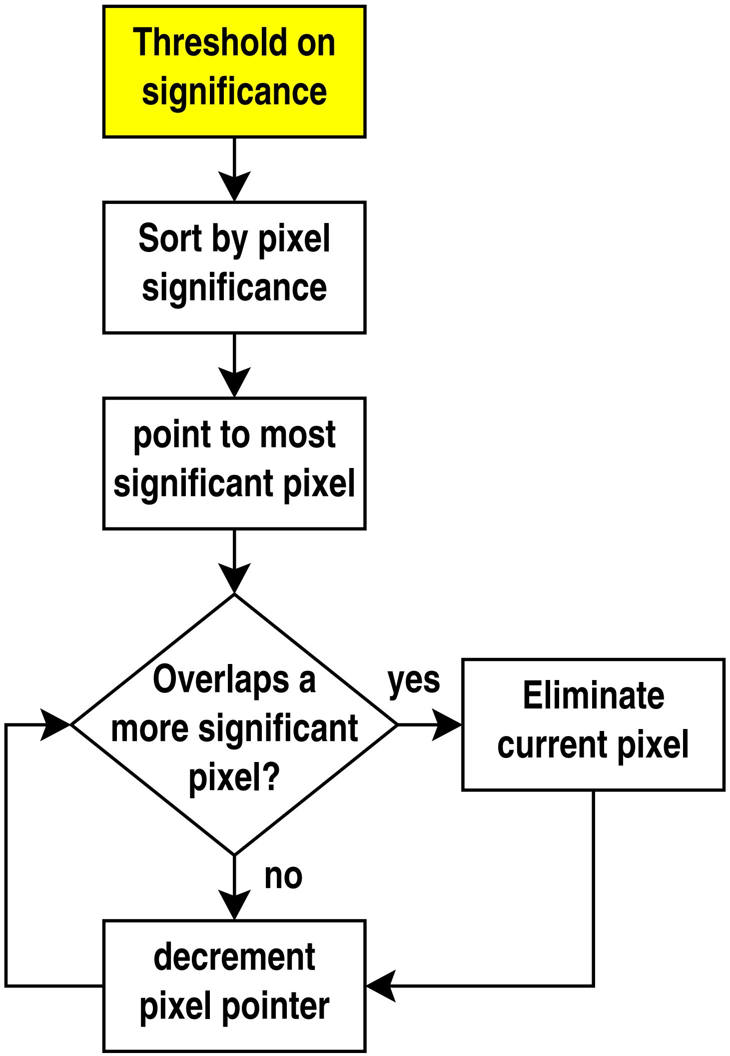
Simple example



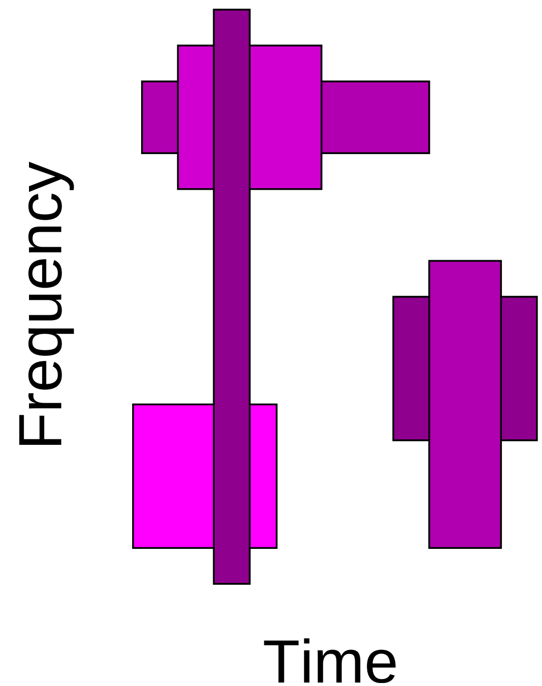
Consider the three events shown above.

Event Selection Algorithm

Selection algorithm



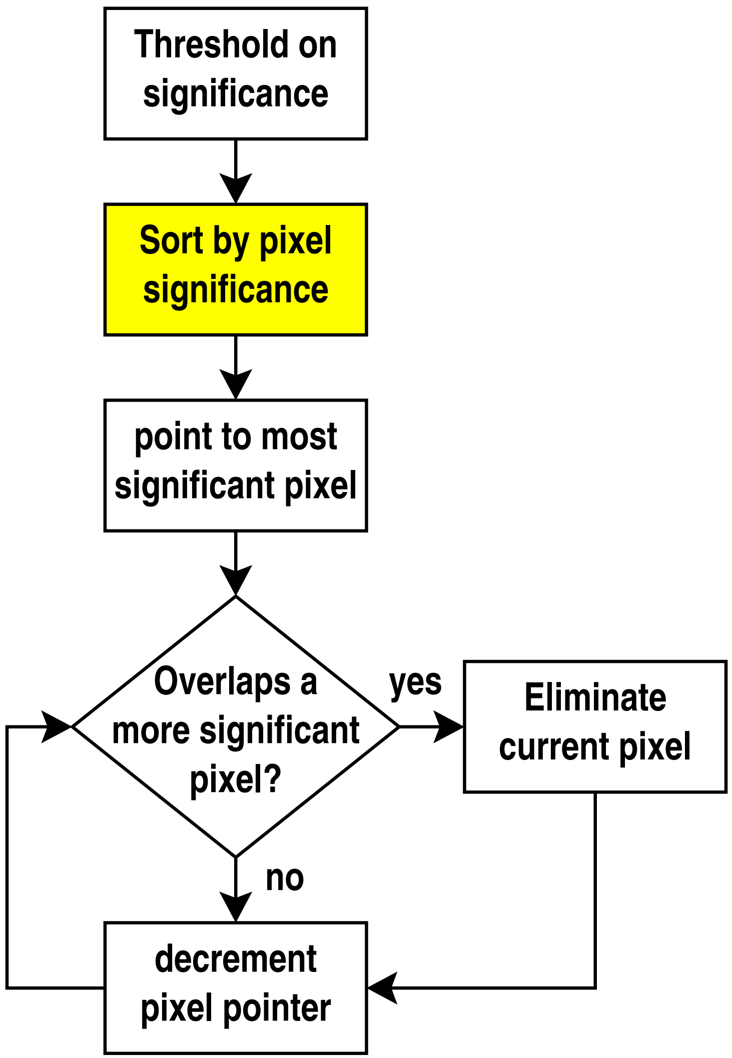
Simple example



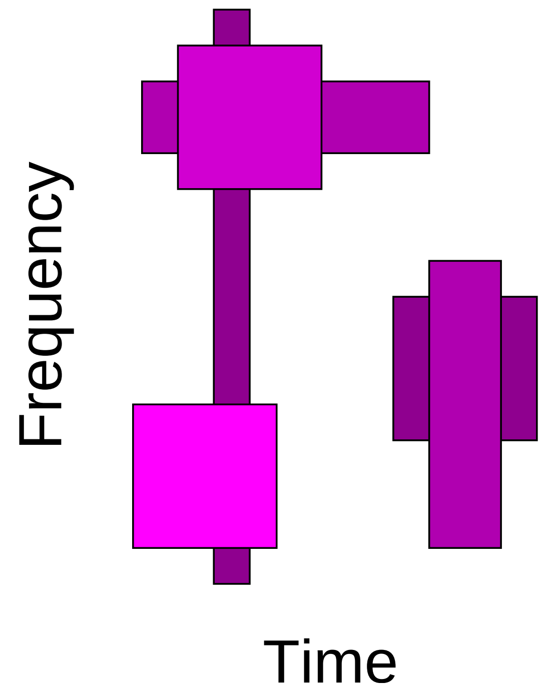
Problem: Thresholding yields multiple pixels per event.

Event Selection Algorithm

Selection algorithm



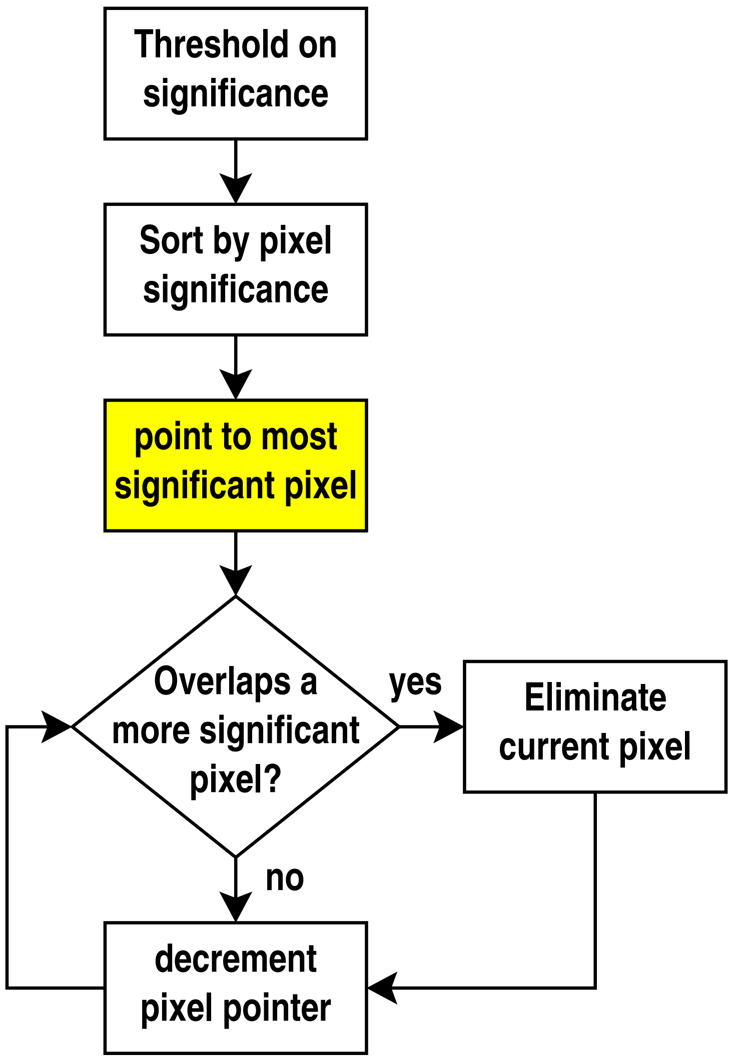
Simple example



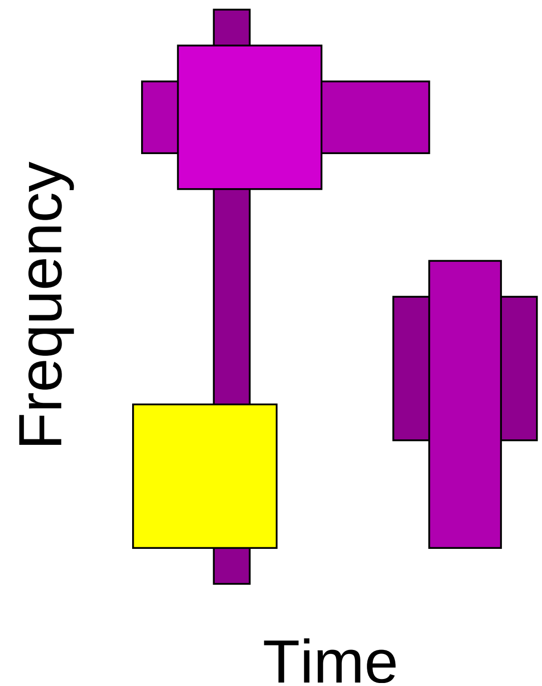
Sort pixels by significance.

Event Selection Algorithm

Selection algorithm



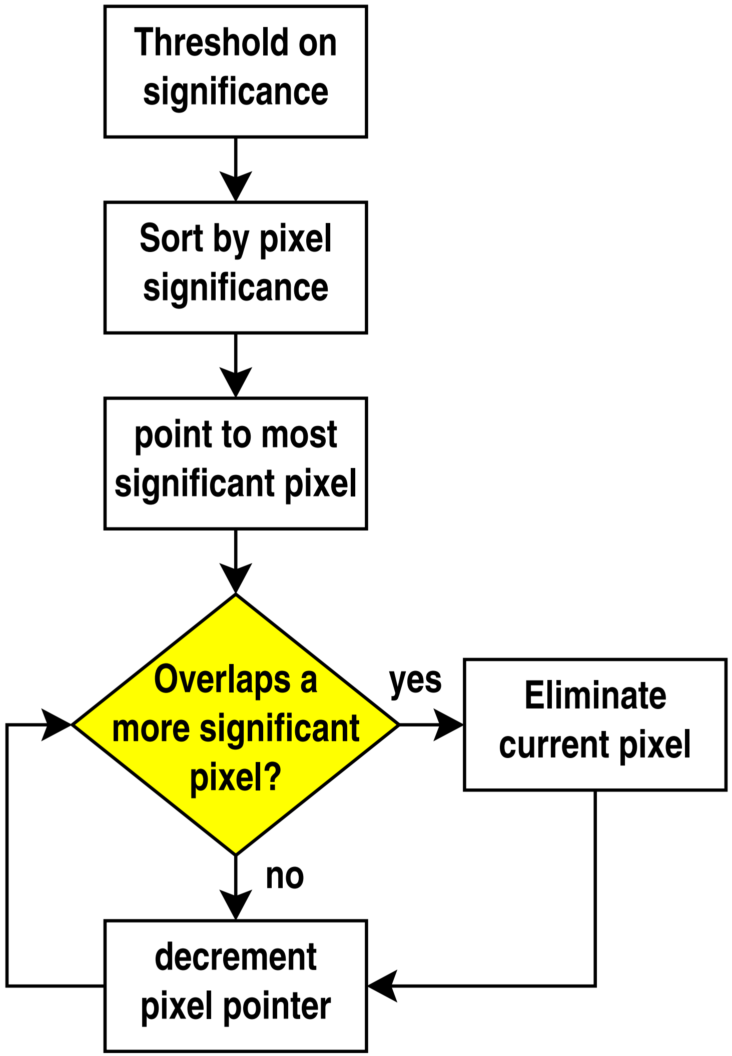
Simple example



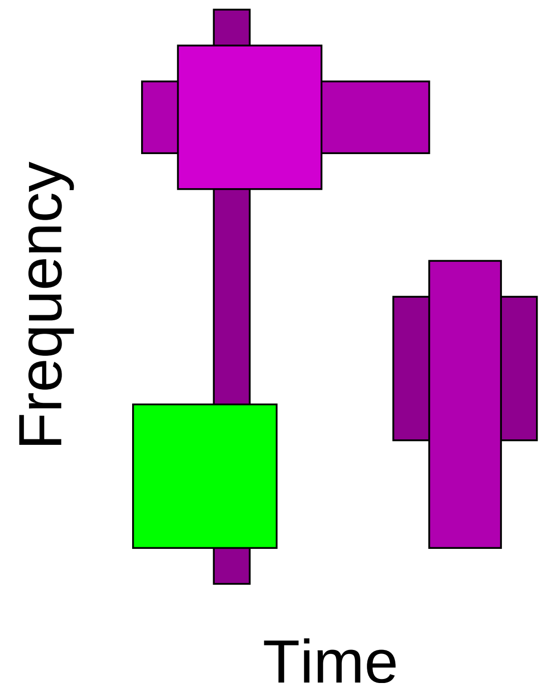
Start with the most significant pixel.

Event Selection Algorithm

Selection algorithm



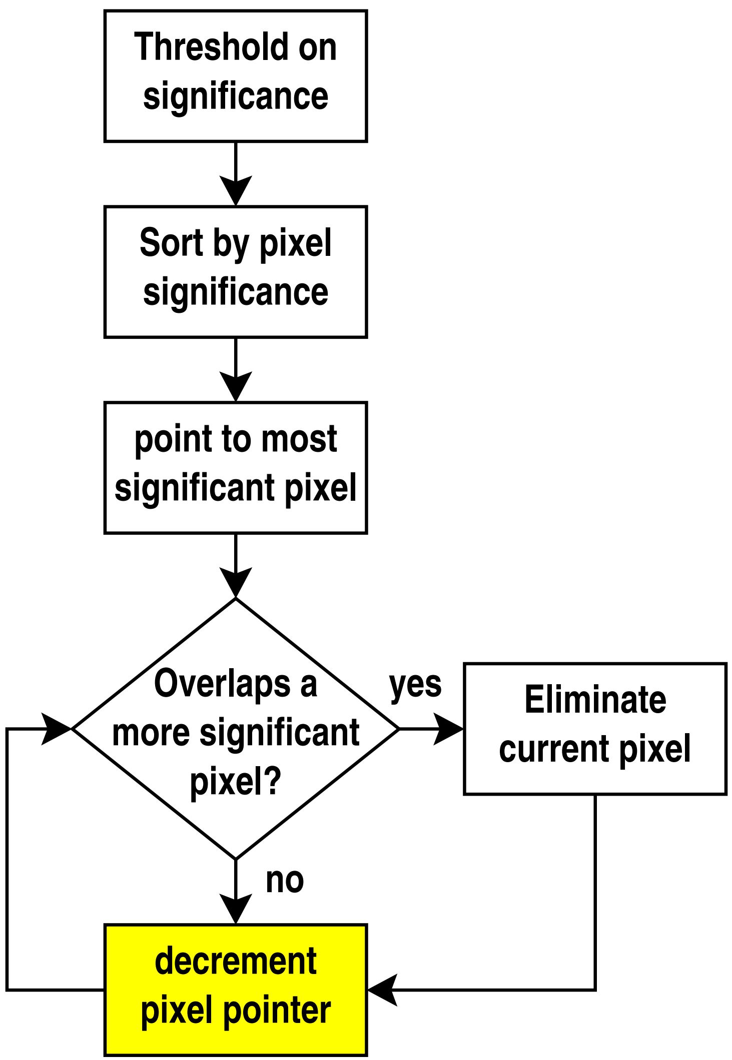
Simple example



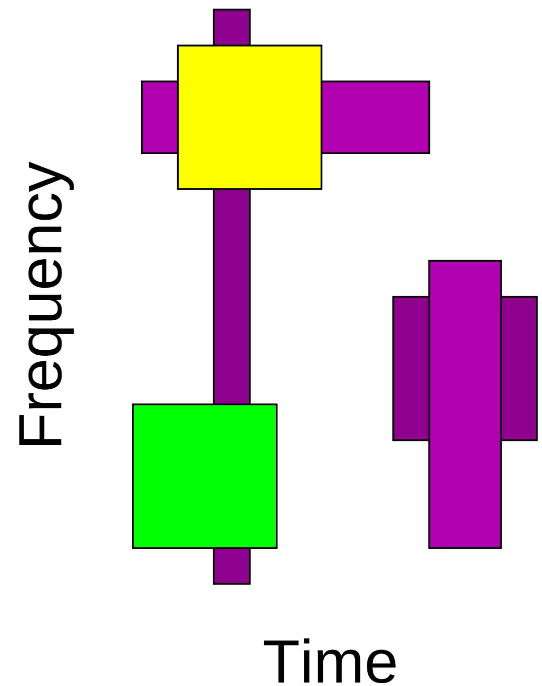
Does it overlap with a more significant pixel? No, keep it.

Event Selection Algorithm

Selection algorithm



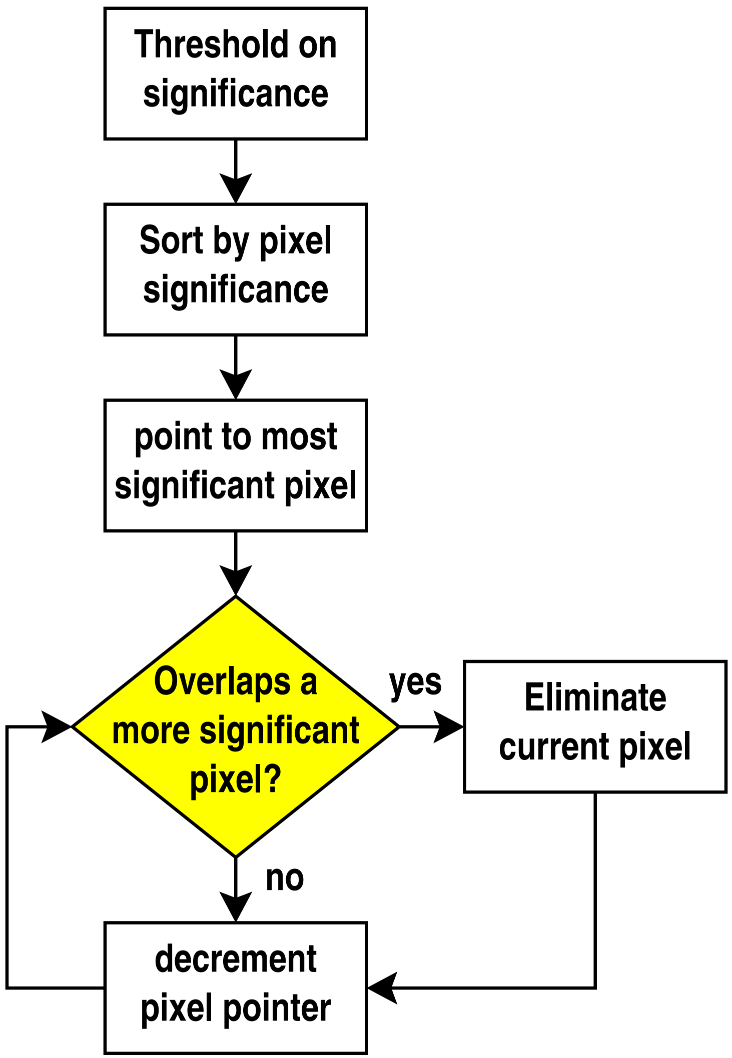
Simple example



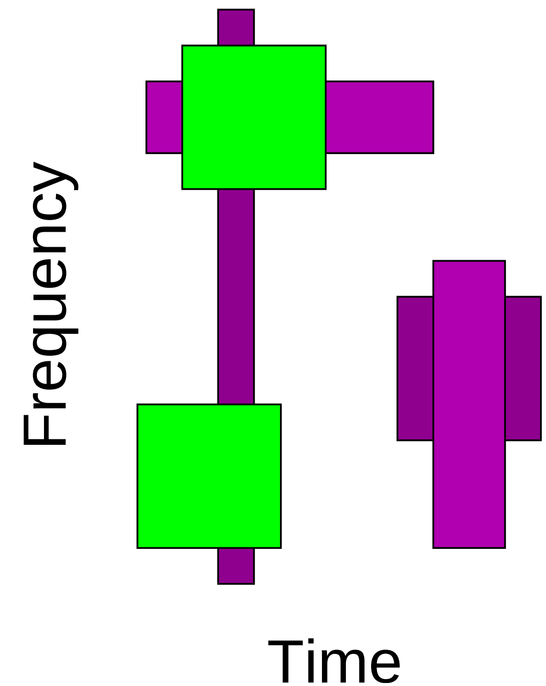
Consider the next significant pixel.

Event Selection Algorithm

Selection algorithm



Simple example

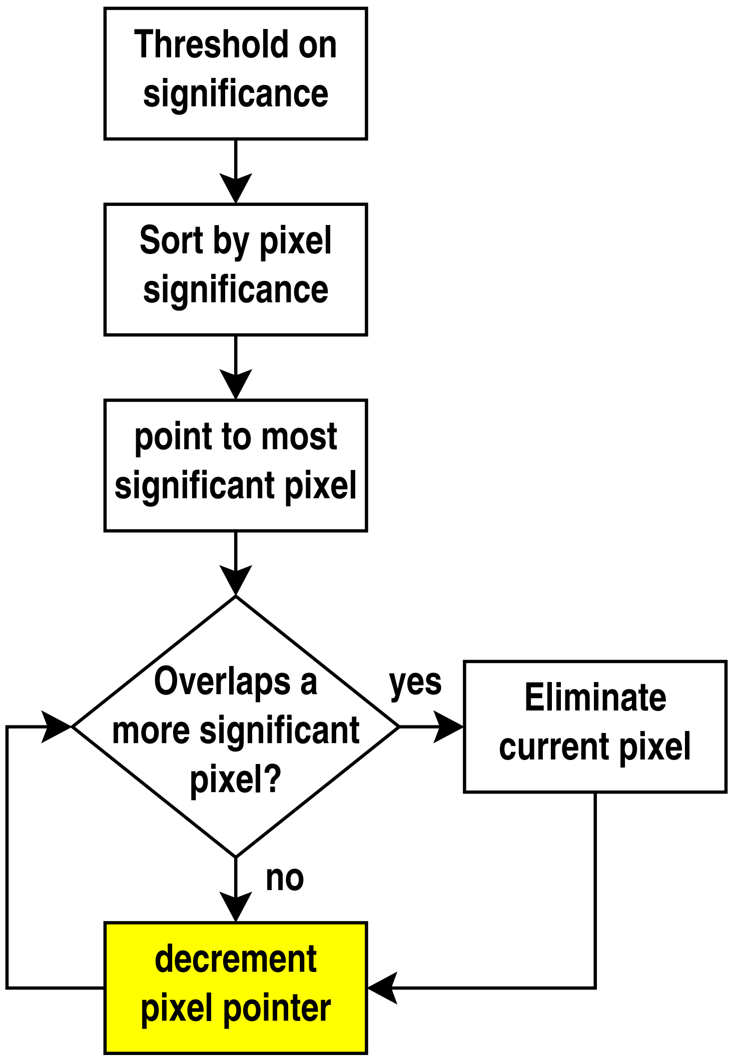


Does it overlap with a more significant pixel? No, keep it.

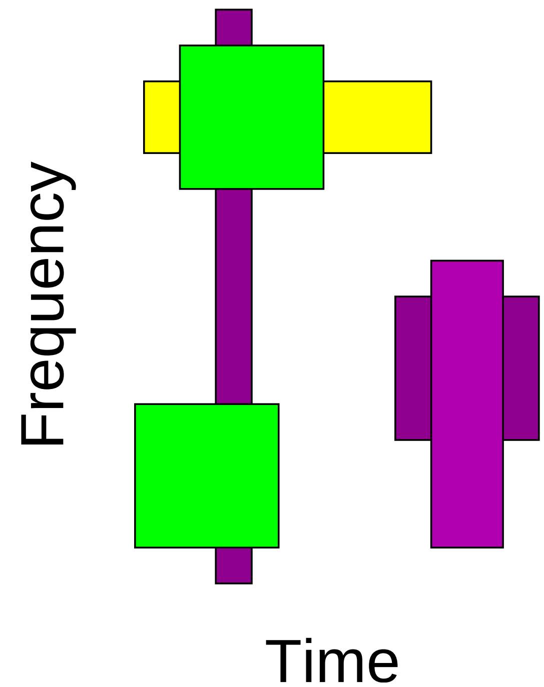
LIGO

Event Selection Algorithm

Selection algorithm



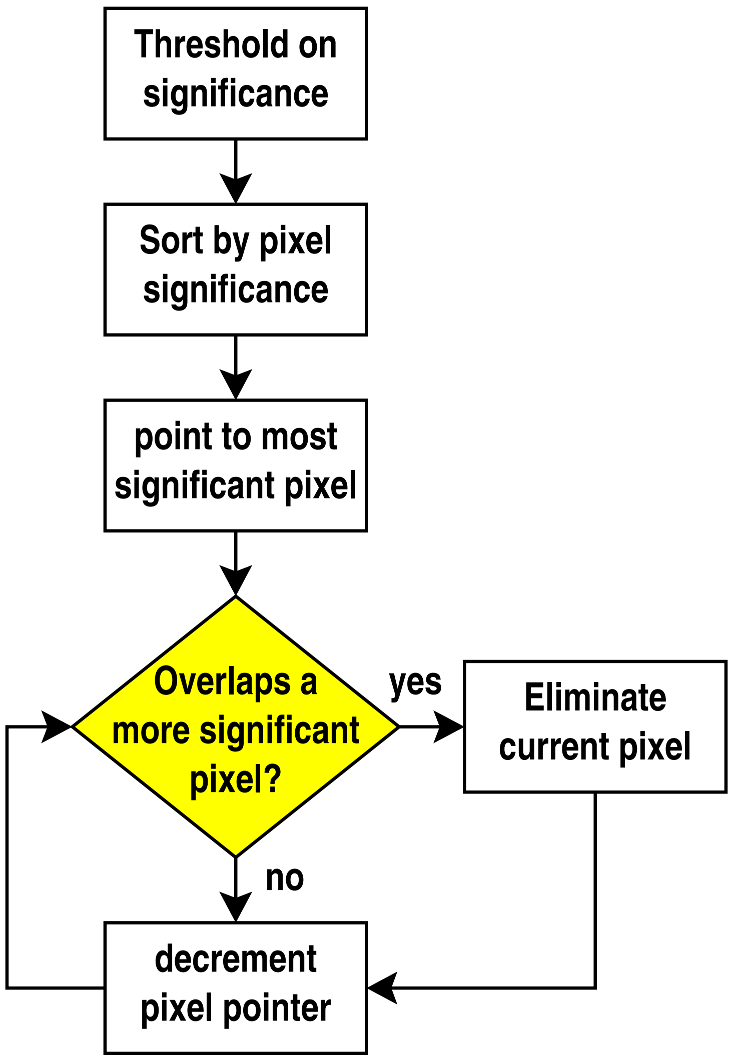
Simple example



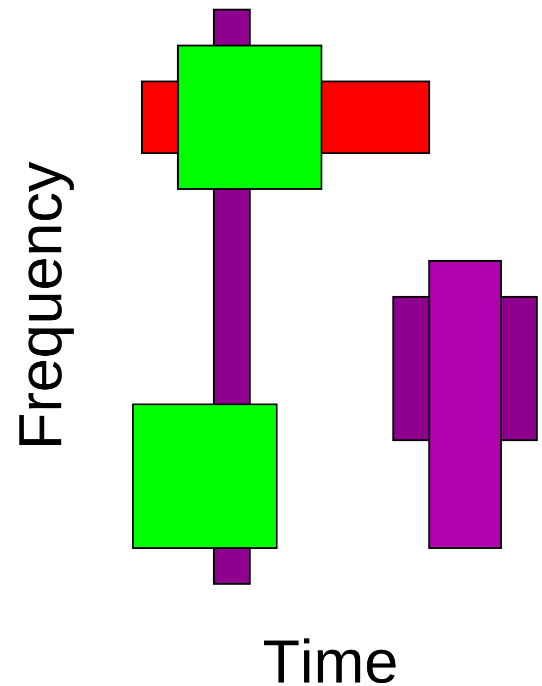
Consider the next significant pixel.

Event Selection Algorithm

Selection algorithm



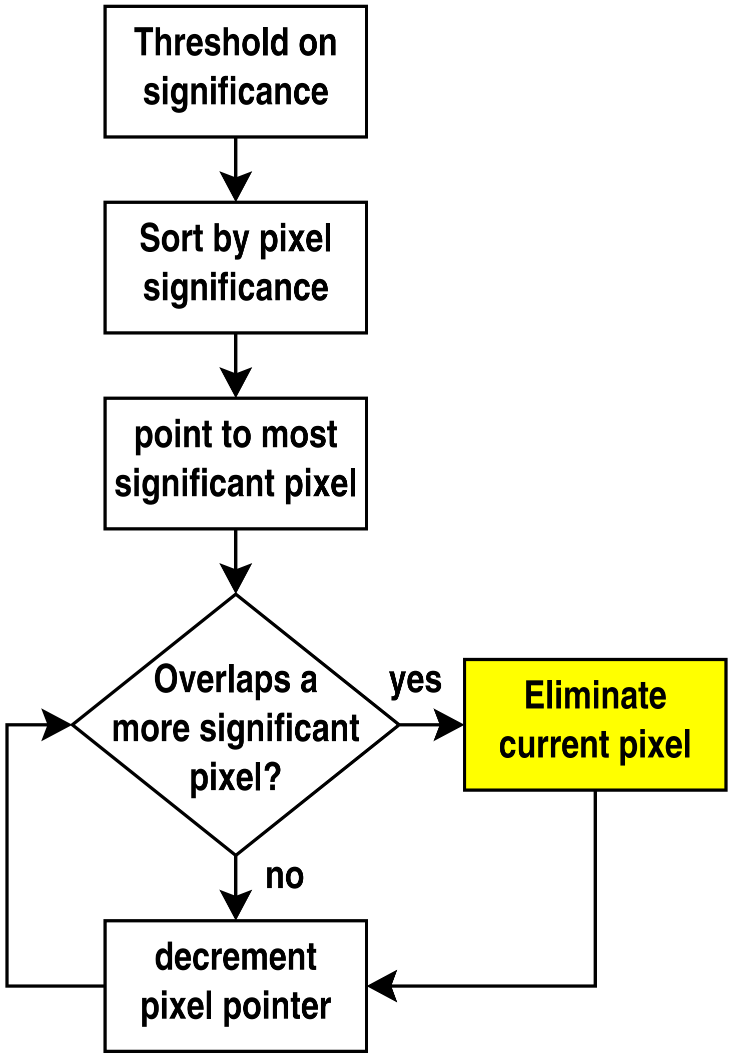
Simple example



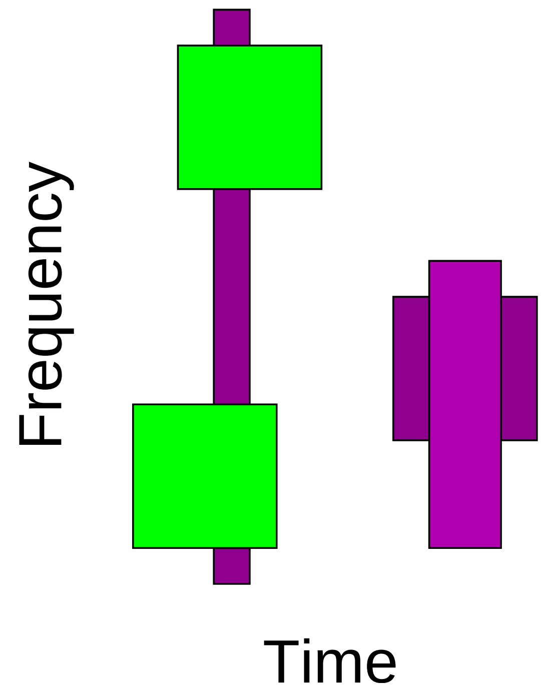
Does it overlap with a more significant pixel? Yes!

Event Selection Algorithm

Selection algorithm



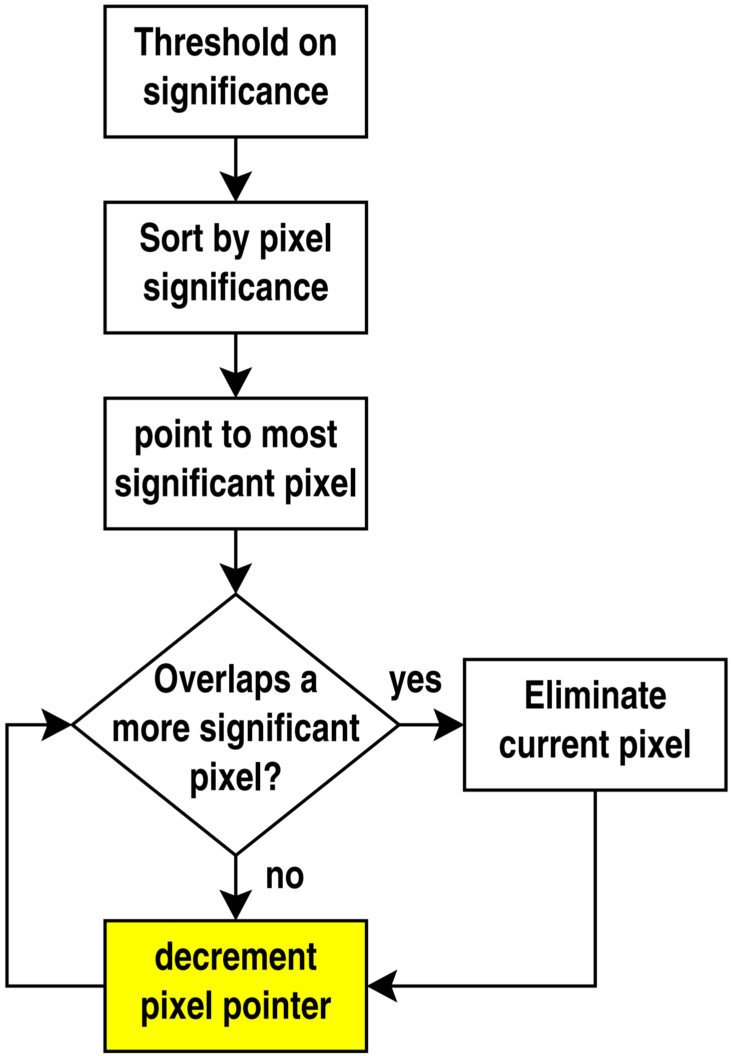
Simple example



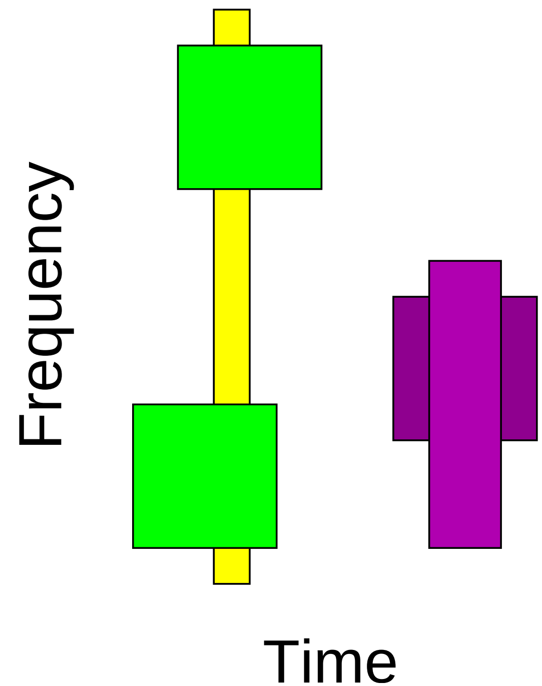
Remove this pixel.

Event Selection Algorithm

Selection algorithm



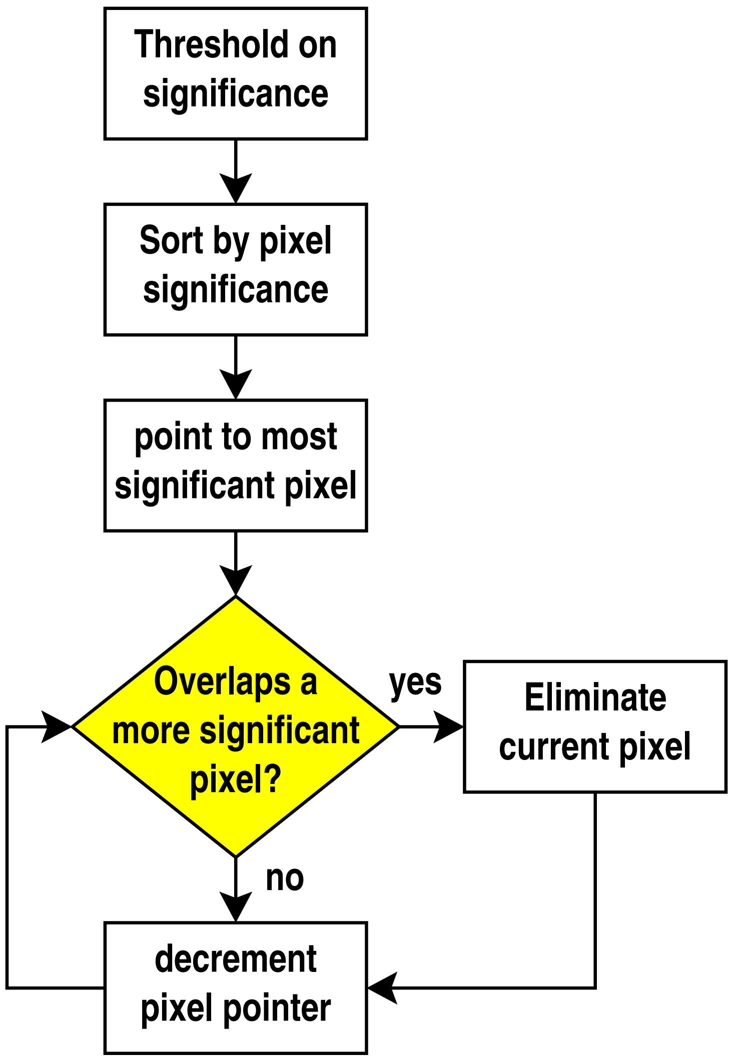
Simple example



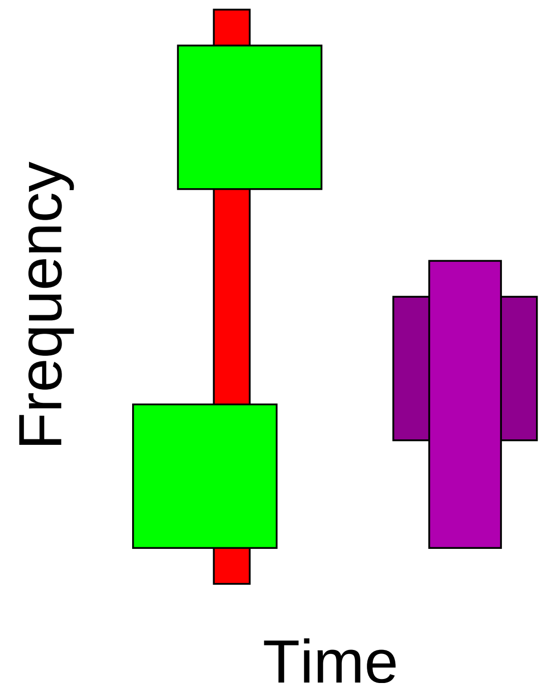
Consider the next significant pixel.

Event Selection Algorithm

Selection algorithm



Simple example

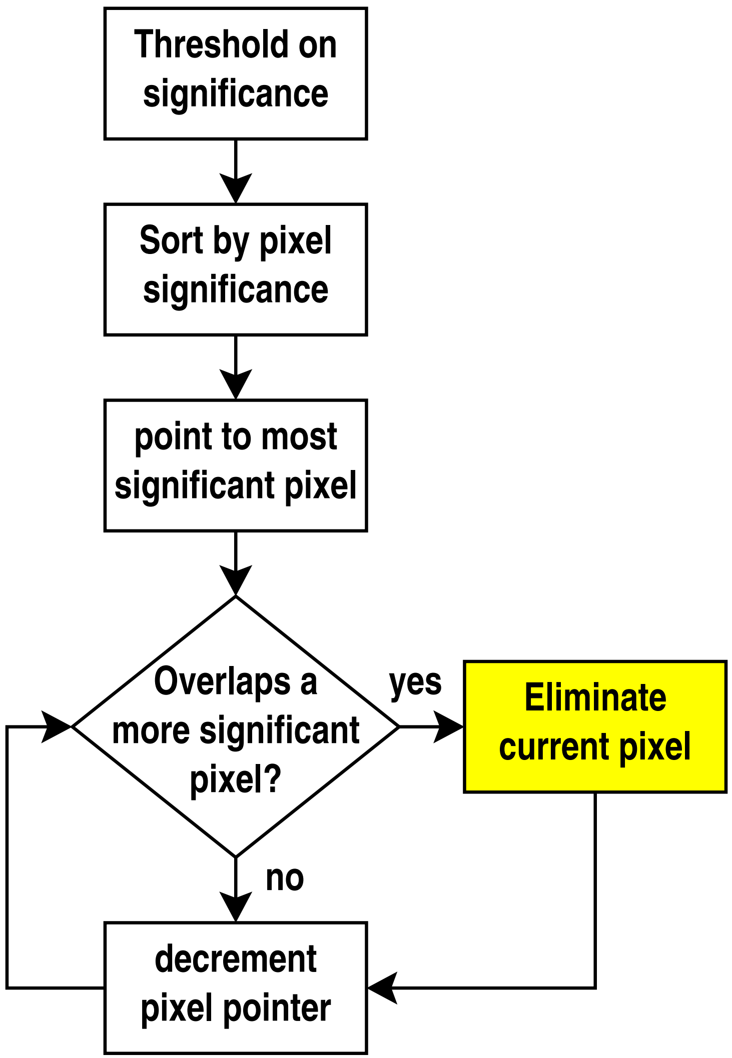


Does it overlap with a more significant pixel? Yes!

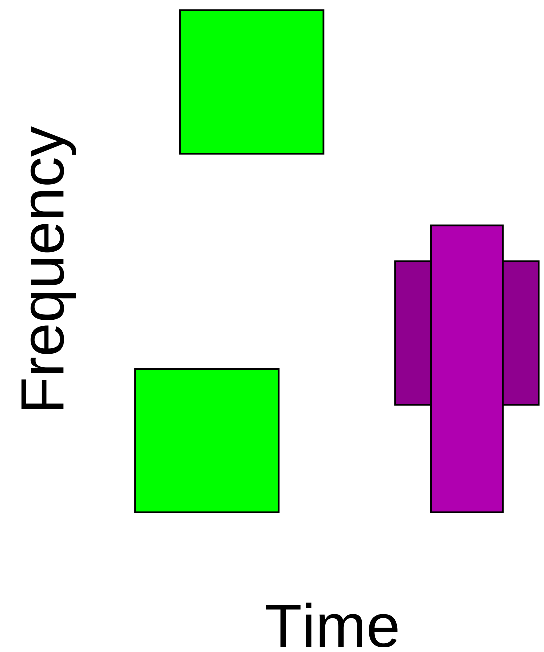
LIGO

Event Selection Algorithm

Selection algorithm



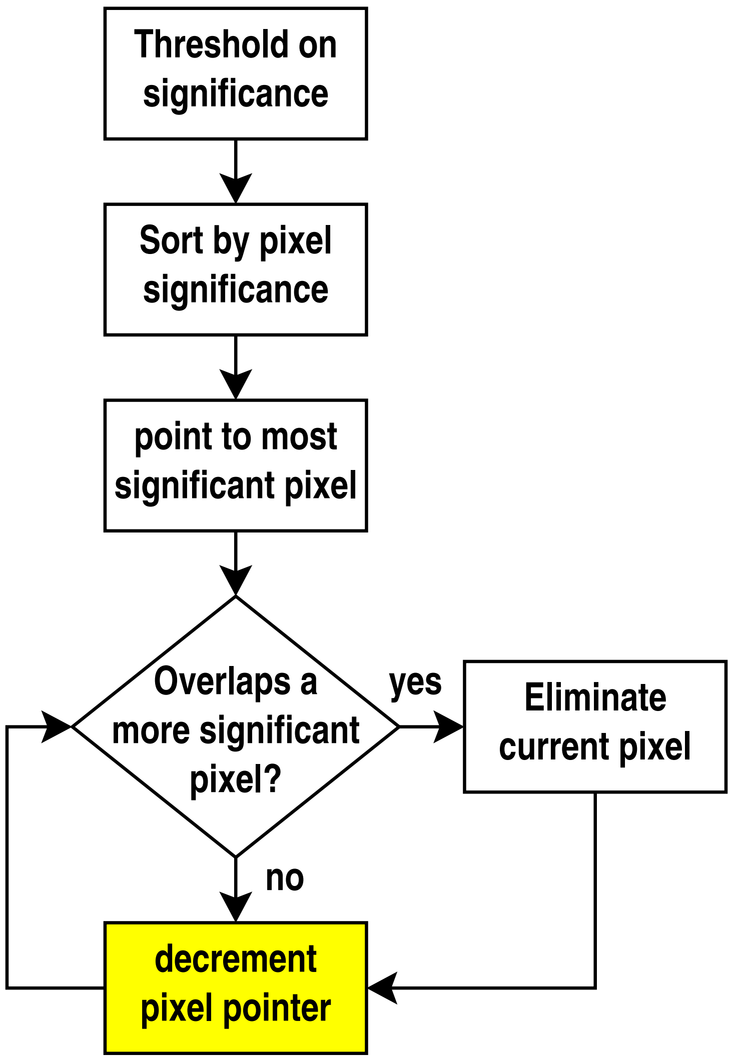
Simple example



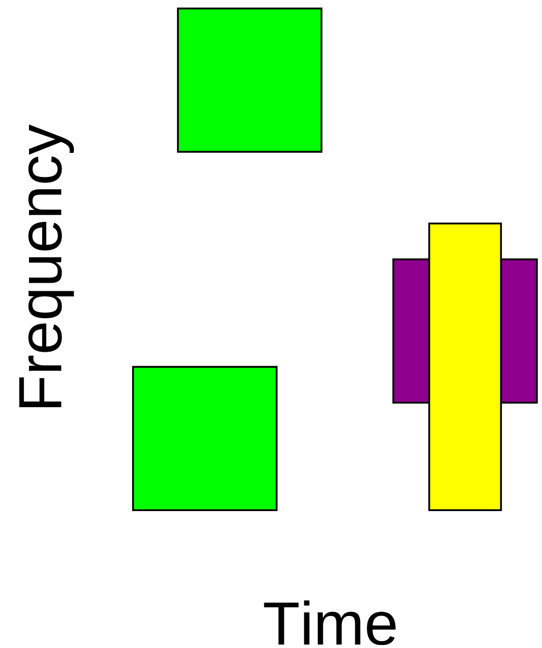
Remove this pixel.

Event Selection Algorithm

Selection algorithm



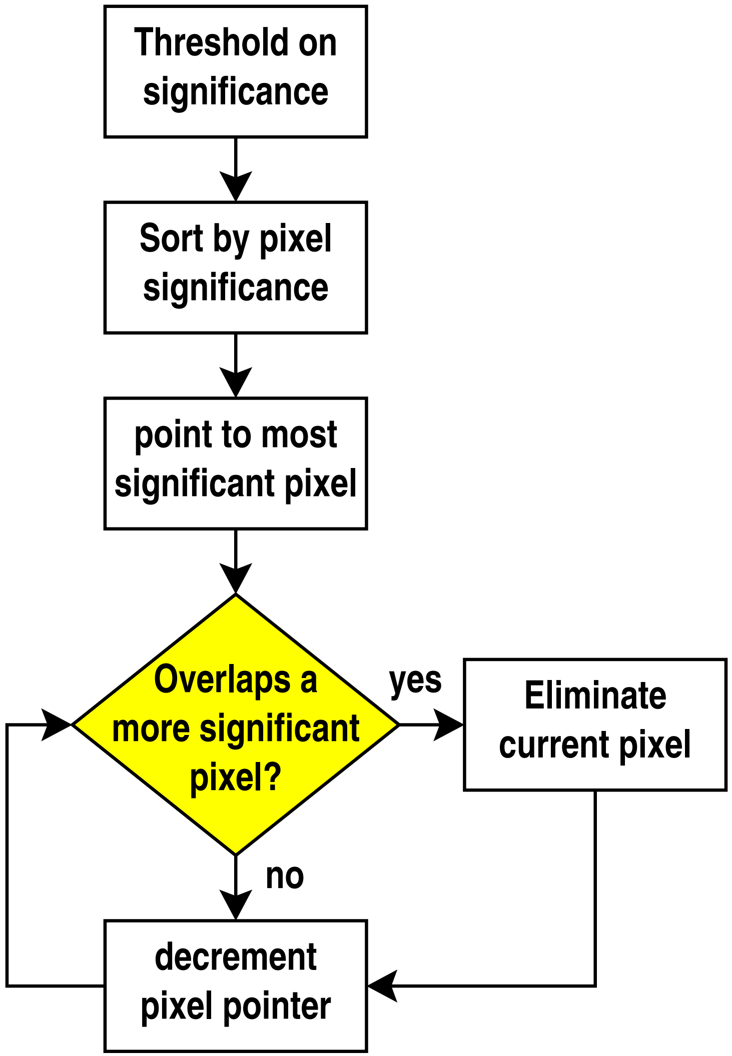
Simple example



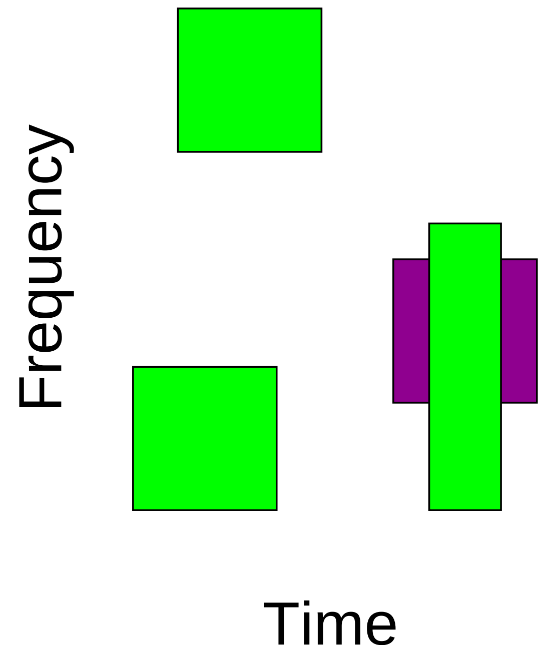
Consider the next significant pixel.

Event Selection Algorithm

Selection algorithm



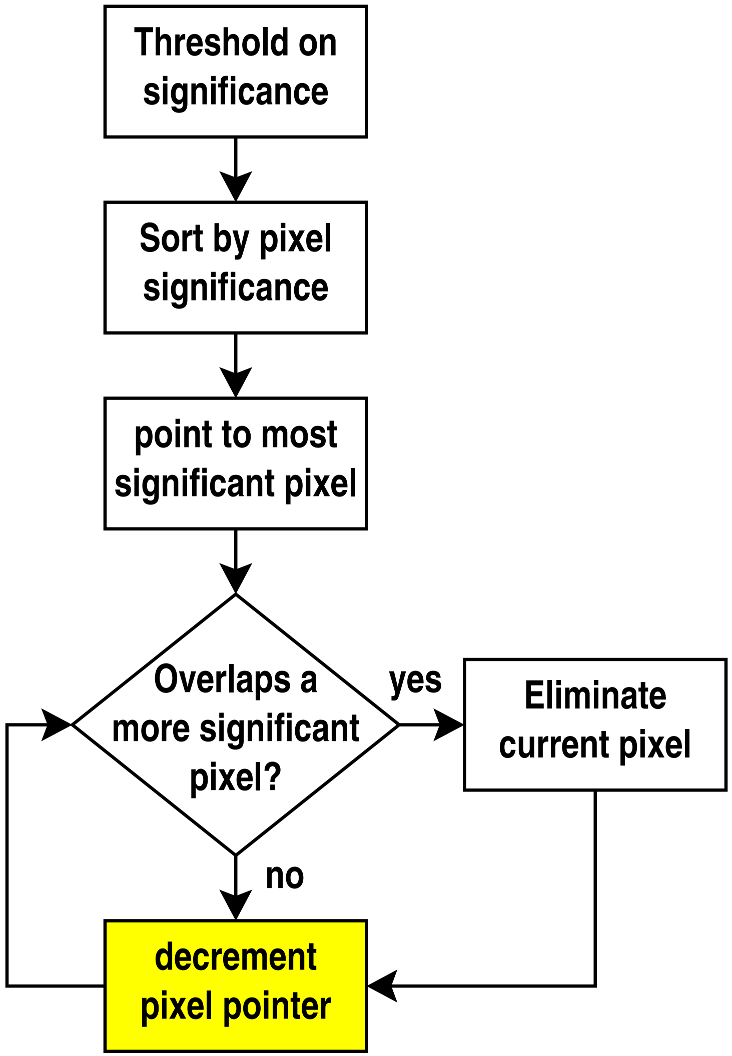
Simple example



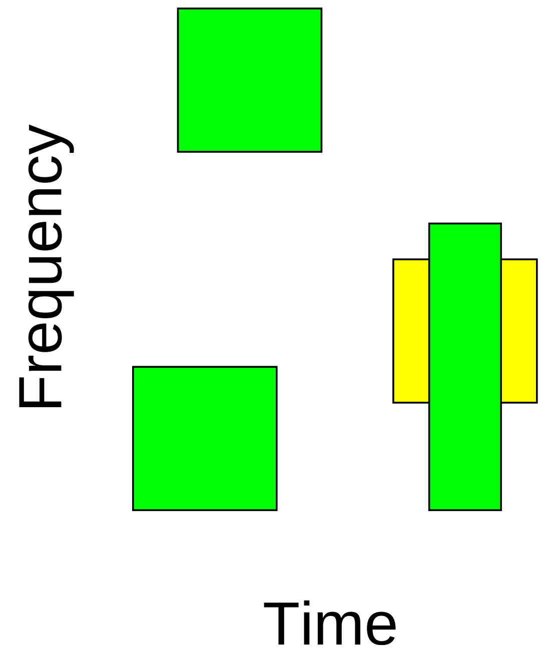
Does it overlap with a more significant pixel? No, keep it.

Event Selection Algorithm

Selection algorithm



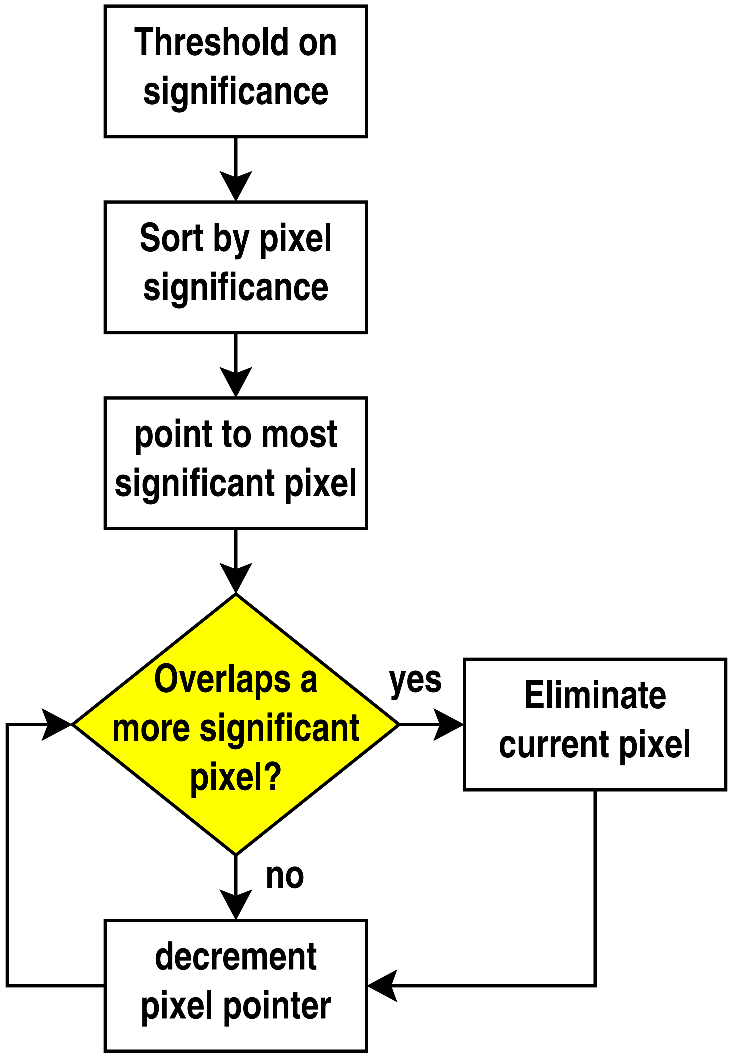
Simple example



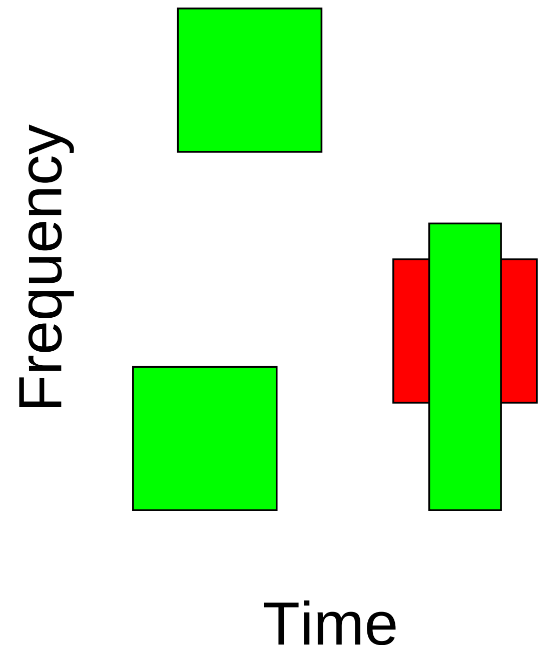
Consider the next significant pixel.

Event Selection Algorithm

Selection algorithm



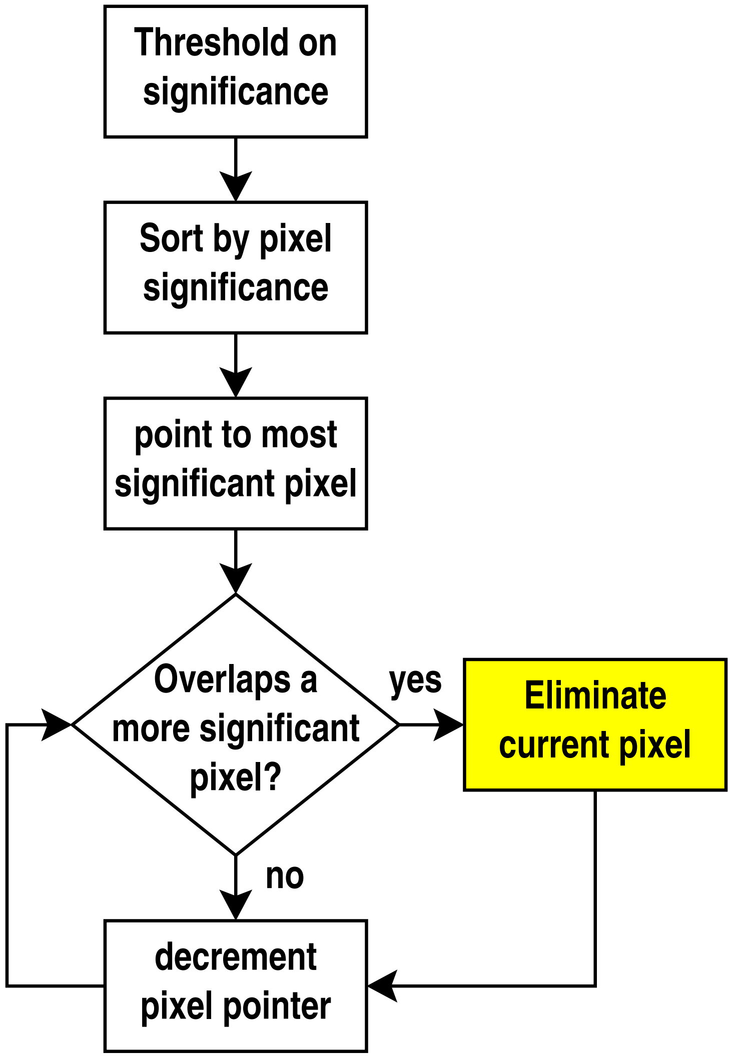
Simple example



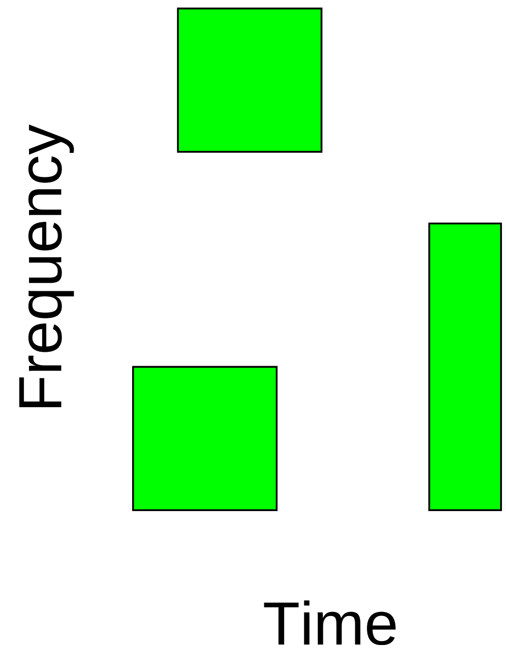
Does it overlap with a more significant pixel? Yes!

Event Selection Algorithm

Selection algorithm



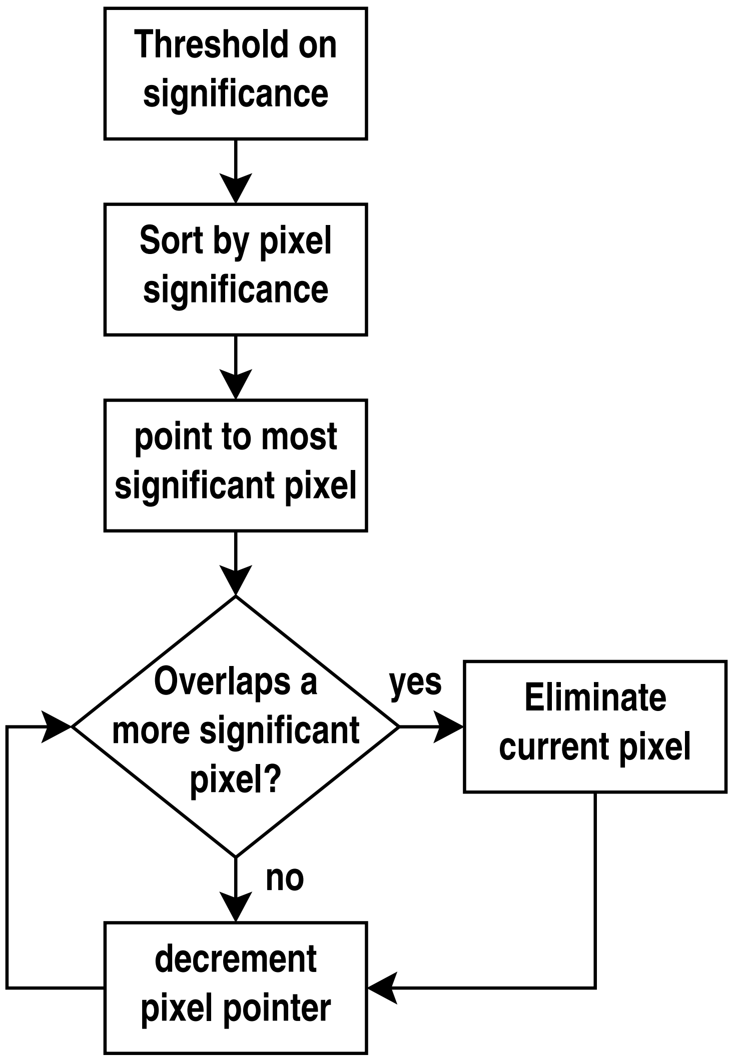
Simple example



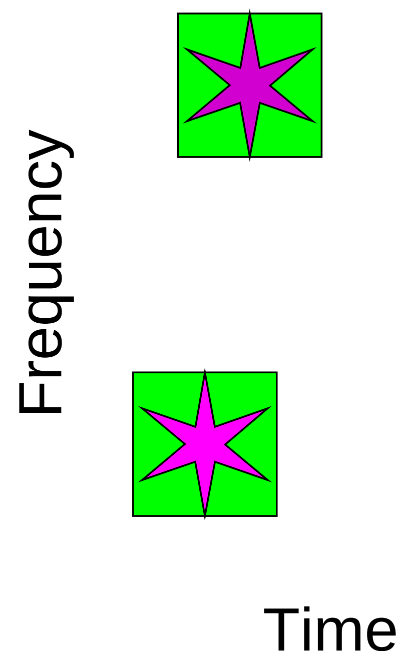
Remove this pixel.

Event Selection Algorithm

Selection algorithm

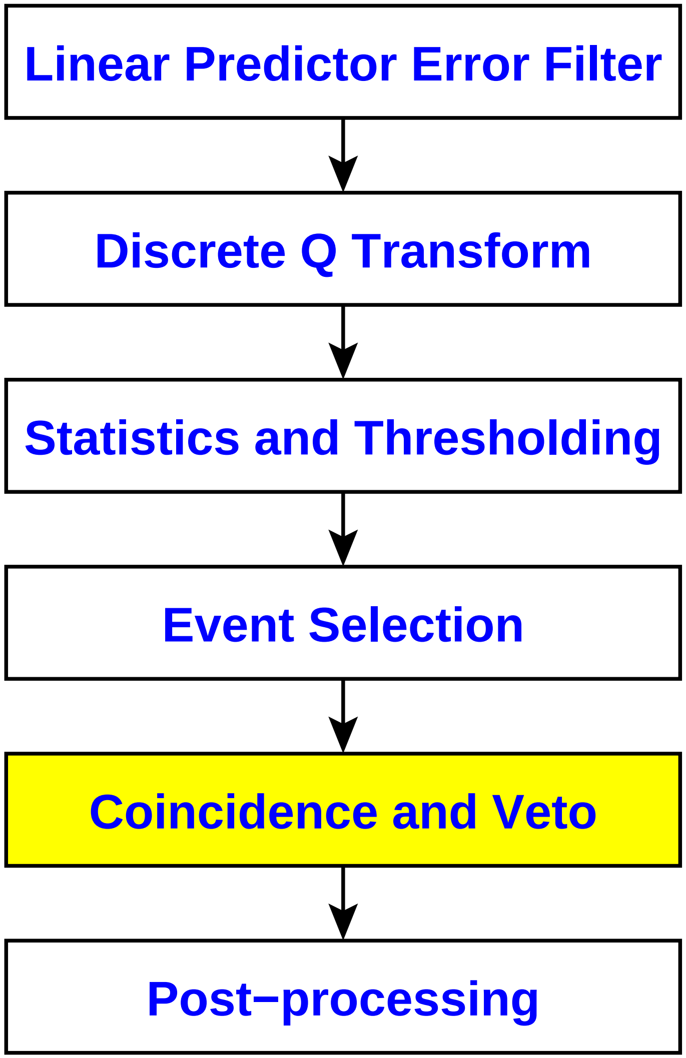


Simple example



Report the remaining pixels, which best estimate the true event parameters!

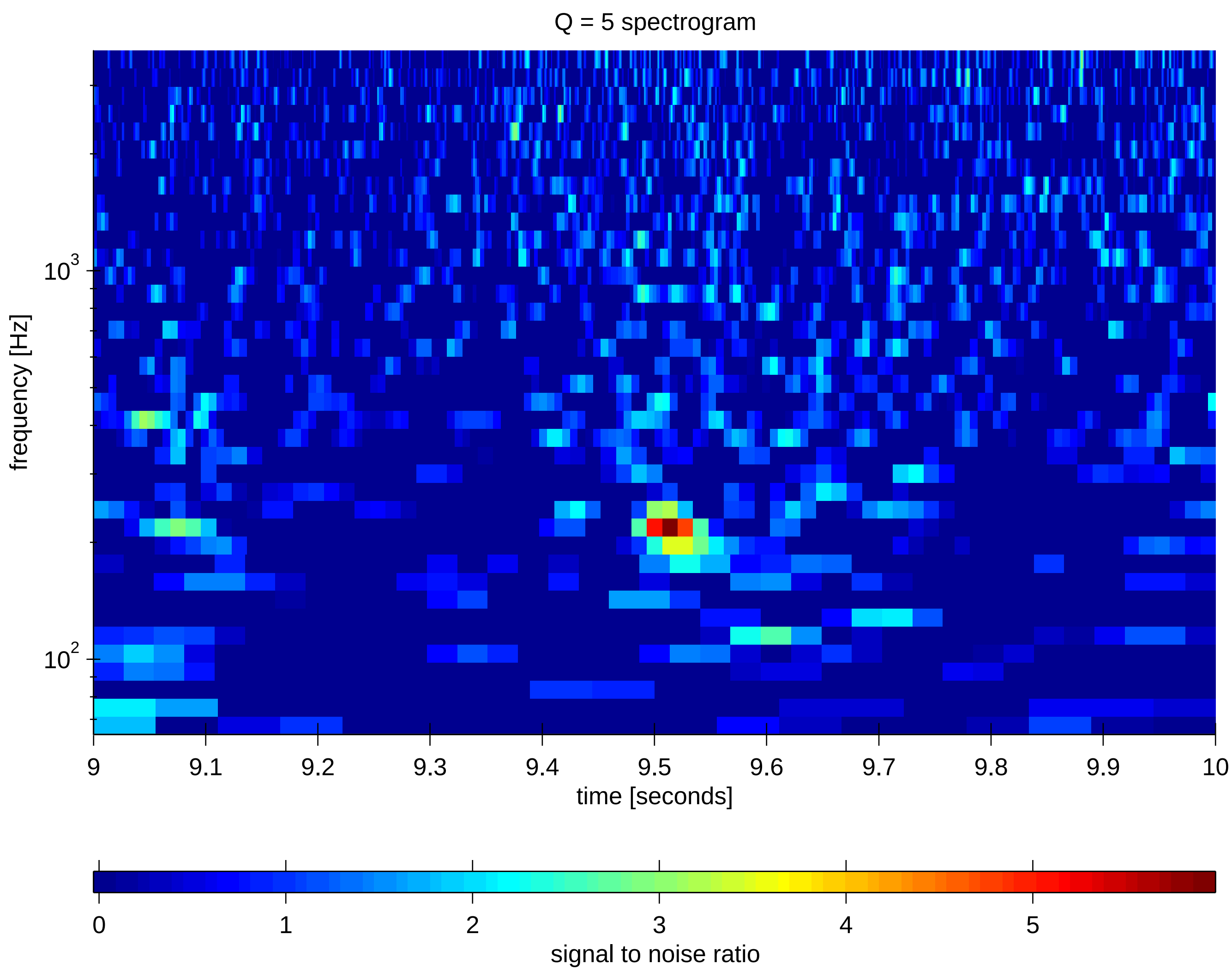
Coincidence Testing and Vetoing



- Event parameters allow time, frequency, and Q coincidence cuts
- Set significance threshold for desired coincident false rate (under construction)
- Study detection efficiency vs. dead area (under construction)
- Simplified tuning allows more powerful veto search

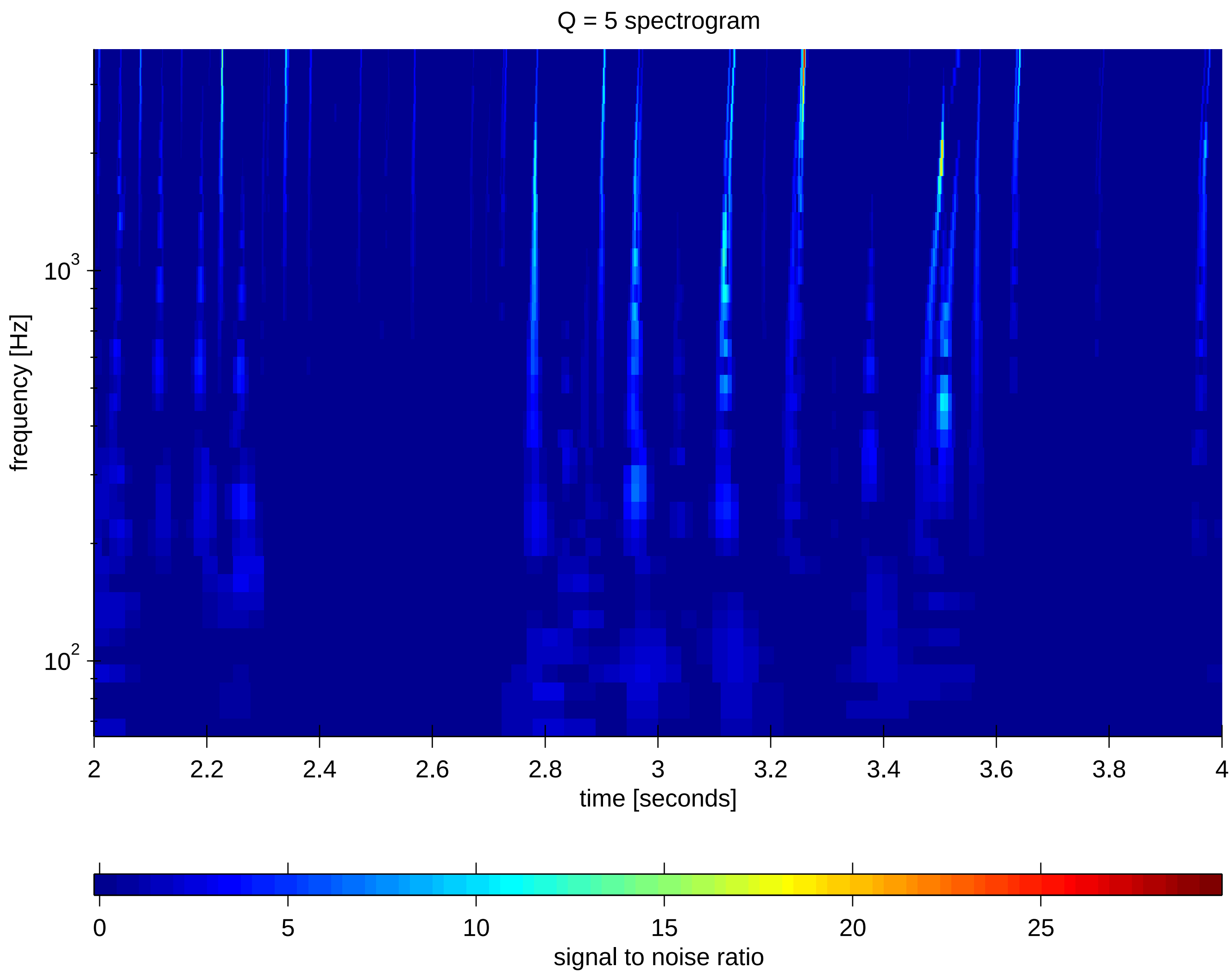
Veto Examples

Hardware injection seen in H1:LSC-AS_I



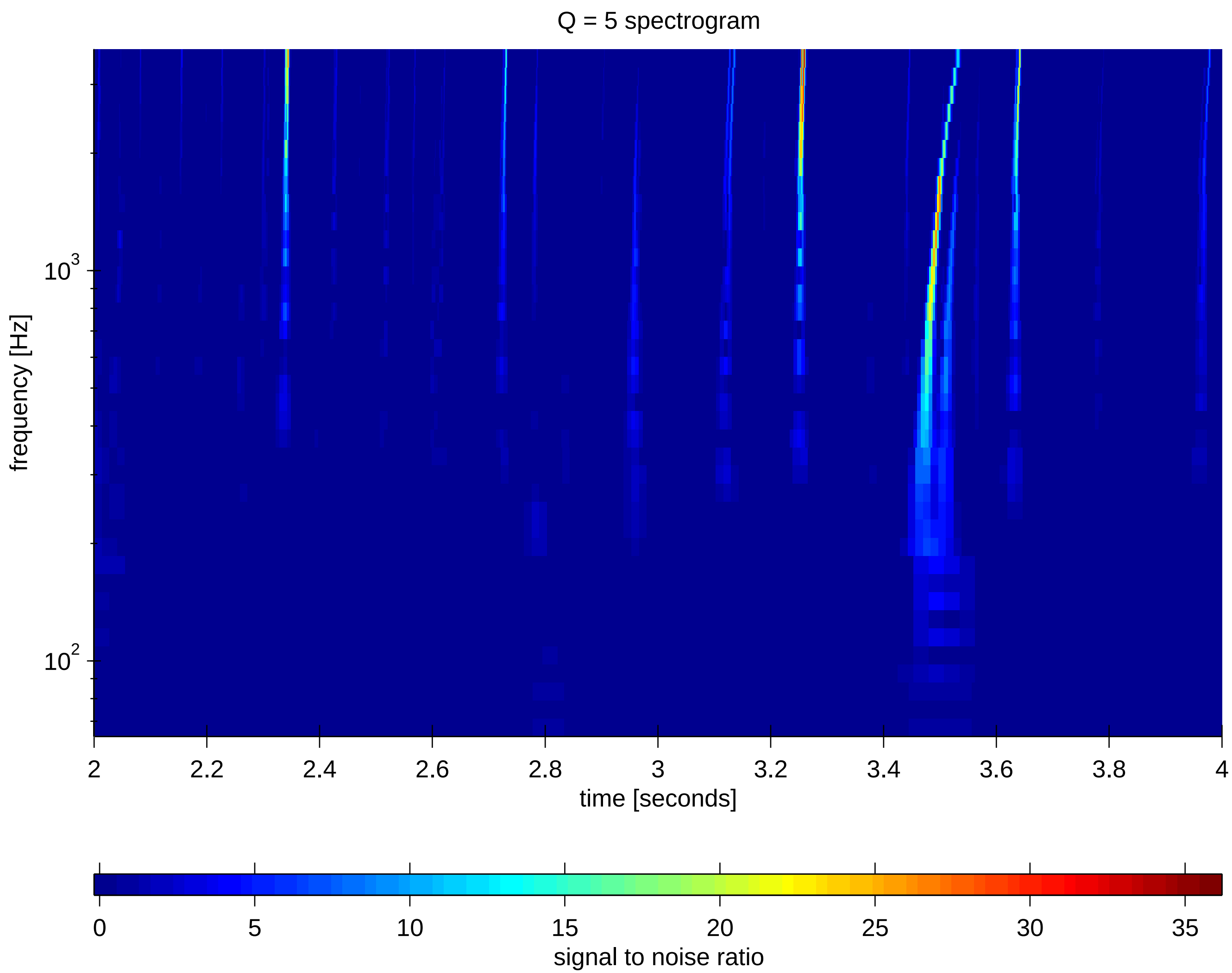
Veto Examples

Glitches seen in H1:LSC-AS_Q

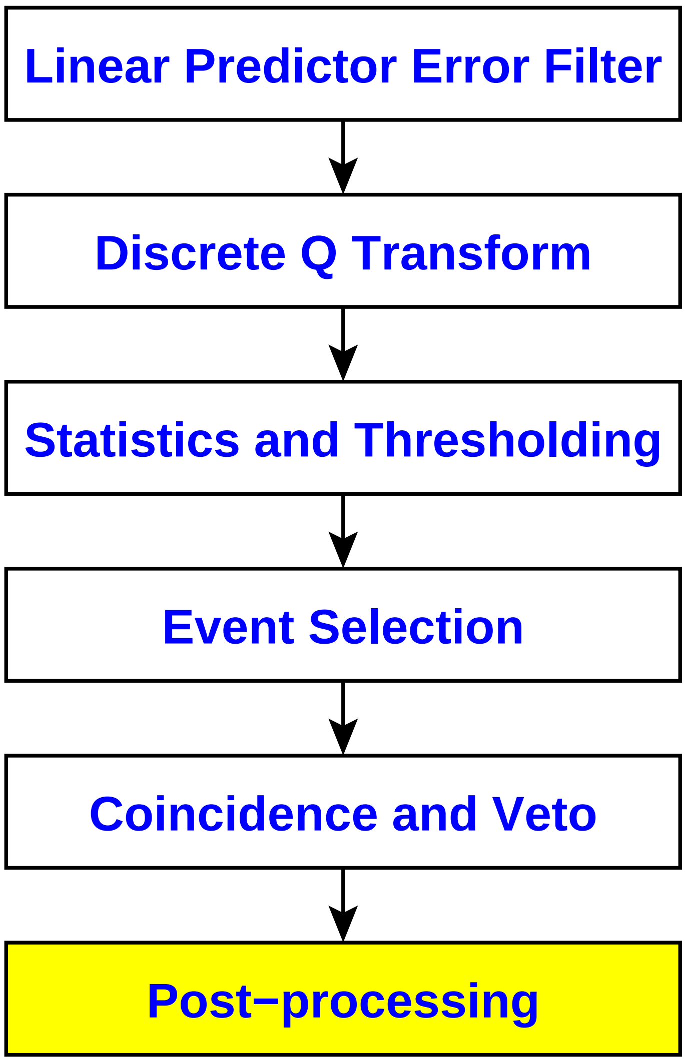


Veto Examples

Glitches seen in H1:LSC-POB_Q



Post-Processing?



- Not yet implemented, but possible options include:
- Parameter estimation
  - Use calibrated data
  - Burst DSO
- Waveform consistency test
  - Cross-correlation
- Amplitude consistency test
  - Use time delay
- Heirarchical search
  - Adaptive search for best pixel match

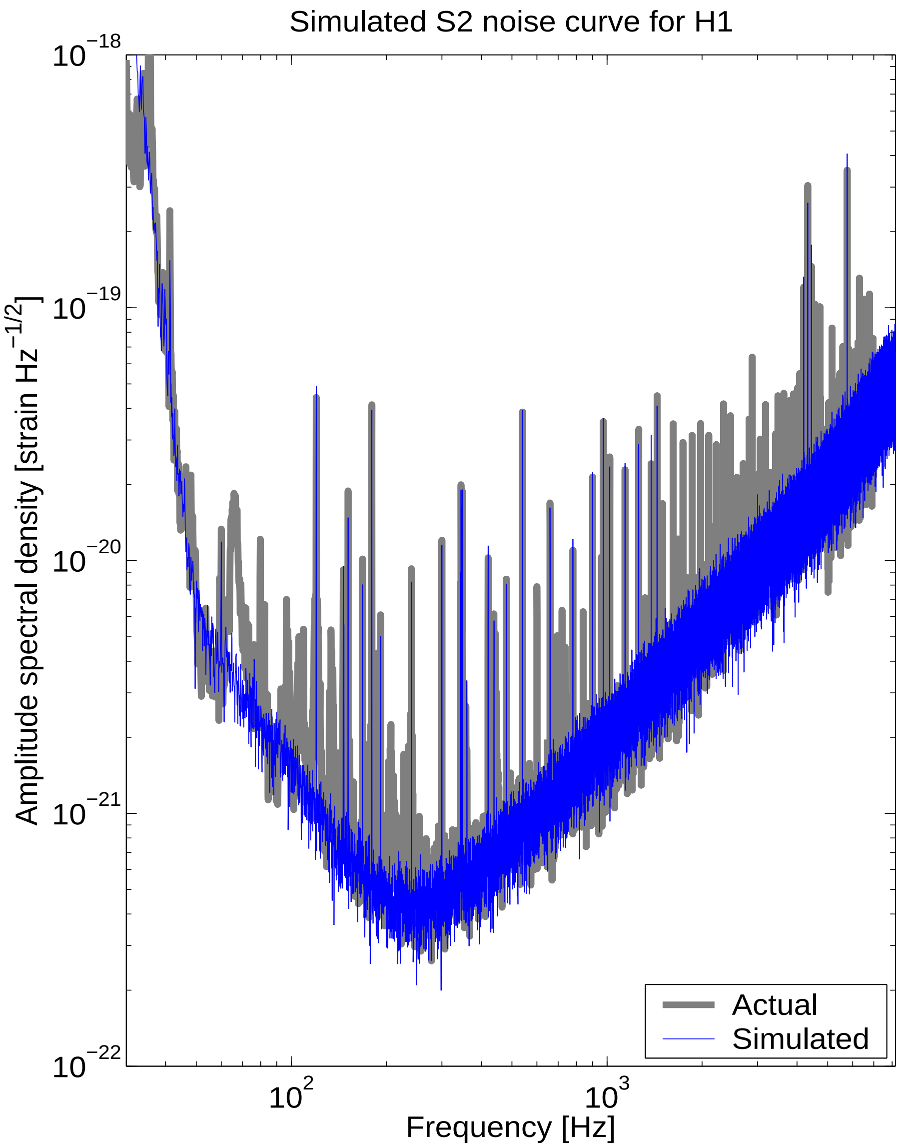
LIGO

Tuning the Q Pipeline

- Linear predictor error filtering greatly simplifies tuning
- Reasonable choices exist for most parameters
- Independent parameters:
  - Frequency band
  - Targeted range of Q
  - Maximum SNR loss due to pixel mismatch
  - Coincidence window duration and bandwidth
  - Triple coincidence false rate
- Dependent parameters:
  - Linear predictor error fi lter order
  - Linear predictor error fi lter training time
  - Data block duration
  - Time-frequency pixel overlap
  - Signifi cance threshold (under construction)

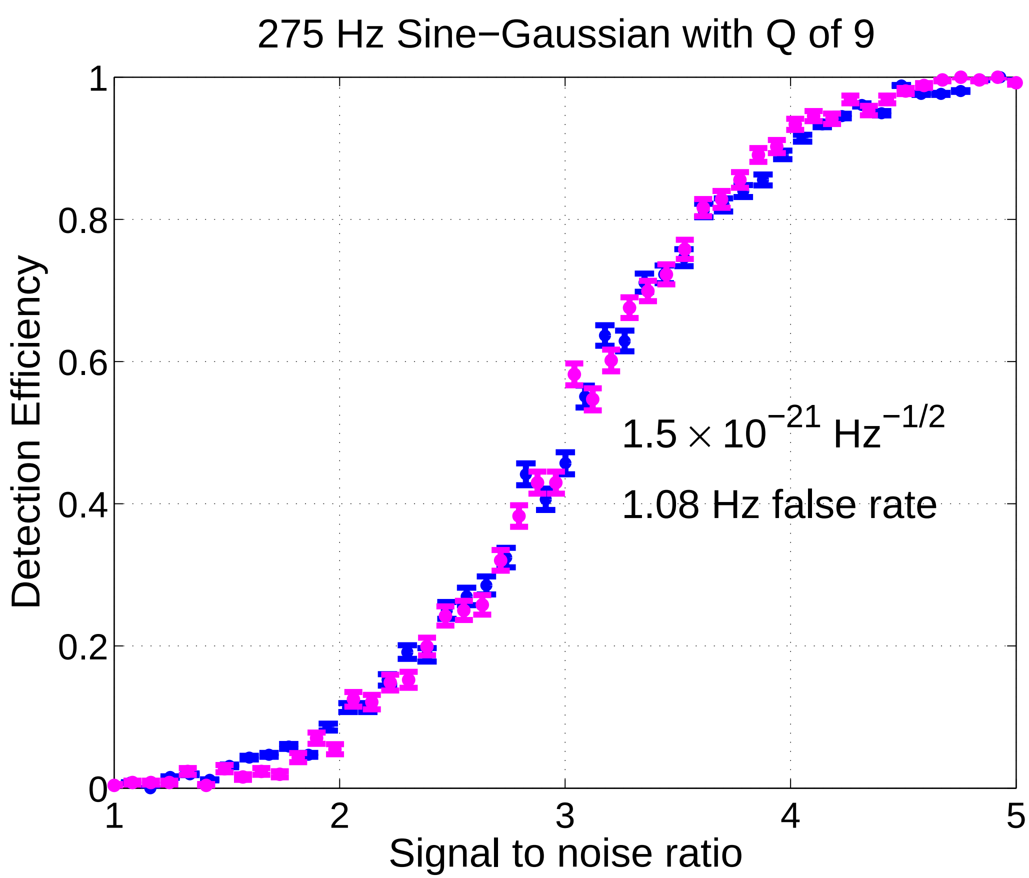
Simulated Data

- Simulated S2 H1 noise
- Shaped Gaussian white noise
- Included major lines
- Random injections
  - Sine-gaussians
- Caveat: No glitches



Detection Efficiency

Detection efficiency for simulated S2 H1 data

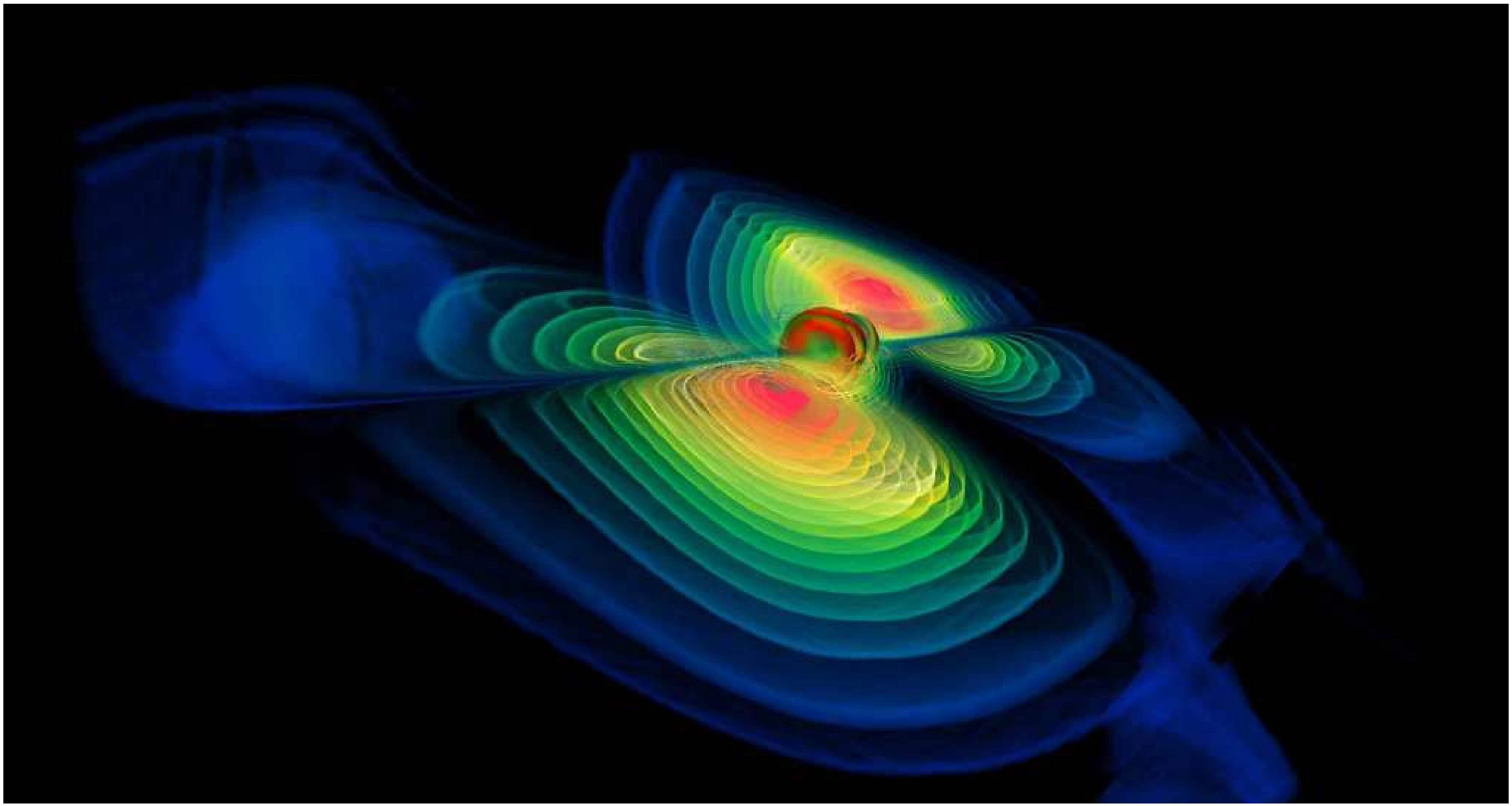


Emperical

Predicted

Black Hole Mergers

A potential target source for the Q pipeline?



LIGO

Black Hole Merger Model

- Equal mass black holes with no spin
- Optimally oriented with isotropic emission
- Fraction of rest mass energy emitted, $\epsilon = 0.01$
- Detectable amplitude signal to noise ratio, $\rho = 5$
- Dimensionless Kerr spin parameter, $a = 0.9$
- Energy distributed uniformly in frequency between the ISCO and QNM frequencies.

$$f_{\text{ISCO}} \simeq 2 \times 10^3 \left(\frac{M}{M_\odot} \right)^{-1} \text{Hz}$$

$$f_{\text{QNM}} \simeq 10^4 \left(\frac{M}{M_\odot} \right)^{-1} \text{Hz}$$

Black Hole Merger Energy

- Energy carried by a gravitational-wave burst:

$$4\pi r^2 \int_{-\infty}^{+\infty} \frac{c^3}{16\pi G} \left| \frac{d}{dt} h(t) \right|^2 dt = \epsilon\, 2Mc^2$$

isotropic emission — gravitational-wave energy flux — total radiated energy

- Detectable signal energy:

$$\int_{-\infty}^{+\infty} \left| \frac{d}{dt} h(t) \right|^2 dt = 4\pi^2 \int_{-\infty}^{+\infty} f^2 \left| \tilde{h}(f) \right|^2 df$$

$$= 4\pi^2 \langle f^2 \rangle h_{\text{rss}}^2 = 4\pi^2 \rho^2 \langle f^2 \rangle \langle S_h \rangle$$

Black Hole Merger Range

- Characteristic frequency:

$$\langle f^2\rangle = 2\int_0^\infty f^2|\tilde{\psi}(f)|^2\,df \simeq f_{\mathrm{ISCO}}f_{\mathrm{QNM}}$$

- Characteristic noise:

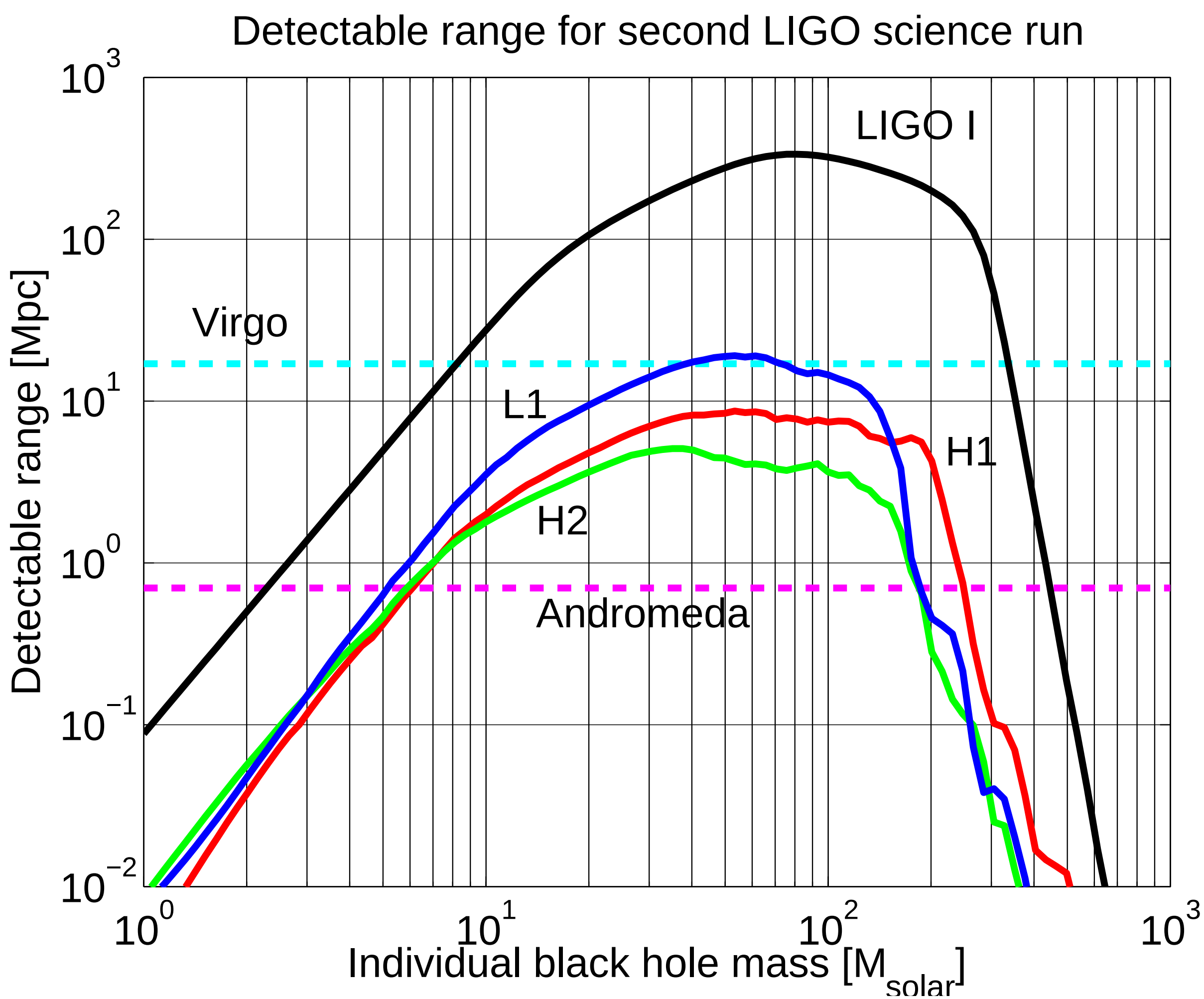
$$\langle S_h\rangle = \left[\int_0^\infty \frac{2|\tilde{\psi}(f)|^2}{S_h(f)}df\right]^{-1} \simeq \frac{f_{\mathrm{QNM}} - f_{\mathrm{ISCO}}}{f_{\mathrm{ISCO}}f_{\mathrm{QNM}}}\left(\int_{f_{\mathrm{ISCO}}}^{f_{\mathrm{QNM}}}\frac{df}{f^2S_h(f)}\right)^{-1}$$

- Detectable distance:

$$r = 2\times 10^{-19}\left(\frac{1}{\langle f^2\rangle\langle S_h\rangle}\right)^{1/2}\left(\frac{M}{M_\odot}\right)^{1/2}\ \mathrm{Mpc}$$

Black Hole Merger Range

Predicted from published detector noise spectra for the second LIGO science run



LIGO

Black Hole Merger Search

- Control room figure of merit for the burst search
- Matched filter search for inspirals and ringdowns in proximity to candidate burst events
  - Smaller data set allows deeper search
    - lower detection threshold
    - finer sampling of the template space
    - increased template space dimensionality (sky position, polarization, spin, etc.)
  - Candidate burst constrains astrophysical parameters
    - decreased template space volume
  - Astrophysical intrepretation for candidate bursts
- Hardware injection of full coalescence waveforms
  - end to end test of the pipeline

Summary

- Linear predictor error filtering greatly simplifies statistical analysis
- The discrete Q transform achieves near optimal time-frequency detection
- The Q pipeline provides a simple, computationally efficient, robust technique for near optimal time-frequency detection of gravitational wave bursts and detector characterization
- The merger phase of binary black hole coalescences is a potential target for the Q pipeline.

Current Status

- Implementation
  - LPEF: Matlab, DMT
  - DQT: Matlab
  - Event selection: C++
- Linear predictor error filters
  - S2 burst analysis
    - Data conditioning
    - Parameter estimation
    - Post-processing
    - Externally triggered search

See http://ligo.mit.edu/~shourov/

LIGO

Future Plans

- Apply to S2 data
- Pipeline Tuning
- Implementation
  - LDAS / LAL / BurstDSO
- Linear predictor error fi lters
  - Recursive least squares
  - Apply to other searches?
- Post-processing
  - Waveform and Amplitude consistency
  - Parameter estimation
- Black hole mergers
  - Burst fi gure of merit
  - Triggered search for inspirals and ringdowns

DQT Derivation

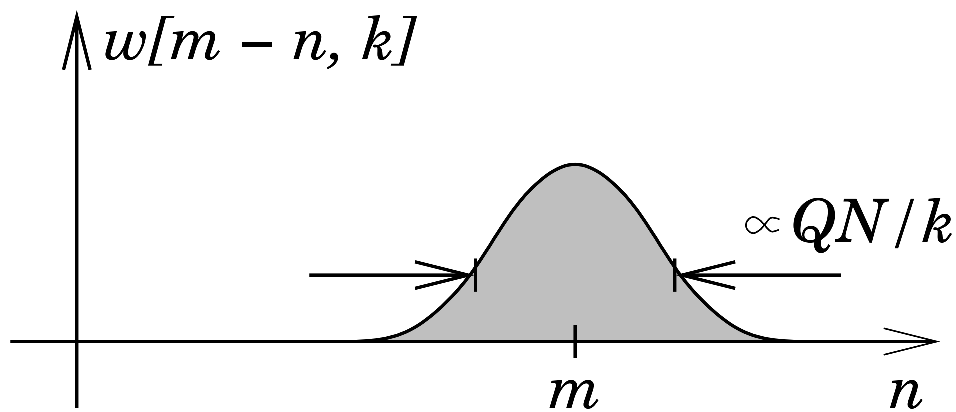
$$\tilde{X}[k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi nk/N}$$

Start with the discrete Fourier transform.

DQT Derivation

$$X_Q[m,k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi nk/N} \; w[m-n,k]$$

Introduce a shifted and scaled time domain window.



LIGO

DQT Derivation

$$X_Q[m,k] = \sum_{n=0}^{N-1} x[n]e^{-i2\pi nk/N} \; w[m-n,k]$$

rename

$$v[n,k]$$

LIGO

DQT Derivation

$$X_Q[m,k] = \sum_{n=0}^{N-1} \quad v[n,k] \quad w[m-n,k]$$

For constant k, this is a convolution in time.

DQT Derivation

$$X_Q[m,k] = \sum_{n=0}^{N-1} \quad v[n,k] \quad w[m-n,k]$$

Introduce the Fourier space:

$$\tilde{X}[l,k] = \sum_{n=0}^{N-1} x[n,k] e^{-i2\pi nl/N}$$

LIGO

DQT Derivation

$$\tilde{X}_Q[l,k] \quad = \quad \tilde{V}[l,k] \qquad \tilde{W}[l,k]$$

Convolution becomes multiplication.

DQT Derivation

$$\tilde{X}_Q[l,k] \quad = \qquad \tilde{V}[l,k] \qquad \tilde{W}[l,k]$$

frequency shift property

$$\tilde{V}[l,k] = \tilde{X}[l+k]$$

$$v[n,k] = x[n]e^{-i2\pi nk/N}$$

DQT Derivation

$$\tilde{X}_Q[l,k] \quad = \qquad\qquad \tilde{X}[l+k] \qquad \tilde{W}[l,k]$$

Inverse Fourier Transform yields...

DQT Derivation

$$X_Q[m,k] = \sum_{l=0}^{N-1} \tilde{X}[l+k] \quad \tilde{W}[l,k] \quad e^{-i2\pi ml/N}$$

the fast discrete Q-transform

Except for choice of window.

Index