China Winter School on LISA MLDC – Lecture 2

Soumya D. Mohanty

The University of Texas Rio Grande Valley

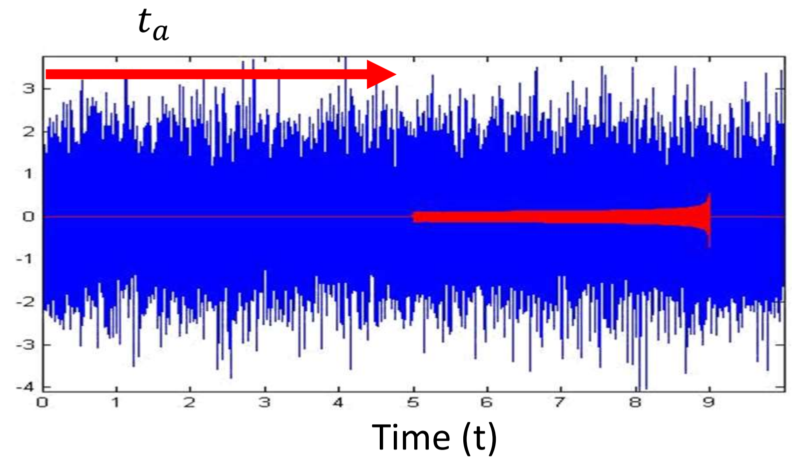
Plan

- Matched filtering
  - Special case: sinusoid signal (amplitude and initial phase parameters)
- Matched filtering in Fourier domain
- Lab exercises

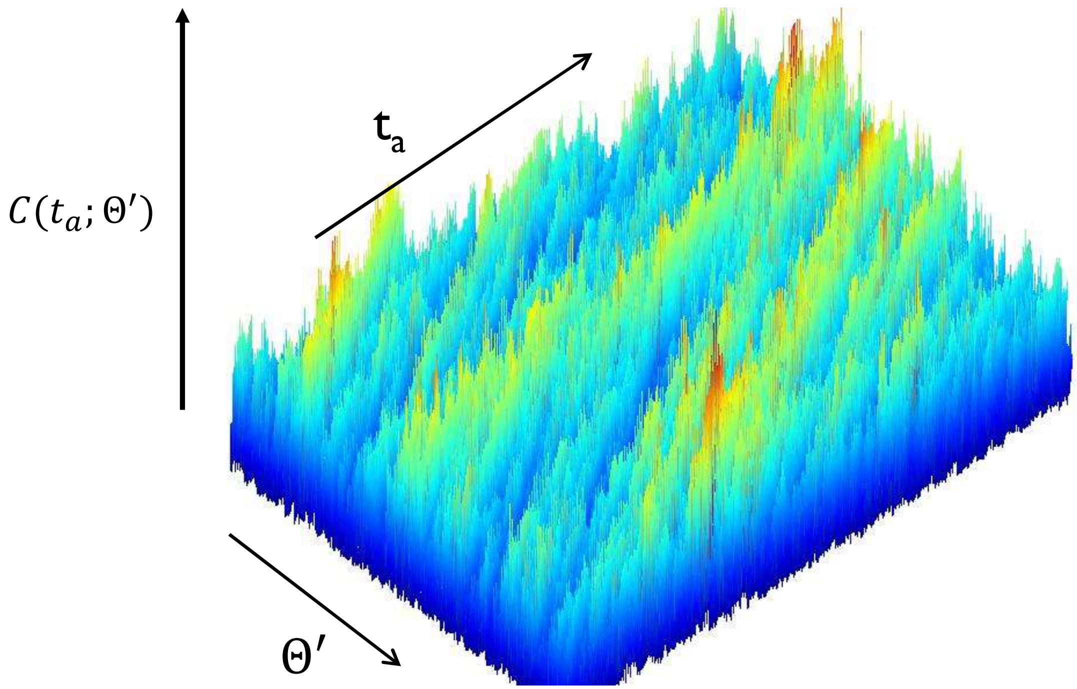
Least-squares fitting

- Given data: $(x_i, y_i); i = 0,1, \dots, N-1$
  - $\bar{y} = (y_0, y_1, \dots, y_{N-1})$
- Assumption: $y_i = f(x_i) + n_i$
  - n_i: Noise (random variable)
  - n_i, n_j: statistically independent if $i \neq j$
  - (Mean) $E[n_i] = 0$
  - (Variance) $E\left[n_i^2\right] = \sigma^2$
- Least squares fitting: $\min_{\Theta} \|\bar{y} - \bar{s}(\Theta)\|^2$
  - $\|\bar{x}\|^2 = \sum_{i=0}^{N-1} x_i^2$
  - Example: $\bar{s}(\Theta) = a\bar{x} + b$
- $\|\bar{y} - \bar{s}(\Theta)\|^2 = \|\bar{y}\|^2 - 2\langle \bar{y}, \bar{s}(\Theta) \rangle + \|\bar{s}(\Theta)\|^2$

Matched filtering



- Independent variable x is time t
  - $y(t) = s(t;\Theta) + n(t)$
  - $y_i = y(t_i)$
- $\Theta = \{t_a, \Theta'\}; s(t;\Theta) = s(t - t_a; \Theta')$
- Least-squares fitting over the time of arrival parameter is a filtering operation
  - $\langle \bar{y}, \bar{s} \rangle = \left(\frac{1}{\delta t}\right) \int_0^T dt\, y(t) s(t - t_a; \Theta') \ = C(t_a;\ \Theta')$: Correlation of data with signal
  - For fixed Θ', $\min_{\mathrm{t_a}} \|\bar{y} - \bar{s}\|^2$ is equivalent to $\max_{t_a} \langle \bar{y}, \bar{s} \rangle$
  - Convolution theorem: $\mathrm{F}[C(t)](f) = F[y(t)](f) \times F[s(t;\Theta')](f)$
  - $F[a(t)](f) = \int_{-\infty}^{\infty} dt\, a(t) e^{-2\pi i f t}$ (Fourier transform)
  - Caution: very loose notation! Assumes $T \gg signal\ length$



Finding the global maximum of $C(t_a; \Theta')$ is a difficult optimization problem

- Large number of maxima
- Becomes worse as the number of parameters increases

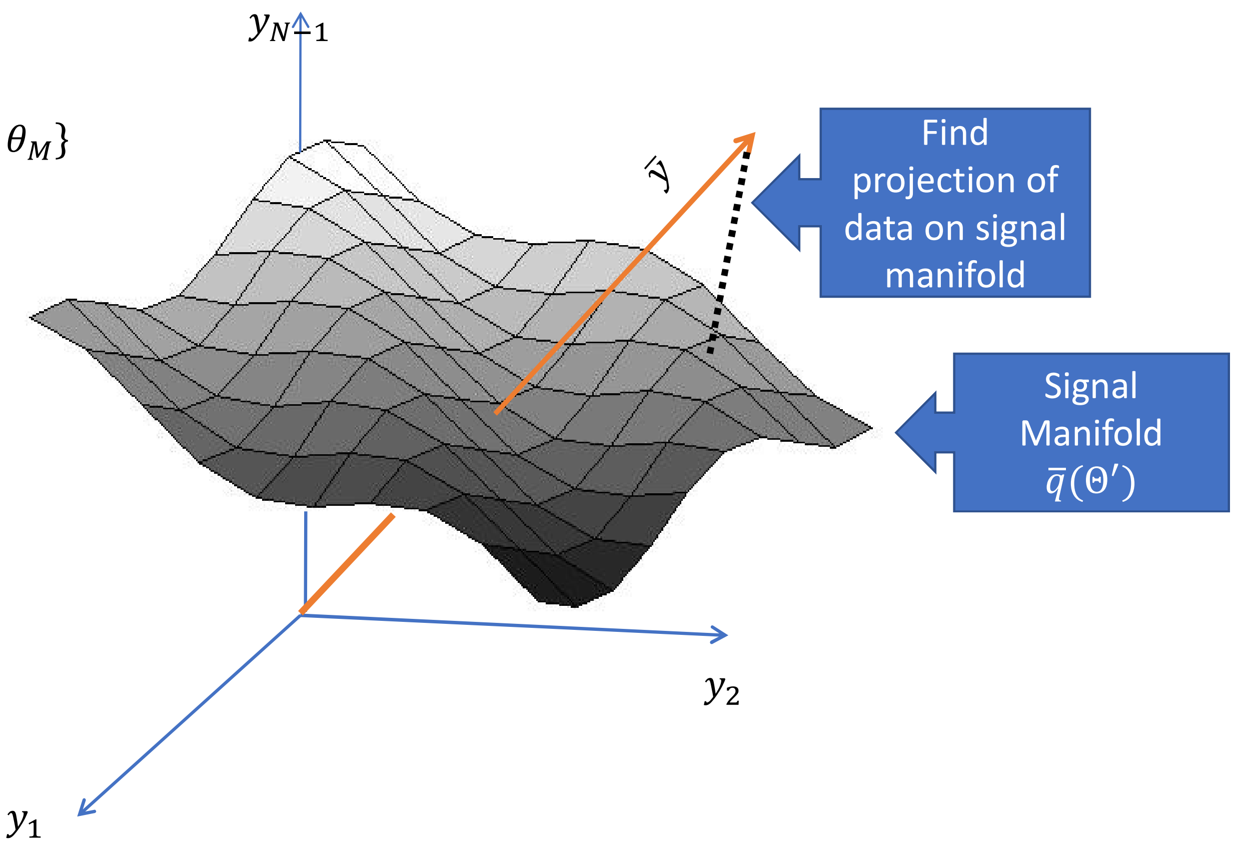
Computational cost of matched filtering is generally very high for LIGO data analysis

Least squares: general case

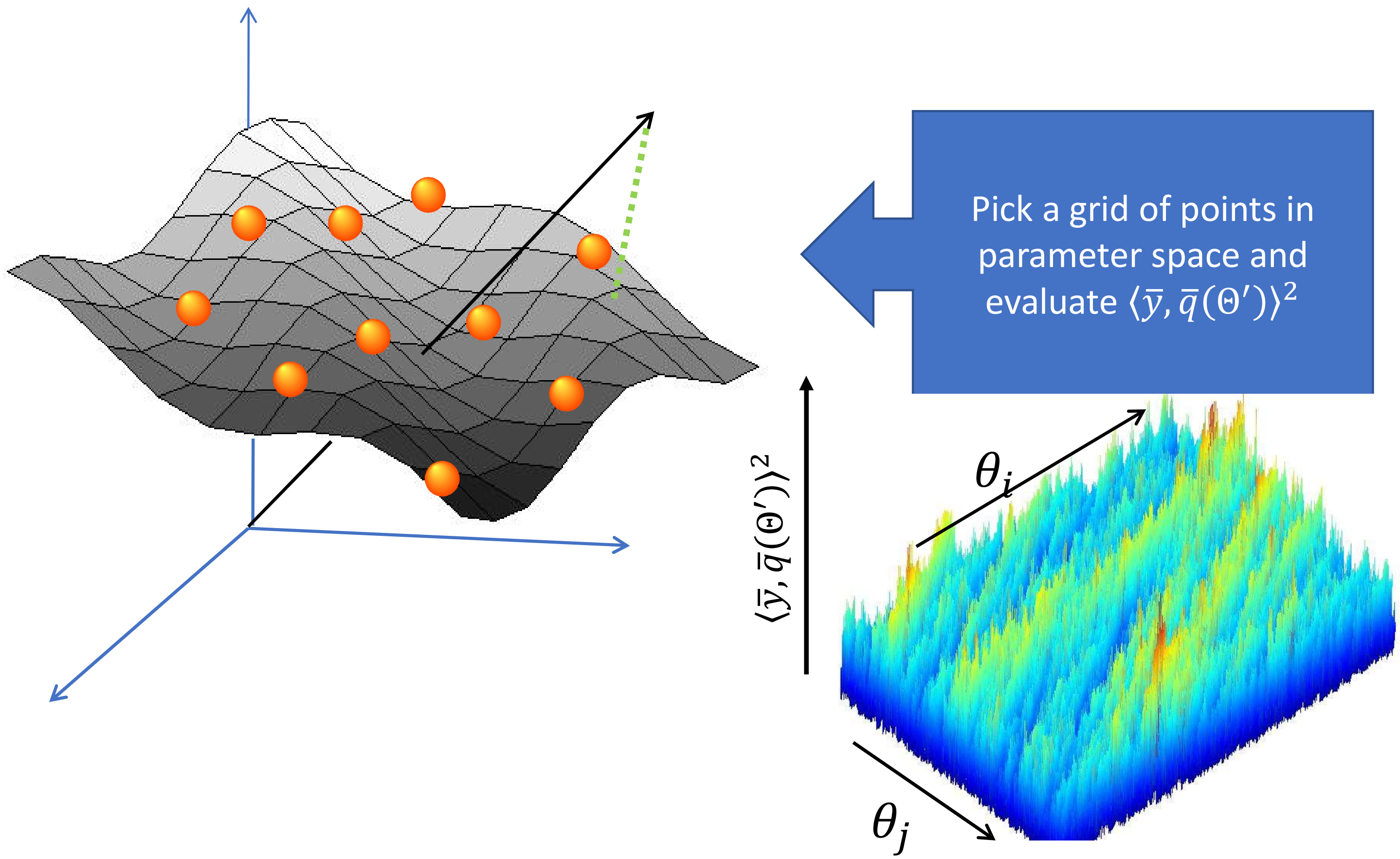
- $\|\bar{y} - \bar{s}(\Theta)\|^2 = \|\bar{y}\|^2 - 2\langle \bar{y}, \bar{s}(\Theta)\rangle + \|\bar{s}(\Theta)\|^2$
- $\langle \bar{y}, \bar{s}(\Theta)\rangle = \sum_{i=0}^{N-1} y_i s_i(\Theta)$
  - This is simply the Euclidean inner product of vectors
- One of the parameters that is always present for all GW signals is the signal amplitude (depends on distance r as $\frac{1}{r}$)
- $\bar{s}(A; \Theta') = A\bar{q}(\Theta')$
- To make the definition of A unique, let $\|\bar{q}(\Theta')\| = 1$.
- $\|\bar{y} - \bar{s}(\Theta)\|^2 = \|\bar{y}\|^2 - 2A\langle \bar{y}, \bar{q}(\Theta')\rangle + A^2$
- $\min_{\Theta}\|\bar{y} - \bar{s}(\Theta)\|^2 = \min_{\Theta'} \min_A(\|\bar{y}\|^2 - 2A\langle \bar{y}, \bar{q}(\Theta')\rangle + A^2)$
- Solution of inner minimization: $A = \langle \bar{y}, \bar{q}(\Theta')\rangle$
- $\min_{\Theta}\|\bar{y} - \bar{s}(\Theta)\|^2 = \min_{\Theta'} \left(-\langle \bar{y}, \bar{q}(\Theta')\rangle^2\right) \Rightarrow \max_{\Theta'}\langle \bar{y}, \bar{q}(\Theta')\rangle^2$

Geometrical picture

$\Theta' = \{\theta_1, \theta_2, \dots, \theta_M\}$



Grid-based search



Pick a grid of points in parameter space and evaluate $\langle \bar{y}, \bar{q}(\Theta') \rangle^2$

Amplitude and initial phase

- $\bar{s}(A,\phi_0;\Theta') = A\bar{q}(\phi_0;\Theta') = AN(\Theta')\sin(\phi(t;\Theta')+\phi_0)$
  - $N(\Theta')$: normalization constant
- $\bar{s}(A,\phi_0;\Theta') = A\,N(\Theta')\cos(\phi_0)\sin\big(\phi(t;\Theta')\big) + AN(\Theta')\sin(\phi_0)\cos(\phi(t;\Theta'))$
- $\bar{s}(A,\phi_0;\Theta') = X\bar{q}_0(\Theta') + Y\bar{q}_1(\Theta')$
- Assume: $\langle \bar{q}_0, \bar{q}_1\rangle \approx 0$
- $\|\bar{y} - \bar{s}(A,\phi_0;\Theta')\|^2 = \|\bar{y}\|^2 - 2X\langle \bar{y},\bar{q}_0\rangle - 2Y\langle \bar{y},\bar{q}_1\rangle + X^2 + Y^2$
- $\min_{\Theta'}\min_{A,\phi_0}\|\bar{y} - \bar{s}(A,\phi_0;\Theta')\|^2 = \min_{\Theta'}\min_{X,Y}\|\bar{y} - (X\bar{q}_0(\Theta') + Y\bar{q}_1(\Theta'))\|^2$
- Solution: $X = \langle \bar{y},\bar{q}_0\rangle,\ \ \mathrm{Y} = \langle \bar{y},\bar{q}_1\rangle$
- $\max_{\Theta'}\,[\langle \bar{y},\bar{q}_0(\Theta')\rangle^2 + \langle \bar{y},\bar{q}_1(\Theta')\rangle^2]$
- Special case: $\phi(t;\Theta') = \omega t;\quad \Theta' = \omega \rightarrow \max_\omega \left|\sum_{k=0}^{N-1} y_k e^{-i\omega k\delta t}\right|^2$
- Discrete Fourier transform: $\tilde{y}_p = \sum_{k=0}^{N-1} y_k e^{-2\pi i pk/N} \approx \left(\frac{1}{\delta t}\right)\int_0^T dt\ y(t)e^{-2\pi i\left(\frac{p}{T}\right)k\left(\frac{T}{N}\right)} = \left(\frac{1}{\delta t}\right)\int_0^T dt\ y(t)e^{-i\omega k\delta t} \approx \sum_{k=0}^{N-1} y_k e^{-i\omega k\delta t}$

Amplitude, initial phase, time of arrival

- $\max_{\Theta'} \max_{t_a} [\langle \bar{y}, \bar{q}_0(t - t_a; \Theta') \rangle^2 + \langle \bar{y}, \bar{q}_1(t - t_a;\ \Theta') \rangle^2]$
  - $\langle \bar{y}, \bar{q}_\alpha(t - t_a; \Theta') \rangle$: Filtering → The whole process is often called "matched filtering" also in GW data analysis
  - Correct name: Least squares fitting
- Example: Linear chirp signal
- $s(t; A, t_a, f_0, f_1, \phi_0) = \begin{cases} 0;\ \ t \notin [t_a, t_a + L] \\ A\, N(f_0, f_1) \sin(2\pi(f_0(t - t_a) + f_1(t - t_a)^2) + \phi_0) \end{cases}$
- Tasks (Track 1&2):
  - Generate data with this signal and noise
  - Noise: just a sequence of independently drawn values from a Gaussian random number generator
  - Implement matched filtering (whole process) to find this signal

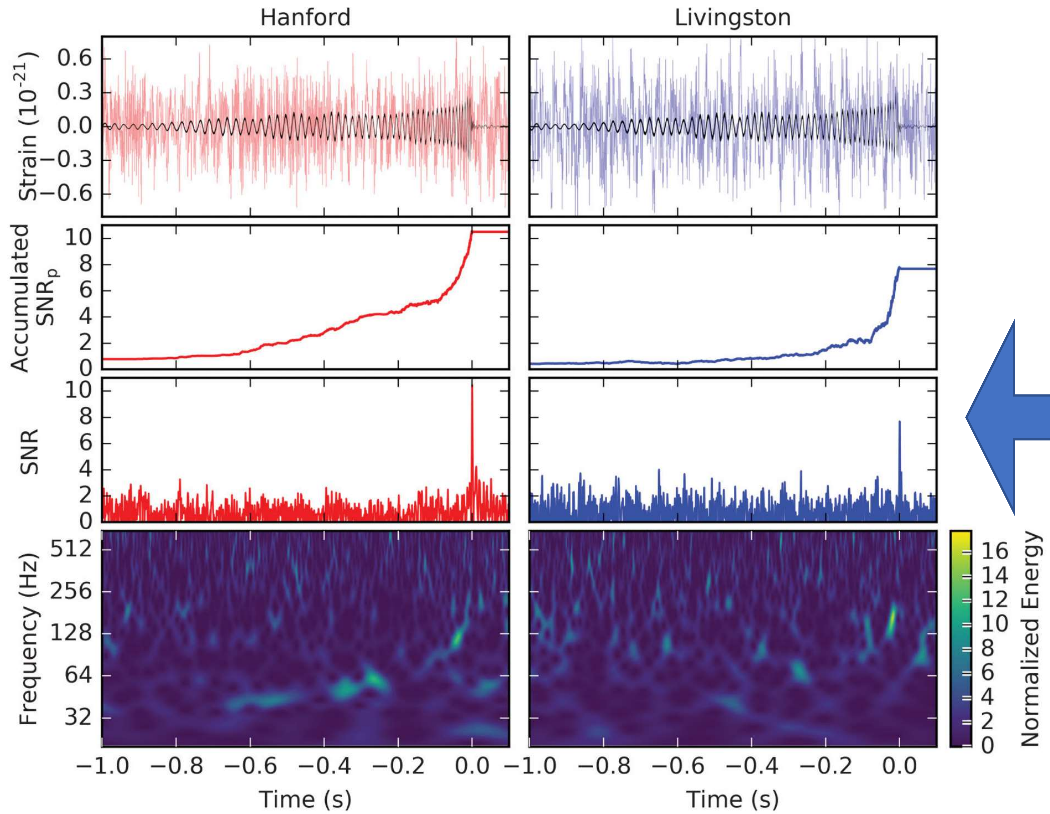
Efficient implementation of filtering

- $\langle \bar{y}, \bar{s}(t_a;\Theta')\rangle = \left(\frac{1}{\delta t}\right)\int_0^T dt\, y(t)s(t-t_a;\Theta') \;= C(t_a;\;\Theta')$: Correlation of data with signal
- For fixed Θ', $\min_{\mathrm{t_a}}\|\bar{y}-\bar{s}\|^2$ is equivalent to $\max_{t_a}\langle \bar{y}, \bar{s}\rangle$
- Convolution theorem: $\mathrm{F}[C(t)](f) = F[y(t)](f) \times F[s(t;\Theta')](f)$

- In terms of Discrete Fourier Transform: $F \rightarrow DFT$
- Inverse DFT: $y_k = \left(\frac{1}{N}\right)\sum_{p=0}^{N-1} \tilde{y}_p e^{2\pi i\, kp/N}$
- $\sum_{p=0}^{N-1} e^{-2\pi i \quad /N} \sum_{l=0}^{N-1} y_l s_{l-p} = (\frac{1}{N^2}) \sum_{r=0}^{N-1}\sum_{s=0}^{N-1} \tilde{y}_r\, \tilde{s}_s{}^* \sum_{p=0}^{N-1}\sum_{l=0}^{N-1} e^{-2\pi i \quad /N} e^{2\pi irl/N} e^{-2\pi i\, s(l-p)/N}$
- $= \left(\frac{1}{N^2}\right)\sum_{r=0}^{N-1}\sum_{s=0}^{N-1} N^2 \delta_{sr}\delta_{ks}\, \tilde{y}_r \tilde{s}_s{}^* = \tilde{y}_k \tilde{s}_k{}^*$
- (Circular shift of signal: $s_{k<0} = s_{N+k}$)
- Multiply (sample by sample) DFT of data and (complex conjugate) of DFT of signal → take inverse DFT
- Use Fast Fourier Transform (FFT) implementation of DFT

GW 151226

PRL **116,** 241103 (2016) PHYSICAL REVIEW LETTERS



Matched filter output for the filter with best matching mass parameters