GLRT for Gaussian noise

GLRT for Gaussian noise in GW data analysis

- Gaussian stationary noise is the main model used in all GW data analysis algorithms
  - Extra steps are needed (e.g., line removal) to deal with the effects of non-Gaussian and non-stationary noise
- We will look at some of the principal forms of GLRT in Gaussian stationary noise that appear across all GW data analysis
  - Unknown amplitude
  - Unknown time of arrival
  - Unknown initial phase
  - Unknown amplitude, time of arrival, and initial phase

GLRT for Gaussian noise: Starting point
(See Chapter 1.3 and 1.4 of textbook)

- Data :

$$\bar{y} = \bar{s}(\Theta) + \bar{n};$$

$\bar{n}$: realization of zero mean Gaussian noise
Θ: set of signal parameters

- Let

$$\langle \bar{z}, \bar{y} \rangle = \bar{z}^T \mathbf{C}^{-1} \bar{y},$$

where $\mathbf{C}$ is the covariance matrix of the noise, and

$$\|\bar{z}\|^2 = \langle \bar{z}, \bar{z} \rangle.$$

- Then, the GLRT is

$$\mathrm{L}_G(\bar{y}) = \max_{\Theta} \ln L\,(\bar{y};\Theta) = \max_{\Theta} \left[-\frac{1}{2} \|\bar{y} - \bar{s}(\Theta)\|^2 + \frac{1}{2} \|\bar{y}\|^2 \right]$$

$$L_G(\bar{y}) = \max_{\Theta} \left(\langle \bar{y}, \bar{s}(\Theta) \rangle - \frac{1}{2} \|\bar{s}(\Theta)\|^2 \right)$$

Amplitude normalization

- Convenient normalization of signals:

$$\bar{s}(\Theta) = \frac{\|\bar{s}(\Theta)\|}{\|\bar{s}(\Theta)\|}\bar{s}(\Theta) = \overbrace{\|\bar{s}(\Theta)\|}^{vector\ length}\ \overbrace{\frac{\bar{s}(\Theta)}{\|\bar{s}(\Theta)\|}}^{unit\ vector} = \|\bar{s}(\Theta)\|\bar{q}(\Theta) = A\,\bar{q}(\Theta)$$

$\mathrm{A} = \|\bar{s}(\Theta)\|;\ \ \|\bar{q}(\Theta)\| = 1 \Rightarrow \bar{q}(\Theta)$ is the unit norm signal

- Any overall factor in $\bar{s}(\Theta)$ is now absorbed in A
  - Note: A **is simply the SNR for the Likelihood Ratio test**
- If this overall factor was in the set of unknown parameters, the set now becomes

$$\Theta = \{A, \Theta'\}$$

where Θ' denotes all remaining parameters

  - $\bar{q}(\Theta)$ now depends only on Θ':

$$\bar{q}(\Theta) \rightarrow \bar{q}(\Theta')$$

- **Then**

$$\mathrm{L}_G(\bar{y}) = \max_{\Theta}\left(\langle\bar{y}, \bar{s}(\Theta)\rangle - \frac{1}{2}\|\bar{s}(\Theta)\|^2\right) \rightarrow L_G(\bar{y}) = \max_{A,\Theta'}\left(A\langle\bar{y}, \bar{q}(\Theta')\rangle - \frac{1}{2}A^2\right)$$

Unknown amplitude
(Also see Appendix C.1 of textbook)

$$L_G(\bar{y}) = \max_{A,\Theta'}\left(A\langle\bar{y},\bar{q}(\Theta')\rangle - \frac{1}{2}A^2\right) = \max_{\Theta'}\left(\max_{A}\left(A\langle\bar{y},\bar{q}(\Theta')\rangle - \frac{1}{2}A^2\right)\right)$$

- Solution of inner minimization:

$$A = \langle\bar{y},\bar{q}(\Theta')\rangle$$

- Hence

$$L_G = \max_{\Theta'}\langle\bar{y},\bar{q}(\Theta')\rangle^2$$

- (From now, $\Theta' \rightarrow \Theta$ the set of parameters **besides amplitude**)

$$L_G = \max_{\Theta}\langle\bar{y},\bar{q}(\Theta)\rangle^2$$

Inner product for stationary noise

$$L_G = \max_{\Theta} \langle \bar{y}, \bar{q}(\Theta\,) \rangle^2$$

- White Gaussian Noise

$$\langle \bar{x}, \bar{y} \rangle \to \sum_{k=0}^{N-1} x_k y_k$$

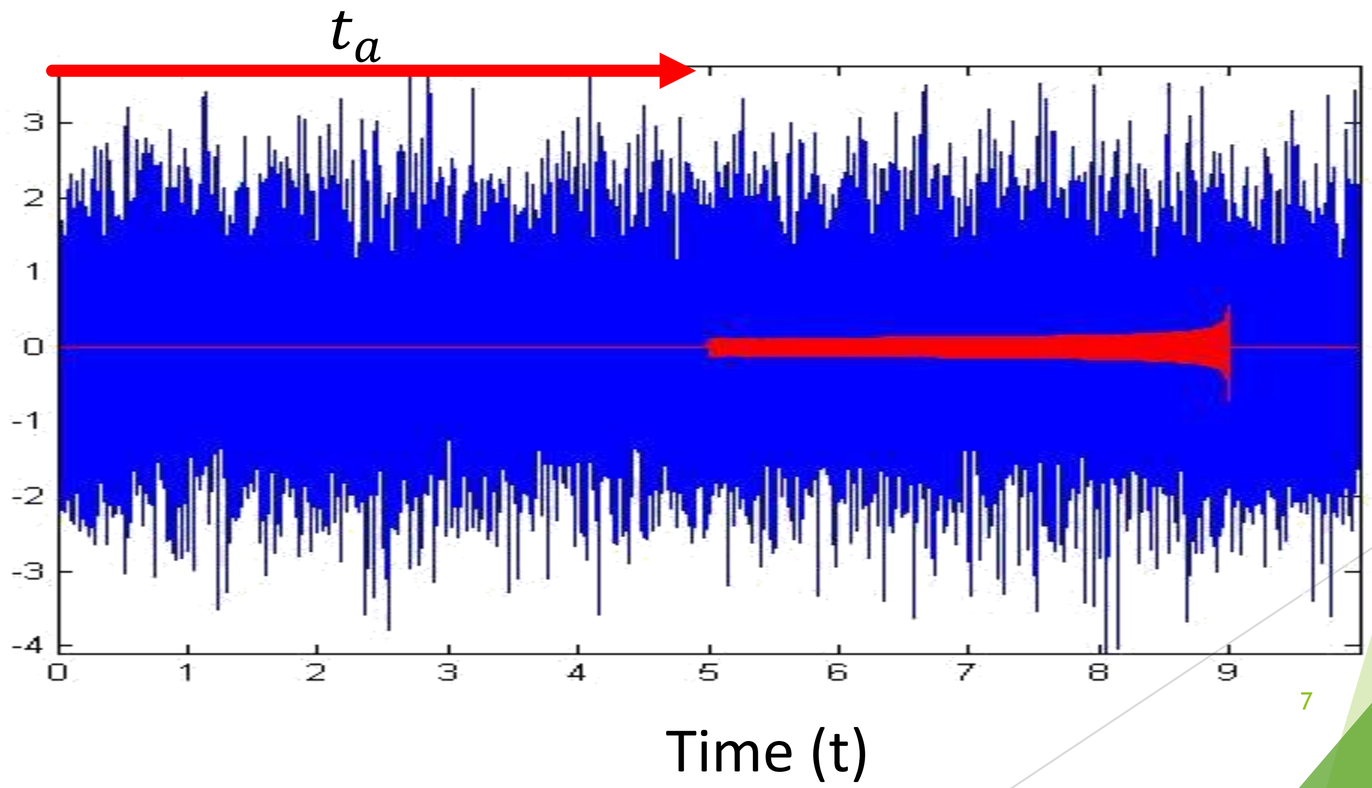
- Stationary Gaussian noise with Power Spectral Density (PSD) $S_n(f)$

$$\langle \bar{x}, \bar{y} \rangle \to \frac{\Delta}{N} \tilde{x} (\tilde{y}^{\dagger} \;./\; \bar{S}_n{}^T)$$

Where $\tilde{x} = F\bar{x}$ is the DFT

Unknown time of arrival

- Signal start time ("time of arrival") parameter: t_a
- $q(t; t_a, \Theta') = q^{(0)}(t - t_a; \Theta')$
- $q^{(0)}(t; \Theta')$: unit norm signal at $t_a = 0$



Matched filtering

- $L_G = \max_{\Theta} \langle \bar{y}, \bar{q}(\Theta) \rangle^2 = \max_{\Theta', t_a} \langle \bar{y}, \bar{q}(t_a, \Theta') \rangle^2$ uses inner product $\langle \bar{y}, \bar{q}(t_a, \Theta') \rangle$
- Obtaining $\langle \bar{y}, \bar{q}(t_a, \Theta') \rangle$ as a function of t_a is a **filtering** operation

$$\langle \bar{y}, \bar{q}(t_a, \Theta') \rangle = \bar{y}\mathbf{C}^{-1}\,\bar{q}^T(t_a, \Theta') = \bar{z}\bar{q}^T(t_a, \Theta')$$

$$= \sum_{k=0}^{N-1} z_k q_k(t_a, \Theta') = \frac{1}{\delta t}\delta t \sum_{k=0}^{N-1} z_k q^{(0)}(t_k - t_a, \Theta')$$

$$\approx \left(\frac{1}{\delta t}\right) \underbrace{\int_0^T dt\, z(t) q^{(0)}(t - t_a; \Theta')}_{Correlation} = \frac{1}{\delta t} \underbrace{\int_0^T dt\, z(t) Q^{(0)}(t_a - t\,; \Theta')}_{Convolution}$$

$$= \textbf{Filtering } z(t) \text{ with } \textbf{impulse response } Q^{(0)}(t) = q^{(0)}(-t)$$

- Filtering done with filter that “matches” the signal $\rightarrow$ **Matched filtering**

Important operations on signals

Convolution of two analog signals,

$f(t)$ and $g(t)$:

$$z(t) = (f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$

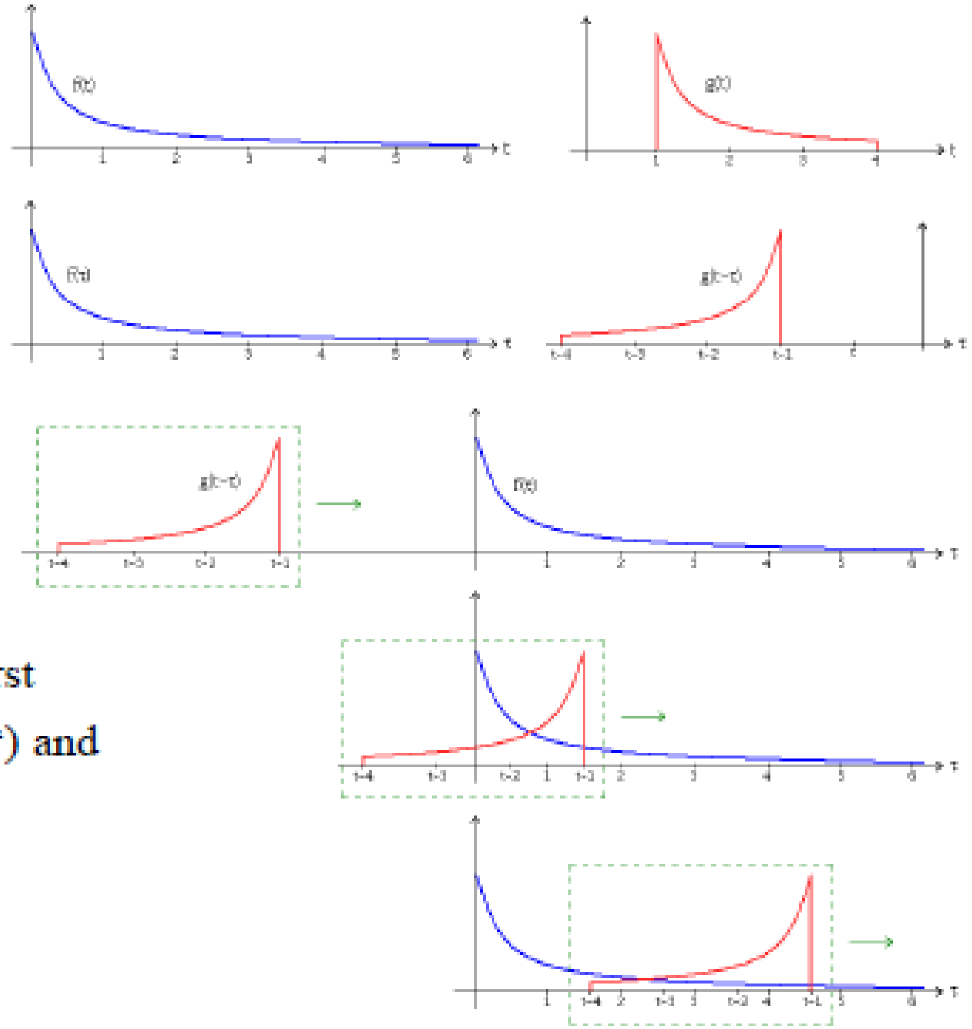
$$= \int_{-\infty}^{\infty} f(t-\tau)g(\tau)d\tau$$

Correlation:

$$z(t) = (f \wedge g)(t) = \int_{-\infty}^{\infty} f(\tau)g(\tau - t)d\tau$$

$$= \int_{-\infty}^{\infty} f(\tau + t)g(\tau)d\tau$$

In convolution, one of the functions is first flipped into its mirror image $g(t) \rightarrow g(-t)$ and then a correlation is computed



Efficient implementation of matched filtering

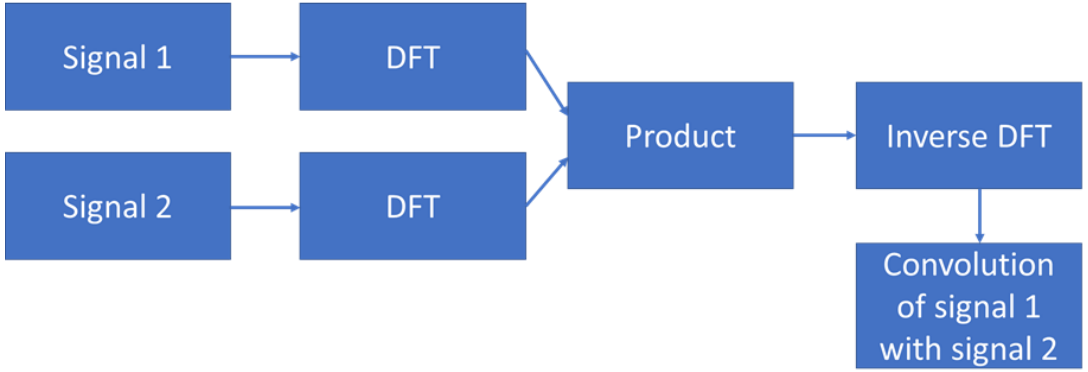
- Since $q^{(0)}(t;\Theta')$ is finite in length, the filtering operation is FIR filtering
- $\Rightarrow \langle \bar{y}, \bar{q}\ (t_a, \Theta')\rangle$ can be implemented efficiently using FFT based correlation

1. Divided (sample by sample) FFT of data $\tilde{y}$ by PSD $\rightarrow \tilde{z}$
2. Multiply (sample by sample) $\tilde{z}$ and (complex conjugate) of FFT of template (having $t_a = 0$)
3. Take inverse FFT

$$F^{-1}\left[(\tilde{y}./\bar{S}_n{}^T) .* (\tilde{q}^{(0)})^*\right] \rightarrow \langle \bar{y}, \bar{q}_a(t_a, \Theta')\rangle$$

for $t_a = k\Delta, k = 0,1,\ldots,N-1$

$$\mathrm{DFT}[\bar{x} * \bar{y}] = \tilde{x}.* \tilde{y} \Rightarrow \bar{x} * \bar{y} = \mathrm{DFT}^{-1}[\tilde{x}.* \tilde{y}]$$



Unknown amplitude and initial phase

$$\bar{s}(\Theta) = A\,\bar{q}(\Theta) \text{ where}$$

$$\bar{q}(\Theta) = \bar{q}(\phi_0, \Theta')$$

with

$$q_k(\phi_0, \Theta') = N(\Theta') \sin(\phi(t_k; \Theta') + \phi_0)$$

- Sinusoidal signal
  - $s(t) = A \sin(2\pi f_0 t + \phi_0)$
  - Parameters: A, f_0, ϕ_0
- Linear chirp signal
  - $s(t) = A \sin(2\pi(f_0 t + f_1 t^2) + \phi_0)$
  - Parameters: A, f_0, f_1, ϕ_0
- Sine-Gaussian signal
  - $s(t) = A \exp\left(-\frac{(t-t_0)^2}{2\sigma^2}\right) \sin(2\pi f_0 t + \phi_0)$
  - Parameters: $A, t_0, \sigma, f_0, \phi_0$

Unknown amplitude and initial phase

$$q_k(\phi_0, \Theta') = N(\Theta')\sin(\phi(t_k;\Theta') + \phi_0)$$

$$\Rightarrow q_k(\phi_0, \Theta') = \underbrace{N(\Theta')\sin(\phi(t_k;\Theta'))}_{q_{0,k}} \underbrace{\cos(\phi_0)}_{X} + \underbrace{N(\Theta')\cos(\phi(t_k;\Theta'))}_{q_{1,k}} \underbrace{\sin(\phi_0)}_{Y}$$

- $N(\Theta')$: normalization constant such that $\|\bar{q}(\phi_0, \Theta')\| = 1$
- Therefore,

$$\bar{q}(\phi_0, \Theta') = X\,\bar{q}_0(\Theta') + Y\,\bar{q}_1(\Theta')$$

$$X^2 + Y^2 = 1$$

Unknown amplitude and initial phase

- From the GLRT for unknown amplitude,

$$L_G = \max_{\Theta} \langle \bar{y}, \bar{q}(\Theta\ ) \rangle^2$$
$$= \max_{\Theta'} \max_{\substack{X,Y \\ X^2+Y^2=1}} \left(X\langle \bar{y}, \bar{q}_0(\Theta') \rangle + Y\langle \bar{y}, \bar{q}_1(\Theta') \rangle \right)^2$$

- The quantity to be maximized is of the form

$$(X\langle \bar{y}, \bar{q}_0(\Theta') \rangle + Y\langle \bar{y}, \bar{q}_1(\Theta') \rangle)^2 = (A_1 n_1 + A_2 n_2)^2 = (\bar{A}.\hat{n})^2$$

Where $\hat{n} = (X, Y)$ is a unit vector since $X^2 + Y^2 = 1$

Unknown amplitude and initial phase

- Solution: $(\bar{A}.\hat{n})^2$ is maximized when the unit vector $\hat{n}$ points along vector $\hat{A} = \bar{A}/|\bar{A}| \Rightarrow$

$$X = \frac{\langle \bar{y}, \bar{q}_0 \rangle}{\sqrt{\langle \bar{y}, \bar{q}_0 \rangle^2 + \langle \bar{y}, \bar{q}_1 \rangle^2}}, \mathrm{Y} = \frac{\langle \bar{y}, \bar{q}_1 \rangle}{\sqrt{\langle \bar{y}, \bar{q}_0 \rangle^2 + \langle \bar{y}, \bar{q}_1 \rangle^2}}$$

$$\Rightarrow L_G = \max_{\Theta'} [\langle \bar{y}, \bar{q}_0(\Theta') \rangle^2 + \langle \bar{y}, \bar{q}_1(\Theta') \rangle^2]$$

- (From now, $\Theta' \rightarrow \Theta$ the set of parameters **besides amplitude and initial phase**)

$$\mathrm{L}_G = \max_{\Theta} [\langle \bar{y}, \bar{q}_0(\Theta) \rangle^2 + \langle \bar{y}, \bar{q}_1(\Theta) \rangle^2]$$

Monochromatic signal in WGN

- Special case: Monochromatic signal in WGN:
$$q(t;\Theta) \to q(t;\omega) = N(\omega)\sin(\omega t + \phi_0)$$
- N samples and Uniform sampling $\Rightarrow t_k = k\Delta, k = 0,1,\dots,N-1$
- WGN $(\sigma^2 = 1) \Rightarrow \langle \bar{x}, \bar{z}\rangle = \sum_{k=0}^{N-1} x_k z_k$ and $\|\bar{x}\|^2 = \langle \bar{x}, \bar{x}\rangle = \sum_{k=0}^{N-1} x_k^2$

$$\therefore \|\bar{q}(\omega)\|^2 = 1 \Rightarrow \langle \bar{q}(\omega), \bar{q}(\omega)\rangle = N^2(\omega)\sum_{k=0}^{N-1}\sin^2(\omega t_K)$$

$$\approx \frac{N^2(\omega)}{\Delta}\int_0^{T=(N-1)\Delta} dt\ \sin^2\omega t \approx \frac{N^2(\omega)}{2\Delta}T = 1 \Rightarrow N(\omega) = \sqrt{\frac{2\Delta}{T}}$$

$\Rightarrow$ Normalization factor $N(\omega) = N_0$ is independent of ω

$$\Rightarrow q_{1,k}(\omega) - iq_{0,k}(\omega) = N_0(\cos\omega t_k - i\sin\omega t_k) = N_0 e^{-i\omega t_k} = N_0 e^{-i\omega k\Delta}$$

Monochromatic signal in WGN

- Now, $A^2 + B^2 = |A - i\, B|^2$

$$\Rightarrow \langle \bar{y}, \bar{q}_0(\omega)\rangle^2 + \langle \bar{y}, \bar{q}_1(\omega)\rangle^2 = |\langle \bar{y}, \bar{q}_1(\omega)\rangle - i\langle \bar{y}, \bar{q}_0(\omega)\rangle|^2$$

$$= \left|\underbrace{\langle \bar{y}, \bar{q}_1(\omega) - i\, \bar{q}_0(\omega)\rangle}_{\langle \bar{x}, \bar{z}\rangle = \sum_{k=0}^{N-1} x_k z_k}\right|^2 = \left|\sum_{k=0}^{N-1} y_k \underbrace{\left(q_{1,k}(\omega) - i\, q_{0,k}(\omega)\right)}_{q_{1,k}(\omega) - i q_{0,k}(\omega) = N_0 e^{-i\omega k\Delta}}\right|^2$$

$$\mathrm{L}_G = \max_{\omega}\left|\sum_{k=0}^{N-1} y_k e^{-i\omega k\Delta}\right|^2 \text{(Ignoring overall constant factors)}$$

- Performing the search for the maximum over a regularly spaced grid in ω given by $\omega_p = 2\pi\, p/(N\Delta)$,
- $\Rightarrow$ L_G for monochromatic sinusoid in WGN: Compute the magnitude of the DFT of the data and find the frequency with the largest peak

Unknown Amplitude, initial phase, and time of arrival

$$\mathrm{L}_G = \max_{\Theta}[\langle \bar{y}, \bar{q}_0(\Theta)\rangle^2 + \langle \bar{y}, \bar{q}_1(\Theta)\rangle^2]$$

- $\Theta = \{ t_a, \Theta' \}$

$$\mathrm{L}_G = \max_{\Theta'}\left[\max_{t_a}(\langle \bar{y}, \bar{q}_0(t_a, \Theta')\rangle^2 + \langle \bar{y}, \bar{q}_1(t_a, \Theta')\rangle^2)\right] = \max_{\Theta'} \lambda(\Theta')$$

1. For a given Θ', Evaluate $\langle \bar{y}, \bar{q}_p(t_a, \Theta')\rangle$, $p = 0, 1$, using FFT based method $\rightarrow$ Two output time series
2. Square the samples of each time series
3. Add the two resulting time series
4. Find the maximum of this time series $\rightarrow$ get $\lambda(\Theta')$

Template bank implementation of GLRT

$$L_G = \max_{\Theta} \langle \bar{y}, \bar{q}(\Theta) \rangle^2$$

Grid search over parameters Θ

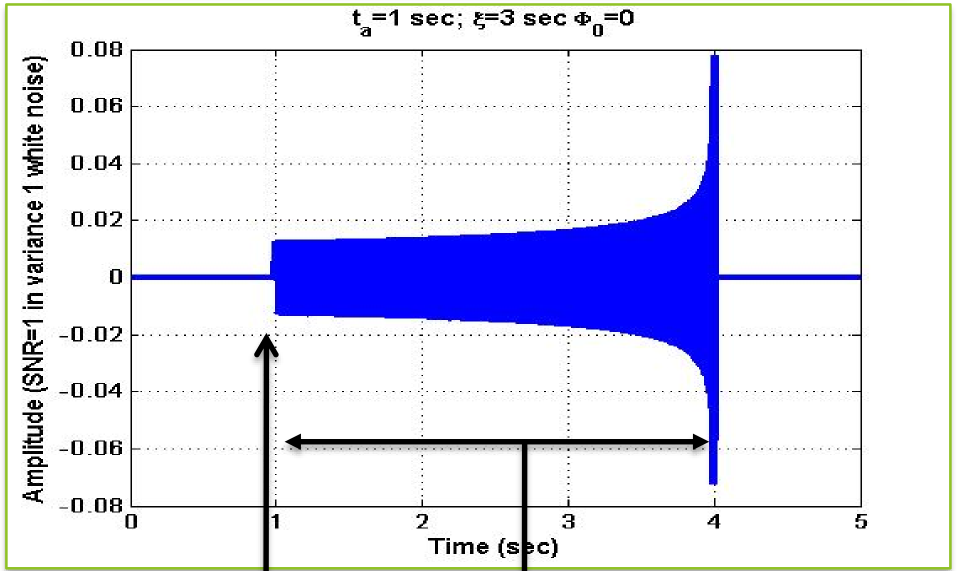
- Fix a grid of points in the parameter space (e.g., regularly spaced)
- For each point in the search grid, compute

$$\lambda(\Theta) = \langle \bar{y}, \bar{q}(\Theta) \rangle^2$$

  - $\bar{q}(\Theta)$ is called a template waveform
  - Unknown initial phase: $\bar{q}_0$, $\bar{q}_1$ are called quadrature templates
  - The set of template waveforms associated with the grid is called a template bank
- Find the grid point that gives the maximum value of $\lambda(\Theta)$
- Grid search over time of arrival parameter is efficient but not for other parameters

Example: Grid search Newtonian Binary Inspiral signal

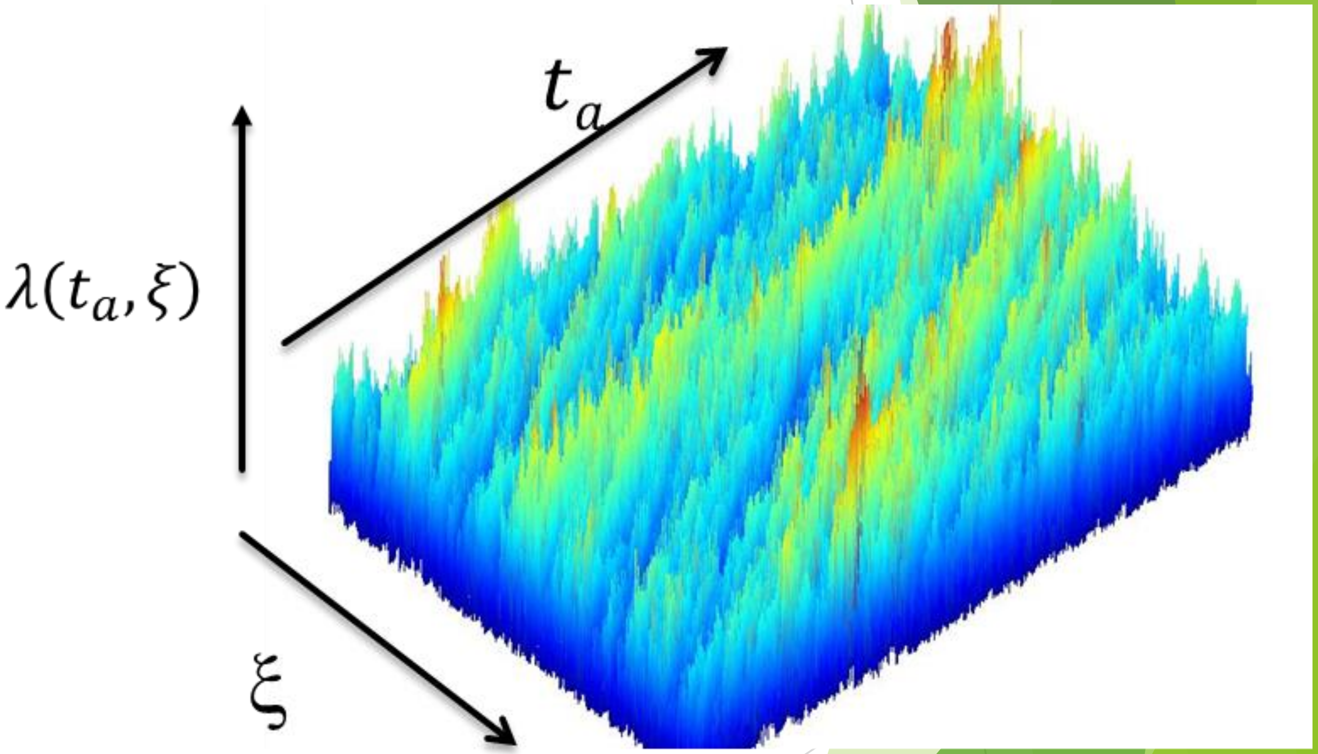
Newtonian inspiral template



t_a : time of arrival

Φ_0 : Initial phase

ξ : Chirp time



Due to noise in the data, $\lambda(\Theta)$ is a random field over Θ

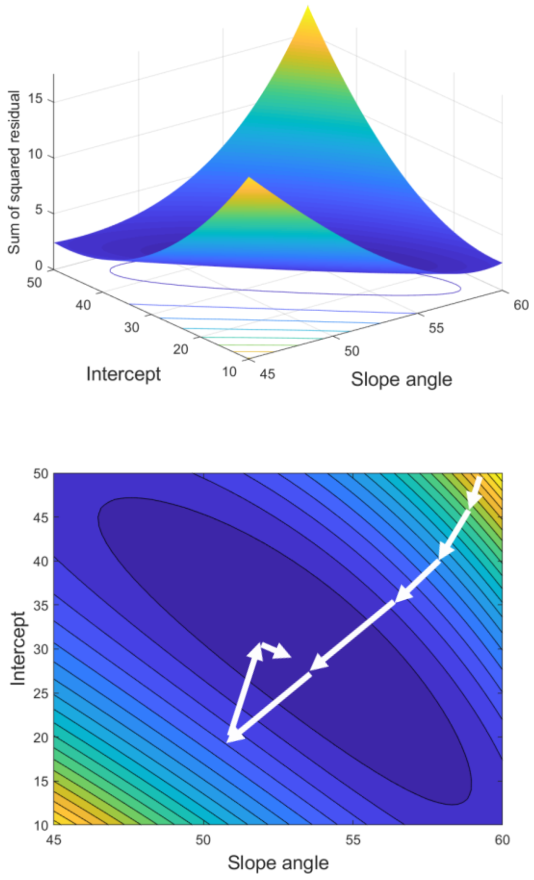
Summary: GLRT for Gaussian stationary noise

- General form of the GLRT obtained for some common cases
- Most GW data analysis algorithms are designed for the case of Gaussian stationary noise (but must be enhanced for real data)
- In GW data analysis,

Extrinsic parameters	Intrinsic parameters
• Parameters such as A, ϕ_0, t_a that can be maximized analytically or efficiently	• Parameters that must be maximized over numerically and/or are challenging to maximize

LINEAR MODELS

Log-likelihood for Gaussian noise and linear parameters: Convex function



MLE or GLRT in Gaussian noise

- Parameters appearing linearly can be maximized algebraically (e.g., See Appendix C.2 of textbook)
  - Unknown amplitude and initial phase:

$$A\,q(\Phi(t)+\phi_0) = A\cos\phi_0 \sin\Phi(t)) + A\sin\phi_0\cos\Phi(t)$$

$$= X\sin\Phi(t) + Y\cos\Phi(t) \Rightarrow X, Y\text{: Linear parameters}$$

- For linear models with large number of parameters, greedy methods (e.g., steepest descent) work well

Intrinsic parameters

- General form of GLRT in GW data analysis:

$$L_G = \max_{\substack{intrinsic\\parameters}} \max_{\substack{extrinsic\\parameters}} \langle \bar{y}, \bar{q}(\Theta_{extrinsic}, \Theta_{intrinsic}) \rangle^2$$

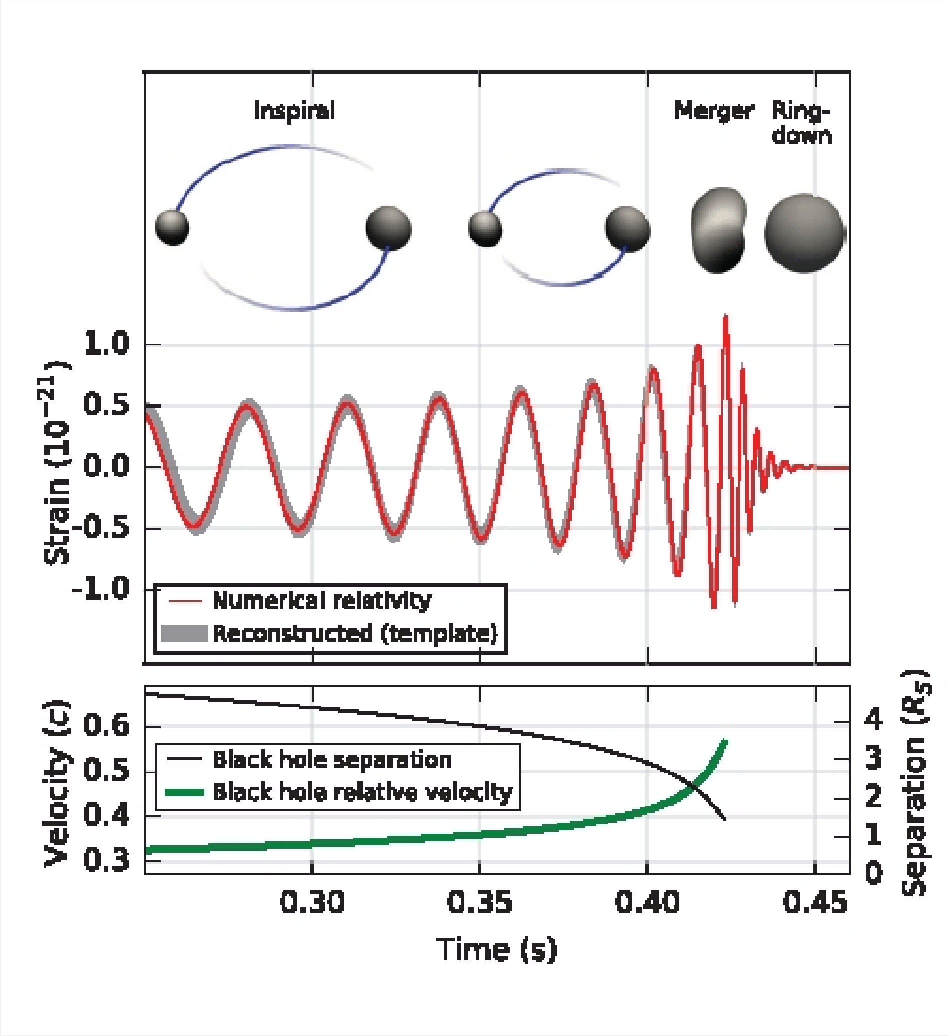
$$= \max_{\substack{intrinsic\\parameters}} \lambda(\Theta_{intrinsic})$$

- The maximization of $\lambda(\Theta_{intrinsic})$ must be done using numerical optimization methods
- In most GW data analysis problems, this is a highly challenging step because grid search becomes very expensive (or impossible with current computers)

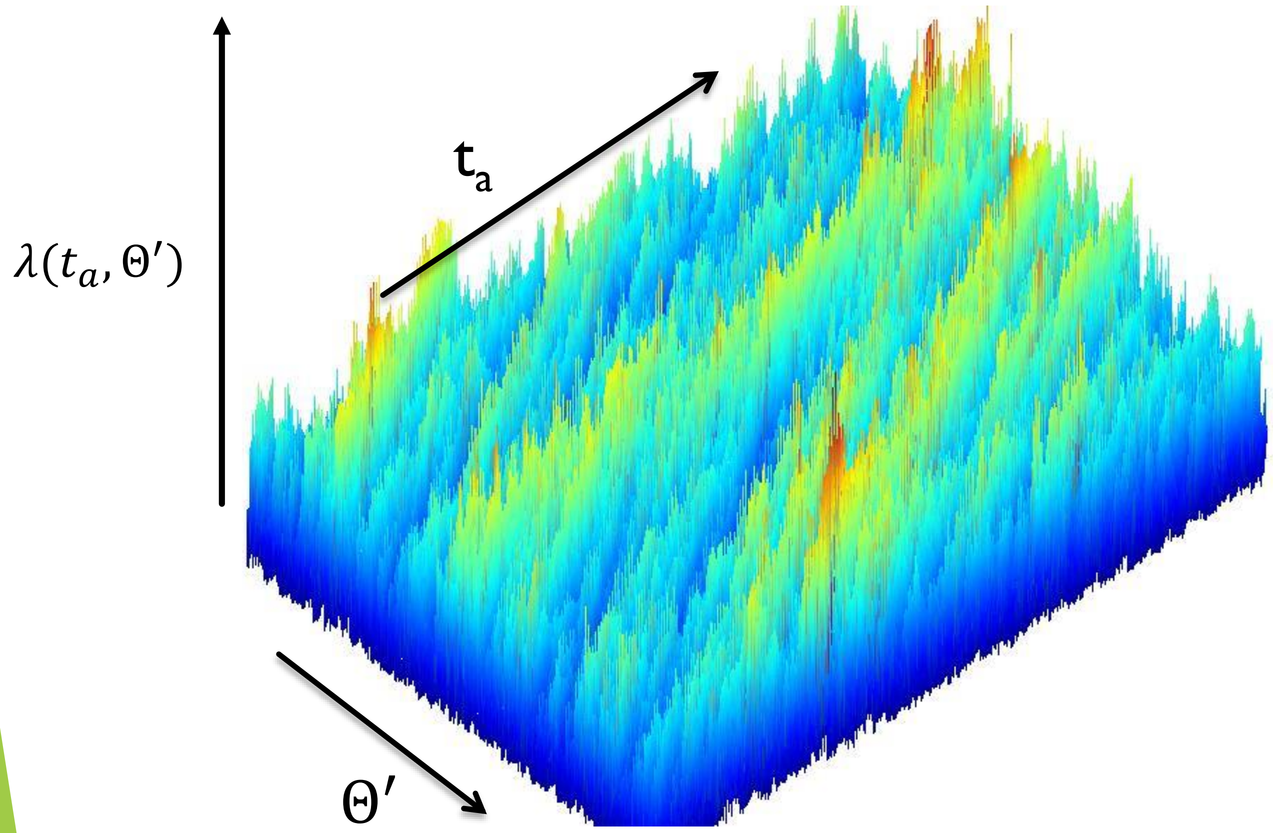
BINARY INSPIRAL SEARCH

Network analysis $\Theta_{intrinsic}$:

- ❑ mass of each component (2)
- ❑ sky location (2)
- ❑ spin of each component (6)
- ❑ Optimization required in 4 to 10 dimensional space
- ❑ Lower mass binaries last longer in the detector frequency band $\Rightarrow$ cost of evaluating $F(\Theta_{intrinsic})$ becomes higher



Binary inspiral search



The numerical optimization problem is

1. Intrinsically difficult
   - Large number of maxima
   - Becomes worse as the number of parameters increases
2. Computationally expensive
   - Binary inspiral network analysis for ground-based detectors grid based search: $\approx 10^8$ points in $\Theta_{intrinsic}$ space with $\approx 10^7$ floating point operations per point (1 hour segments) $\Rightarrow$ 0.3 Tflops to just keep up with the incoming data rate
   - Computational bottleneck $\Rightarrow$ current searches follow a sub-optimal approach $\Rightarrow$ Lower sensitivity $\Rightarrow$ Reduced rate of detections