CBC gravitational wave data analysis
the way it is done by the LSC and VIRGO

Frédérique MARION

2011 APCTP International School on Numerical Relativity and Gravitational Waves
JULY 28 ~ AUGUST 3, 2011 at APCTP POHANG in KOREA
Compact binary coalescences

GW Detectors

CBC Detection & Science

Search strategies

Source population

Multi-messenger astronomy

Waveforms

Analysis techniques
Outline

- **What, why, how?**
  - A reminder of what we are looking for
    - Sources
    - Waveforms
  - The basic search technique
    - Matched filtering

- **How, in real life**
  - Exploring the parameter space
  - Dealing with background
    - Multi-detector analysis
    - Background estimation
    - Vetoes
    - Detection statistics

- **Tying it all together**
  - A typical search pipeline
  - A brief review of LIGO-Virgo CBC searches
Chapter 1

What, why, how?
Evolution of ground-based GW detectors

1st generation interferometric detectors
- Initial LIGO, Virgo, GEO600
  - Enhanced LIGO, Virgo+

2nd generation detectors
- Advanced LIGO, Advanced Virgo, GEO-HF, LCGT

3rd generation detectors
- Einstein Telescope, US counterpart to ET

Unlikely detection
- Science data taking
  - First rate upper limits
  - Set up network observation

Improved sensitivity
- Lay ground for multi-messenger astronomy

Likely detection
- Routine observation
  - GW astronomy

Thorough observation of Universe with GW

→ GW astronomy
The target sources

- **Final evolution stage of compact binary systems**
  - Systems like PSR1913+16 reaching coalescence of the two stars

- **System may involve**
  - Neutron stars (NS)
  - Black holes (BH)
    - For ground based detectors, stellar mass black holes
    - Advanced detectors: up to intermediate mass black holes
    - Super-massive BH: lower frequency, space based detectors, pulsar timing

**Gravitational waves**

 Orbit decays by radiating energy as Gravitational Waves

**Strain versus time: chirp**

Chirp ends at frequency inversely proportional to total mass

\[ f_{ISCO} = \frac{2.8M_\odot}{M} \quad 1600 \text{ Hz} \]
What makes CBC promising sources?

- **We know “a lot” about the sources**
  - Such systems do exist
    - Although rates are uncertain and low...
  - The emitted waveform is known with some accuracy

- **General Relativity**
  - Test theory in strong field
  - Test/constrain alternative gravity theories

- **Astrophysics**
  - Measure merger rates
    - As a function of parameters
  - Inform source distribution
  - Study effect of matter in BNS/NSBH waveform
  - Short, hard GRBs
    - Confirm or rule out merger progenitor

- **Cosmology**
  - CBC inspirals as standard sirens
    - Independent measurement of Hubble constant

Many of these require combining information from gravitational wave, electromagnetic and/or particle observations ➔ Multi-messenger astronomy

[Ref.1]
Source population: rates

- Uncertain, even for most confident predictions
- BNS
  - Extrapolations from observed Galactic binary pulsars
    - Small sample, few parameters
  - From population-synthesis models
    - Observational and theoretical constraints on parameters
- NS-BH and BBH
  - From population-synthesis models
    - Many open questions for BBH
  - Extrapolation from extra-Galactic X-ray binaries
    - IC10 X-1 and NGC300 X-1
    - Formed by a 20-30 black hole and a massive Wolf-Rayet star
    - Should evolve into a binary black hole with $M_{\text{chirp}} \approx 15 M_\odot$
- Still an open point whether BNS or BBH dominate

GW rate upper limits and measurements likely to be key input to further constrain parameters
Source population: expected event rates (I)

- Expected number of detected events

\[ N = \mathcal{R} \, N_G \, T \]

- Astrophysical rate
  - Per galaxy or per unit volume

- Number of observable galaxies
  - Volume probed
  - Depends on detector sensitivity
    - Horizon distance: distance at which an optimally located and oriented source would produce a signal to noise ratio of 8
    - Relies on the ability of the analysis to extract signal from noise at that threshold

[Ref. 2]
Source population: expected event rates (II)

<table>
<thead>
<tr>
<th>IFO</th>
<th>Source</th>
<th>$\tilde{N}_{\text{low}}$ yr$^{-1}$</th>
<th>$\tilde{N}_{\text{re}}$ yr$^{-1}$</th>
<th>$\tilde{N}_{\text{pl}}$ yr$^{-1}$</th>
<th>$\tilde{N}_{\text{up}}$ yr$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>NS-NS</td>
<td>$2 \times 10^{-4}$</td>
<td>0.02</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>NS-BH</td>
<td>$7 \times 10^{-5}$</td>
<td>0.004</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BH-BH</td>
<td>$2 \times 10^{-4}$</td>
<td>0.007</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IMRI into IMBH</td>
<td>&lt; $0.001$</td>
<td></td>
<td></td>
<td>0.01$^c$</td>
</tr>
<tr>
<td></td>
<td>IMBH-IMBH</td>
<td>$10^{-4}$</td>
<td></td>
<td></td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>Advanced</td>
<td>NS-NS</td>
<td>0.4</td>
<td>40</td>
<td>400</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>NS-BH</td>
<td>0.2</td>
<td>10</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BH-BH</td>
<td>0.4</td>
<td>20</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IMRI into IMBH</td>
<td></td>
<td></td>
<td></td>
<td>$10^b$</td>
</tr>
<tr>
<td></td>
<td>IMBH-IMBH</td>
<td></td>
<td></td>
<td></td>
<td>$300^c$</td>
</tr>
</tbody>
</table>

Realistic rates do get substantial for advanced detectors

[Ref. 3]
Source population: parameters

- **Masses**
  - What is the component mass distribution in BBH systems?
  - Especially, how large?
    - \( \sim 20 \, M_\odot \) in systems from isolated binary evolution
    - \( \sim 30 \, M_\odot \) in systems from dynamical formation in dense stellar environments
  - Intermediate mass black holes?

- **Spin**
  - Small for neutron stars
  - Not well constrained for black holes, can be high

- **Distribution**
  - Mergers may occur away from host galaxies in case of high kicks
Phases of the evolution

- **Inspiral**
  - The realm of post-Newtonian expansions
  - Accurate analytical computations, under assumption of adiabatic evolution of quasi-circular orbit
  - Valid up to innermost stable circular orbit
  - \( f_{\text{ISCO}} = \frac{2.8 M_\odot}{M} \times 1600 \text{ Hz} \)
  - Length \( \sim 34 \left( \frac{M}{M_\odot} \right)^{-5/3} \left( \frac{f_0}{40 \text{ Hz}} \right)^{-8/3} \) s

- **Plunge/Merger**
  - The realm of numerical relativity
  - Duration \( << 1 \text{ s} \)

- **Ringdown**
  - The realm of black hole perturbation theory
  - Relaxation of perturbed final black hole
  - Duration \( << 1 \text{ s} \)
Inspiral (I)

- **PN approximants are found under various avatars**
  - Choice of parameter used in PN expansion / method used to solve differential equations ➔ different families of waveforms
  - Time domain or frequency domain
    - Frequency domain uses the stationary phase approximation

- **Known up to order PN2.5 for the amplitude and order PN3.5 for the phase**
  - Current searches use **restricted waveforms**
    - Amplitude kept at Newtonian order
    - Good enough for detection, at least for initial detectors
    - May cost some accuracy in parameter estimation, especially for high mass systems
  - **Spin effects** appear from order PN1.5 (spin-orbit) and PN2 (spin-spin)
    - Expected to be negligible for NS, may be significant for BH
    - Spins not aligned with orbital momentum cause orbital plane to precess, leading to more complex waveforms

[Ref. 4] [Ref. 5]
Inspiral (II)

From A. Buonanno

High order terms matter more for higher mass systems

\[ M = (1.4 + 1.4)M_\odot \]
\[ f_{in} = 40 \text{ Hz}; \ f_{fin} = 1570 \text{ Hz} \]
\[ \chi = |S|/m^2 \]

<table>
<thead>
<tr>
<th>Number of cycles</th>
<th>Number of useful cycles:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newtonian:</td>
<td>16034</td>
</tr>
<tr>
<td>1PN:</td>
<td>+441</td>
</tr>
<tr>
<td>1.5PN</td>
<td>-211</td>
</tr>
<tr>
<td>Spin-orbit:</td>
<td>+65.7\chi_1 + 65.7\chi_2</td>
</tr>
<tr>
<td>2PN</td>
<td>+9.9</td>
</tr>
<tr>
<td>2.5PN</td>
<td>-11.7 + 9.2\chi_1 + 9.2\chi_2</td>
</tr>
<tr>
<td>3PN:</td>
<td>+2.6</td>
</tr>
<tr>
<td>3.5PN:</td>
<td>-0.9</td>
</tr>
<tr>
<td>Number of useful cycles:</td>
<td>247.8</td>
</tr>
<tr>
<td></td>
<td>+24.0</td>
</tr>
<tr>
<td></td>
<td>-20.0</td>
</tr>
<tr>
<td></td>
<td>6.2\chi_1 + 6.2\chi_2</td>
</tr>
<tr>
<td></td>
<td>+1.5</td>
</tr>
<tr>
<td></td>
<td>-2.3 + 0.8\chi_1 + 0.8\chi_2</td>
</tr>
<tr>
<td></td>
<td>+0.6</td>
</tr>
<tr>
<td></td>
<td>-0.2</td>
</tr>
</tbody>
</table>

\[ M = (15 + 15)M_\odot \]
\[ f_{in} = 40 \text{ Hz}; \ f_{fin} = 147 \text{ Hz} \]
\[ \chi = |S|/m^2 \]

<table>
<thead>
<tr>
<th>Number of cycles</th>
<th>Number of useful cycles:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newtonian:</td>
<td>302</td>
</tr>
<tr>
<td>1PN:</td>
<td>+39</td>
</tr>
<tr>
<td>1.5PN</td>
<td>-37</td>
</tr>
<tr>
<td>Spin-orbit:</td>
<td>+11.7\chi_1 + 11.7\chi_2</td>
</tr>
<tr>
<td>2PN</td>
<td>+3.3</td>
</tr>
<tr>
<td>Spin-spin:</td>
<td>-1.7\chi_1 \chi_2</td>
</tr>
<tr>
<td>2.5PN</td>
<td>-6.2 + 3.6\chi_1 + 3.6\chi_2</td>
</tr>
<tr>
<td>3PN:</td>
<td>+2</td>
</tr>
<tr>
<td>3.5PN:</td>
<td>-0.8</td>
</tr>
<tr>
<td>Number of useful cycles:</td>
<td>10.7</td>
</tr>
<tr>
<td></td>
<td>+4.0</td>
</tr>
<tr>
<td></td>
<td>-6.2</td>
</tr>
<tr>
<td></td>
<td>1.9\chi_1 + 1.9\chi_2</td>
</tr>
<tr>
<td></td>
<td>+0.8</td>
</tr>
<tr>
<td></td>
<td>-0.4\chi_1 \chi_2</td>
</tr>
<tr>
<td></td>
<td>-2.3 + 0.8\chi_1 + 0.8\chi_2</td>
</tr>
<tr>
<td></td>
<td>+1.2</td>
</tr>
<tr>
<td></td>
<td>-0.5</td>
</tr>
</tbody>
</table>
Inspiral (III)

Inspiral phase ends at \( f_{\text{ISCO}} = \frac{2.8M_\odot}{M} \) 1600 Hz

<table>
<thead>
<tr>
<th>( m_1 / m_2 )</th>
<th>( f_{\text{ISCO}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4 / 1.4 M_\odot</td>
<td>1600 Hz</td>
</tr>
<tr>
<td>12.5 / 12.5 M_\odot</td>
<td>180 Hz</td>
</tr>
</tbody>
</table>

- **Low mass systems**
  - Inspiral ends out of the detector bandwidth
  - What happens after inspiral can safely be ignored

- **High mass systems**
  - Inspiral ends within the detector bandwidth
  - Merger and ringdown carry significant signal to noise ratio and must be taken into account
- **Boring for BBH**
  » But a great success of numerical relativity!
  » Smooth transition from inspiral to ringdown

- **Rich waveform for BNS and NS-BH**
  » Not essential for detection
  » Observation with advanced detectors (possibly in narrow-band configuration) may allow to constrain NS equation of state

### BNS (from J.Read)

![BNS waveform](image)

**Small compactness (stiff EOS)**

- Sudden shut down because of the tidal disruption

**Moderate compactness/stiffness**

- Tidal disruption and quasi normal mode

**Large compactness (soft EOS)**

- No tidal disruption and quasi normal mode
Ringdown

- Perturbed black hole relaxes to final equilibrium state by emitting gravitational waves
  - Superposition of quasinormal modes, dominated by fundamental mode, \( l = m = 2, n = 0 \)
  - Frequency and quality factor depend on mass and spin of final black hole
  
  \[ f_0 = \frac{1}{2\pi} \frac{c^3}{GM} g(\hat{a}) \quad Q = 2(1 - \hat{a})^{-9/20} \]
  \[ g(\hat{a}) = 1 - 0.63(1 - \hat{a})^{3/10} \quad \hat{a} = Jc/GM^2 \]

- Amplitude \( A \) depends on \( f_0, Q \) and fraction \( \epsilon \) of final black hole mass radiated as GW
  - \( \epsilon \) scales with \( \eta^2 \)
  \[ \eta = m_1 m_2 / (m_1 + m_2)^2 \]
  - Symmetric mass ratio
  - \( \epsilon \sim 1\% \) for equal mass systems
**Significance of inspiral, merger, ringdown signals for LIGO/Virgo, depending on system mass**

- **We need analytical waveforms**
  - Computational cost of numerical simulations to generate long and accurate waveforms over full parameter space is prohibitive
  - Truly analytical or obtained solving ordinary differential equations
Inspiral-Merger-Ringdown (II)

- **Effective one body approach** [Ref. 6]
  - Uses PN theory and NR computations
  - Plunge is smooth continuation of adiabatic inspiral
  - Very short transition from merger to ringdown
  - Model parameters tuned to NR simulations of non-spinning binaries with mass ratios 1:1 – 4:1

- **Phenomenological IMR waveforms**
  - Hybrid waveforms obtained matching PN and NR waveforms
  - Approximated with analytical, phenomenological waveforms in frequency domain

\[ A_{\text{eff}}(f) e^{i \Psi_{\text{eff}}(f)} \]

\[ A_{\text{eff}}(f) = \begin{cases} 
(f/f_{\text{merg}})^{-7/6} & \text{if } f < f_{\text{merg}} \\
(f/f_{\text{merg}})^{-2/3} & \text{if } f_{\text{merg}} \leq f < f_{\text{ring}} \\
\omega \mathcal{L}(f, f_{\text{ring}}, \sigma) & \text{if } f_{\text{ring}} \leq f < f_{\text{cut}} 
\end{cases} \]

\[ \Psi_{\text{eff}}(f) \] expansion in powers of \( f \)
The signal (I)

Signal sensed by detector is a combination of two polarizations

\[ h(t) = F_+(\theta, \phi, \psi) h_+(t) + F_\times(\theta, \phi, \psi) h_\times(t) \]

\( F_+ \) and \( F_\times \): detector response functions depend on sky location \((\theta, \phi)\) and polarization angle \(\psi\)

\[
F_+ = -\frac{1}{2}(1 + \cos^2 \theta) \cos 2\phi \cos 2\psi - \cos \theta \sin 2\phi \sin 2\psi
\]

\[
F_\times = \frac{1}{2}(1 + \cos^2 \theta) \cos 2\phi \sin 2\psi - \cos \theta \sin 2\phi \cos 2\psi
\]

\( F_+ \) and \( F_\times \) can be approximated as constant over length of CBC signals in ground based detectors
The signal (II)

$h_+ \text{ and } h_\times$ are obtained from post-Newtonian developments
Up to 2.5PN order in amplitude:

$$h(t) = \sum_{k=1}^{N} \sum_{m=0}^{5} A_{k,m/2}(t) \cos(k \varphi(t) + \varphi_{k,m/2})$$

with $A_{k,m/2}(t) \propto (2\pi M f(t))^{(m+2)/3}$

$\varphi(t)$ is the orbital phase of the binary system

$f(t) = \frac{1}{2\pi} \frac{d\varphi}{dt}$ is the instantaneous orbital frequency

» h(t) is a linear combination of harmonics of the orbital phase
  – k runs over harmonics, m/2 is PN order in amplitude

» Restricted waveforms: all terms with $k \neq 2$ are neglected
  – other harmonics of the orbital frequency are ignored

[Ref.21]
The signal (III)

The restricted waveform at the detector can be written:

\[ h(t) = \frac{1}{D_{\text{eff}}} \frac{M_{\text{pc}}}{D_{\text{eff}}} A(t) \cos(\varphi(t) - \varphi_0) \]

with \( D_{\text{eff}} = \frac{D}{\sqrt{F_d^2 (1+\cos^2 \nu)^2 / 4 + F_x^2 \cos^2 \nu}} \) the effective distance

(distance of an optimally located and oriented source that would produce the same signal strength)

\[ A(t) = A f(t)^{2/3} \]

At Newtonian order:

\[ f(t) = f_0 \left(1 - \frac{t}{\tau_0}\right)^{-3/8} \quad \varphi(t) = \frac{16\pi f_0 \tau_0}{5} \left[1 - \left(\frac{f}{f_0}\right)^{-5/3}\right] \]

\[ \tau_0 = \frac{5}{256} M^{-5/3} (\pi f_0)^{-8/3} \] time from frequency \( f_0 \) to coalescence

\( \mathcal{M} \) is called the chirp mass

\[ \mathcal{M} = \mu^{3/5} M^{2/5} \quad M = m_1 + m_2 \quad \mu = m_1 m_2 / M \]

\[ = \eta^{3/5} M \quad \eta = \mu / M \] symmetric mass ratio
Inspiral signal length

Length $\sim 34 \left( \frac{M}{M_\odot} \right)^{-5/3} \left( \frac{f_0}{40 \text{ Hz}} \right)^{-8/3} \text{ s}$

Higher mass systems give shorter signals

At a given total mass, asymmetric systems give longer signals than equal mass systems

<table>
<thead>
<tr>
<th>$m_1 / m_2$</th>
<th>$f_0$</th>
<th>10 Hz</th>
<th>30 Hz</th>
<th>50 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4/1.4 $M_\odot$</td>
<td>984s</td>
<td>52.5s</td>
<td>13.5s</td>
<td></td>
</tr>
<tr>
<td>3.0/3.0 $M_\odot$</td>
<td>277s</td>
<td>14.8s</td>
<td>3.8s</td>
<td></td>
</tr>
<tr>
<td>10.0/10.0 $M_\odot$</td>
<td>37.2s</td>
<td>2.0s</td>
<td>0.5s</td>
<td></td>
</tr>
<tr>
<td>19.0/1.0 $M_\odot$</td>
<td>196s</td>
<td>10.5s</td>
<td>2.7s</td>
<td></td>
</tr>
</tbody>
</table>

Time-frequency spectrogram

Simulated inspiral signal injected in detector noise

Length increases if detector bandwidth starts at lower frequency
Matched filtering (I)

» Construct a filtered signal

\[ S = \int_{-\infty}^{\infty} s(t)Q(t)dt \]

calibrated detector output

filter chosen to optimize the signal to noise ratio (SNR)

» \( S \) can also be written in the frequency domain

\[ S = \int_{-\infty}^{\infty} \tilde{s}(f)\tilde{Q}^*(f)df \]

» If the detector output is noise + some signal

\( s(t) = n(t) + h(t) \) with \( h(t) = \alpha T(t - t_0) \)

\( T(t) \): normalized expected signal entering detector bandwidth at time \( t = 0 \)

» The expectation value of the signal \( S \) is

\[ <S> = \int_{-\infty}^{\infty} <\tilde{s}(f)> \tilde{Q}^*(f)df = \int_{-\infty}^{\infty} \tilde{h}(f)\tilde{Q}^*(f)df \]
The noise is defined as:

\[ N = S - < S > = \int_{-\infty}^{\infty} \tilde{n}(f) \tilde{Q}^*(f) df \]

\[ < N > = 0 \quad \text{but} \quad < N^2 > = \int_0^{\infty} S_h(f) |\tilde{Q}(f)|^2 df \]

where \( S_h(f) \) is the one-sided noise power spectrum of the detector:

\[ < \tilde{n}(f) \tilde{n}^*(f') > = \frac{1}{2} S_h(|f|) \delta(f - f') \]

We can define an inner product

\[ (A|B) = \int_{-\infty}^{\infty} \tilde{A}(f) \tilde{B}^*(f) S_h(|f|) df \]

and rewrite

\[ < S > = \left( \frac{\tilde{h}}{S_h} |\tilde{Q} \right) \text{ and } < N^2 > = \frac{1}{2} (\tilde{Q}|\tilde{Q}) \]

What is the optimal filter \( Q \) maximizing \( SNR^2 = \frac{<S>^2}{<N^2>} = 2\left( \frac{\tilde{h}}{S_h} \right)^2 (\tilde{Q}|\tilde{Q}) \)?

use property \( (A|B)^2 \leq (A|A)(B|B) \)

\[ (A|B)^2 = (A|A)(B|B) \text{ only if A proportional to B} \]
We choose

$$\tilde{Q}(f) \propto \frac{\tilde{h}(f)}{S_h(|f|)}$$

$$Q(f) = 2 \frac{\tilde{T}(f)}{S_h(|f|)} e^{2\pi if t_0}$$

Signal $S$ for arrival time offset $t_0$ is given by

$$S = \int_{-\infty}^{\infty} \tilde{s}(f) \tilde{Q}^*(f) df$$

$$= 4 \int_{0}^{\infty} \frac{\tilde{s}(f) \tilde{T}^*(f)}{S_h(f)} e^{-2\pi if t_0} df$$

Fourier transform of $S$

$S$ can be easily obtained for all arrival times $t_0$ by means of an FFT

The optimal signal to noise ratio is:

$$SNR^2 = 2\alpha^2 \left( \frac{\tilde{T}}{S_h} \right)$$

If $T$ is normalized such that $\left( \frac{\tilde{T}}{S_h} \right) = \frac{1}{2}$ then $<N^2> = 1$ and $SNR^2 = \alpha^2$

Now that we know the optimal filter, we can redefine the inner product in the more usual way:

$$(a, b) = 4 \Re \int_{0}^{\infty} \frac{\tilde{a}(f) \tilde{b}^*(f)}{S_h(f)} df$$

$$S = (s | T')$$
Matched filtering & likelihood (I)

Likelihood ratio

\[ \Lambda = \frac{P(s|h)}{P(s|0)} = \frac{e^{-(s-h|s-h)/2}}{e^{-|s|/2}} \]

Residual (data – signal) is distributed as noise

Data is distributed as noise

The noise model defines the likelihood function

\[ \log \Lambda = (s|h) - \frac{1}{2}(h|h) \]

Take \( h = \alpha h_0 \) with \( (h_0|h_0) = 1 \)

\[ \log \Lambda = \alpha(s|h_0) - \frac{\alpha^2}{2} \]

Maximize \( \log \Lambda \) by taking derivative with respect to \( \alpha \)

\[ \alpha = (s|h_0) \]

\[ \log \Lambda_{\text{max}} = \frac{1}{2}(s|h_0)^2 = \frac{1}{2} \rho^2 \]
Why is the noise distribution expressed in terms of the scalar product defined earlier?

» Single data point, Gaussian noise
\[ p(n_i) = \frac{1}{\sqrt{2\pi\sigma_i}} e^{-n_i^2/2\sigma_i^2} \]

» Uncorrelated noise data points
\[ p(n_1, n_2, \ldots, n_N) = p(n_1) \ p(n_2) \ldots \ p(n_N) \]
\[ p(s|h) = \frac{1}{(2\pi)^{N/2}} \left( \frac{1}{\prod_i \sigma_i} \right) e^{-\sum_i (s_i-h_i)^2/\sigma_i^2} \]

» Correlated noise
\[ p(n_1, n_2, \ldots, n_N) = \frac{1}{\sqrt{\text{det}(2\pi C)}} e^{-\frac{1}{2} n_i C_{ij}^{-1} n_j} \]

» Noise correlation matrix \( C_{ij} = E[n_i n_j] \)

» For stationary noise, \( C_{ij} \) depends only on \( j - i \)

» Autocorrelation
\[ C_{ij} \rightarrow C_{i(i+m)} = C(m) = E[n_i n_{i+m}] = \frac{1}{N} \sum_i n_i n_{i+m} \]

» Continuous version
\[ C(\tau) = E[n(t)n(t+\tau)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} n(t)n(t+\tau)dt \]

» \( S(f) \) is the Fourier transform of \( C(\tau) \)

» In the frequency domain
\[ n_i C_{ij}^{-1} n_j \rightarrow 4 \int_0^\infty \frac{\tilde{n}(f)\tilde{n}^*(f)}{S(f)} df \equiv (n|n) \]
Chapter 2
How, in real life
Matched filtering in practice (I)

- The FFT allows to extract $S$ for all possible arrival times
  » Easy to maximize SNR over $t_0$

[Ref.10]
The phase of the chirp signal is unknown

\[ h(t) = A \left[ h_c(t) \cos \Phi + h_s(t) \sin \Phi \right] \]

cosine and sine phases of the waveform

» The SNR has to be maximized over all possible values of \( \Phi \)

Filter with \( T_{0^\circ} \) and \( T_{90^\circ} \) and take quadratic sum

\[ S^2 = \sqrt{S_{0^\circ}^2 + S_{90^\circ}^2} \]

Noise has a \( \chi^2 \) distribution with 2 degrees of freedom \( p(\rho) = \rho e^{-\rho^2/2} \)

Signal has a non-central \( \chi^2 \) distribution ⇒ Gaussian distribution if signal strong enough
Matched filtering is optimal

- If the noise is Gaussian, the matched filtering provides the optimal statistic
  - Selecting events by setting a threshold on the SNR $\rho > \rho^*$ guarantees the lowest false alarm probability for a given detection probability

False alarm probability

$$\int_{\rho^*}^{\infty} \rho e^{-\rho^2/2} d\rho = e^{-\rho^*^2/2}$$

Detection probability for signal with SNR $\rho_s$

$$\frac{1}{\sqrt{2\pi}} \int_{\rho^*}^{\infty} e^{-(\rho-\rho_s)^2/2} d\rho$$
Matched filtering requires waveform knowledge

Waveform depends on parameters

- As many as 17 parameters for generic waveform
- 15 if circular orbit assumed (no eccentricity)
- 9 if components assumed non-spinning

Parameters for non-spinning system on circular orbit

- Component masses \( m_1, m_2 \) or \( M, \eta \)
- Sky location of source \( \theta, \phi \)
- Distance to source \( D \)
- Orientation of binary system
  - inclination \( i \), polarization \( \psi \), coalescence phase \( \phi_0 \)
- Coalescence time \( t_0 \)

In the simplest case, we need to explore the 2-dimensional mass space
Scanning the parameter space (I)

With properly normalized templates, the filtered SNR for template $u$ is

$$SNR = \langle h, u \rangle,$$

or rather

$$\max_{\phi_c, t_c} \langle h, u(\theta) e^{i(2\pi ft_c + \phi_c)} \rangle$$

$$\lambda = (\phi_c, t_c, \theta)$$

with $\theta$ the intrinsic parameters.

If data containing signal $u(\lambda_1)$ is filtered with a template with different parameters $u(\lambda_2)$ the fraction of the optimal SNR recovered is given by the ambiguity function $A(\lambda_1, \lambda_2) = (u(\lambda_1), u(\lambda_2))$ maximized over extrinsic parameters, i.e. is given by the match:

$$M(\theta_1, \theta_2) = \max_{\phi_c, t_c} \langle u(\theta_1), u(\theta_2) e^{i(2\pi ft_c + \phi_c)} \rangle$$
Scanning the parameter space (II)

Map of match around point in parameter space

Define minimal match
acceptable SNR loss ➞ isomatch contour
Scanning the parameter space (III)

» From the match, define a metric on the parameter space

\[ g_{ij}(\theta) = -\frac{1}{2} \frac{\partial^2 M(\theta, \Theta)}{\partial \Theta^i \partial \Theta^j} \bigg|_{\Theta=\theta} \]

» In the regime \(1-M \ll 1\) the match can be approximated by

\[ M(\theta, \theta + \delta\theta) \sim 1 - g_{ij} \delta\theta^i \delta\theta^j \]

[Ref.8]

» Instead of the masses \(m_1, m_2\), it is more convenient to use as parameters:

\[ \tau_0 = \frac{5}{256} M^{-5/3} (\pi f_0)^{-8/3} \eta^{-1} \]

\[ \tau_1 = \frac{5}{192} M^{-1} (\pi f_0)^{-2} \left( \frac{743}{336\eta} + \frac{11}{4} \right) \]

» For matches above \(\sim 95\%\), isomatch contours are ellipses

» In the \(\tau_0, \tau_1\) space, the metric components \(g_{ij}\) are constant at 1PN order, and have small variations at higher order.
Scanning the parameter space (III)

» Each isomatch contour defines a region of the parameter space which overlaps with the template in the center with a match better than some value $M$

» The template in the center can be used to search for signals in that region of the parameter space, at the price of a controlled loss of SNR ($< 1 - M$)

» Templates should be placed over the parameter space in order to
  - Achieve coverage of space (no holes)
  - Preserve computational efficiency: keep number of templates as low as possible

$\Rightarrow$ Take into account variations of ellipse size and orientation across the parameter space

[Ref.9]
CBC Low Mass search, \( M_{\text{tot}} \in [2 ; 35M_\odot] \)

> Request 0.97 minimum match
> Low Mass region is more densely populated than high mass region
  - Related to duration and number of cycles of the templates

Coming back to \( m_1, m_2 \) variables

High mass templates are short and contain few cycles

> It is easier to match different templates in the high mass region

\( \tau_0 \approx 14\text{s} (\sim 1100 \text{ cycles}) \)

\( \tau_0 \approx 0.17 \text{s} (\sim 14 \text{ cycles}) \)
Template banks in practice (I)

Examples of template bank size (a week in S6/VSR2)

Low Mass [2; 35 $M_\odot$]
High Mass [25; 100 $M_\odot$]

- Template bank size varies between 1,000 and 12,000 templates depending on type of search and detector.
Examples of template bank size (a week in S6/VSR2)

Low Mass [2; 35 M☉]

Template bank size varies between 1,000 and 12,000 templates depending on type of search and detector.

Differences between detectors ↔ different shapes of noise power spectrum.
Template banks in practice (III)

Examples of template bank size (a week in S6/VSR2)

**Low Mass [2; 35 \( M_{\odot} \)]**

- **Template bank size** varies between 1,000 and 12,000 templates depending on type of search and detector.
- **Differences between detectors** ↔ different shapes of noise power spectrum.
- **Fluctuations in time** ↔ non-stationary detector noise.

Estimate power spectrum from the median of fifteen 256s data segments.
Why is spin a difficult problem?

- **Waveforms**
  - Inspiral signal from PN-theory available for spinning components
  - Reliable inspiral-merger-ringdown waveforms for spinning components not available over full parameter space
    - Needed as spin expected mostly for black holes ➔ high mass systems

- **Scanning the parameter space with template banks becomes more complex with additional dimensions**

- **Taking additional parameters explicitly into account is not always the best strategy**
  - Non-spinning templates have some ability to capture spinning signals
  - Extra degrees of freedom increase background
    - Detection efficiency may increase, but false alarm rate as well
  - Need to identify regions of parameter space where searching explicitly for spinning sources is warranted
Background is not Gaussian!

- **Coincidences**
  - Reduce false alarm rate by requiring coincident triggers in several detectors
  - Allows to estimate the non-Gaussian background from the data themselves

- **Instrumental vetoes**
  - Check for anomalies in detector behavior, statistically associated with excess triggers

- **Signal based vetoes**
  - Check trigger internal consistency with expected CBC signal

- **Special case of targeted searches**
  - e.g. GRB
  - Estimate background “off source”
Coincidences

- Require coincident triggers in 2 or more detectors
  - Check parameter consistency within allowed « windows » $\Delta t$, $\Delta M$, $\Delta \eta$
  - Smaller coincidence windows $\Rightarrow$ larger reduction of FAR
    - Window size depends on the resolution with which each detector is able to determine those parameters
    - $\Delta t$ must allow for time of flight between detectors
      - LIGO Hanford – LIGO Livingston: 10 ms
      - Virgo – LIGO: 30 ms

Timing precision: typically a few ms

Chirp mass very well determined

$\eta$ less precisely determined
Coincidences (II)

- Fixed coincidence windows are not optimal

Parameters are correlated

Errors on parameters vary across the parameter space

[Ref.12]
Coincidences (III)

- Use ellipsoids to define coincidences
  » Builds in correlation and accuracy variation
  » One tunable scale parameter
- Achieves background reduction of a factor 10

[Ref.13]
Background estimation (I)

A zero-lag trigger (true coincidence)

Detector 1

Detector 2
Background estimation (II)

Background estimated by shifting detector data in time

A time-slide trigger (accidental coincidence)

\[ \Delta T \]

\[ t \]

\[ \Delta T \] much larger than GW travel time and analysis method autocorrelation time

⇒ several seconds typically
• Background distribution obtained by counting number of coincident triggers found in each slide experiment
  » 100 time slide experiments performed routinely
  » Works well for distant sites, not so well for co-located detectors because of correlated background

• Compare distribution of zero-lag triggers to background estimated with time slides

**Histogram of coincident triggers versus combined SNR**

Region where outlier triggers would appear

Zero-lag triggers

Background distribution (time-slides)
Background estimation limitations (I)

- **Number of available time slides**
  - Depends on total analyzed time and slide step

- **Different time slides are not completely independent**
  - They all reuse the same data
  - OK if most significant events are due to random coincidences of relatively quiet (common) events in each detector
  - Not OK if most significant events are due to uncommon (loud) event in detector 1, coincident with very common events in detector 2

- **Background non-stationarity**
  - Averaging over time slides tends to wash out the structure

*Somewhat arbitrary (fraction of) run ≡ stretch of data reasonably homogeneous in sensitivity and glitchiness*
Background estimation limitations (II)

- False alarm rate estimated with R time slides for 2 detectors
  \[ FAR = \alpha_1 \alpha_2 \Delta t \]

- Uncertainty on FAR
  \[ \sigma_{FAR}^2 = \alpha_1 \alpha_2 \Delta t \left( \frac{1}{R} + \alpha_1 \Delta t + \alpha_2 \Delta t \right) \]

  Expected term if time slides independent

  Uncertainty on single detector trigger rate estimation propagates to \( \sigma_{FAR}^2 \)

- As \( R \) increases, \( \sigma_{FAR}^2 \) saturates

- For a given FAR, sigma smallest if \( \alpha_1 \) and \( \alpha_2 \) similar

[Ref. 19]
Background estimation limitations (III)

- FAR estimated for a (blind) hardware injection during S6/VSR3 runs
  - Was considered as a potential detection until un-blinded
  - Much work went into estimating FAR beyond upper limit given by 100 time slides

  Presence of signal biases FAR estimate
  - Didn’t prevent from reaching FAR of \((1 / 7000\) years) – with trials factor
  - Might be more of an issue when numerous signals are present in the data
    - But then we won’t care so much about requiring very low FARs
Signal based vetoes: $\chi^2$ test (I)

- Basic idea: look at how the SNR is distributed across the detector bandwidth and check whether this is consistent with what is expected from a true signal.

$$\text{SNR}^2 = 4 \int_0^\infty \frac{|\tilde{h}_S(f)|^2}{S_h(f)} \, df \sim A \int_0^{f_{\max}} \frac{f^{-7/3}}{S_h(f)} \, df$$
Signal based vetoes: $\chi^2$ test (II)

» The matched filter integral can be written as a sum over distinct frequency bands

$$(a, b) = \sum_{j=1}^{p} (a, b)_j \quad \text{with} \quad (a, b)_j = 4 \Re \int_{\Delta f_j} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_h(f)} \, df$$

» The frequency intervals are chosen so that for a true signal the SNR is uniformly shared among the frequency bands

$${\Delta f}_j \quad \text{such that} \quad (T, T)_j = \frac{1}{p} (T, T) \quad \text{or} \quad \int_{\Delta f_j} \frac{f^{-7/3}}{S_h(f)} \, df = \frac{1}{p} \int_{0}^{f_{\text{max}}} \frac{f^{-7/3}}{S_h(f)} \, df$$

» The filtered SNR can be written as

$$\rho = \sum_{j=0}^{p} \rho_j \quad \text{with} \quad \rho_j = (h, Q)_j$$

» A discriminating statistics is built

$$\chi^2 = p \sum_{j=0}^{p} (\Delta \rho_j)^2 \quad \text{with} \quad \Delta \rho_j = \rho_j - \frac{\rho}{p}$$

[Ref.15]

» If the noise is stationary and Gaussian, the $\chi^2$ has a $\chi^2$-distribution with p-1 degrees of freedom both for noise and for true signals

» Excess noise is expected to produce $\chi^2$ values which are outliers with respect to the Gaussian noise/signal distribution
Signal based vetoes: $\chi^2$ test (III)

- $\chi^2$ distribution for true signals in practice
  - Large SNR events tend to show larger $\chi^2$ values than expected from the naive distribution
  - An effect of using template banks
  - The slight mismatch between the signal and the template is enough to evidence differences between the expected SNR frequency distribution and the measured one $\Rightarrow$ high $\chi^2$
  - The cut used to eliminate background must allow some quadratic dependence of the $\chi^2$ on the SNR
  - Apply threshold on variable
    \[
    \xi^2 = \frac{\chi^2}{p(1+\delta^2 \rho^2)}
    \]

- Tuning
  - Adjust $p$, $\delta$ and threshold not to reject true signals
  - Cut must be loose enough to be robust with respect to missing features in the templates
    - Spin
    - Ringdown

[Ref.16]
Signal based vetoes: $\chi^2(t)$

- Look at $\chi^2(t)$: “$r^2$ veto”
  - Use as discriminating variable the time spent by $\chi^2(t)$ above some threshold in some time window prior to the measured coalescence time

[Ref.16]
Other signal based vetoes

- **Bank $\chi^2$**
  - Check consistency of observed SNR across different templates in bank
  - Glitches typically cause high SNR in many templates
  - Signals give a well-prescribed distribution of SNR across template bank

- **Autocorrelation $\chi^2$**
  - Check consistency of SNR time series against characteristic shape expected for signals

- **Coherent tests**
  - Coherent analysis allows checking consistency of signal across detectors
  - Coincidence test on $(M, \eta)$ already enforces some coherence
  - Null stream with no GW contribution can further test coherence
Instrumental vetoes (I)

- **Signal based vetoes are powerful but**
  - They are usually computationally expensive
  - They do not provide any feedback on the detector

- **Identify anomalies in the detector behavior/environment statistically coincident with CBC triggers**
  - Ideally, understand origin of bad behavior and fix it
  - Help clean up the background by eliminating the corresponding triggers
  - Data quality flags and vetoes
What are the criteria for a good instrumental veto?

- **Efficiency** ($\varepsilon$): eliminate false triggers, especially triggers with high SNR
  - $\varepsilon =$ fraction of triggers which are flagged
  - High efficiency is good!

- **Use percentage (UP)**: they should be associated often enough with triggers
  - UP = fraction of DQ segments used to flag at least one trigger
  - High use percentage is good!

- **Dead time** ($d$): they should not eliminate a large fraction of data
  - $d =$ fraction of science time which is flagged
  - Low dead time is good!

- **Safety**: they should not eliminate true signals!
  - Checked with hardware injections
  - Safety is crucial!
Instrumental veto categories (I)

- Vetoes are categorized according to severity, statistical correlation and dead time
  - Category 1
    - Data not suitable for being analyzed
    - e.g.: detector not at operating point; missing data…

Thermal Compensation System failure in V1
- **Vetoes are categorized according to severity, statistical correlation and dead time**
  - **Category 1**
    - Data not suitable for being analyzed
    - e.g.: detector not at operating point; missing data…
  - **Category 2**
    - Well understood instrumental problems
    - Strong statistical correlation
    - Usually low dead time, high efficiency
    - e.g.: overflow in sensing and control system
Category 2 veto example

Overflow of sensing and control system in H2:

Time-frequency representation of two overflows as seen in the GW channel

Single detector triggers from CBC search

Time window marked by the original DQ flag

Veto window tuned for CBC search

[Ref. 18]
Instrumental veto categories (III)

- **Vetoes are categorized according to severity, statistical correlation and dead time**
  - **Category 1**
    - Data not suitable for being analyzed
    - e.g.: detector not at operating point; missing data…
  - **Category 2**
    - Well understood instrumental problems
    - Strong statistical correlation
    - Usually low dead time, high efficiency
    - e.g.: overflow in sensing and control system
  - **Category 3**
    - Suspected instrumental problems
    - Positive statistical correlation, but not well understood
    - Dead time can be large
    - Includes ad-hoc vetoes based on auxiliary channels
    - e.g.: high seismic activity, strong wind
Category 3 veto example

Seismic glitches in V1:

⇒ Two seismic glitches of the same amplitude will not have the same impact on the GW channel

[REF 21]
Instrumental veto categories (IV)

- Vetoes are categorized according to severity, statistical correlation and dead time
  - **Category 1**
    - Data not suitable for being analyzed
    - e.g.: detector not at operating point; missing data…
  - **Category 2**
    - Well understood instrumental problems
    - Strong statistical correlation
    - Usually low dead time, high efficiency
    - e.g.: overflow in sensing and control system
  - **Category 3**
    - Suspected instrumental problems
    - Positive statistical correlation, but not well understood
    - Dead time can be large
    - Includes ad-hoc vetoes based on auxiliary channels
    - e.g.: high seismic activity, strong wind
  - **Category 4**
    - Poorly understood, weak but positive correlation
    - May veto whole noisy epochs
Instrumental veto categories (V)

Vetoes are categorized according to severity, statistical correlation and dead time

» Category 1
- Data not suitable for being analyzed
- e.g.: detector not at operating point; missing data…

» Category 2
- Well understood instrumental problems
- Strong statistical correlation
- Usually low dead time, high efficiency
- e.g.: overflow in sensing and control system

» Category 3
- Suspected instrumental problems
- Positive statistical correlation, but not well understood
- Dead time can be large
- Includes ad-hoc vetoes based on auxiliary channels
- e.g.: high seismic activity, strong wind

» Category 4
- Poorly understood, weak but positive correlation
- May veto whole noisy epochs

Data not analyzed

Look for possible detections

Look for possible detections and compute rate upper limits

Used only for follow-ups
Typical performance of instrumental vetoes

Example: VSR2 – Low latency CBC search
SNR distribution after applying vetoes (single detector)

These numbers are provided as examples. They are subject to large variations between different runs or different detectors.
Detection statistics

- Rank triggers according to their significance
  - Rank signals higher than background

- Matched filter SNR $\rho$ is the ideal detection statistic in Gaussian, stationary noise
  - But noise is neither Gaussian nor stationary

- Replace $\rho$ by a better statistic
  - Get as close to ideal case as possible
Detection statistic: effective SNR

- Fold $\chi^2$ parameter into ranking statistic
  
  » Down-weight triggers with high $\chi^2$ value

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

- For coincident search over several detectors, use combined effective SNR

$$\rho_c = \sqrt{\sum_{\text{det}} \rho_{\text{eff,det}}^2}$$
Acknowledging differences in background

- **Coincidence types**
  - For 3 detectors for instance, triggers can occur as double or triple coincidences
  - Triple coincidences have less background than doubles

- **Differences across parameter space**
  - Not all regions of the mass space bring the same background

---

![Graph showing Combined effective SNR vs. triggers into different mass bins]

- Fewer high mass templates in template bank, but they are short and tend to pick up glitches with higher (effective) SNR
- Divide triggers into different mass bins

[Ref. 23]
Detection statistic: false alarm rate

- Gather triggers in categories (coincidence type / mass bin)
  - For instance if data available from 3 detectors, triples + 3 types of doubles
    - H1L1V1, H1L1, H1V1, L1V1
  - For instance (low mass search), low, medium and high chirp mass bins

- Rank each trigger according to its false alarm rate in its own category

$$\text{FAR} = \frac{\sum_i n_i(\rho_c \geq \rho_c, \text{trigger})}{\sum_i T_i}$$

- Take trials factor into account to compute combined FAR

$$\text{cFAR} = \text{FAR} \times N_{\text{bins}} \times N_{\text{coinc}}$$

- Any intermediate statistic ranking triggers loudness/significance can be mapped into a FAR
Detection statistic: IFAR and beyond

- **Inverse of combined FAR used as detection statistics**
  - (c)FAR in units of yr$^{-1}$, IFAR in units of yr
  - Guarantees that all categories bring the same background

**Example: Result of the S5 1yr Low Mass Search for analyzed triple time**

Categories not all equally sensitive to GW signals (e.g. if detectors have different sensitivities)
- Weight IFAR with probability to observe signal with higher IFAR
- **Likelihood statistic**

Region of interest for detection
Detection checklist (I)

- **What are the interesting candidates for detection?**
  - Triggers surviving the coincidence test and all veto cuts
  - Triggers standing above background (low false alarm rate)
    ➔ Detection candidates are submitted to a detection checklist for review

- **Goals of the candidate follow-up review**
  - Perform sanity checks of the analysis results
  - Reject candidates that do not match minimal quality criteria

- **Checks mostly qualitative**
  - A place for testing new ideas/techniques that are not mature enough to be part of the main analysis pipeline
  - Require human expertise and training

- **What candidate follow-ups cannot do**
  - Increase significance of a detection candidate
  - Ranking statistic remains the key criterion to identify possible detections
Detection checklist (II)

Check for the presence of Data Quality flags or KW based Vetoes
Status of the interferometers
Check for environmental or instrumental causes
Check Electronic Log Book and Glitch Reports

Examine candidate appearance \((\text{time-freq, SNR} \& \chi^2 \text{ time series})\)
Trigger SNR distribution in analyzed chunk

Consistency of candidate parameters
Trigger coherence
Segmentation stability
Calibration stability

Try to identify the cause of the trigger
Check for non stationarities in the data, characterize the background locally
Compare triggers between detectors
Test the robustness of the detection

➡ Requires a lot of information about search triggers, GW channel, auxiliary channels...

See Chunglee Kim’s lectures on parameter estimation
Setting upper limits on CBC rate

- Use loudest event statistic in a Bayesian approach
- Probability that all signal events have SNR below some value $\rho$:
  \[ P(\rho|\mu) = e^{-\mu\epsilon(\rho)} \quad \text{(signal Poisson distributed)} \]
  \[ \mu = RT \quad \text{with} \ R \ \text{the event rate and} \ T \ \text{the observation time} \]
  \[ \epsilon(\rho) : \text{signal detection efficiency with SNR threshold} \ \rho \]
- Neglecting the background, posterior probability distribution for $\mu$:
  \[ P(\mu < \mu_p|\rho_{max}) = N^{-1} \int_0^{\mu_p} d\mu \ p(\mu) \ p(\rho_{max}|\mu) \quad \text{with} \ p(\rho|\mu) = dP(\rho|\mu)/d\rho \]
  Solve $p = P(\mu < \mu_p|\rho_{max})$ for $\mu_p$ to get 100p% CL upper limit
  With uniform prior $p(\mu)$ \[ R_{90\%} = \frac{3.890}{T\epsilon(\rho_{max})} \]
- Background can be taken into account to get better upper limit
- Results from previous searches can be incorporated as priors $p(\mu)$
- Account for uncertainties on detector calibration, waveforms...
The role of injections

- **Simulated GW signals can be added to detector data in different ways, with various goals**
  
- **Hardware injections**
  - Physically injected in the detector by acting on the interferometer mirrors
    - Can be done coherently across detectors
  - Test analysis end-to-end, from detector calibration to search pipeline
  - Test safety of instrumental vetoes
  - Invasive, so number limited

- **Blind hardware injections**
  - Secretly injected by a small team of people
  - Test detection process

- **Software injections**
  - Added offline to detector calibrated data, thousands can be performed
  - Used to tune search pipelines – signal/background separation
  - Used to compute detection efficiencies as a function of distance
    - Volume (or cumulative luminosity) probed, for rate upper limits
Chapter 3

Tying it all together
CBC coincident search pipeline (I)

One data set per interferometer
The waveform depends on the system parameters

⇒ Scan the mass space with template banks

Example for « Low Mass » search
(2-35 Msun)
Match filter:

⇒ keep triggers above some SNR threshold
CBC coincident search pipeline (IV)

Require coincidence between two or more detectors

- Time and mass parameters
- Allows reducing the false alarm rate
- Allows estimating the background by applying time shifts to the data
CBC coincident search pipeline (V)

Identify rare and weak events in detector noise that is non-Gaussian and not stationary:

⇒ Need to apply vetoes

- Signal based vetoes
  Check consistency between measured and expected signals (e.g. $\chi^2$ test)

- Instrumental vetoes
  Eliminate poor data quality times due to artifacts in detector or environment
CBC coincident search pipeline (VI)

- Surviving coincident triggers are ranked according to a detection statistic.
- Triggers that stand above the background are submitted to a detection checklist.
Search strategies

- **All sky blind searches**
  - Scan as much as the parameter space as possible
    - Low mass, high mass, ringdown

- **Externally triggered searches**
  - BNS or NSBH mergers likely progenitors or short, hard GRBs
    - Specific search around times of short GRBs
  - Other possibilities can be explored in the future
    - High energy neutrinos
    - Time delays with GW key input parameter

- **Searches triggering electromagnetic follow-up**
  - Low latency required
LIGO-Virgo S5/VSR1 low mass searches (I)

- Long observation periods at design sensitivities
- Sources with total mass between $2 - 35 \, M_\odot$ and minimum component mass of $1 \, M_\odot$
  - Search can be based on inspiral part of signal
  - 2PN SPA templates used

<table>
<thead>
<tr>
<th>Component Masses ($M_\odot$)</th>
<th>BNS</th>
<th>BHNS</th>
<th>BBH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizon Distance (Mpc)</td>
<td>~ 30</td>
<td>~ 50</td>
<td>~ 90</td>
</tr>
<tr>
<td>Cumulative Luminosity ($L_{10}$)</td>
<td>370</td>
<td>1600</td>
<td>8300</td>
</tr>
</tbody>
</table>
LIGO-Virgo S5/VSR1 low mass searches (II)

- **Search parameters**
  - 3D coincidence test \((t, M, \eta)\)
  - \(\chi^2\) and \(r^2\) consistency tests
  - Detection statistics
    - IFAR based on combined effective SNR for LIGO only searches
    - Likelihood for LIGO-Virgo search

---

[Ref. 24]
LIGO-Virgo S6/VSR2-3 low mass search

- Improvements over S5-VSR1 searches
  - New data, from more sensitive detectors
  - Search for total mass between 2 – 25 M\(_\odot\)
    - 25 – 35 M\(_\odot\) better covered by high mass search
  - 3.5 PN SPA templates used
  - Detection statistic: IFAR based on combined “new SNR”
    - “new SNR” improves over effective SNR
    - Avoid “over-ranking” events with very low \(\chi^2\)

- No GW candidate found, improved upper limits will follow
LIGO S5 high mass search

- Sources with total mass between 25 – 100 $M_\odot$ and minimum component mass of 1 $M_\odot$
  - Inspiral, merger and ringdown matter
- IMR waveforms used
  - EOBNR waveforms used as templates
  - EOBNR and Phenomenological IMR waveforms used as injections for efficiency computations
- Pipeline similar to low mass search
  - Coincidence test, consistency tests, detection statistics
  - Different tuning

[Ref. 25]
LIGO S4 ringdown search

- **S4:** one month of data, below design sensitivity

- **Frequency band 50 Hz – 2 kHz**
  - Mass range 11 – 440 M\(_\odot\) for black hole with \(\hat{a} = 0.9\) oscillating in fundamental mode

- **Ringdown search**
  - Template bank covering \((f_0, Q)\) space
  - 50 Hz \(\leq f_0 \leq 2\) kHz \(2 \leq Q \leq 20\)
  - \(\hat{a} = 0\) \(\hat{a} = 0.994\)
  - Multi-detector coincidences in \((t, f_0, Q)\)
  - Amplitude consistency test for co-located detectors H1 and H2
  - Detection statistic: combination of individual detector SNRs

- **Set rate upper limit for frequency band 70 – 140 Hz**
  - Mass range 85 – 390 M\(_\odot\) (sensitivity to 85 Mpc)
  - \(R_{90\%} = 3.2 \times 10^{-5} \text{ yr}^{-1} \text{ Mpc}^{-3} = 1.6 \times 10^{-3} \text{ yr}^{-1} L_{10}^{-1}\)

[Ref. 26]
Multiple messengers

- **Electromagnetic counterparts to compact mergers**
  - Require matter, so not always expected: none for BBH but worth checking!
  - Beamed emissions: $\gamma$, X-ray, optical, radio
    - Afterglows expected to be less beamed than GRB more likely to detect
  - Isotropic emission
    - Faint transient powered by radioactive decay

- **Neutrinos**
  - Low energy $\nu$ and high energy $\nu$
    - detectable in the local Universe

- **Benefits from multi-messenger astronomy**
  - Increase confidence in GW detection
  - Improve sensitivity of GW detectors
  - Get astrophysical context
  - Pinpoint source location
  - Break parameter degeneracy
GRB triggered searches

- Search data around times of GRBs observed by γ-Xray satellite based instruments
  - Search in [-5 s, +1 s] window around GRB time
- Similar to an all sky, low mass search, with some differences
  - Lower background
    - Direction of potential source known
    - Time delays between detectors known
    - Tighter coincidence possible
    - Smaller amounts of data searched
  - Lower threshold possible
    - More sensitive search
- Background estimated from off-source data
- S5/VSR1 GRB searches used likelihood statistic
LIGO-Virgo S5/VSR1 GRB triggered searches (I)

- 22 short GRBs with enough data from at least 2 detectors
- No detection candidate
- Distance lower limits derived

The interesting case of GRB 070201

» Not a merger in M31

[Ref. 27]
LIGO-Virgo S5/VSR1 GRB triggered searches (II)

- Look at all GRBs as a whole
- False alarm probability distribution for on-source data consistent with off-source data
**LIGO-Virgo S6/VSR2-3 GRB triggered searches**

- **Move to coherent analysis pipeline**
  - Computationally affordable because direction of potential source is known
  - N-detector coincident SNR has 2N noise degrees of freedom
    - N detectors x 2 phases
  - Coherent SNR has 4 noise degrees of freedom
    - 2 polarizations x 2 phases
    - Equivalent for non-degenerate 2-detector network, superior for >2-detector network
  
  \[
  \rho_{\text{c}o\text{i}n\text{c}}^2 = \sum_X \rho_X^2 = \sum_X \frac{(s^X | h_0)^2 + (s^X | h_{\pi/2})^2}{(\sigma_X)^2}
  \]

  \[
  \rho_{\text{c}o\text{i}n\text{c}}^2 = \sum_{X,Y} \sum_{i=0,\pi/2} \left( s^X \left| \frac{h_i}{\sigma_X} \right) \left[ \delta_{XY} \right] \left( s^Y \left| \frac{h_i}{\sigma_Y} \right) \right
  \]

  \[
  \rho_{\text{c}oh}^2 = \sum_{X,Y} \sum_{i=0,\pi/2} \left( s^X \left| \frac{h_i}{\sigma_X} \right) \left[ f_+^X f^Y + f_0^X f_0^Y \right] \left( s^Y \left| \frac{h_i}{\sigma_Y} \right) \right
  \]

  \[
  f_{+,-,x} \text{ depend on detectors response functions and sensitivities}
  \]

  [Ref. 28]

- **Null stream consistency**
  - Null SNR \( \rho_{\text{null}}^2 = \rho_{\text{c}o\text{i}n\text{c}}^2 - \rho_{\text{c}oh}^2 \)
  - Should be small for signals
    - \( \chi^2 \) distributed with (2N-4) d.o.f.
  - Can be large for incoherent noise transients

[Graph showing background and simulated signals]
LIGO-Virgo S6/VSR3 low latency search

- **Uses MBTA pipeline**
  - Multi-band template analysis, low latency oriented pipeline
  - Typical latency for trigger generation 2-3 minutes

- **Rapid sky-localization and online data quality check**
  - Typical latency for alerts (with human cross-check) 30-40 minutes

- **Only triple-coincident candidates with low FAR are sent for follow-up**
  - FAR estimated locally from single detector trigger rates (no time slides)
● The computing cost of a matched filter search based on a template bank is due to
  » The number of templates $\propto$ detector bandwidth
  » The size of the FFT involved in the matched filtering operation
    - Template duration $\propto$ dominated by the low frequency evolution
    - Sampling frequency $\propto$ imposed by the high frequency content of the signal

● The analysis can be split in a few bands (two or three)

\[
\int_{f_{\min}}^{f_{\max}} \tilde{h}(f)\tilde{Q}^*(f)df = \int_{f_{\min}}^{f_1} \tilde{h}(f)\tilde{Q}^*(f)df + \int_{f_1}^{f_2} \tilde{h}(f)\tilde{Q}^*(f)df + \int_{f_2}^{f_{\max}} \tilde{h}(f)\tilde{Q}^*(f)df
\]

[Ref.11]
MBTA (II)

- **Build one bank of *real* templates per frequency band**
  - Less templates in each bank
  - Short templates in high frequency band
  - Data can be downsampled for the low frequency bands filtering
  ⇒ Less and shorter FFTs

- **Filter data with each template bank**
  - Complex filtered signal (phase and quadrature) for each template
MBTA (III)

- Build a bank of virtual templates on the full frequency band
  » To each virtual template associate a real template in each frequency band

- Add coherently the filtered signals
  » Interpolate low frequency band results
  » Apply time delays and phase offsets between frequency bands
    - Take signal evolution into account
  » Conditional combination
    - If SNR exceeds some threshold in at least one of the bands
    - Built-in hierarchy

- Final threshold is applied on combined signal
Sky localization for low latency search

- **Online sky localization uses**
  - Timing information from 3 detectors
    - Time single detector signals at reference frequency where SNR is largest to improve accuracy
  - Amplitude information
    - Helps breaking symmetry w.r.t. detector plane
  - Galaxy catalog information

- **Accuracy remains modest**
  - Tens of square degrees typical
References (I)

Grav. 27, 173001 (2010)
References (II)

   LSC, Tuning matched filter searches for compact binary systems (2007)
[17] F.Robinet et al., Class. Quant. Grav. 27, 194012 (2010)
[19] M.Was et al., Class. Quant. Grav. 27, 194014 (2010)
[22] P.Brady et al., Class. Quantum Grav. 21 (2004)
   LSC, Phys. Rev. D 80 , 047101 (2009)