VESF school 2013

Searching for compact binary coalescence signals

Marie Anne Bizouard LAL, Orsay, France



VESF School on gravitational waves, neutrino and multi-wavelength EM observations: The new frontier of Astronomy

Outline

- Chapter 1: Data analysis
 - Overview of GW searches
 - Methods/techniques
 - Challenges with real data
 - Source parameter inference
- Chapter 2: Results
 - CBC search results with LIGO-Virgo
 - Detection scenarios with advanced detectors

Data analysis overview



Let's start with the beginning: Sources

- Compact binary systems: Neutron stars (NS) and/or black holes (BH)
- What can LIGO-Virgo detect? The last minutes of the coalescence, the merger and the ring-down for a certain regime of masses [1 M_{\odot} 400 M_{\odot}]



GW searches zoology



CBC searches: transient signal searches for LIGO-Virgo!

Methods summary



Network of ground based detectors



GW: what they are

• 2 polarizations



- Linear combination of + and x polarizations. Can be linear, circular or elliptical polarized.
- For instance a CBC wave is circularly polarized if traveling face on. In the other cases, it will be elliptically polarized.

Response of a GW interferometer



- How to detect the path of a GW?
 - \rightarrow GW induces a differential change of the arms' length
- \rightarrow light phase shift measurement

$$h_{det}(t) = F_+ \times h_+(t) + F_\times \times h_\times(t)$$



- Directional detector
- Directional sensitivity depends on polarization in a certain (+,×) basis

Network sky coverage



• GW detectors' readout system provides at any instant an estimate of strain: a quantity that is sensitive to arms' length difference: $h \sim \Delta L/L$

→ Digitized discrete time series: raw(t) (sampled at 16384 Hz or 20000 Hz) and synchronized with GPS clocks.

→ Calibration of raw(t): apply a frequency dependent factor [in reality this is a bit more complicated ...] $v_{1:h_16384Hz at 988224632.125}$

→ h(t) time series that is detector noise plus all hypothetical GW signals h(t) = n(t) + GW(t)

Question: nature of the noise?



• Detector monitoring: ~1000 auxiliary channels recorded at different sampling rate (environment/control monitoring)

 \rightarrow detector characterization effort to disentangle genuine GW signal from noise

GW detectors sensitivities



- Best noise spectrum achieved by LIGO Hanford, LIGO Livingston and Virgo
- Non white, non smooth ... and non stationary ...

CBC horizon distance

Another way to represent the sensitivity: distance at which an optimally oriented and located BNS (equal mass) system is detected with SNR=8

$$D = \frac{1}{8} \left(\frac{5\pi}{24c^3}\right)^{1/2} (G\mathcal{M})^{5/6} \pi^{-7/6} \sqrt{4 \int_{f_{low}}^{f_{high}} \frac{f^{-7/3}}{S_n(f)} df}$$



Frequency and time domain GW data representations

Fourier transform:
$$\widetilde{x}(f) = \int_{-\infty}^{\infty} dt \, x(t) e^{-i2\pi f t}$$

 $x(t) = \int_{-\infty}^{\infty} df \, \widetilde{x}(f) e^{i2\pi f t}$

Time series x_j with N samples at times $t_j = t_0 + j \times \Delta t$

 \rightarrow Discrete Fourier transform:

 $\Delta f = \frac{f_{sampling}}{N}$

$$\widetilde{x}_k := \sum_{j=0}^{N-1} x_j e^{-i2\pi jk/N}$$
$$x_j = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \widetilde{x}_k e^{i2\pi jk/N}$$

• Efficient algorithm to compute discrete Fourier transform: Fast Fourier Transform (FFT)

Power Spectral Density (PSD) estimation

 $\begin{aligned} &\text{PSD} = \text{Fourier transform of the auto-correlation function of the data} \\ &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t) x(t+\tau) dt \ e^{-2\pi i f \tau} d\tau = \tilde{x}(f) \tilde{x}^*(f) = |\tilde{x}(f)|^2 & \text{Wiener-Khinchin} \\ & \text{theorem} \end{aligned}$

When data has infinite extend in time domain, PSD estimate

$$\lim_{T \to \infty} \frac{1}{T} |\tilde{x}_T(f)|^2$$

In reality: finite amount of data \rightarrow true PSD is convolved with the Fejèr kernel (Fourier transform of a square function) \rightarrow bias of estimators

Estimators:

Simplest estimator (periodogram): FFT the data \rightarrow square each frequency component.

Averaged periodogram: to reduce variance of periodograms

Windowed data periodogram: to reduce spectral leakage (data are not periodic!). Tapered window

Welch approach: average of periodograms computed over overlapping windowed data segments

Time frequency representation

- Transient signal \rightarrow localized in time and frequency
- Many time frequency transforms (spectrogram, WignerVille, Wavelet, ...).
- Massive use of time-frequency map for (un-modelled) searches and detector characterization



CBC pipeline processing core: match filtering



Matched filter is the optimal filter to maximize the SNR in presence of additive noise. Detector's output is:

$$x(t) = n(t) + h(t)$$

To extract h(t) one filters x(t). The simplest linear filter is correlation

$$\mathcal{C}(t) = \int x(t')k(t-t')dt'$$

k(t) is the impulse response function of the filter $(x(t) = \delta(t)$ $\Rightarrow C(t) = k(t))$ $C(t) = \int \tilde{x}(f)\tilde{k}^*(f)e^{2\pi i f t}df$

which is just the inverse Fourier transform of $\tilde{x}(f)\tilde{k}^*(f)$

Matched filtering

Now we need to find k(t) that maximizes the signal-to-noise ratio (SNR),

$$\begin{aligned} x(t) &= n(t) + h(t) \Rightarrow \mathcal{C}(t) = \mathcal{N}(t) + \mathcal{H}(t) \\ \rho(t) &= \frac{\mathcal{C}(t)}{\sqrt{\mathcal{N}^2(t)}} \\ \overline{\mathcal{N}^2(t)} &= \int df |\tilde{k}(f)|^2 S_n(f) \end{aligned}$$

In absence of noise, one can show that $\rho(t)$ is bounded,

$$\begin{split} \rho^2(t) &<= \int df \frac{|\tilde{h}(f)|^2}{S_n(f)} \\ \rho(t) \text{ is maximal when } \tilde{k}(f) \propto \frac{\tilde{h}^*(f)}{S_n(f)} \\ \mathcal{C}(t) &= 4 \int_0^\infty \frac{\tilde{x}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t} df \end{split}$$

Matched filter : summary



- Phase coherence is more important than amplitude matching
- Need to build a bank of template that will cover the full parameter space
- \rightarrow Filter over the full template bank
- → Threshold on C(t) → trigger generation

• SNR is simply:
$$\rho^2(t) = \int_{-\infty}^{\infty} df \frac{|\tilde{h}(f)|^2}{S_n(f)}$$

In practise, integrals computed over $[f_l, f_h]$ f_l : after "seismic wall" f_h : when signal stops and/or $f_{sampling}^{20}$

- Analyse "good" segments of data
- Fourier inverse over overlapping segments of 256 s (avoid wrap around problems)
- PSD computed over longer segments (2048 s)



Source parameters vs signal parameters

- Inspiral source parameters
 - Masses (m_1, m_2)
 - Spins $(\vec{S_1}, \vec{S_2})$
 - Coalescence time (t_c)
 - Orbital phase at coalescence (ϕ_c)
 - Inclination of orbital plane (ι)
 - Sky location (RA, δ)
 - Distance (d)
 - Eccentricity

- ightarrow negligible for low mass $\left(M_{tot} < 10 M_{\odot}
 ight)$
- → maximizing analytically(2 quadratures filtering)
- \rightarrow affects amplitude only
- \rightarrow negligible for standard formation scenario
- Intrinsic parameters: masses, spins and coalescence time
 - \rightarrow 2+2*3 = 8 unknown parameters
 - → Templates bank is going to be huge!
 - → search split in 2 \neq mass regimes (\neq approximations, \neq physics, spins effects)

Compact binary spins

- Observations:
 - NS are found weakly spinning.
 - Stellar mass BH spins are almost maximal.
- Spin orientation and precession:
 - For isolated compact binary it is likely that spins are nearly aligned to the orbital angular momentum (tilt angle depends on SN kicks distribution)
 - For dynamically formed CB: no a priori spins alignment. If mis-aligned → spin-orbit & spin-spin couplings will cause the spins to precess around the total angular momentum.
 - \rightarrow maximal effects expected on the phase evolution.
 - \rightarrow face-on (minimal impact) vs edge-on (maximal impact)



Waveforms zoology

- Which mass regime (low mass, high mass, mass ratio)?
- Can we neglect the spins?

	Approximations	Spins
Low mass $(M_{tot} < 12 M_{\odot})$	BNS case SNR dominated by the inspiral phase. Waveforms accurate until PN approx. breaks	For initial detectors: spin could be neglected. Advanced LIGO/Virgo: spin 0.015 - 0.1 \rightarrow 3%-25% mismatch. Aligned spins is still OK?
High mass $(M_{tot} > 12 M_{\odot})$	For masses higher than 50 Msun, merger and ring-down contribute dominantly → need full waveform inspiral + merger + ring-down (IMR) Use of NR waveforms for merger.	BH spins is likely maximal Spin precession effects can be large. Spins cannot be neglected for Advanced LIGO/Virgo

Template parameter space: masses



The more massive the system the lower the GW frequency at merger

$$f_{\rm ISCO} \sim \frac{4 \rm kHz}{M_{tot}} \qquad f_{\rm merger} \sim \frac{15 \rm kHz}{M_{tot}}$$

- → For BNS waveforms are inside LIGO/Virgo band
- \rightarrow BBH merge inside LIGO/Virgo band

→ Mtot>100 Msun : only merger+ring-down

Numerical relativity breakthrough to describe the full inspiral + merger+ring-down waveforms

It's now possible to accurately calculate final stages of inspiral, merger and ring-down.

→ can construct "hybrid" waveforms:



Detection vs parameter estimation

- Complete description of the signal requires to explore the full parameter space: masses (2 dof) + spins (6 dof) + NR inputs
- Template search --> computationally limited
- Spinning waveforms have degeneracies
 - \rightarrow for detection:
 - A small inefficiency with 2dof template bank is acceptable
 - Use spinning waveforms to estimate the loss of efficiency for spinning CBC

 \rightarrow for GW candidate parameter estimation, use the most complete waveforms (up to 15 parameters to estimate)

Template bank construction

- Goal: pave the mass parameter space with waveform such that any (m1,m2) system can be detected with a minimal loss of efficiency
- Distance between 2 neighboring templates:
 - Compute the match filtering M of T1 and T2: 1-M gives the loss of SNR.
 - Define an acceptable minimal match.
 - Several algorithms (metric based, stochastic, hash cell, ...) have been developed.



Waveforms template

• Restricted waveforms with 3.5 PN corrections: $h(f) = C f^{-7/6} e^{-i\Psi(f)}$

$$\begin{split} \Psi(f;M,\eta) &= 2\pi f t_C - 2\phi_C - \pi/4 \\ &+ \pi \left[\frac{38\,645}{756} - \frac{65}{9}\eta \right] \left[1 + 3\ln\left(\frac{v}{v_0}\right) \right] + \left\{ \frac{11\,583\,231\,236\,531}{4\,694\,215\,680} - \frac{640}{3}\pi^2 - \frac{6\,848}{21}\left(\gamma + \ln(4\,v)\right) \right. \\ &+ \left(-\frac{15\,335\,597\,827}{3\,048\,192} + \frac{2\,255}{12}\pi^2 \right) \eta + \frac{76\,055}{1\,728}\eta^2 - \frac{127\,825}{1\,296}\eta^3 \right\} v^6 \qquad v = (\pi M f)^{1/3} \\ &+ \left. \pi \left[\frac{77\,096\,675}{254\,016} + \frac{378\,515}{1\,512}\eta - \frac{74\,045}{756}\eta^2 \right] v^7 \right\}, \end{split}$$

- Model complete (include also spin-orbit (1.5 PN) and spin-spin (2PN) effects
- Waveforms stop at ISCO.
- OK for low mass CBC

Waveforms template

- Restricted waveforms with 3.5 PN corrections.
- Effective One Body approach (EOB): describe in a non perturbative way the transition from the adiabatic inspiral to the unstable plunge.



- Restricted waveforms with 3.5 PN corrections.
- Effective One Body approach (EOB): describe in a non perturbative way the transition from the adiabatic inspiral to the unstable plunge.
- EOB matched with NR (EOBNR): mismatch with NR < 0.2%



Parameter space: masses



Coincidence analysis: each IFO data **H1** H2 L1 Stream will be processed separately Generate Generate Generate Generate Template Template Template Template Bank Bank Bank Matched Matched Matched Matched Filter Filter Filter **Ellipsoidal Coincidence Test** (Correlations Between Mass, Time Parameters) χ² and χ^2 and χ^2 and other signal other signal other signal other signal based vetoes based vetoes based vetoes based vetoes **Ellipsoidal Coincidence Test** with DQ vetoes

Bank

Filter

 χ^2 and







Coincidence between 2, 3 or 4 detectors (time & mass space coincidence)

 \rightarrow false alarm rate reduced

 \rightarrow time offset triggers to increase the effective livetime of the search for an accurate background estimation





- Remaining coincident triggers are ranked according to a detection statistic (ex: combined SNR, weighted likelihood, ...)
- Outstanding triggers (low p-value) are studied individually

Waveform consistency tests: χ^2 test

Time

10[°]

10 SNR

- - Weighted SNR

$$\rho_{\text{new}} = \begin{cases} \rho, & \chi^2 \le n_{\text{dof}} \\ \frac{\rho}{\left[\left(1 + \frac{\chi^2}{n_{dof}}^{4/3}\right)/2\right]^{1/4}}, \ \chi^2 > n_{\text{dof}} \end{cases}$$

04/15/13

10

10

10

10

10' يج

Waveform consistency tests: r² test

• Use as discriminating variable the time spent by $\chi^2(t)$ above some threshold in some time window prior to the measured coalescence time.



Gaining confidence in a signal candidate

- How do we know whether a signal in the data is a real GW?
 - Consistency with a source model (see previous examples)
 - Define a performant ranking statistic
 - Estimate p-value (coincidence/consistency in multiple detectors)
 - Absence of instrumental problems at the time of the signal
 - Validation of instrument response (candidate follow-up)
 - Association with a known astrophysical object (parameter estimation)

Coincidence/consistency tests – background estimation



- Signals should arrive at consistent times:
 - LIGO Hanford vs Livingston: ± 10 ms
 - LIGO vs Virgo: ± 27 ms
- Signals should have consistent properties:
 - Same or similar template in all detectors
 - Consistent frequency, amplitude, ...

Coincidence/consistency tests – background estimation

- CBC search asks for coincidence in at least 2 detectors
 - time coincidence
 - Mass coincidence (M and η)
- Use ellipsoids in the parameter space
 - To take into account correlation between parameters and parameters accuracy
 - Ellipsoids built using parameter space metric
 - One "distance" parameter to tune





Background estimation

• Mass dependence of the FAR: high mass templates that have less cycles brings loudest SNR triggers.



 \rightarrow define several mass regimes to define the FAR and the ranking statistic based on FAR.

Background estimation





Example (low mass LIGO-Virgo):



- Minimal data quality cuts: ask for periods when IFOs are "locked" and in "Science". No ADC saturation ...
- \rightarrow not enough to suppress all noise transients







- Instrumental vetoes based IFO slow monitoring (low power, electronic failure, etc)
- Instrumental vetoes based on statistical properties of coincidence between the GW channel and auxiliary channels



Virgo VSR2-3

- Statistical properties:
 - Efficiency (ε): eliminate false triggers, especially those with high SNR.
 Fraction of triggers which are flagged
 - Use percentage (UP): veto segments should always eliminate at least 1 trigger. Fraction of vetoes used to veto at least 1 trigger.
 - Dead time (dt): fraction of science time that is vetoed
 - Safety: vetoes should never suppress a real GW events. This is checked using hardware injected signals (force/current applied to a mirror to produce a differential motion equivalent to the effect of a GW)
- Auxiliary channels are "selected" according to several criteria:
 - High ε/dt, high UP, safety OK
- According to their statitical properties, vetoes belongs to different categories (CAT1, 2, 3)



CBC low mass search: Virgo vetoes eliminate a fraction of the loudest coincident triggers

Parameter estimation

- Parameters error estimate: Fisher matrix method
- We have a model → Bayesian approach provides well understood PDFs and degeneracies as a function of SNR (posterior distribution combines the information in the data with the prior information).
- The posterior distribution of a set of parameters $\vec{\theta}$ given a model H is given by Bayes' theorem:

$$p(\vec{\theta}|\{d\}, H) = \frac{p(\vec{\theta}|H)p(\{d\}|\vec{\theta}, H)}{P(\{d\}|H)}$$



• MCMC are rather computationally slow but accurate. Present implementations include fulling precessing spinning waveforms (15 parameters).

Example: mass estimation

- NS-BH inspiral signal injected into LIGO & Virgo
- Parameters extracted via a Bayesian MCMC including the effects of spins
- Masses are reconstructed correctly but the chirp mass is better reconstructed since it governs the phase evolution of the waveform.



Example: sky location & distance estimation

VESF

- Sky location is reconstructed but nowhere near as well as can be done with EM telescopes.
- Luminosity distance can be extracted from the amplitude of one knows the inclination angle 1.
- The 2 polarizations of the waveform depend on the inclination angle differently, so they can be disentangled if we have 3 or more detectors, oriented differently.
- CBC are standard sirens to get luminosity distance → measure Hubble constant (with 10% accuracy)





Results



LIGO – Virgo Runs



S6/VSR2-3 LIGO Virgo sensitivities



S6/VSR2-3 low mass CBC search

Phys. Rev. D 85, 082002 (2012)

- Search for 2-25 M_{\odot} total mass CBC
- PN restricted waveforms
- No evidence for a GW signal
- 90% upper limits on the events rate

NSNS: 1.3 x 10⁻⁴ Mpc⁻³ yr⁻¹ NSBH: 3.1 x 10⁻⁵ Mpc⁻³ yr⁻¹ BHBH: 6.4 x 10⁻⁶ Mpc⁻³ yr⁻¹

Still 2 orders of magnitude above "realistic" rate



S6/VSR2-3 high mass CBC search

Phys. Rev. D 87, 022002 (2013)



- Components

- EOBNR waveforms template (no spins)
- Maximum sensitive distance of the detectors over this period for a (20,20) M_{\odot} coalescence was 300 Mpc
- 90% confidence level merger rate limit of 3.3 $10^{-7}\,{
 m Mpc^{-3}\,yr^{-1}}$ for an equal mass 19-28 M_{\odot}

GRBs

Astrophys. J. 760, 12 (2012)

- BNS and NSBH merger are suspected to be short GRB progenitors
- Study more than 300 (long & short) GRBs since 2003
- Short GRB search: coherent CBC triggered pipeline
- 2009-2010 results: exclusion distance of 16 Mpc for BNS and 28 Mpc for NSBH.
- Assuming all GRBs emit the same amount of energy, one can derive an exclusion limit on the cumulative number as a function of their redshift.
 - \rightarrow with advanced detectors, we will be sensitive to what EM observations say.

GRB 051103

Astrophys. J. 755, 2 (2012)

- Short hard-γ spectrum GRB overlapping M81 (3.6 Mpc) observed few days before S5.
- Progenitors: NS-NS/BH or SGR giant flares
- CBC and burst searches: no evidence for a GW signal:

-->Merger progenitor is excluded at 98% CL. -->SGR hypothesis can't be ruled out (weak GW emission).

• Given the importance of GRB/GW association, and the rather near possible host galaxy this is a significant non detection result

GRB 071103

Ap.J. 681(2):1419–1430 (2008)

- Short GRB whose error box overlapped spiral arm of M31 (770 kpc away)
- LIGO hanford H1 (4km) & H2 (2km) were operating and sensitive up to 35 Mpc and 15 Mpc.
- Null results. Exclude CBC progenitor in M31 as source with 99% confidence
- Can't exclude SGR in M31

Low latency CBC searches

VESF

A & A 541, A155 (2012)

- Low latency search pipeline development (Goal: send alerts to telescopes within few minutes. GW inspiral signal may be detected first)
- Low mass search (1-35 M_{\odot}), including source sky position estimation.
- Search prototype tested at the end of S6/VSR3 with alerts sent to telescopes.
- 3 events have been selected. 1 has been followed-up by our telescope partners (FAR: 1/6.4 days)
- Latency: ~minutes
- Sky location: not so great (tens of deg²).
 - Using catalogues of galaxies help (but no galaxy complete catalogues for advanced detectors horizon distances)
 - Including spins may improve the accuracy

Electromagnetic follow-ups to GW triggers ("LOOC-UP")

Analyze GW data promply to identify possible event candidates and reconstruct their apparent sky position \rightarrow send alerts to telescopes

Preparation for advanced detectors

- Low latency searches: streaming mode, sky location (parameter estimation) and candidate significance estimation is less than a few minutes to send alerts.
- Pipeline developments
- Waveforms:

- Template placement algorithms for spinning waveforms
- Phenomenological spinning IMR waveforms gain study
- Long term projects (after discovery): astrophysics with BH and NS
 - Test of GR (measure deviation to GR)
 - Tidal disruption of NS near merger
 - EOS study

Bright future with advanced LIGO and advanced Virgo

Expected rate with 10 times more sensitive detectors

		-	2		
IFO	$Source^{a}$	$\dot{N}_{ m low}$	$\dot{N}_{ m low}$ $\dot{N}_{ m re}$		$\dot{N}_{ m max}$
		$\rm yr^{-1}$	$\rm yr^{-1}$	${ m yr}^{-1}$	$\rm yr^{-1}$
Initial	NS-NS	2×10^{-4}	0.02	0.2	0.6
	NS-BH	$7 imes 10^{-5}$	0.004	0.1	
	BH-BH	$2 imes 10^{-4}$	0.007	0.5	
	IMRI into IMBH			$< 0.001^{b}$	0.01^{c}
	IMBH-IMBH			10^{-4d}	10^{-3e}
Advanced	NS-NS	0.4	40	400	1000
	NS-BH	0.2	10	300	
	BH-BH	0.4	20	1000	
	IMRI into IMBH			10^b	300^c
	IMBH-IMBH			0.1^d	1^e

Promising ... but when?

- 2015: A 3 month run with the two-detector H1L1 network at early aLIGO sensitivity (40 80 Mpc BNS range). Virgo in commissioning at ~ 20 Mpc with a chance to join the run.
- 2016-17: A 6 month run with H1L1 at $80-120\,{\rm Mpc}$ and Virgo at $20-60\,{\rm Mpc}.$
- 2017-18: A 9 month run with H1L1 at $120-170\,{\rm Mpc}$ and Virgo at $60-85\,{\rm Mpc}.$
- 2019+: Three-detector network with H1L1 at full sensitivity of 200 Mpc and V1 at 65 115 Mpc.
- 2022+: Four-detector H1L1V1+LIGO-India network at full sensitivity (aLIGO at 200 Mpc, AdV at 130 Mpc).

	Estimated				Number	% BNS Localized		
	Run	Burst Range (Mpc)		BNS Range (Mpc)		of BNS w		vithin
Epoch	Duration	LIGO	Virgo	LIGO	Virgo	Detections	$5 deg^2$	20deg^2
2015	3 months	40 - 60		40 - 80		0.0004 - 3		-
2016 - 17	6 months	60 - 75	20-40	80-120	20 - 60	Q.006 - 20	2	5-12
2017 - 18	9 months	75 - 90	40-50	120 - 170	60-85	0.04 - 100	1-2	10-12
2019 +	(per year)	105	40 - 70	200	65 - 130	0.2 - 200	3-8	8-28
2022 + (India)	(per year)	105	80	200	130	0.4 - 400	17	48

First discovery in 2016?

Need to be lucky for EM follow-up ?