LIGO Computational challenges in gravitational-wave data analysis

- Lightning introduction to gravitational waves, the detectors, and the astrophysical sources.
- Techniques for detection and study of Compact Binary Coalescence (CBC)
- Template banks
- Waveforms
- Challenges ahead.

Alan Weinstein, Caltech LIGO-G1300617 ROM in GR Workshop, Caltech, 6-7 June 2013



"Colliding Black Holes" National Center for Supercomputing Applications (NCSA)







Gravitational Waves

Static gravitational fields are described in General Relativity as a curvature or warpage of space-time, changing the distance between space-time events.



Shortest straight-line path of a nearby test-mass is a ~Keplerian orbit.

If the source is moving (at speeds close to c), *eg,* because it's orbiting a companion, the "news" of the changing gravitational field propagates outward as gravitational radiation – a wave of spacetime curvature







A NEW WINDOW ON THE UNIVERSE



The history of Astronomy: new bands of the EM spectrum opened → major discoveries! GWs aren't just a new band, they're a new spectrum, with very different and complementary properties to EM waves.

- Vibrations of space-time, not in space-time
- Emitted by coherent motion of huge masses moving at near light-speed; not vibrations of electrons in atoms
- Can't be absorbed, scattered, or shielded.

GW astronomy is a totally new, unique window on the universe





Interferometric detection of GWs













Advanced LIGO schedule







Advanced GW Networks

- Gravitational-wave astronomy is greatly enhanced by having a multiplicity of interferometers distributed over the globe.
 - » GW interferometry, 'Aperture synthesis'
- Advantages include:
 - » Source localization in near real time
 - » Enhanced network sky coverage
 - » Maximum time coverage a fraction of the detectors are always listening'
 - » Detection confidence coincidence
 - » Source parameter estimation
 - » Polarization resolution







LIGO The Advanced GW Detector Network GEO600 (HF) , 🛹 Advanced LIGO Hanford 1 KAGRA Advanced LIGO 🐻 Advanced Livingston Virgo LIGO-India



EM counterparts to GW sources (if any) are short-lived and faint





LISA

Three spacecraft in orbit about the sun, with 5 million km baseline The center of the triangle formation will be in the ecliptic plane 1 AU from the Sun and 20 degrees behind the Earth.







LISA (NASA/JPL, ESA) may fly in the next 10 years!





Electromagnetic waves

- over ~16 orders of magnitude
- Ultra Low Frequency radio waves to high energy gamma rays

Gravitational waves

- over ~8 orders of magnitude
- Terrestrial + space detectors







Data analysis algorithms must be optimal!

- These detectors are "precision measurement science" writ large!
- They are marvels of engineering, and arguably the most precise measuring devices ever built - and at great expense (several \$100M for the network).
- But it is needed, because our signals are so weak (and so far, not yet detected!).
- There is no excuse for losing even a small amount of SNR due to imperfect data analysis; it is imperative to use techniques that are as close to optimal as possible, given finite / practical computing resources.





LIGO GW sources for ground-based detectors: The most energetic processes in the universe



- <u>Coalescing</u> Compact Binary Systems: Neutron Star-NS, Black Hole-NS, BH-BH
- Strong emitters, well-modeled,
- (effectively) transient



Credit: Chandra X-ray Observatory

Asymmetric Core Collapse Supernovae

- Weak emitters, not well-modeled ('bursts'), transient

- Cosmic strings, soft gamma repeaters, pulsar glitches also in 'burst' class

Cosmic Gravitationalwave Background

- Residue of the Big Bang, long duration

- Long duration, stochastic background



Spinning neutron stars

- (effectively) monotonic waveform
- Long duration



GWs from coalescing compact binaries (NS/NS, BH/BH, NS/BH)





• Neutron star – neutron star (Centrella et al.)



Tidal disruption of neutron star

A unique and powerful laboratory to study strong-field, highly dynamical gravity and the structure of nuclear matter in the most extreme conditions







Unmodeled, short-duration (<~ 1 s) GW Bursts





Magnetar flares / storms

Supernova collapse



High-mass binary merger and ringdown







Gravitational waves from Big Bang







Pulsars and continuous wave sources



Pulsars in our galaxy

»non axisymmetric: 10-4 < ε < 10-6
»science: EOS; precession; interiors
»"R-mode" instabilities
»narrow band searches best

Radiation

R-modes

21

Magnetic

Field

 $h = \frac{4\pi^2 G}{c^4} \frac{I f_{GW}^2}{d} \varepsilon$

 $f_{GW} = 2f_{ROT}$

Pulsar Model Axis



Frequency (Hz)





The CW challenge: All sky and frequency searches for GWs from spinning neutron stars



- Most spinning neutron stars are not observed pulsars; EM dim and hard to find.
- But they all emit GWs in all directions
- Some might be very close and GW-loud!
- Must search over huge parameter space:



- » sky position: 150,000 points @ 300 Hz, more at higher frequency or longer integration times
- » frequency bins: 1/T_{obs} over hundreds of Hertz sensitive detection band
- » df/dt: tens(s) of bins
- » Spin axis inclination and azimuthal angle
- » Binary orbit parameters (if in a binary system)
- This can add up to ~ 10¹⁵ templates or more...
- Computationally limited! Full coherent approach on only a fraction of observing time requires ~100,000 computers (Einstein@Home)

Einstein@Home: the Screensaver

- GEO-600 Hannover [~]
- LIGO Hanford
- LIGO Livingston
- Current search point
- Current search coordinates
- Known pulsars
- Known supernovae remnants

- User name
- User's total credits
- Machine's total credits
- Team name
- Current work % complete







Project stats info

Project name	Users	#last day	▼ Hosts	last day	Teams	🜲 last day	Countries	🜲 last day	Total credit	\$last day
BOINC combined	2,591,677	706	9,146,874	-4,895	98,433	12	273	0	1,525,795,095,184	1,469,151,807
Einstein@Home	342,993	57	4,283,556	4,727	10,698	-1	222	0	80,458,550,275	86,908,255
SETI@Home	1,397,169	482	3,417,140	1,696	61,527	3	233	0	189,671,473,004	64,583,578
World Community Grid	406,774	48	1,739,067	477	22,039	5	224	0	103,086,777,592	71,248,098
Rosetta@Home	360,274	44	1,121,279	139	10,170	0	225	0	23,058,096,753	11,874,586
Climate Prediction	270,011	14	555,789	56	7,690	0	221	0	20,361,132,625	7,346,165
MilkyWay@home	161,532	99	328,502	186	3,584	0	210	0	170,118,702,983	73,157,891
LHC@Home Classic	112,930	49	294,149	54	4,652	3	195	0	1,603,698,281	6,032,892
PrimeGrid	54,547	8	191,648	47	2,609	0	187	0	234,029,966,084	172,811,383
Malaria Control	68,419	10	164,157	47	2,211	0	208	0	3,249,170,284	2,316,289
Spinhenge@home	58,706	0	152,959	0	2,139	0	183	0	2,415,881,597	0
SIMAP	42,933	6	147,503	49	2,314	0	182	0	3,638,340,201	3,590,609
ABC@home	75,111	8	140,449	27	1,832	0	196	0	4,376,657,649	494,415
QMC@Home	49,838	0	130,406	0	2,188	0	177	0	5,159,038,236	0
POEM@HOME	40,585	14	109,330	45	1,467	0	170	0	48,724,355,440	128,086,877
Cosmology@Home	54,686	13	103,964	37	1,802	0	190	0	3,967,898,340	2,891,700
SZTAKI Desktop Grid	36,662	6	99,354	16	1,555	0	177	0	641,690,393	415,325
Docking@Home	32,683	11	86,647	16	1,105	1	142	0	4,261,126,321	3,616,316
Collete Contestuar	22,022	1.5	00 101	EO	1 202		100	0	100 010 040 750	128,026,000

Snapshot circa May 2013





The Gravitational Wave Signal Tableau



courtesy of Peter Shawhan





Summary of Data Analysis Methods



courtesy of Peter Shawhan

GWs from coalescing compact binaries (NS/NS, BH/BH, NS/BH)





• Neutron star – neutron star (Centrella et al.)



Tidal disruption of neutron star

A unique and powerful laboratory to study strong-field, highly dynamical gravity and the structure of nuclear matter in the most extreme conditions







Source Parameters vs. Signal Parameters



Only have to explicitly search over masses and coalescence time ("intrinsic parameters")



- Detection, via matched (Weiner) filtering through template banks (*optimal* technique in Gaussian noise)
- Detection pipeline validation / testing by injections into real, noisy (non-Gaussian, non-stationary) detector data.
- Evaluating detection pipeline sensitivity in real data using injections, evaluating sensitive volume (in Mpc³) as a function of parameter space.
- Parameter estimation: extracting source parameters (all ~15 of them), including 15-dim posterior PDFs.
- In most of these applications, need waveforms for arbitrary values of parameters in target space.





Mass space for template-based search

• The more massive the system, the lower the GW frequency at merger.

• Binary neutron star (BNS) waveforms are in LIGO band during inspiral; merger & ringdown are out-of-band.

 Higher-mass Binary black hole (BBH) waveforms merge in-band

•Above ~few 100 M_{sun}, all LIGO can see is the merger and ringdown







Searching for Known Waveforms



• $h(t) = A(t) \cos(\Psi(t))$

Waveform known well, or fairly well, in some parameterized space

- » e.g. inspiral with 1.4+1.4 M_{\odot}
- » or with 10+1.4 M_{\odot}

This and the next dozen slides are lifted from the excellent lectures at the CGWA Summer School in May 2012 by Peter Shawhan, U Maryland





Phase Evolution of an Inspiral

Accurate knowledge of the phase is crucial for matched filtering •Orbital phase vs. time \rightarrow orbital phase vs. frequency during chirp Post-Newtonian expansion" if spins are negligible: $\Psi(f) = 2\pi f t_c + \frac{3}{128n} (\pi m f)^{-5/3}$ Newtonian $+\frac{5}{96\eta}\left(\frac{743}{336}+\frac{11}{4}\eta\right)(\pi mf)^{-1}$ 1PN Relativistic effect $-\frac{3\pi}{8\eta}(\pi mf)^{-2/3}$ 1.5PN $+\frac{15}{64\eta}\left(\frac{3058673}{1016064}+\frac{5429}{1008}\eta+\frac{617}{144}\eta^2\right)(\pi mf)^{-1/3}$ 2PN + ··· $m = (m_1 + m_2), \quad \eta = \frac{m_1 m_2}{m^2}$ and "chirp mass" is $m\eta^{3/5}$ where

33





Inspiral Phase to 3.5PN

$$\begin{split} \Psi(f;M,\eta) &= 2\pi f t_C - 2\phi_C - \pi/4 \\ &+ \frac{3}{128\eta v^5} \Biggl\{ 1 + \left(\frac{3715}{756} + \frac{55}{9}\eta\right) v^2 - 16\pi v^3 + \left(\frac{15\,293\,365}{508\,032} + \frac{27\,145}{504}\eta + \frac{3085}{72}\eta^2\right) v^4 \\ &+ \pi \left[\frac{38\,645}{756} - \frac{65}{9}\eta\right] \left[1 + 3\ln\left(\frac{v}{v_0}\right) \right] + \Biggl\{ \frac{11\,583\,231\,236\,531}{4\,694\,215\,680} - \frac{640}{3}\pi^2 - \frac{6\,848}{21}\left(\gamma + \ln(4\,v)\right) \\ &+ \left(-\frac{15\,335\,597\,827}{3\,048\,192} + \frac{2\,255}{12}\pi^2 \right) \eta + \frac{76\,055}{1\,728}\eta^2 - \frac{127\,825}{1\,296}\eta^3 \Biggr\} v^6 \\ &+ \pi \left[\frac{77\,096\,675}{254\,016} + \frac{378\,515}{1\,512}\eta - \frac{74\,045}{756}\eta^2 \right] v^7 \Biggr\}, \end{split}$$

• ... where $v = (\pi M f)^{1/3}$





Basic Illustration of Matched Filtering



Filtering compresses signal into (approximately) a δ -function





Matched Filtering





Rewrite correlation integral using Fourier transforms...

$$\Rightarrow C(t) = 4\int_{0}^{\infty} \widetilde{s}(f) \ \widetilde{h}^{*}(f) e^{2\pi i f t} df$$

This is simply the inverse FFT of $\widetilde{s}(f)$ $\widetilde{h}^*(f)$

Computationally efficient way to calculate filter output for a range of times! Can compute h(f) in advance, and s(f) once for many filters.





FFT of data $C(t) = 4 \int_{0}^{\infty} \frac{\widetilde{s}(f)}{S_{n}(f)} \frac{\widetilde{h}^{*}(f)}{S_{n}(f)} e^{2\pi i f t} df$ Noise power spectral density

- Equivalently, "whiten" h(f) and s(f) by dividing by sqrt of noise power spectral density (noise amplitude SD, "ASD")
- IC(t)I automatically maximizes over coalescence phase φ
- Look for maximum of |C(t)| above some threshold \rightarrow trigger









Searching a Full Data Set

Search overlapping intervals to cover science segment, avoid wrap-around effects

Do inverse FFTs on, say, 256 s of data at a time

Estimate power spectrum from bin-by-bin median of fifteen 256-sec segments







Template Banks

- Want to be able to detect any signal in a large parameter space of possible signals
 - » All with different phase evolution
 - » Current astrophysical knowledge gives few clues for masses, spins
- ... but do it with a finite set of templates!
- So build a bank of templates
 - » Make sure there is a "close enough" template for every part of the signal space
 - » Require a minimum overlap between signal and template, e.g. 0.97
- Often can calculate a "metric" which parameterizes the mismatch for small mismatches







- This isn't data reduction, this is data explosion!
- In the end, we find peaks in $|z_i(t)|$ (triggers) and threshold on trigger SNR, greatly reducing the data (especially if it is Gaussian!)
- A low-mass search might have N ~ 10,000 templates, but the templates overlap greatly; only ~10% or fewer "independent" templates.
- One second of data (16,384 samples) lives in a 16384-dim space, mostly noise; the signal space spanned by independent physical templates is only ~1000-dim subspace. Filtering throws away the non-signal-like noise living in the rest of the full space.

41





Template Bank Construction

Templates and Parameter Space in Mass Coordinates













Only a few (red) error ellipses shown, to motivate the template placement.





Different Bank Layout Methods



Hexagonal placement is easily automatable, and generalizable to N dimensions.

However, in higher dim, and/or if the metric is unknown or complicated, stochastic methods can be used to place templates at a specified min match.





Reducing the order of the problem

- Why filter through ~10,000 overlapping templates, when only a fraction of them are "independent"?
- Use ROM methods to decompose the template bank, find the "independent" templates, filter the data only through them, and reconstruct the response to the original templates in the bank afterwards.
- Will this save CPU time (at the expense of extra effort)?
- We think it *will* be necessary, specially as we move to:
 - » (much) longer waveforms (up to 30 min)
 - » (much) larger template banks covering larger parameter spaces (total mass, mass ratio, spins, ...)
- Example: GstLAL: multi-banding, SVD, ...





In-band chirp length







GstLAL_inspiral

- Built from "off the shelf" open-source signal processing tools (Gstreamer).
- Time-domain signal processing to handle long aLIGO waveforms.
- Multi-band / multi-rate filtering.
- Adaptive whitening.





- SVD to reduce the number of filters.
- Real-time trigger generation and adaptive event significance (FAR).
- Capable of very low (~s) latency.





GstLAL_inspiral multi-banding





Cannon et al. (2012, ApJ 748:136) http://dx.doi.org/10.1088/0004-637X/748/2/136





More & better waveforms needed!

- Optimal filtering requires accurate waveforms across the full detection band, phase coherent up to ~25,000 cycles!
- Phase evolution depends on spin as well as masses.
- Spins non-aligned with orbital L will cause precession of the orbital plane, AM at the detectors.
- Binaries with total mass above $\sim 15 M_{\odot}$ will merge in-band.
- We do not yet have fully parameterized waveforms covering precessing spin and merger & ringdown!
- Waveforms from binary neutron stars will be tidally distorted above 500 Hz, disrupted above 1000 Hz.
- What if GR fails in these extremely strong-field, highly dynamical regimes?



Analytic Waveforms are increasingly less accurate for high-mass binaries





- PPN expansion for inspiral phase (in powers of v/c) breaks down as $v/c \rightarrow 1$, and increasingly depends on how the expansion is done.
- Also, black hole spin can have a large effect on the waveform, both in phase evolution and amplitude
- Higher mass systems transition from inspiral to merger and ring-down, in the detection band!





Phase evolution depends on mass ratio and spins



52





Higher mass systems merge in detection band







Effects of tidal disruption of neutron stars near merger







Numerical Relativity to the Rescue !

- It's now possible to accurately calculate final stages of inspiral, merger, and subsequent ringdown
- Can construct "hybrid" waveforms, "stitching" PPN early inspiral to late-inspiral NR, to extend to detection band f_{min}:







Can't we just use NR waveforms?

- They currently only span the last few (tens) of cycles to merger.
 - » Extend backward (to detection band f_{min}) by stitching to PPN analytical waveforms.
- They are very expensive to compute, especially for high mass ratios, high spin.
- We need to cover a large parameter space, 8-dim (masses and spins). This is a large space to sample finely!
- AND we need to have smooth, continuous coverage of that parameter space, especially for parameter estimation.
 - » Use analytical or phenomenological waveform models "tuned" to NR.
- Can we make use of a small number of hybridized NR waveforms that sample the parameter space, and "interpolate" between them using ROM methods?



Class. Quantum Grav. 29 (2012) 124001

P Ajith et al







Analytic Model Tuned Using NR

 "EOBNR": Effective One-Body model, with some parameters adjusted to match NR waveforms







Coarse sampling of 3-dim parameter space (NINJA2)



Figure 2. The mass ratio q and dimensionless spins χ_i of the NINJA-2 hybrid waveform submissions.





Vivien Raymond, http://www.ligo.caltech.edu/~vraymond/>

- Input: strain time series from all detectors
- Stochastically sample from parameter space, compute overlap of signal with data in each detector
- Sample distribution converges to posterior
- Can be computationally expensive
- Takes hours to days, currently





Some conclusions

- GW signals from astrophysical sources come in a huge variety of morphologies, durations, bands, and level of understanding / model-ability.
- They share one common property: by the time they reach us, they are *weak*.
- Pulling these signals out of noisy data, and extracting their astrophysical parameters, is a huge challenge, computationally and intellectually.
- We need all the clever and powerful techniques we can bring to bear on the problem... like ROM.