HPC Challenges in Astrophysics: Large Scale Scientific Computing for Gravitational-wave Detection

Kipp Cannon

Presented at IHPCSS19, Riken, Kobe, July 9, 2019

HPC Challenges in Astrophysics: Large Scale Scientific Computing for Gravitational-wave Detection

Kipp Cannon (観音切符)

Presented at IHPCSS19, Riken, Kobe, July 9, 2019

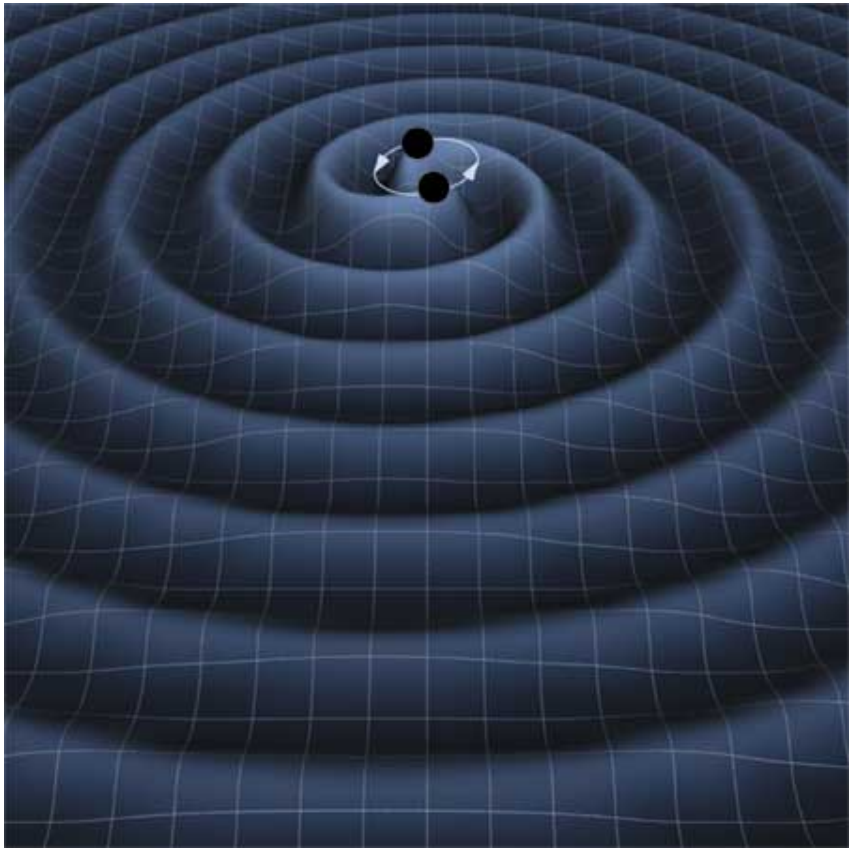
What I Won't Talk About

- PDE solvers.
  - Spacetime/gravity, electromagnetics, matter.
- N-particle solvers.
- Radiation transport codes.
- Lattice QCD.
- Control systems.

Simulation of GW170817 by BAM collaboration...

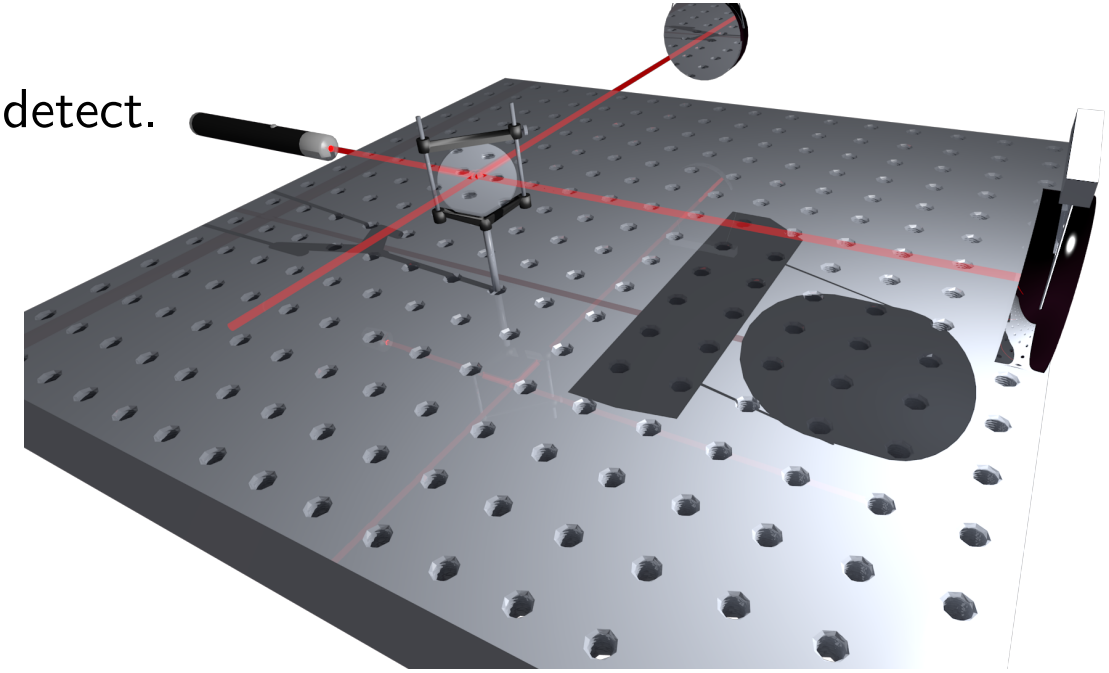
Gravitational Waves

- In electrodynamics, movement of charges and currents can lead to radiation of waves.
- In general relativity (gravity), similar phenomenon: movement of charges (mass) and currents (momentum) can lead to radiation of waves.
- The waves transport energy away from the system.
- Canonical example: spiral pattern of waves radiated by orbiting masses.
- Lycra and drill...



Gravitational-Wave Astronomy

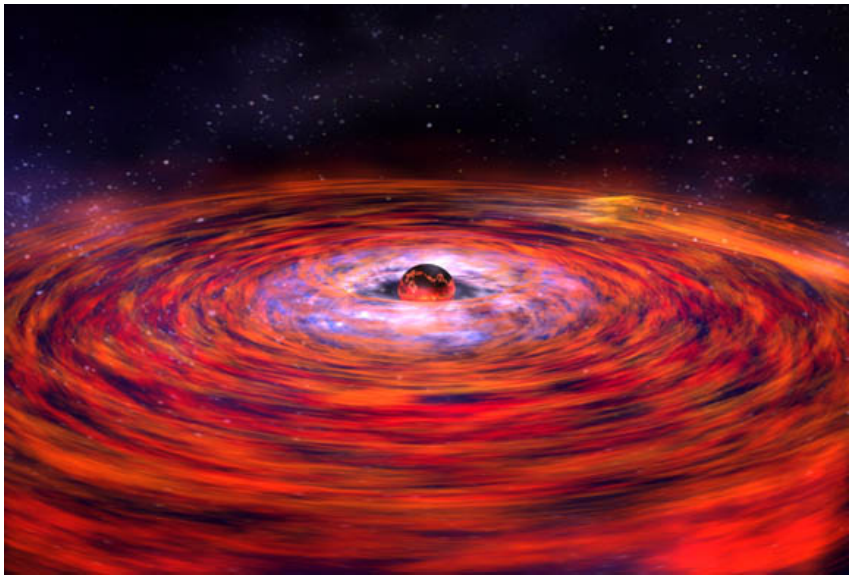
- Gravitational waves are produced by different physical processes than those that produce EM waves: carry different information about their sources.
- Weak coupling of gravity to matter: nearly everything is transparent to GWs.
- Promise to reveal things about nature that are inaccessible by other means.
- But are hard to detect.

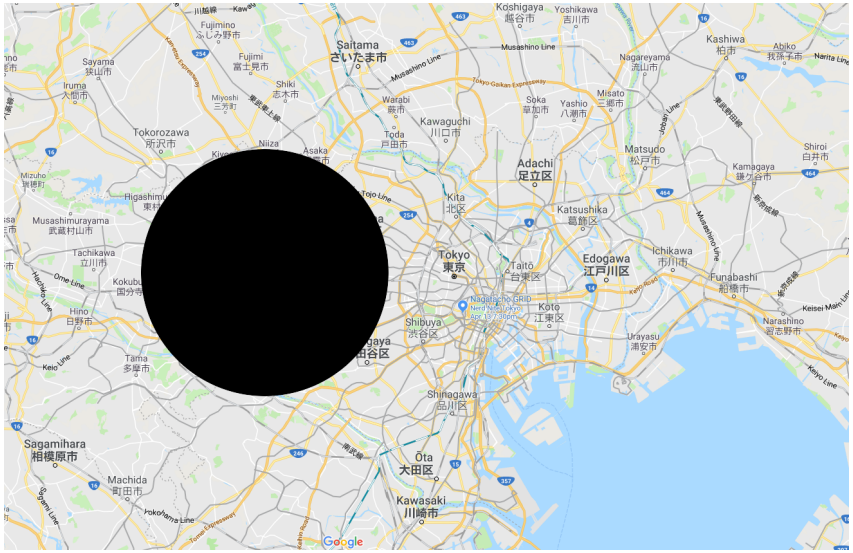


Real Detectors

LIGO Livingston Observatory. Each arm is 4 km long. An O(1) MW laser field resonates in the arm cavities.

Neutron Stars



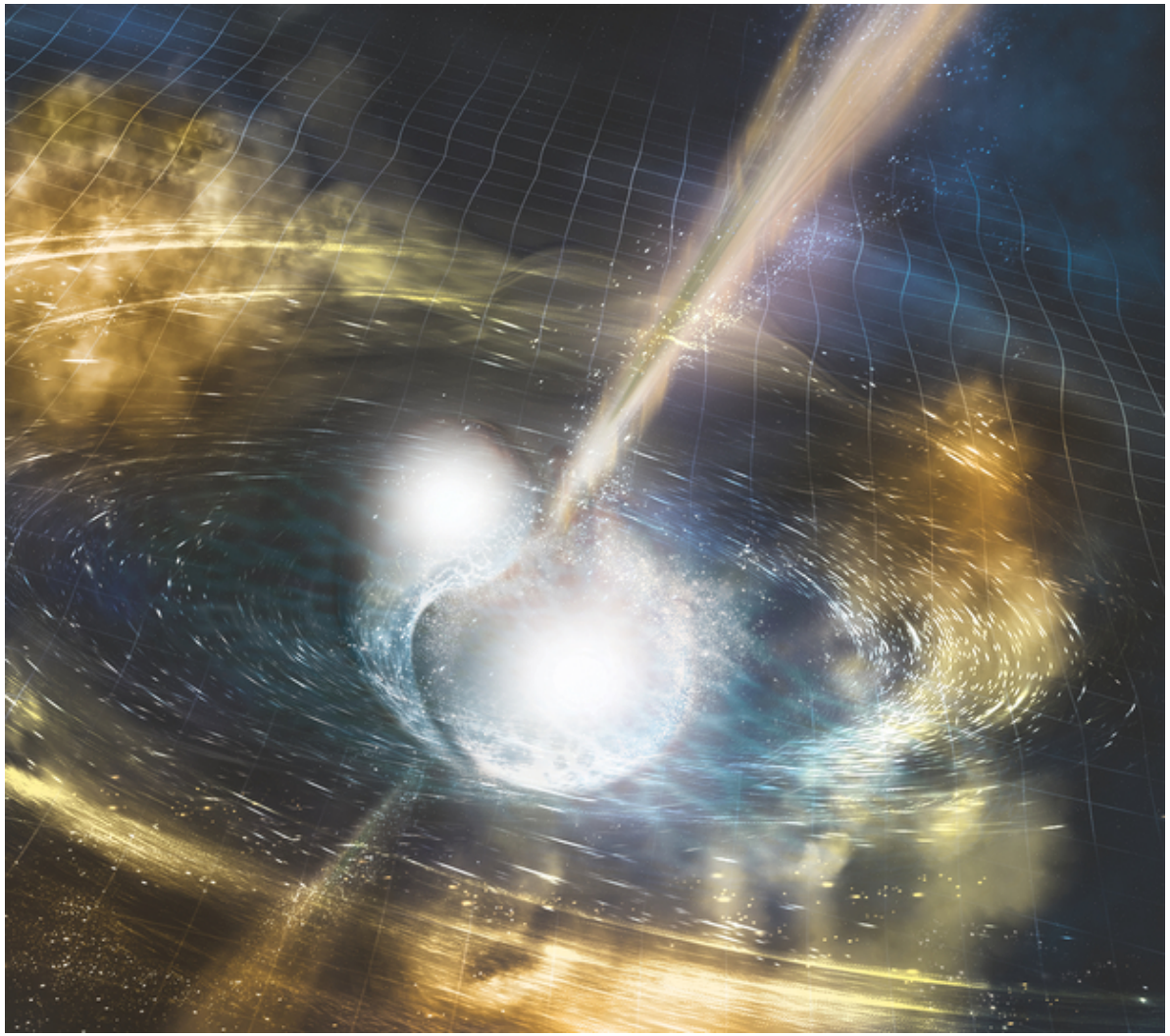


- Collapsed core of massive star.
- ≈1.5x mass of our Sun.

- About $20\,\mathrm{km}$ in diameter.

Simulation of GW170817 by BAM collaboration...

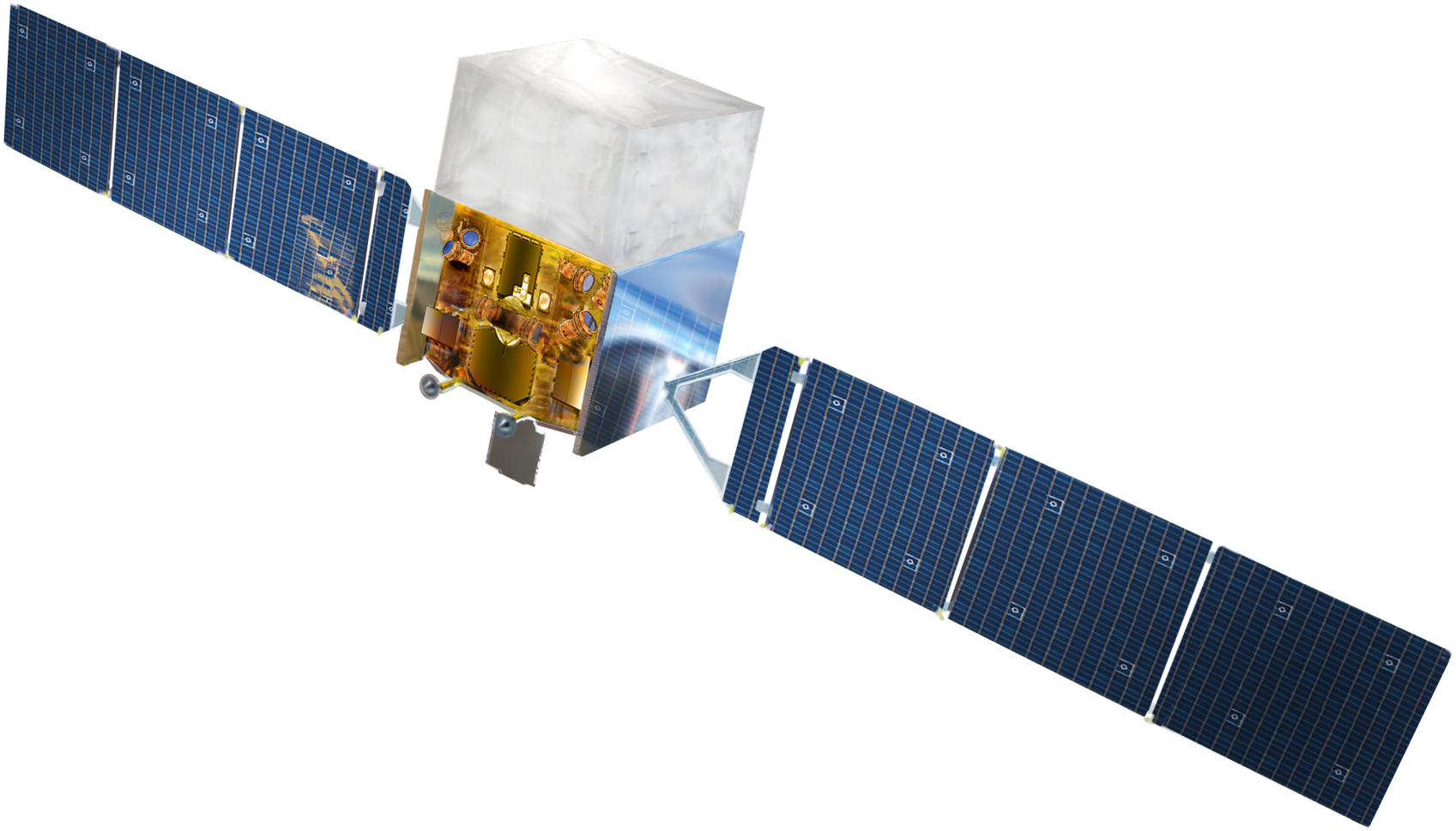
GW170817



Depiction of the origin of GW170817. Aurore Simonnet.

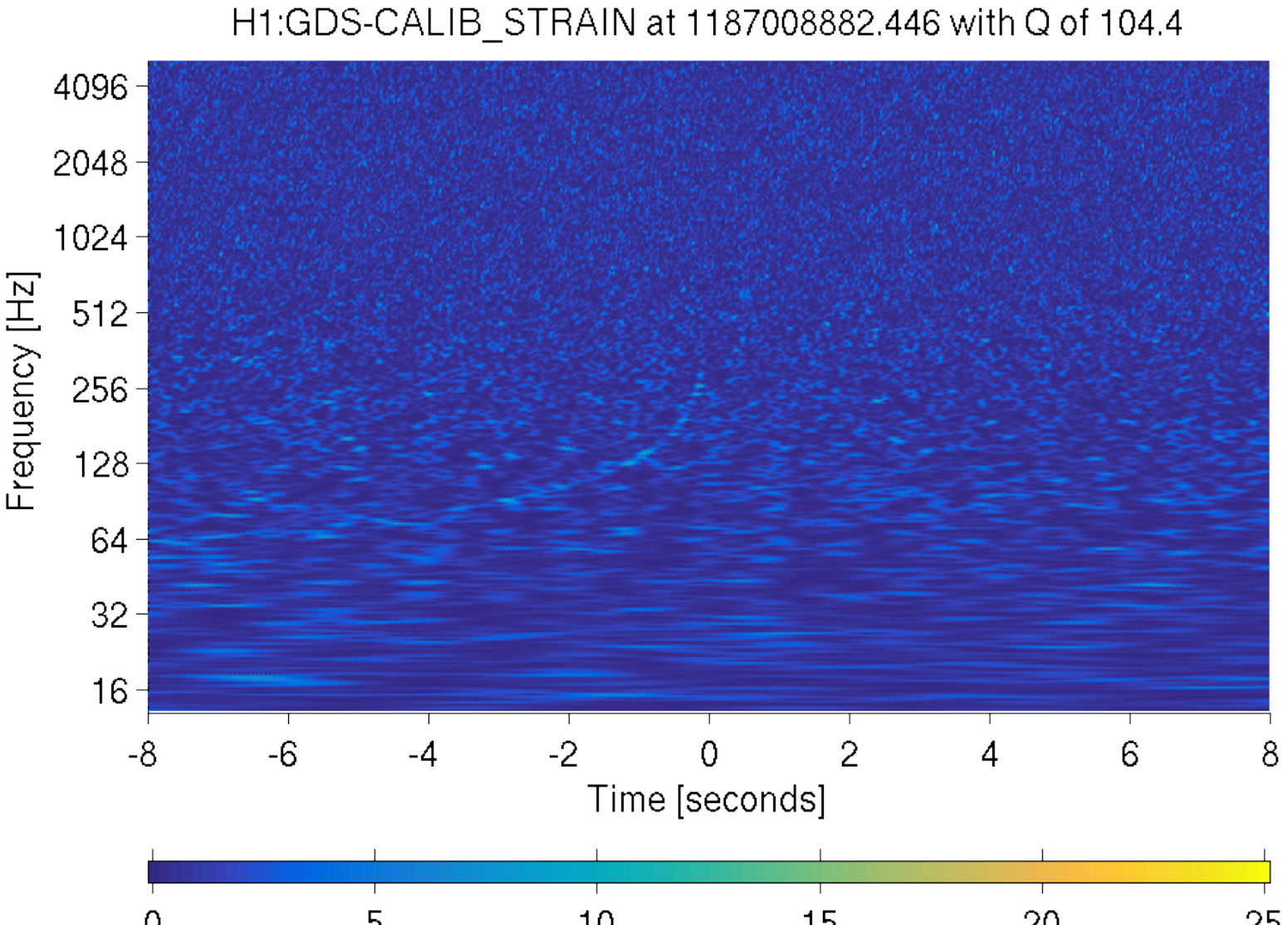
First Report

- γ rays reported by the Fermi GBM as trigger 170817.529 524666471.



Detection

- Followed by a report of a gravitational-wave signal preceding the γ rays by about 2 s.
- Consistent with a neutron star collision.



Optical Counterpart

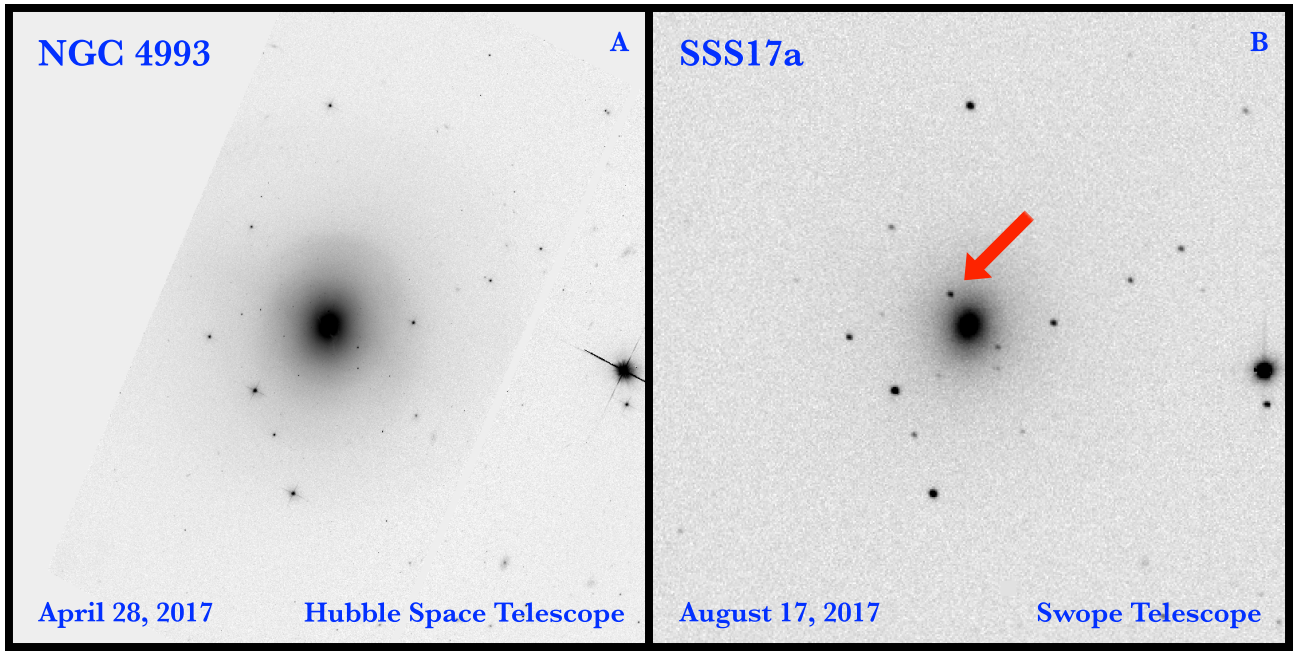
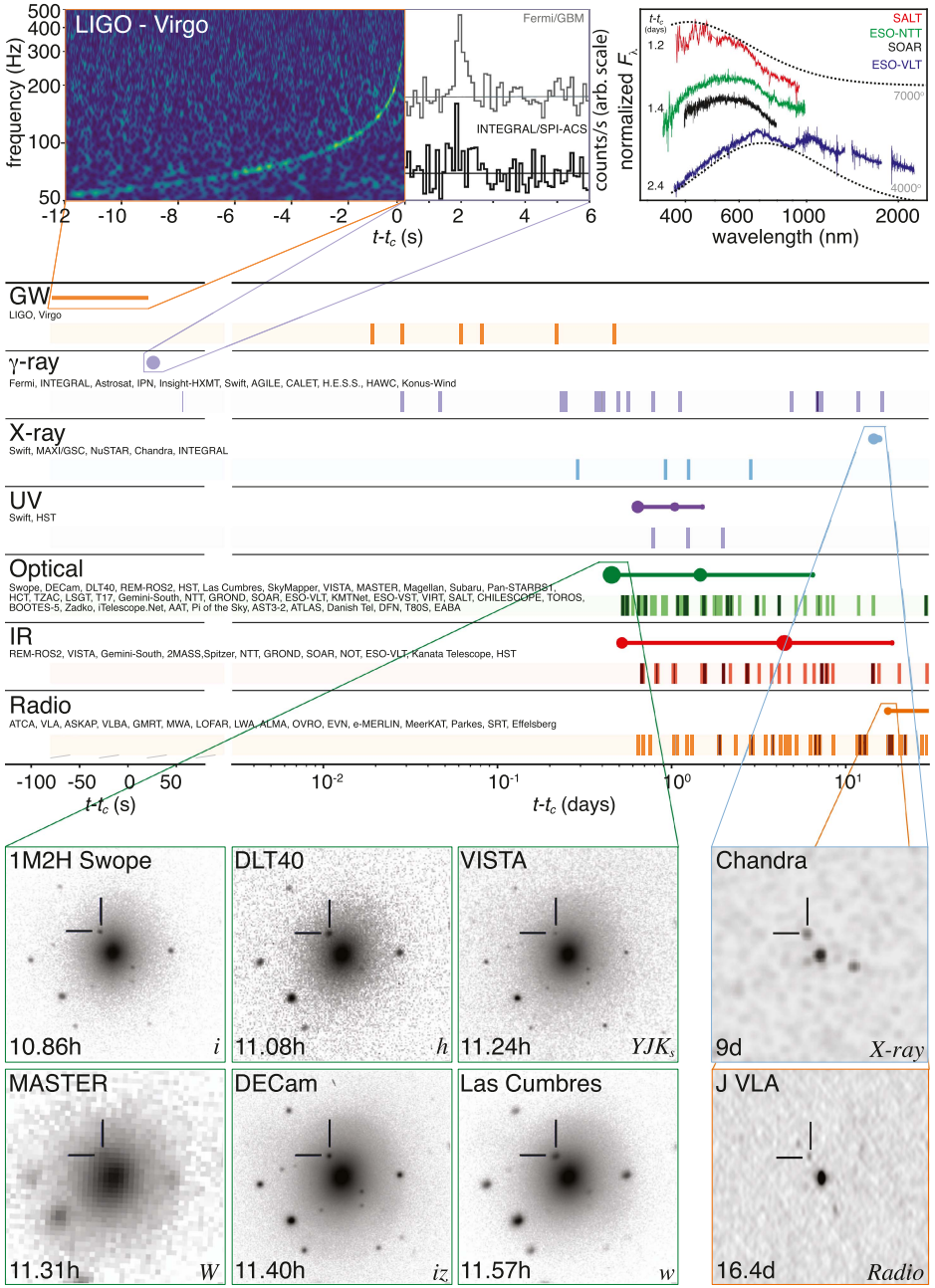


Figure 4 **3′ × 3′ images centered on NGC 4993 with North up and East left.** *Panel A*: *Hubble Space Telescope* F606W-band (broad V) image from 4 months before the GW trigger (*25, 35*). *Panel B*: Swope image of SSS17a. The i-band image was obtained on 2017 August 17 at 23:33 UT by the Swope telescope at Las Campanas Observatory. SSS17a is marked with the red arrow. No object is present in the *Hubble* image at the position of SSS17a (*25, 35*).

D. A. Coulter, *et al.*, "Swope Supernova Survey 2017a (SSS17a), the optical counterpart to a gravitational wave source" Science, October (2017). Photo taken August 18, 08:33 JST, reported 10:05:23 JST.

Vast Follow-up Campaign

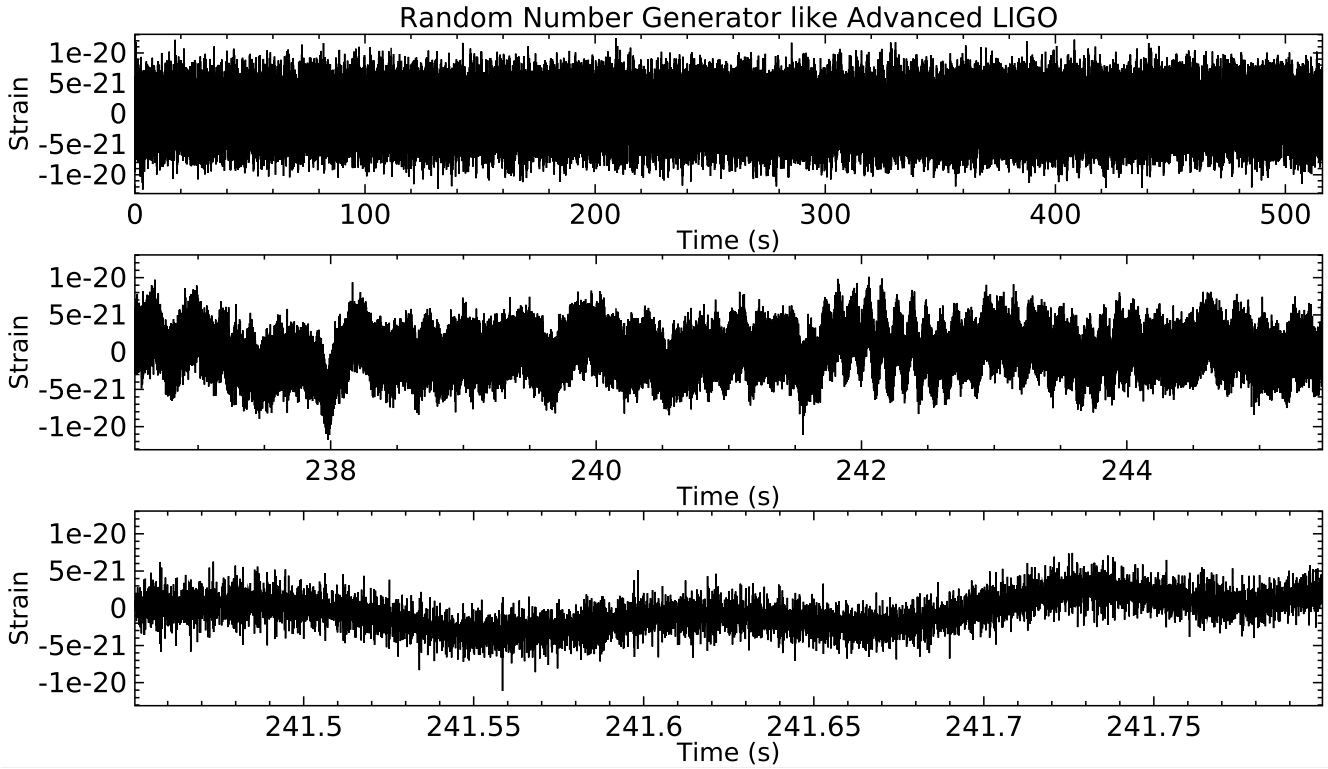


From Astrophys. J. Lett., 848:L12, 2017, arXiv:1710.05883

Behind The Scenes: Technical Introduction

- Gravitational-wave antennas are like radio antennas.
- Observatory provides digitized record of projection of field strength onto detector.
- No (real) detector has infinite bandwidth: something sets a low-frequency cut-off and something sets a high-frequency cut-off.
- For ground-based detectors: seismic noise and shot noise (laser field quantization).
- Compact object merger signals are frequency-swept sinusoids;
- start at low frequency, move to high.

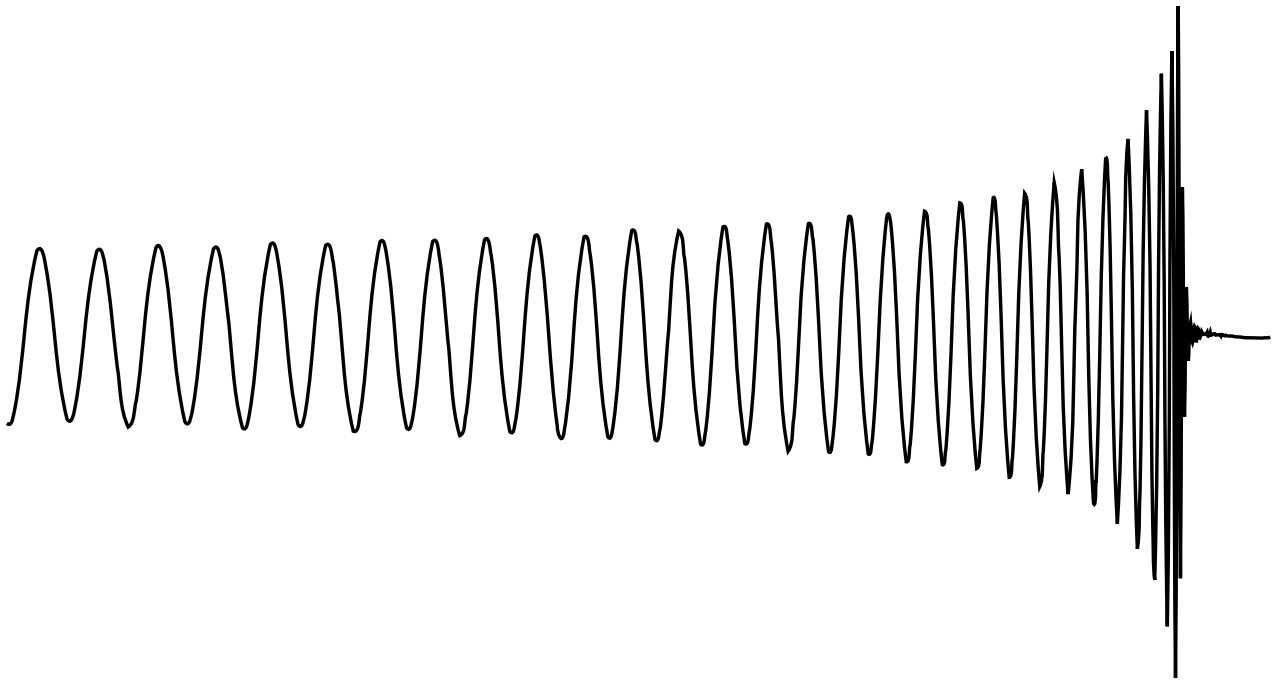
Behind The Scenes: Technical Introduction



- We have audio frequency time series data from O(few) antennas.
- Noise is mostly stationary, coloured, Gaussian noise,
- but also contains non-stationary "glitch" components.

Movie of noise spectrum...

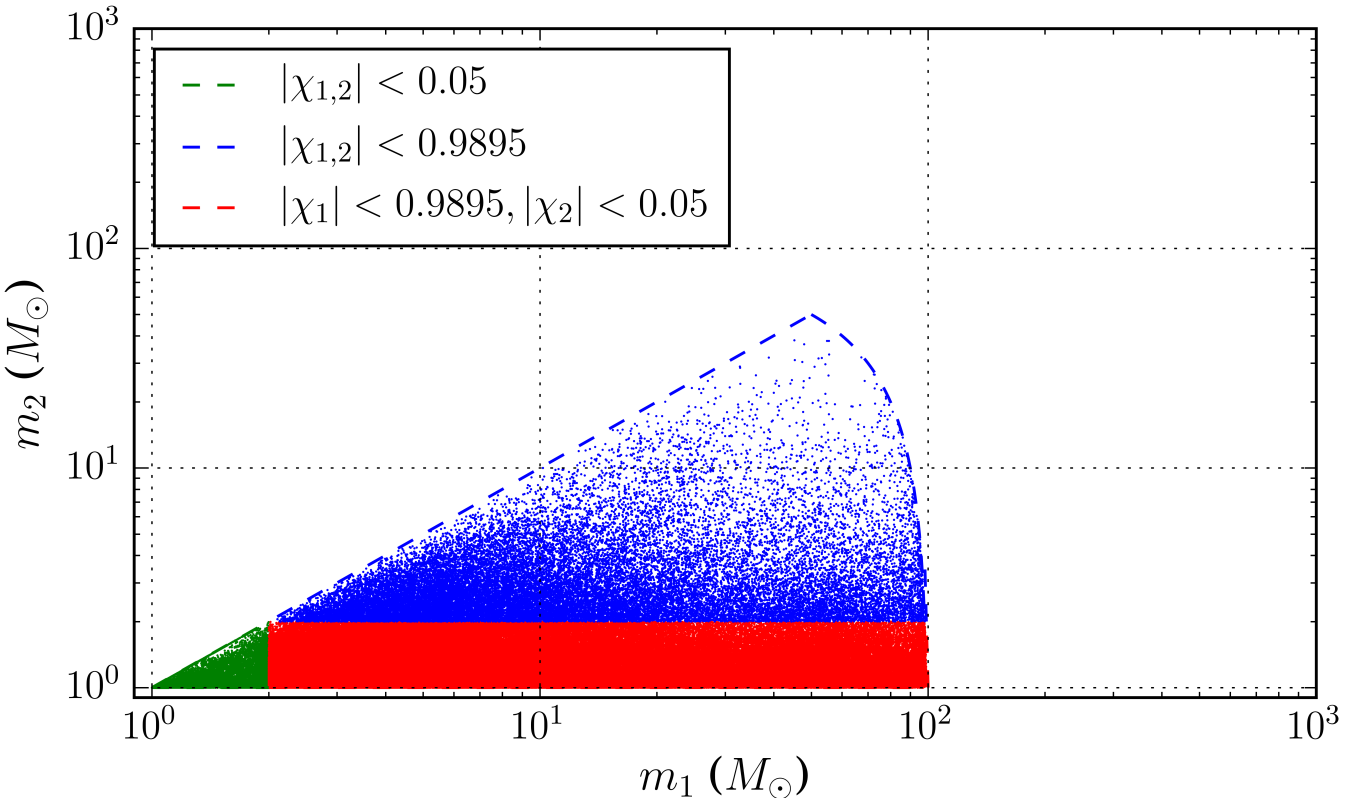
Behind The Scenes: Technical Introduction



- We are looking for signals like the above in the data.

Sound of neutron star collision...
Sound of GW170817...
Graphical depiction of matched filtering by Alex Nitz...

Behind The Scenes: Technical Introduction



Template bank used for O1 search (2015, first discovery of GWs). 249077 waveforms. From S. Caudill, *Techniques for gravitational-wave detection of compact binary coalescence*, Proceedings of 26th European Signal Processing Conference, 2018. ("χ" = mass-weighted projection of spin onto orbital angular momentum)

Cold Hard Facts

- LIGO's data are double-precision sampled at 16384 Hz, Virgo's are single precision at 20000 Hz.
- Compact object merger template banks have O(1 million) templates.
- Longest are O(10) minutes = O(10 million) samples.
- Naive brute-force inner products: $\mathrm{O}(10^{18})$ **FMA operations / second** $\times$
- Lowest hanging fruit: inner products $\rightarrow$ FFT convolution saves 6 orders of magnitude, still need O(1) PFLOPS *sustained* horsepower.
- Also, remember, we need the answer to the question "is this 10 min long signal in the data?" in seconds.
- Bonus: need to repeat the analysis many times with simulations hiding in the data to measure detection sensitivity for astrophysical interpretation of results.

Cold Hard Facts

- You cannot do this by coding up a prototype serial algorithm, debugging it, and then buying a room full of computers and parallelizing the *?!$ out of the code.

Cold Hard Facts

- You cannot do this by coding up a prototype serial algorithm, debugging it, and then buying a room full of computers and parallelizing the *?!$ out of the code.
- Horsepower will not solve this for us.

Cold Hard Facts

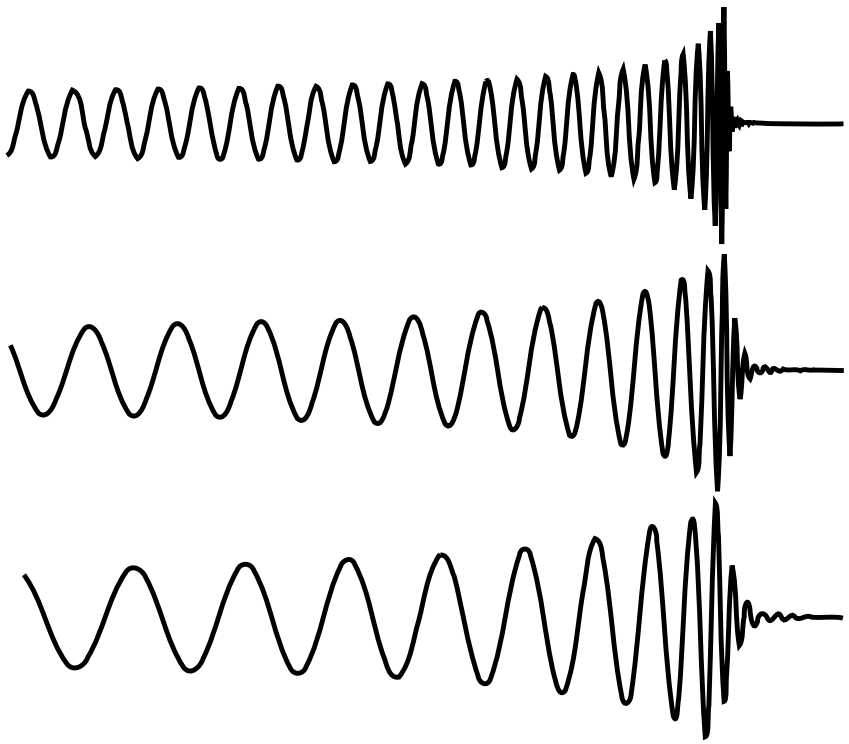
- You cannot do this by coding up a prototype serial algorithm, debugging it, and then buying a room full of computers and parallelizing the *?!$ out of the code.
- Horsepower will not solve this for us.
- "Optimization" will not solve this for us.

Cold Hard Facts

- You cannot do this by coding up a prototype serial algorithm, debugging it, and then buying a room full of computers and parallelizing the *?!$ out of the code.
- Horsepower will not solve this for us.
- "Optimization" will not solve this for us.
- You need to solve this by going back to the science, and changing the question.

Solution

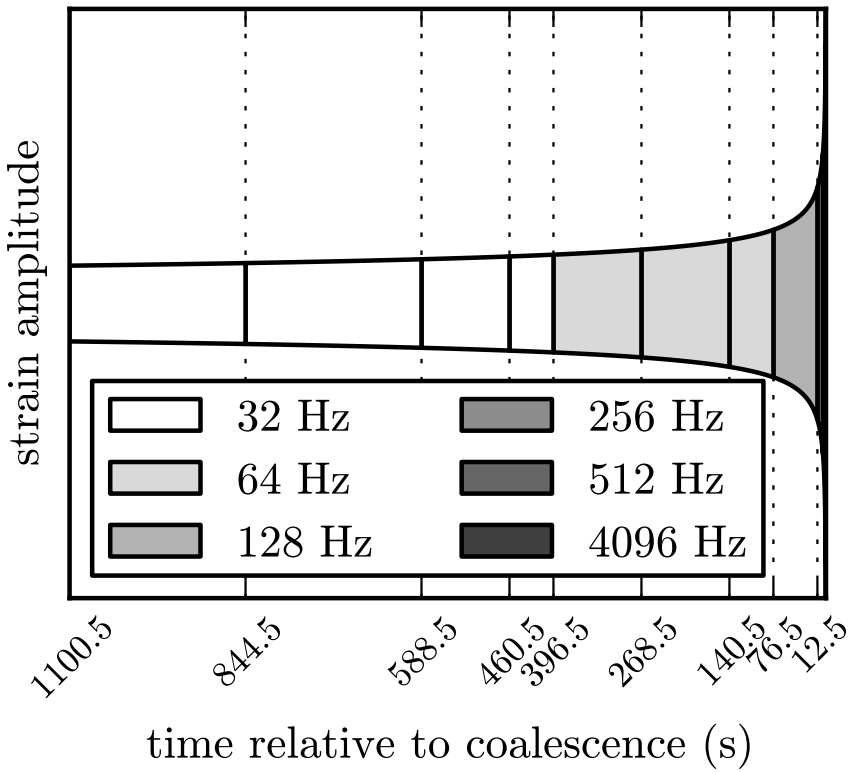
Take advantage of the structure of signals



- Slice templates into fragments, use lower sample rates for earlier parts.
- Use singular-value decomposition of each slice to obtain reduced-order approximation of template bank matrix.

Solution

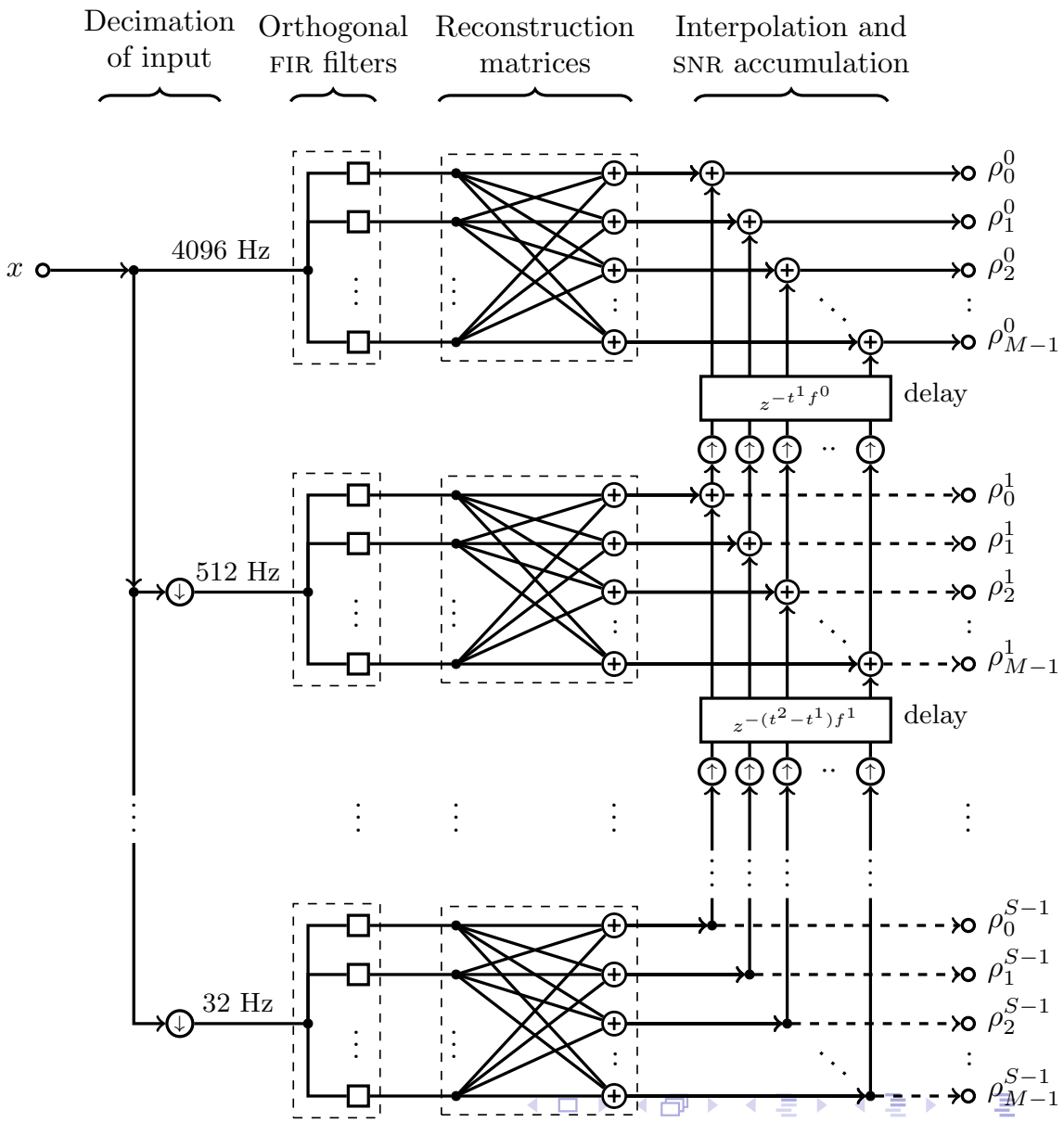
Take advantage of the structure of signals



- Slice templates into fragments, use lower sample rates for earlier parts.
- Use singular-value decomposition of each slice to obtain reduced-order approximation of template bank matrix.

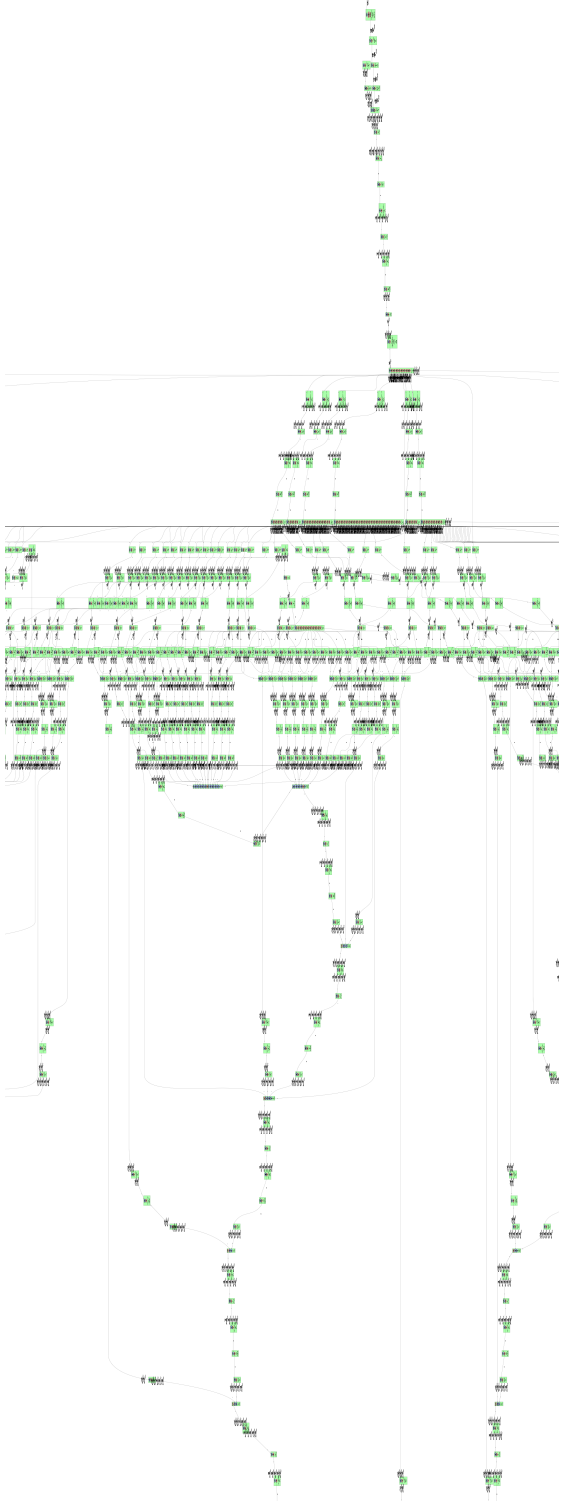
Solution

- Use libre DSP framework to put it together.
- gstreamer

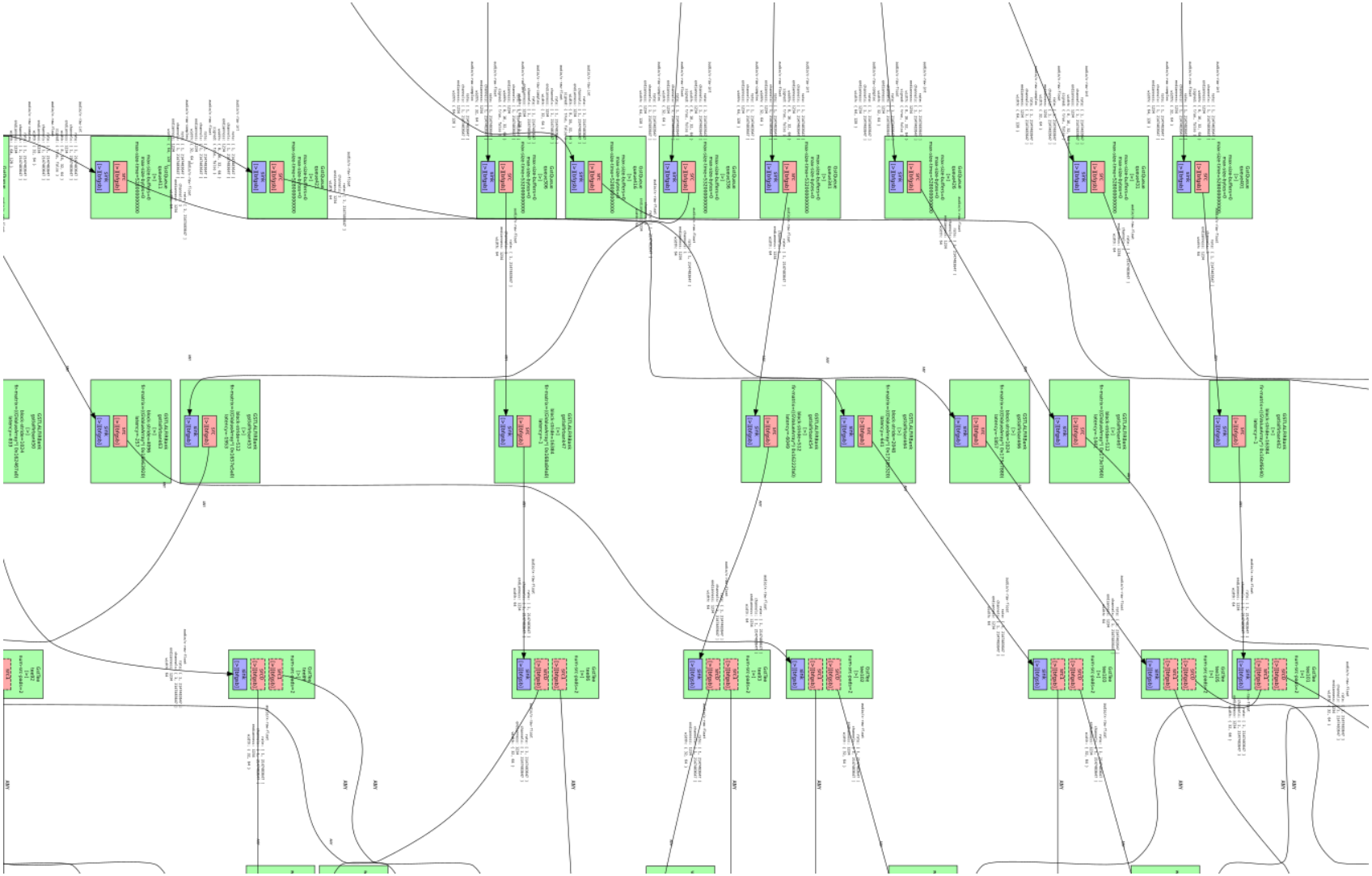


Detection System

- "GstLAL" detection system uses consumer multimedia processing software called "GStreamer".
- $< 20\,\mathrm{s}$ of latency from phase centre to alert database.
- Latency dominated by data calibration and distribution: $\sim 14\,\mathrm{s}$



Detection System



Tip of the Iceberg

- What I have been describing is the system that scans the data and reports *potential* signals.
  - "Candidates".
  - Matched-filter signal-to-noise ratio excursions.
- Each template reports O(1) candidate per second — total candidate identification rate = O(1 MHz).
- Need to pick the interesting ones: purpose-built machine learning code.
- Need to rigorously assess statistical significance.
- Need to repeat analysis with simulated signals hidden in data to measure detection sensitivity.
- Need to interpret results:
  - Infer astrophysical event rate.
  - Infer compact object mass distribution.
  - Infer neutron star equation of state.
  - Test general relativity: was the waveform correct?

Mistakes We Made

- Vastly underestimated computing requirements: "the control room will have a *dedicated workstation* to identify interesting fragments of data".
- Boeing style of software engineering: write a spec, hand to programmers, integrate what they produce, and find GWs.
- Not Invented Here: in-house FFT, in-house RNG, in-house call tracing.
- Hubris: it will take less time to write it myself than to read the manual.
- Over-design: LIGO-Virgo frame file specification. To this day, *nothing* can read the format.
- All I need to know is MatLab: if your only tool is a hammer, every problem looks like a nail.

What Worked

- Agile development: we didn't know how to find signals, and needed software that was easy to reconfigure.
  - More important than raw performance.
  - Yes, Python and SQL have a place in HPC.
- Libre software community participation: do not fix bugs and add features and keep the result in your home directory, *push your fixes upstream*.
- Use existing solutions wherever possible
  - Take the time to research what is available.
  - LIGO-Virgo frame file format, when all we really needed was a tar ball of .wav files.
  - But it must be actively maintained, do not select software that is not seeing regular releases.
- Keep your software up-to-date, track upstream API changes.

Credits

- Gravitational-Wave Open Science Center
  `https://www.gw-openscience.org/about/`
- LIGO Algorithm Library
  `https://git.ligo.org/lscsoft/lalsuite`
- GstLAL `https://git.ligo.org/lscsoft/gstlal`
- Also starring: Python, GStreamer, GSL, Scipy, Numpy, SQLite, libexpat, HDF5, HEALPix, QHull, GCC, ATLAS, FFTW, Matplotlib, emcee, HTCondor, framecpp, Kafka.