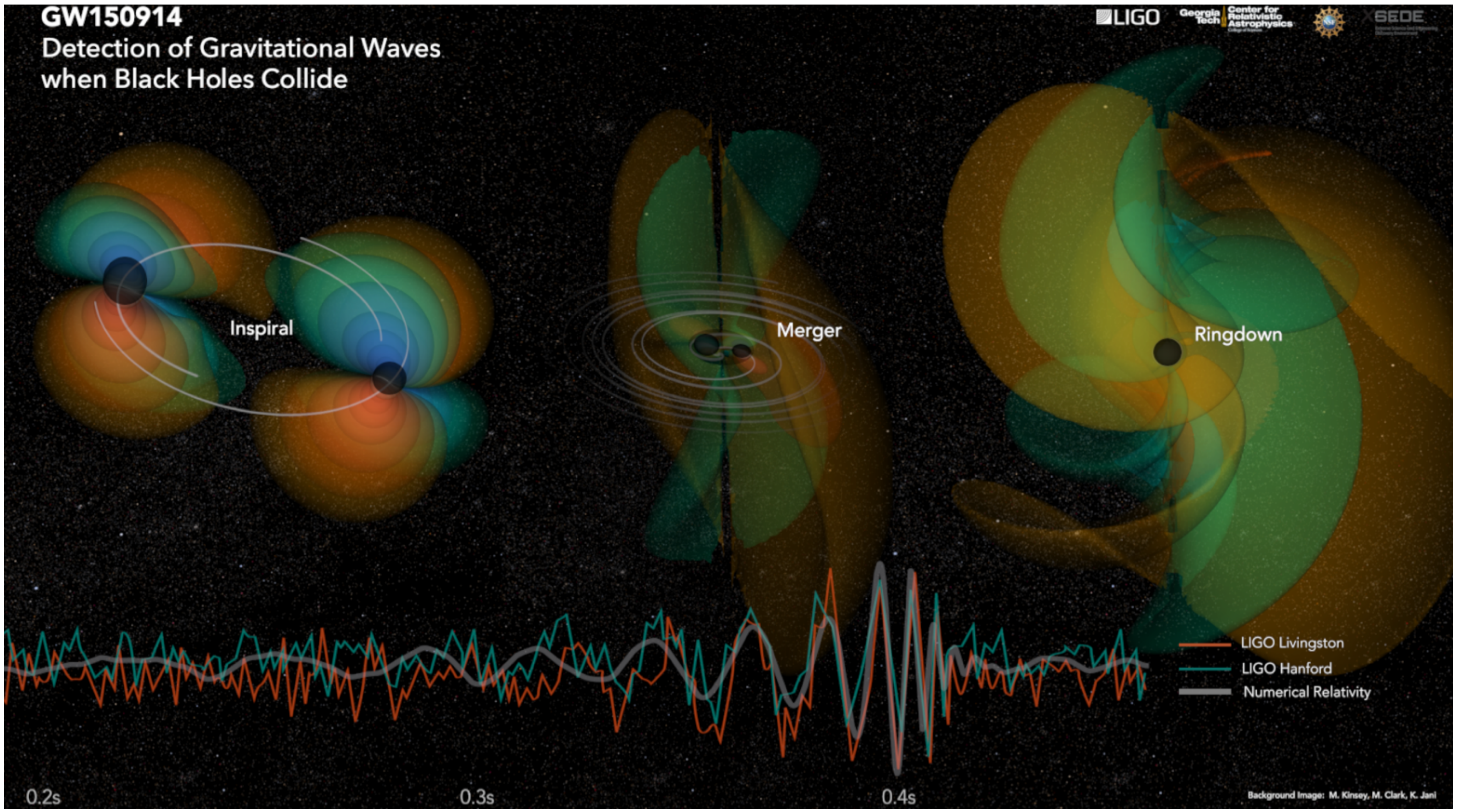
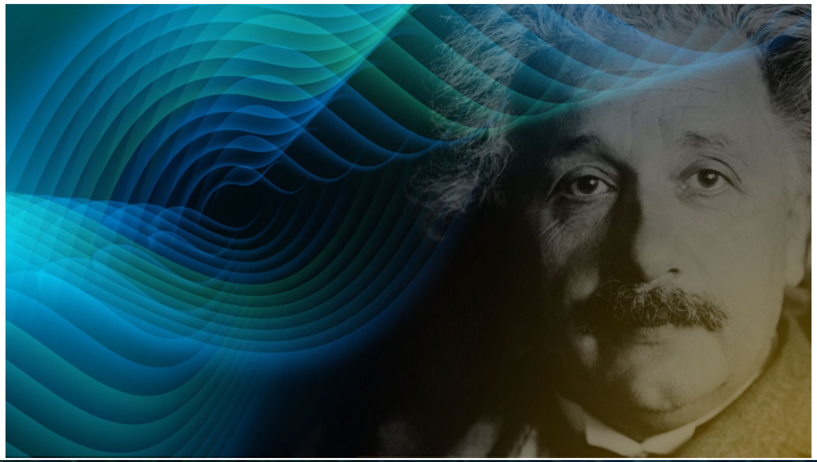
OSG as a “Universal Adapter” for LIGO

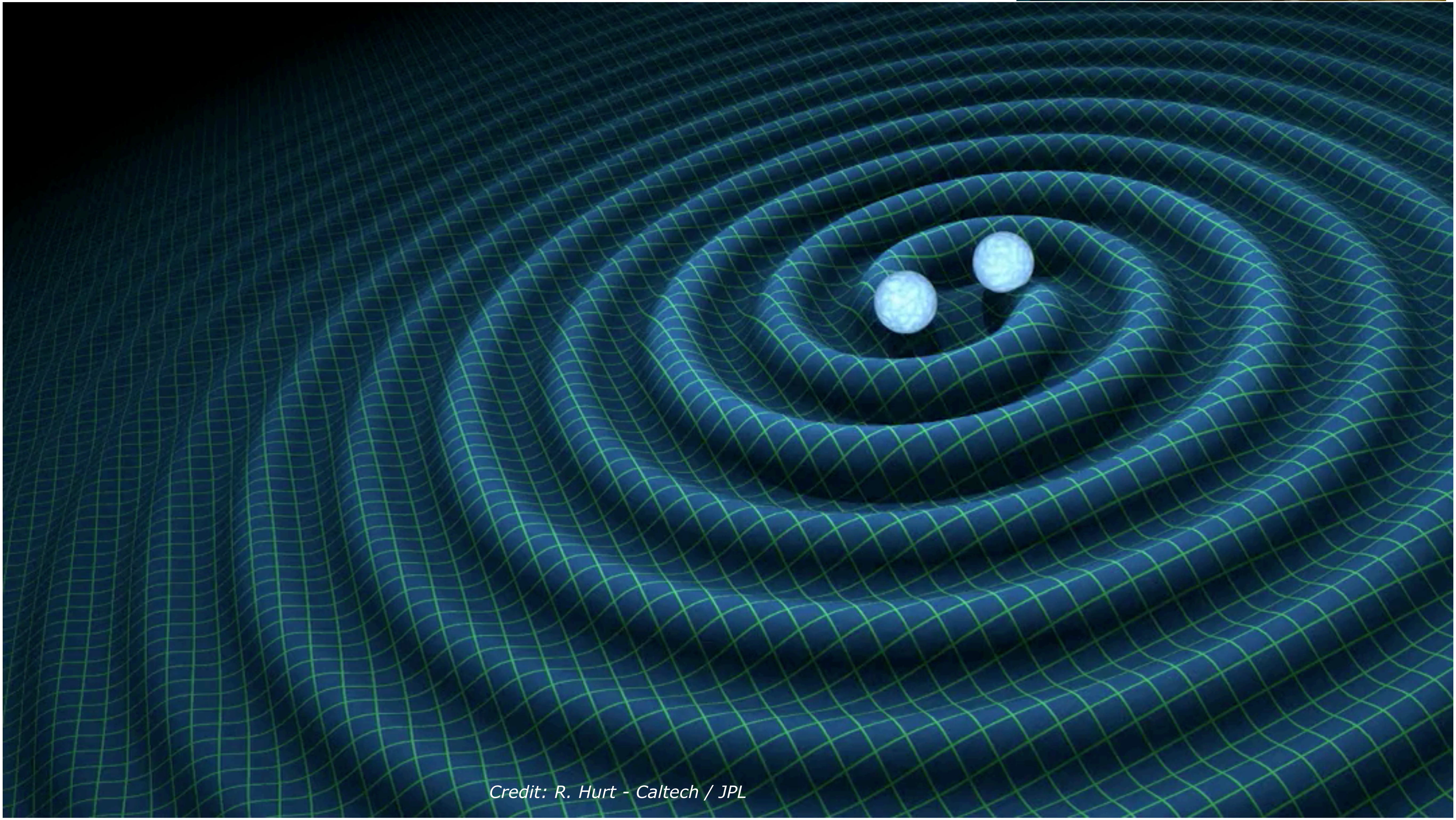


Peter Couvares, LIGO Laboratory - Caltech

May 19, 2016

What are Gravitational Waves?





Credit: R. Hurt - Caltech / JPL

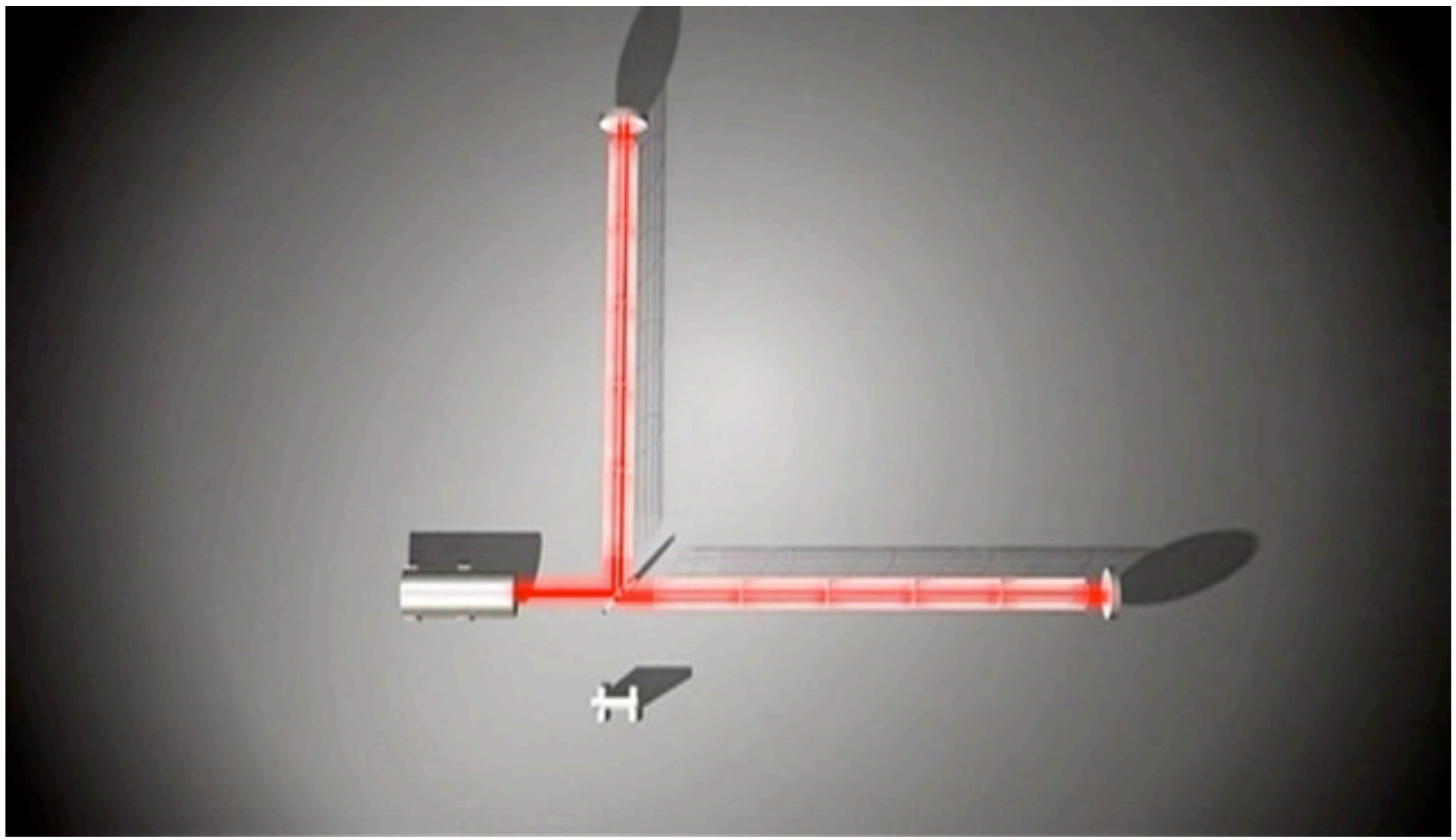
The Challenge of Detecting Gravitational Waves

They are tiny!
A gravitational wave from two merging neutron stars 500 million light years away is:

$$\frac{\Delta L}{L} \sim 10^{-22}$$

Equivalent to measuring distance between the sun and Proxima Centauri to less than the width of a human hair!

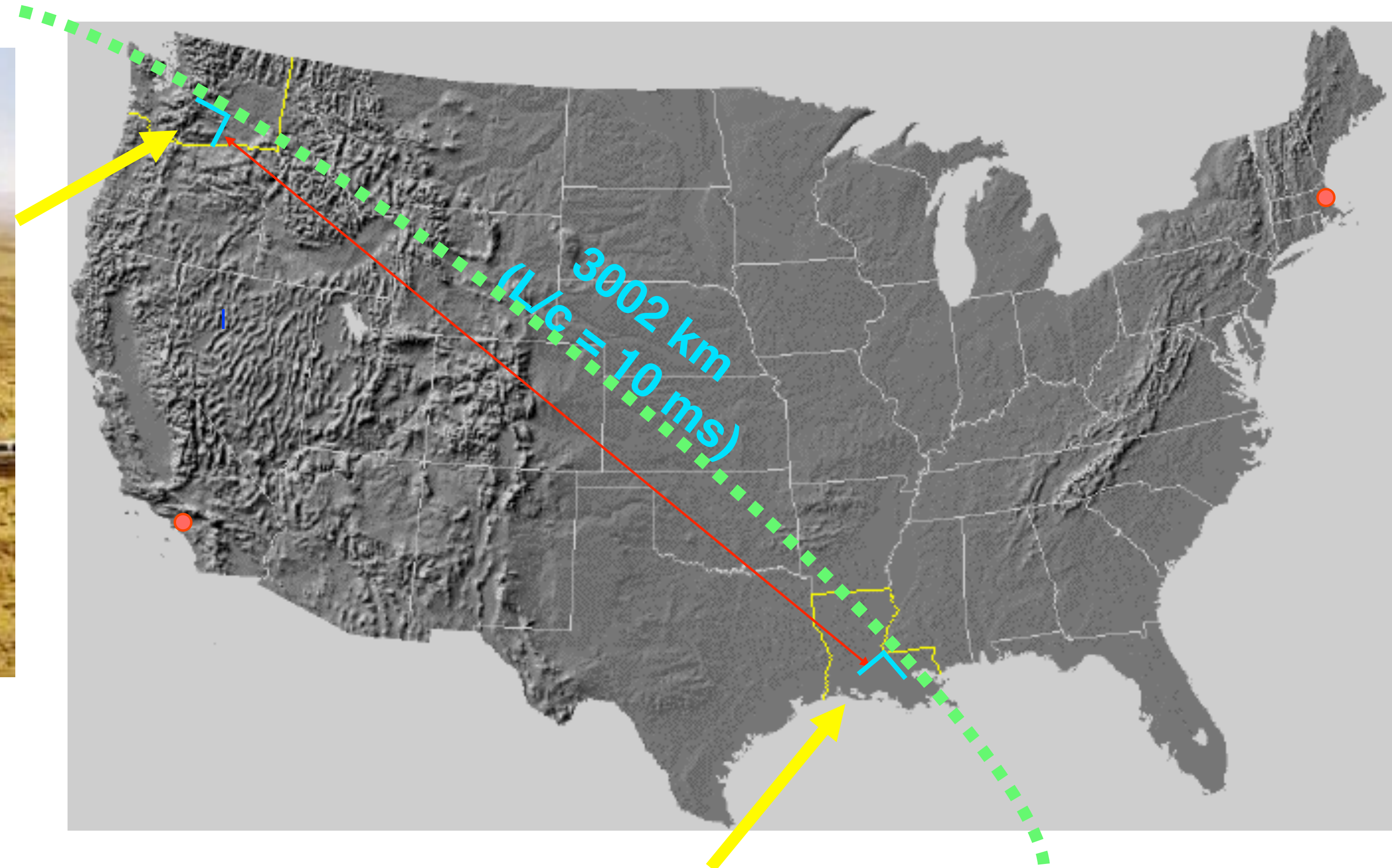
LIGO:
Laser Interferometer Gravitational-wave Observatory



Einstein's messengers,
National Science Foundation video

LIGO: Laser Interferometer Gravitational-wave Observatory

Hanford, WA



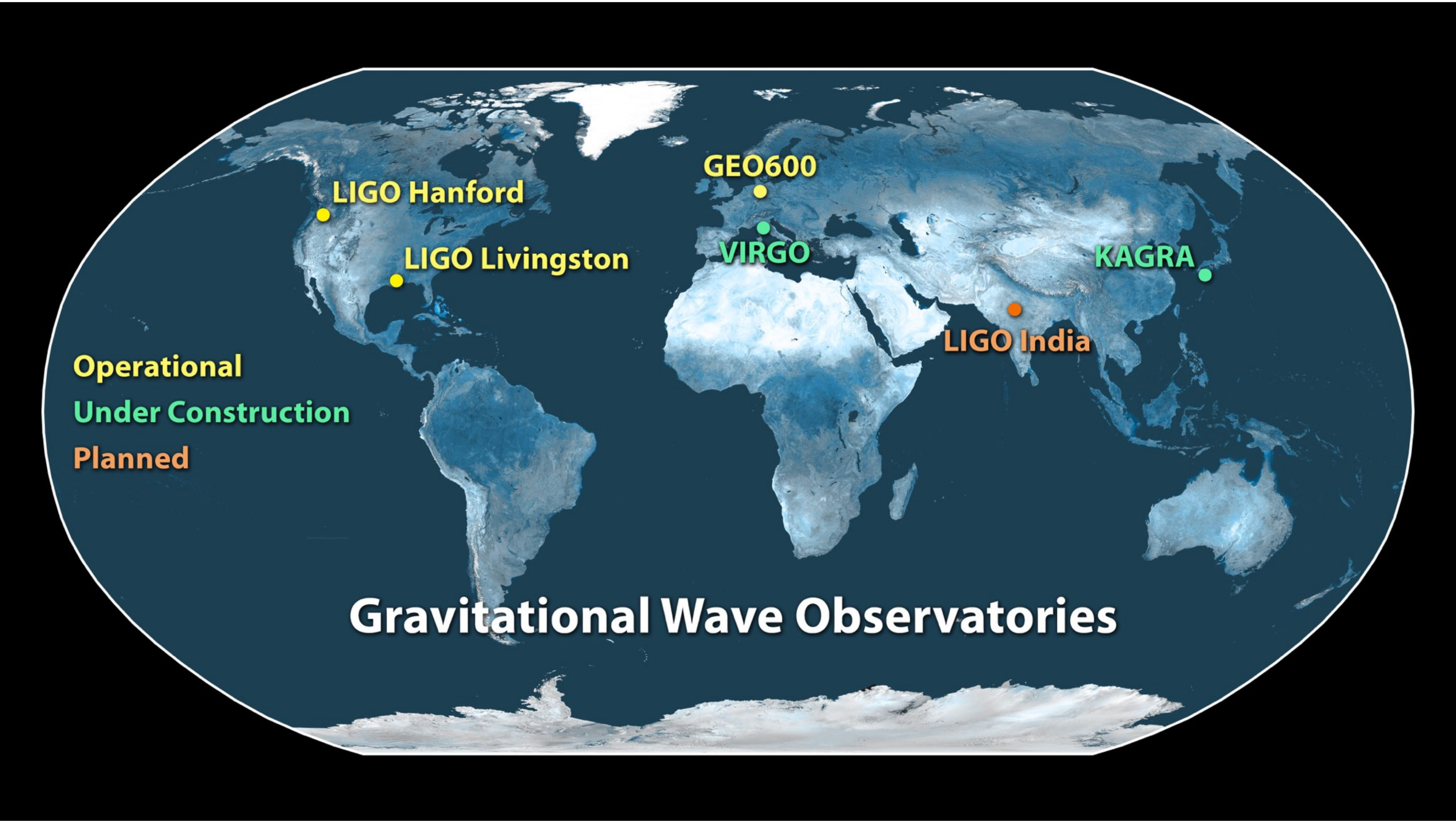
Livingston, LA

LIGO Observatories constructed from 1994-2000
Initial LIGO operated from 2002-2010

The LIGO Scientific Collaboration

~ 1000 members
~ 80 institutions
16 countries

A Global Quest

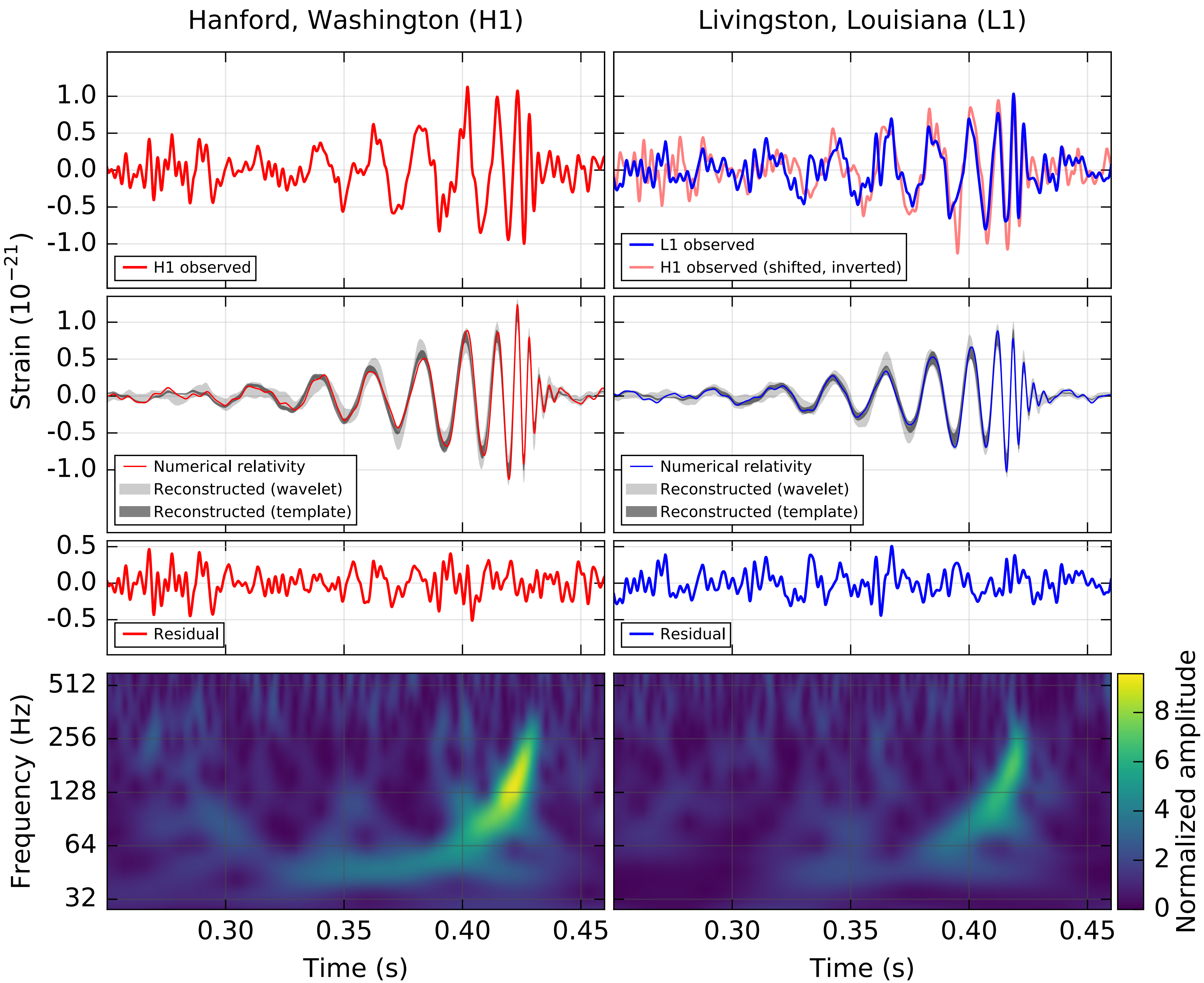


GW150914

band-passed
35-350 Hz

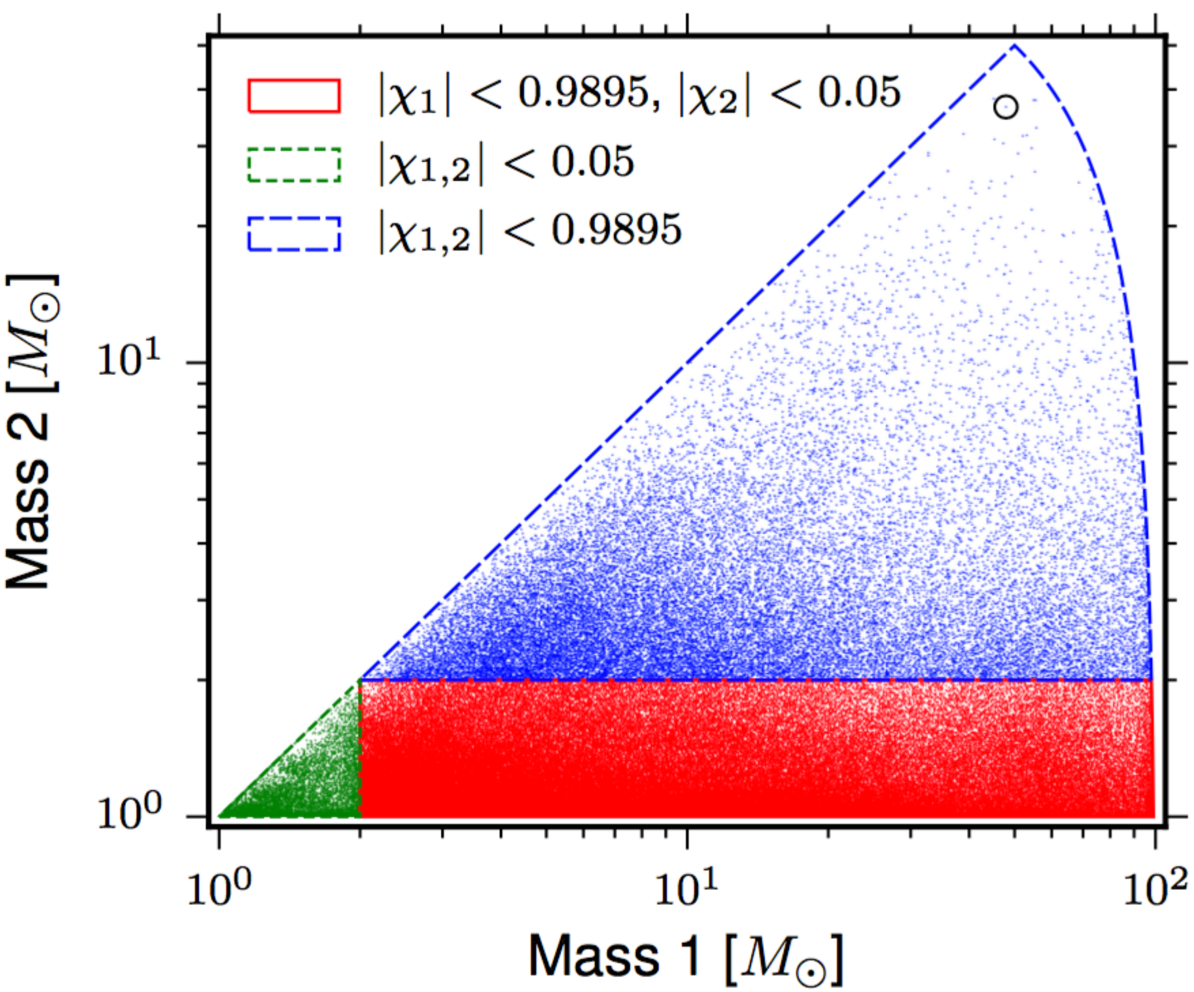
data-NR waveform

Time-Frequency map



Matched Filter Search

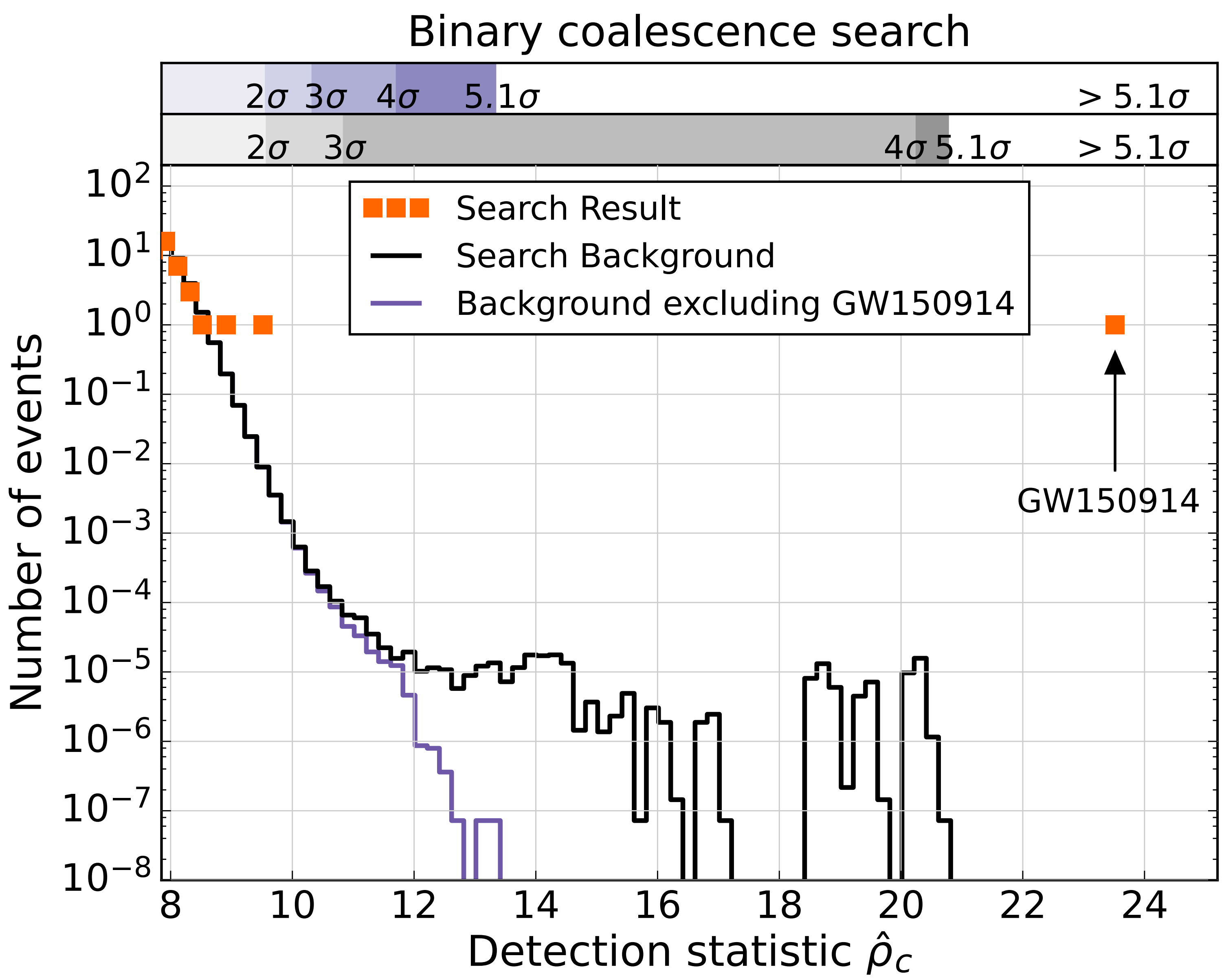
- GW signals from compact binaries are well-modeled.
- Matched filtering of data in the frequency domain used to search for these signals — optimal for modeled signals in noise.
- Since we do not know a priori the parameters of individual binaries we may detect, a bank of template waveforms is generated that spans the astrophysical signal space (mass, spin)
- We need enough templates to match the full range of signals we expect. Template banks are made "dense" enough so that <1% of signals have a matched-filter SNR loss greater than 3% — this requires ~250k templates in a bank!
- Every chunk of time-series data from the detector must be compared to every template. This is the dominant computational cost of our search.



Matched Filter Search

- An SNR threshold identifies “triggers” in each interferometer.
- Chi-square detection statistic tells us how well each trigger matches the template waveform — this is also computationally expensive — performed only on coincident triggers in >1 interferometer.
- Background estimation using “time slides" — tells us the likelihood that triggers of any given strength are to be caused by coincident noise (rather than a signal).
- Multiply ~250k templates in a bank by the number of chunks to analyze, times the number of slides in an observing run — lots of computing
- Embarrassingly parallel — can be parallelized over time and/or templates.
- Search input = primarily time series data (few TB), output = triggers (few MB)

GW150914
Detection Confidence



GW150914: Source

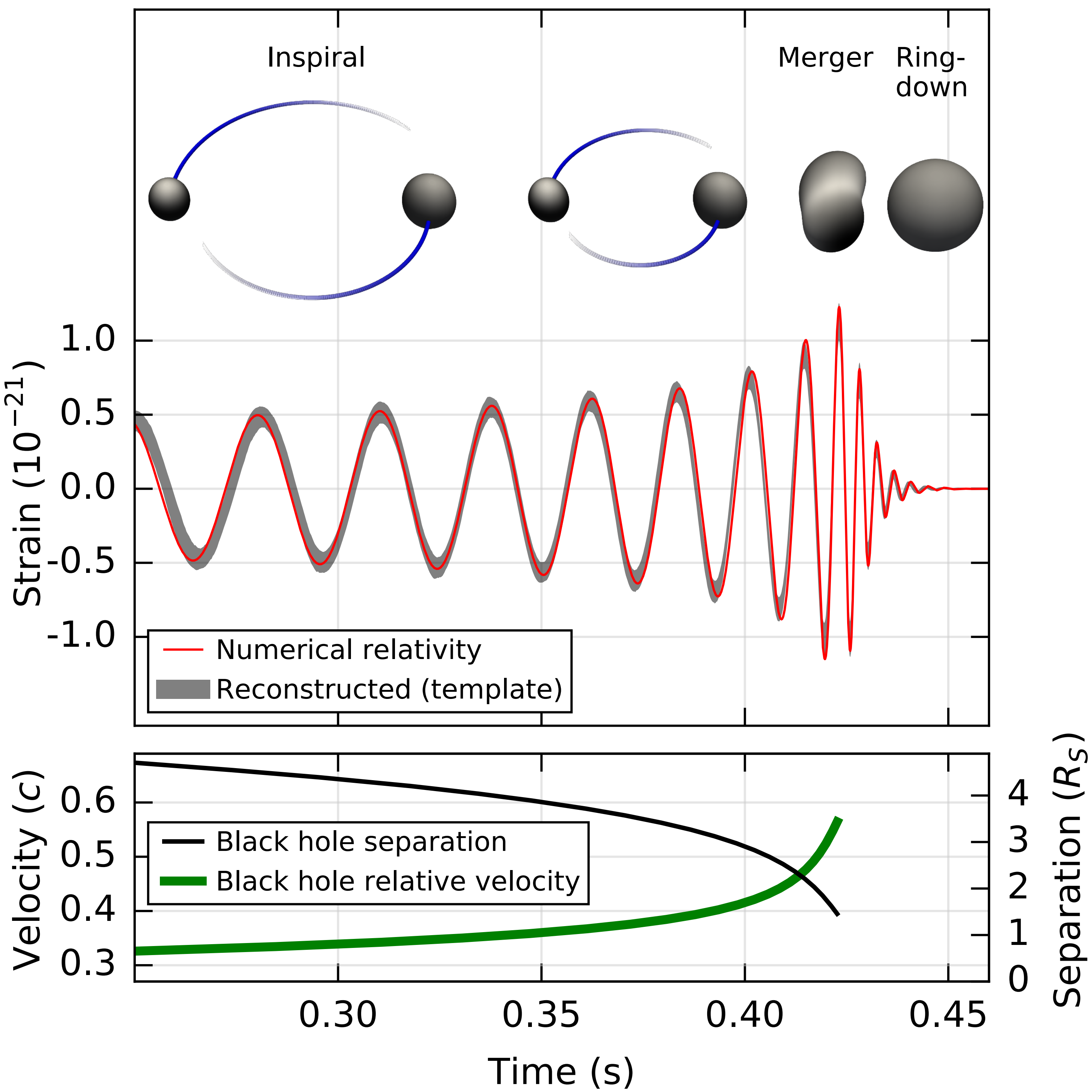
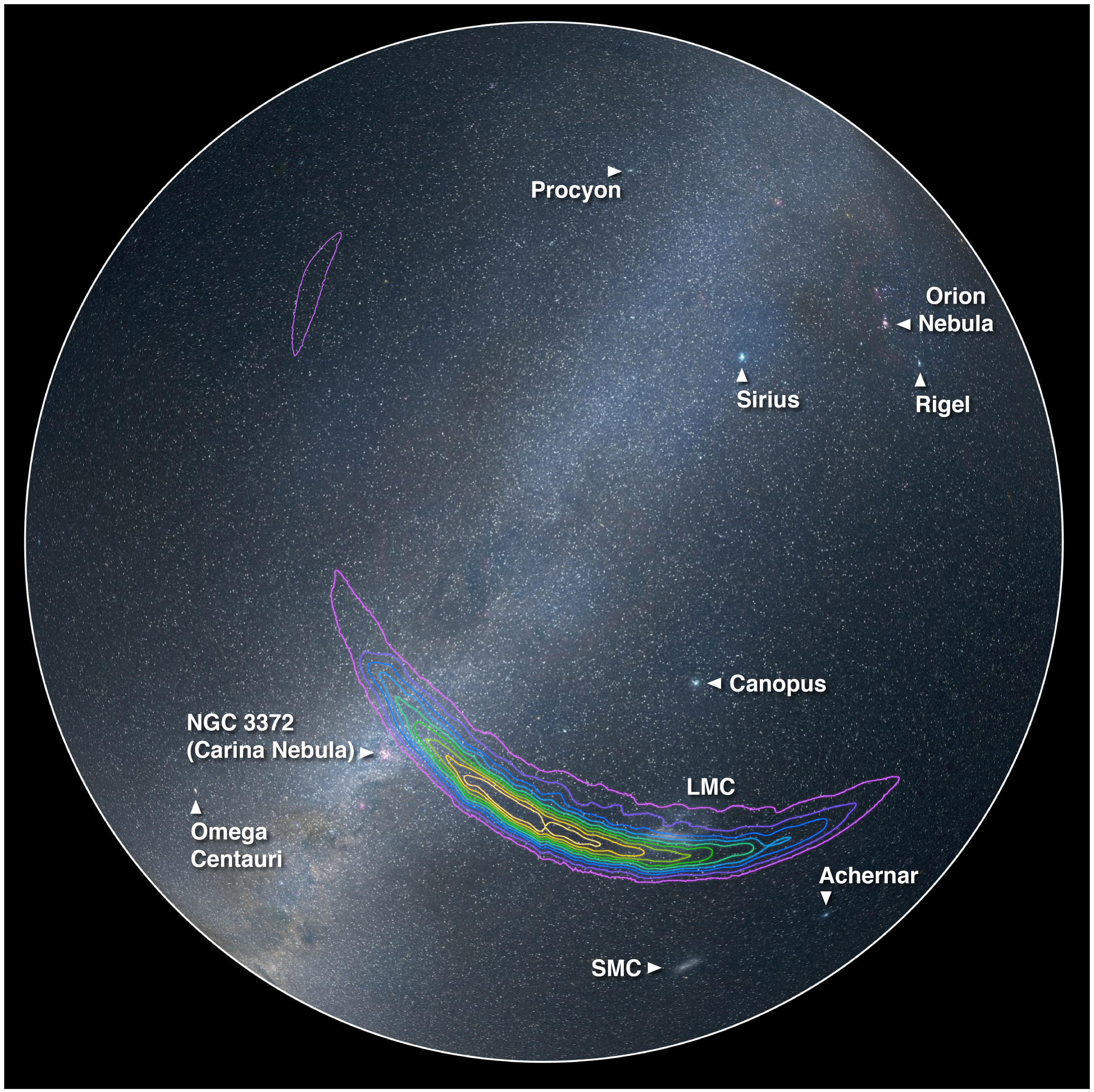


TABLE I. Source parameters for GW150914. We report median values with 90% credible intervals that include statistical errors, and systematic errors from averaging the results of different waveform models. Masses are given in the source frame, to convert to the detector frame multiply by $(1 + z)$ [87]. The source redshift assumes standard cosmology [88].

Primary black hole mass	$36^{+5}_{-4}\ M_{\odot}$
Secondary black hole mass	$29^{+4}_{-4}\ M_{\odot}$
Final black hole mass	$62^{+4}_{-4}\ M_{\odot}$
Final black hole spin	$0.67^{+0.05}_{-0.07}$
Luminosity distance	410^{+160}_{-180} Mpc
Source redshift, z	$0.09^{+0.03}_{-0.04}$

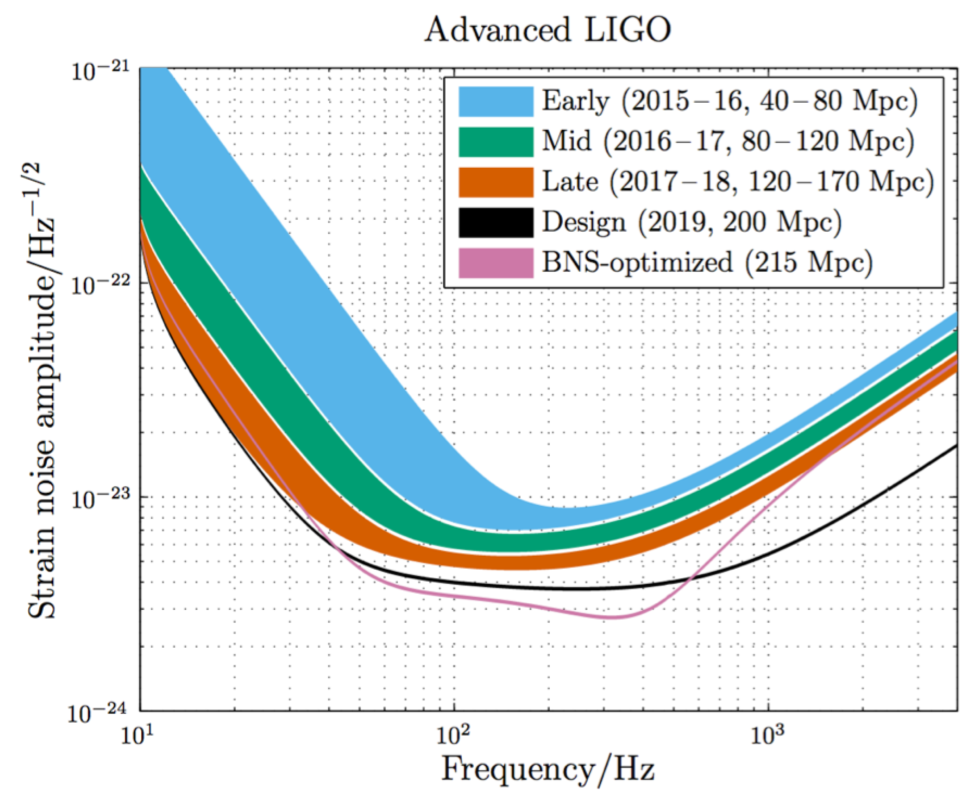
GW150914: Localization



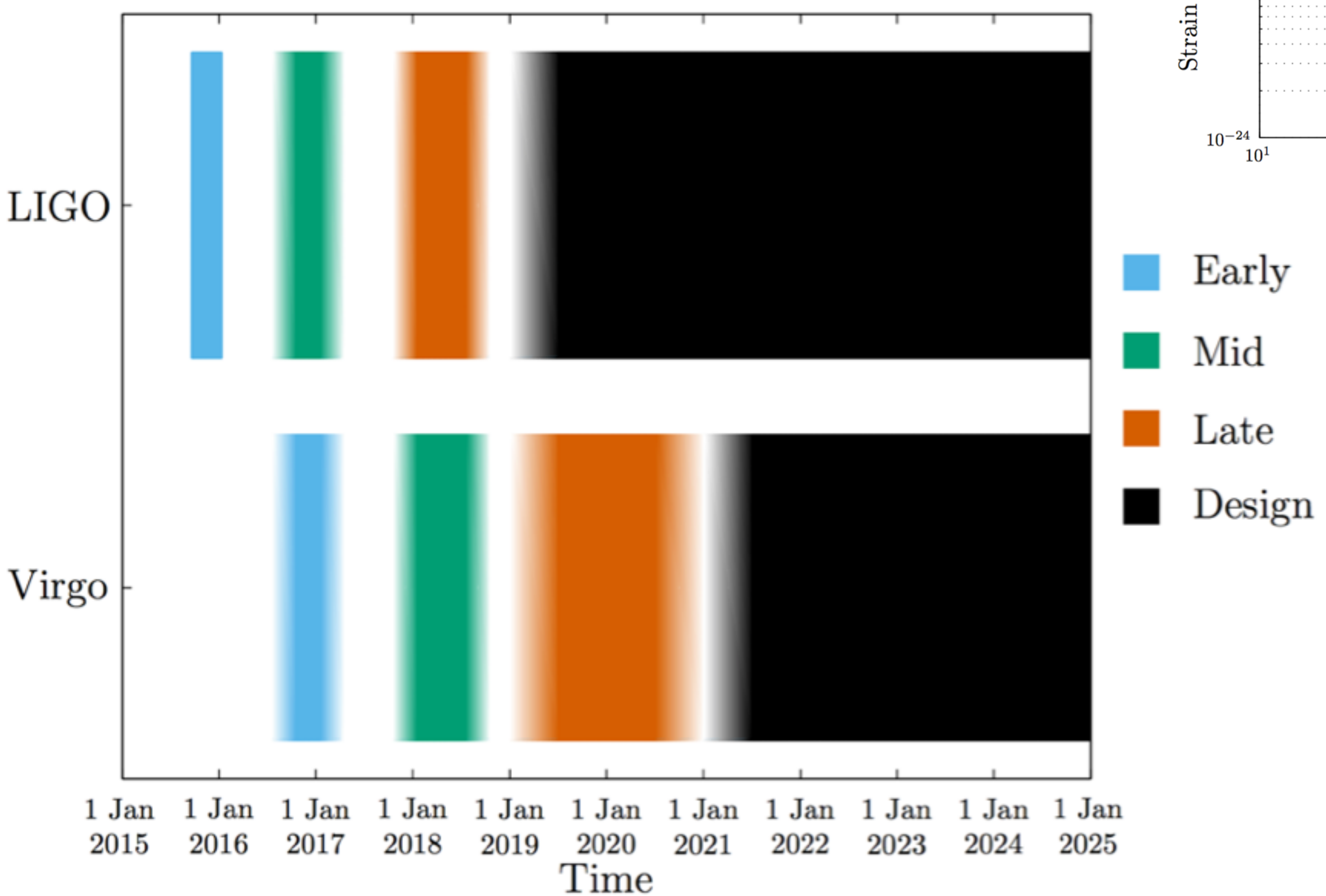
Enabling multi-messenger astronomy with gravitational waves:

- ~60 Partners from 19 countries
- ~150 instruments covering the full spectrum from radio to very high-energy gamma-rays

What is next?



Advanced LIGO
Early (2015 – 16, 40 – 80 Mpc)
Mid (2016 – 17, 80 – 120 Mpc)
Late (2017 – 18, 120 – 170 Mpc)
Design (2019, 200 Mpc)
BNS-optimized (215 Mpc)
Strain noise amplitude/Hz$^{-1/2}$
Frequency/Hz
10^{-21}
10^{-22}
10^{-23}
10^{-24}
10^{1}
10^{2}
10^{3}



LIGO
Virgo
Early
Mid
Late
Design
1 Jan 2015
1 Jan 2016
1 Jan 2017
1 Jan 2018
1 Jan 2019
1 Jan 2020
1 Jan 2021
1 Jan 2022
1 Jan 2023
1 Jan 2024
1 Jan 2025
Time

Computing

- LIGO data is analyzed by a big, distributed collaboration of scientists — the LIGO Scientific Collaboration (LSC).
- Compute-intensive science. 100's of millions of CPU core-hours per year. Almost all of it embarrassingly parallel HTC work.
- Traditionally, data analysis computing performed "in-house" on the LIGO Data Grid (LDG), which consists of dedicated HTC clusters at 7 LSC sites, including the LIGO Laboratory. About half of these LDG cycles are from the U.S., about half from Europe. Most of the U.S. cycles are funded by the NSF.
    - Plus volunteer computing with Einstein @ Home!
- With more computing resources we can do more and better science.

The Problem

- We've increasingly had opportunities to utilize additional dedicated, shared, and opportunistic HPC+HTC resources beyond the LIGO Data Grid, but have had difficulty using them:
    - XSEDE allocations
    - Individual PI clusters
    - Campus clusters (e.g., OrangeGrid @ Syracuse University)
    - HPC centers (e.g., PACE @ Georgia Tech, SciNet in Canada)
    - Opportunistic OSG resources
    - Virgo collaboration resources (CNAF, Lyon, Nikhef clusters in Europe)
    - future: cloud? (Amazon EC2, Azure, etc.)
- Difficult to integrate these into the LIGO Data Grid. LDG sites seem to require 0.5-2 FTEs of dedicated LIGO sysadmins to operate.
- Difficult to run data analysis pipelines on resources that don't look like the LIGO Data Grid.
- **Result: in the past we've either not used these resources, or have developed labor-intensive one-off solutions to utilize them individually and temporarily.**

The Idea

- Use the Open Science Grid as “universal adapter” to allow LSC data analysts to submit their search pipelines via a familiar Condor interface at a local LDG site, but have them seamlessly run on these diverse external resources.

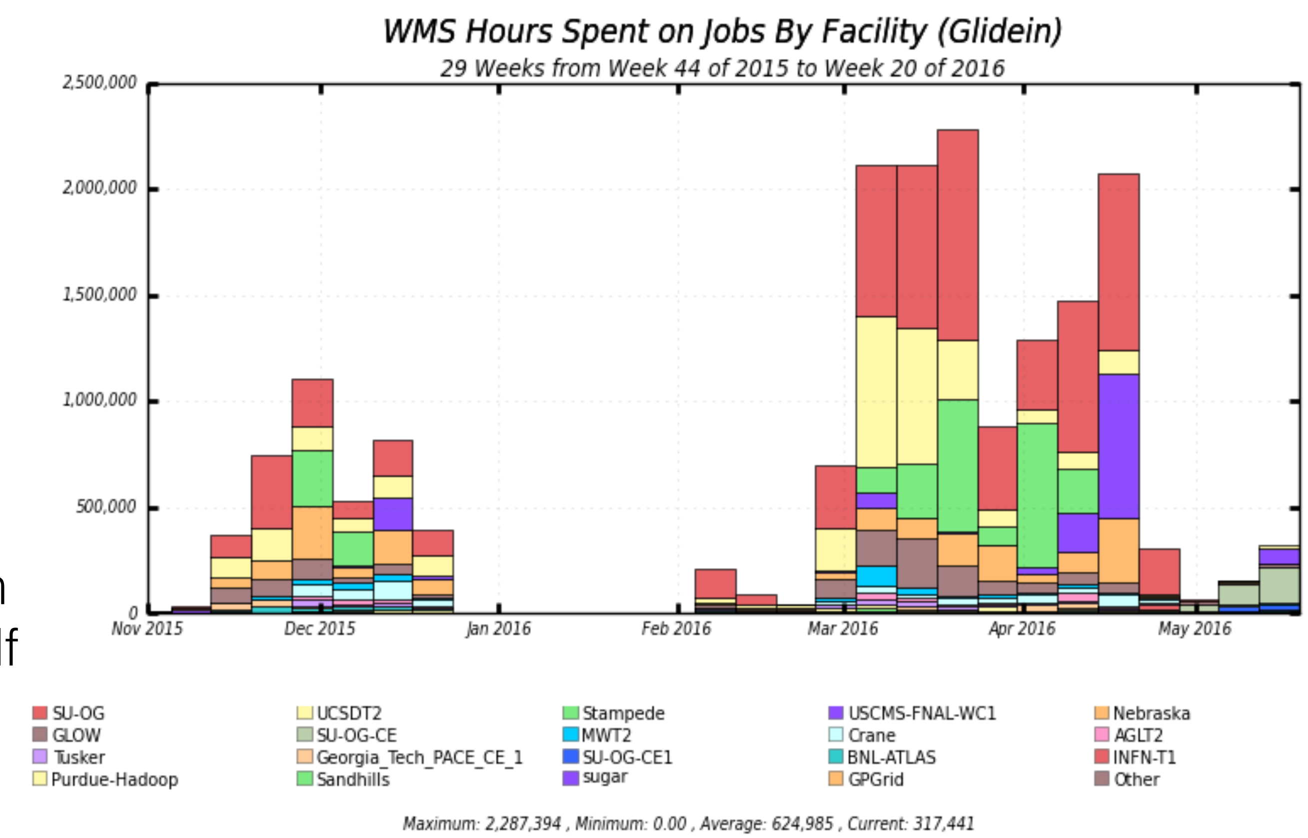
- If our LIGO data scientists can talk to OSG, and OSG can talk to everyone else, we’ve got a solution.

Last Year

- My agenda before a meeting with Brian B at Condor Week last year (found this on Tuesday while preparing this talk):
  - “1) Is my vision sane for ‘OSG as plumbing’ for connecting Condor-based LIGO submit sites to campus compute clusters (and then later maybe even for LIGO-managed compute clusters)?”
- Answer: YES

LIGO Use of OSG

- LIGO's OSG computing contributed directly to the results in the detection paper.
- LIGO analyses have run across 17 different OSG resources.
- **>17 million CPU core-hours from OSG.**
- Approx. 1/3 of the total core-hours were from Syracuse, a site affiliated with LIGO (with half of that allocated in advance on dedicated cores, and half opportunistic on both dedicated and opportunistic cores)
- Approx 1/3 of the total core-hours were from XSEDE (with half of that allocated @ Stampede, half opportunistic @ COMET).
- Approx. 1/3 of total core-hours were from OSG sites unaffiliated with LIGO (all opportunistic).



- ~5TB of input data stored at the Holland Computing Center (HCC) at the University of Nebraska-Lincoln.
- The total data volume distributed to jobs from Nebraska >1PB.
- Data rates from Nebraska storage to worker nodes ~10Gbps sustained. (Recently demonstrated >30Gpbs by accident!)

Today

- OSG was 2nd largest contributor to O1 computing after the Albert Einstein Institute in Hannover, Germany, more than any U.S. LDG site.

- O1 did not yet include dedicated cycles from Georgia Tech, SciNet, or Virgo resources — O2 may include all three via OSG.

- The catch: OSG requires some adaptation by scientists to utilize, vs. just running on the LDG.

- LIGO's most computationally intensive search pipeline (PyCBC) is the only production pipeline running on OSG now — we hope LIGO's second most intensive pipeline (cWB) will be able to use OSG in O2, as well as two other major pipelines (lalinference and BayesWave).

- Data analyst training is an issue! OSG Summer Schools FTW.

What in it for LIGO?

- Access to considerable, previously un-utilized or underutilized resources beyond the LDG.
- Elasticity at periods of high demand.
- If a quick phone call can bring new resources to bear, we want to be able to actually use them.
- Seamless access to cloud resources if/when we should want them.
  - This is still controversial — not clear it's cheaper or a good idea — but the plumbing is there if we decide to try.

What’s in it for LSC data analysts?

- **Cycles** — access to additional computing resources beyond the LDG.
- **Cluster agnosticism** — submit in one place, run wherever there are cycles.
- **Decentralized dependency management** — users can build and deploy their own external dependencies in CVMFS without coordinating with or breaking other pipelines (vs. OS package installs in system paths).
- **You can quickly bring new resources online** — easier to bring your own local (or others' friendly) resources to bear on your searches.
- **Expanding pie** — Those who don’t use OSG benefit as other LSC computing moves onto OSG, and off of the LDG clusters they need.

What's the Catch?

- OSG provides a flexible but much less rich and curated environment than LDG today (this could change, but these are early days)
- Developers have to manage their own dependencies in CVMFS rather than relying on LDG computing staff to preinstall them in OS paths.
- Accounting is still crude.
- Prioritization is harder — can't just ask your favorite LDG site's PI to kick everyone else off so your jobs run.
- Debugging pipeline failures and issues can be trickier — more middleware, more resource providers.

What's yet to be done?

- LDG management of common pipeline dependencies in CVMFS.
  - Complements rather than excludes decentralized search group, pipeline, or user-managed software!
- Usage accounting by search tag, LIGO.ORG username, etc.
- Remove CVMFS deployment bottlenecks by delegating per-search or per-user deployment capabilities.
  - We could solve this ourselves by hosting our own CVMFS repo and building our own solution, but I'd rather someone solve it for us. Are we unique?
- Making more LDG resources available via OSG.
  - Einstein @ Home (BOINC)?

Why OSG Is Awesome

- OSG really is a “universal adapter” to diverse resource types for LIGO: dedicated LIGO CPUs, “friendly” campus clusters, “friendly” PI clusters, opportunistic OSG CPUs, XSEDE allocations, (and in the future, cloud CPUs?)
- Outsourced plumbing (factories, CEs, etc.) + expert help = easy to get started without making a huge labor investment.
- Track record of success — HEP forged a path.
- Friendly, enthusiastic, skilled, **results-oriented**, **flexible** OSG staff. Not hung up on boundaries, processes — focused on science goals.

Challenges Using OSG

- Search pipelines must be “ported” from LIGO Data Grid environment to OSG environment
    - assume lowest-common-denominator OS install, understand external dependencies, build and deploy them into CVMFS.
    - LIGO input data is not local.
    - New checkpointing challenges (Condor stduniv -> application checkpointing)
- Complicated Accounting: currently manual aggregation of two sources of data, with different units and metadata (CPU core hours vs SUs, users, pipelines, etc.)
- May reduce systems administration burden in some ways, but may also “shift” it from sysadmins to “grid admins” — how much can OSG help?
- Complicates our computing model — funding implications — need to be clear that elasticity does not provide the same benefits as in-house computing to meet baseline demand and deliver low-latency computing.

Thank you!

Extras

What Gravitational Waves Can Tell Us About the Universe (A Partial List)

Gravitational waves are an entirely new way to probe the nature of the universe!

Physics

- Is General Relativity the correct theory of gravity?
- How does matter behave under extreme gravity?
- Are black holes really the black holes of General Relativity?

Astrophysics & Astronomy

- What powers short gamma ray bursts, the brightest events in the universe?
- How do stars explode?
- How many stellar mass black holes are there in the universe?
- Do intermediate mass black holes exist? How many are there in the universe?

Cosmology

- Can we detect the residue of the Big Bang?

Black Hole Merger and Ringdown

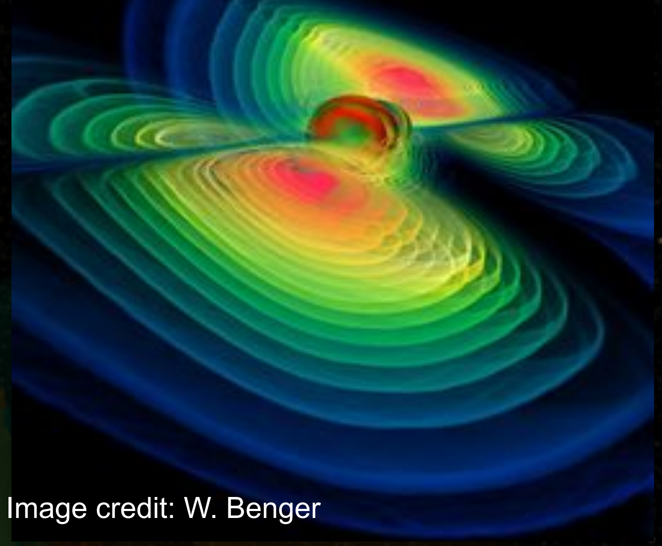


Image credit: W. Benger

Image credit: NASA

Supernovae

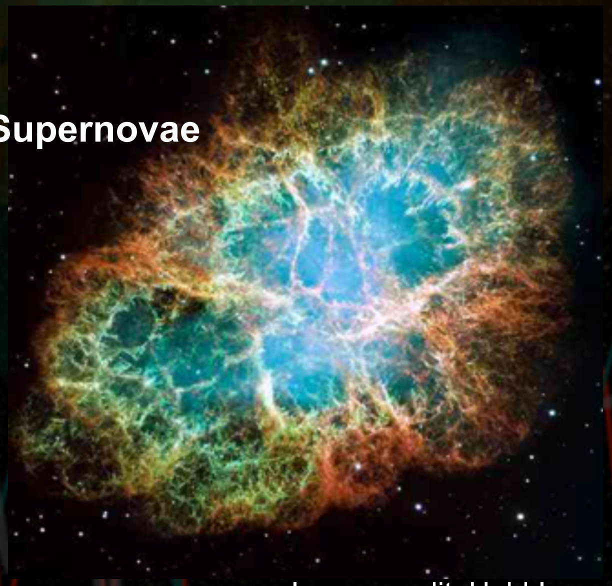


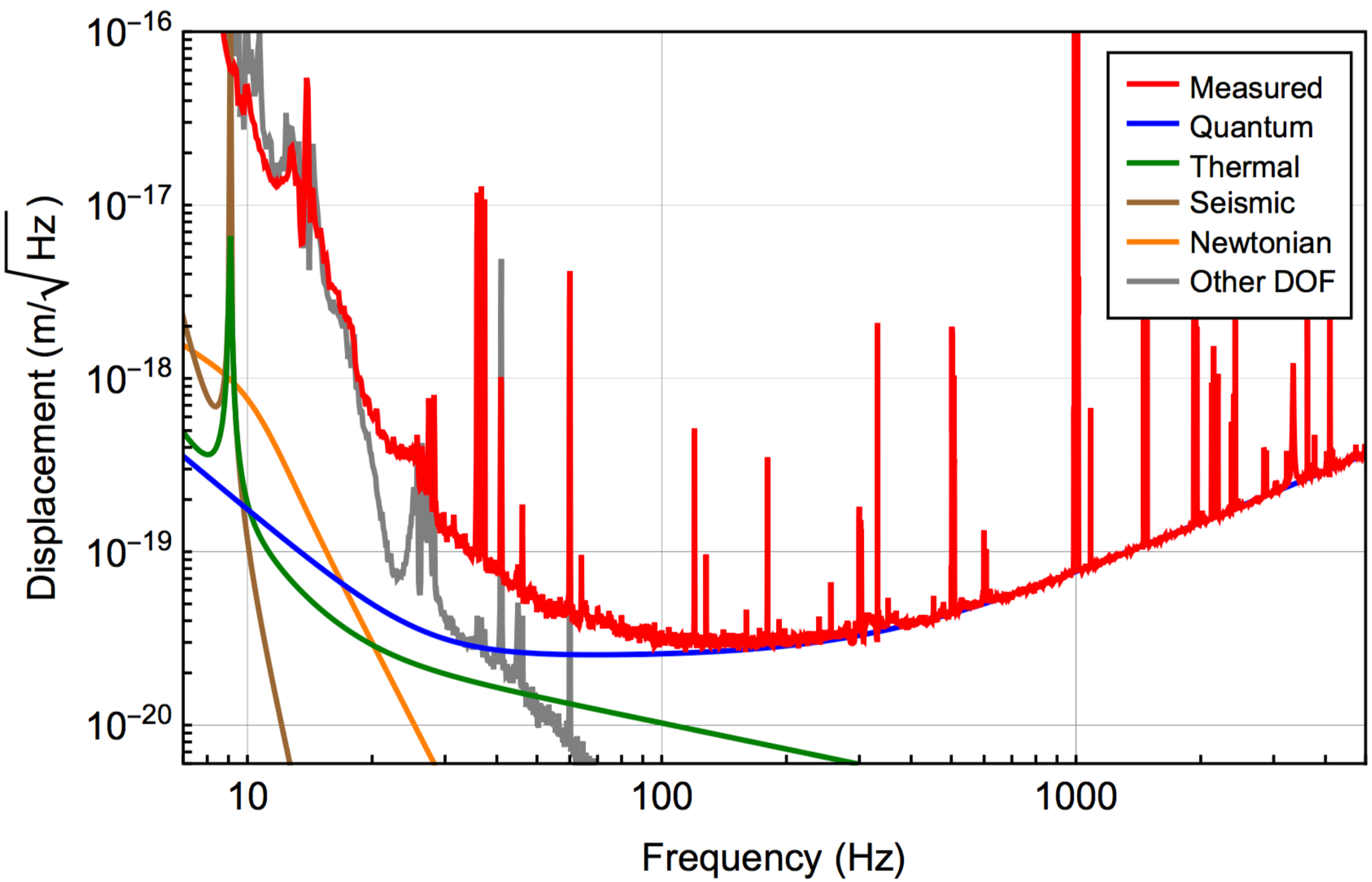
Image credit: Hubble

LIGO:

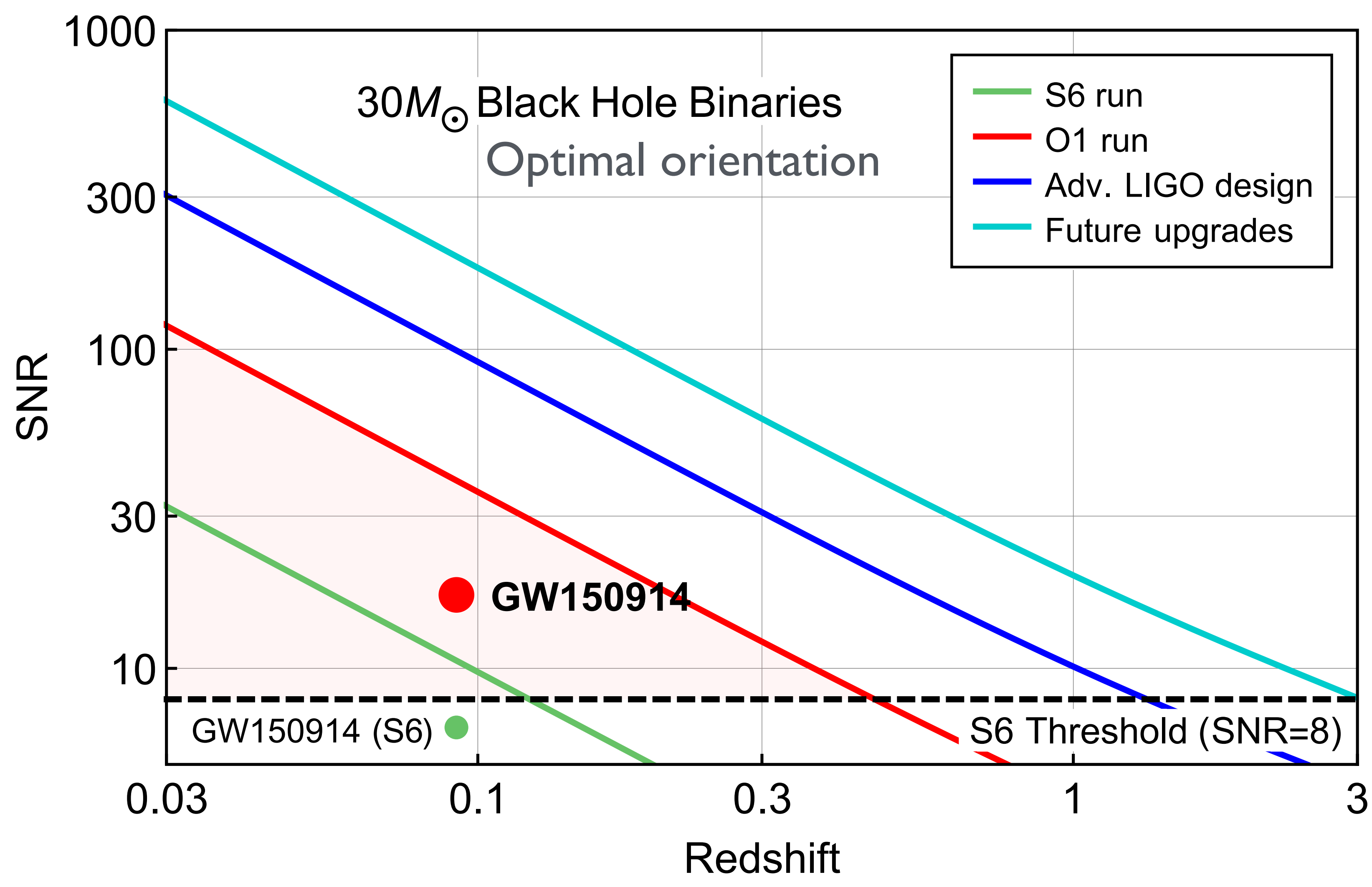
Laser Interferometer Gravitational-wave Observatory

Advanced LIGO First Science Run

(Sept 12, 2015 - Jan 12, 2016)



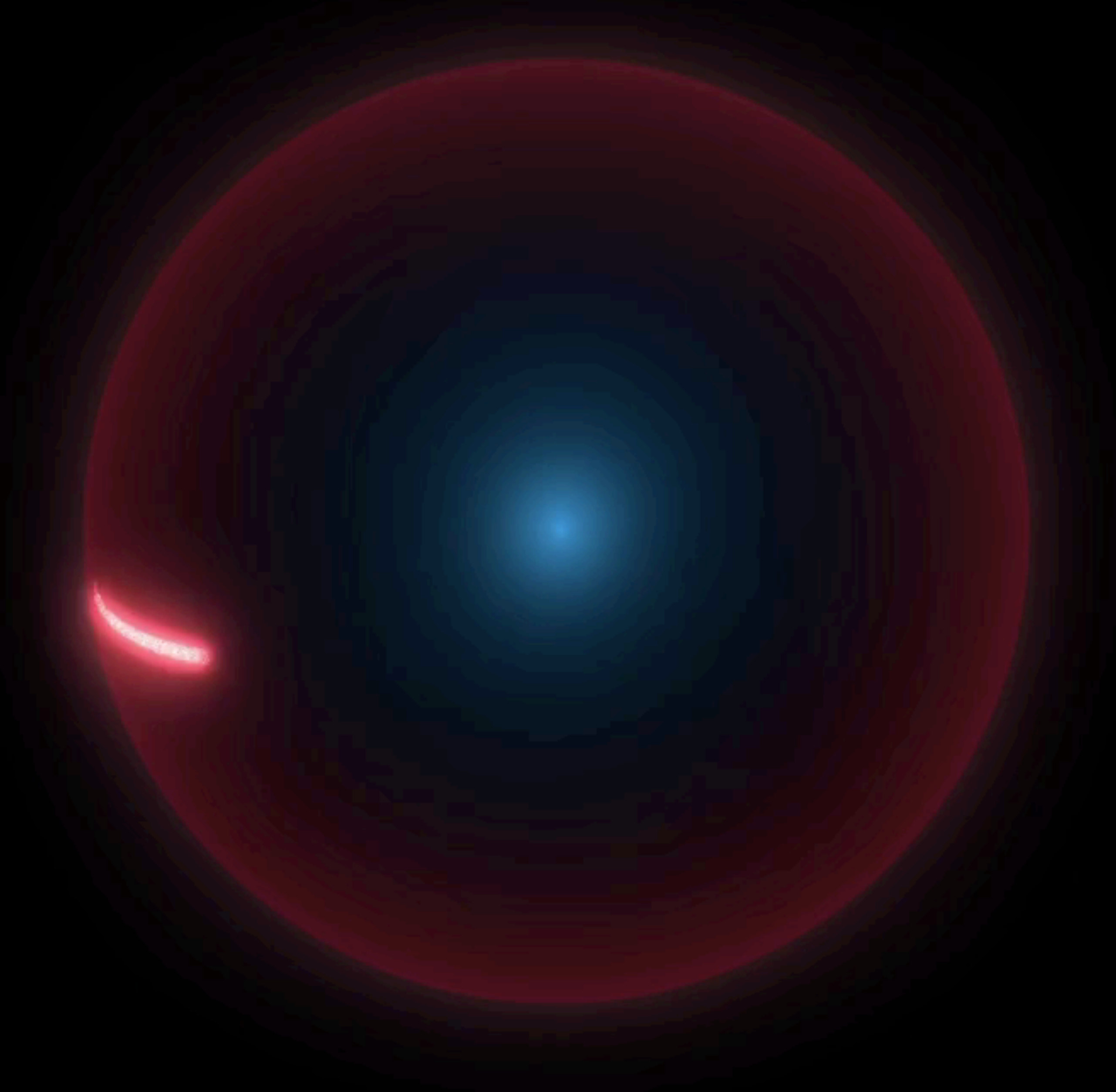
Sensitivity Improvement



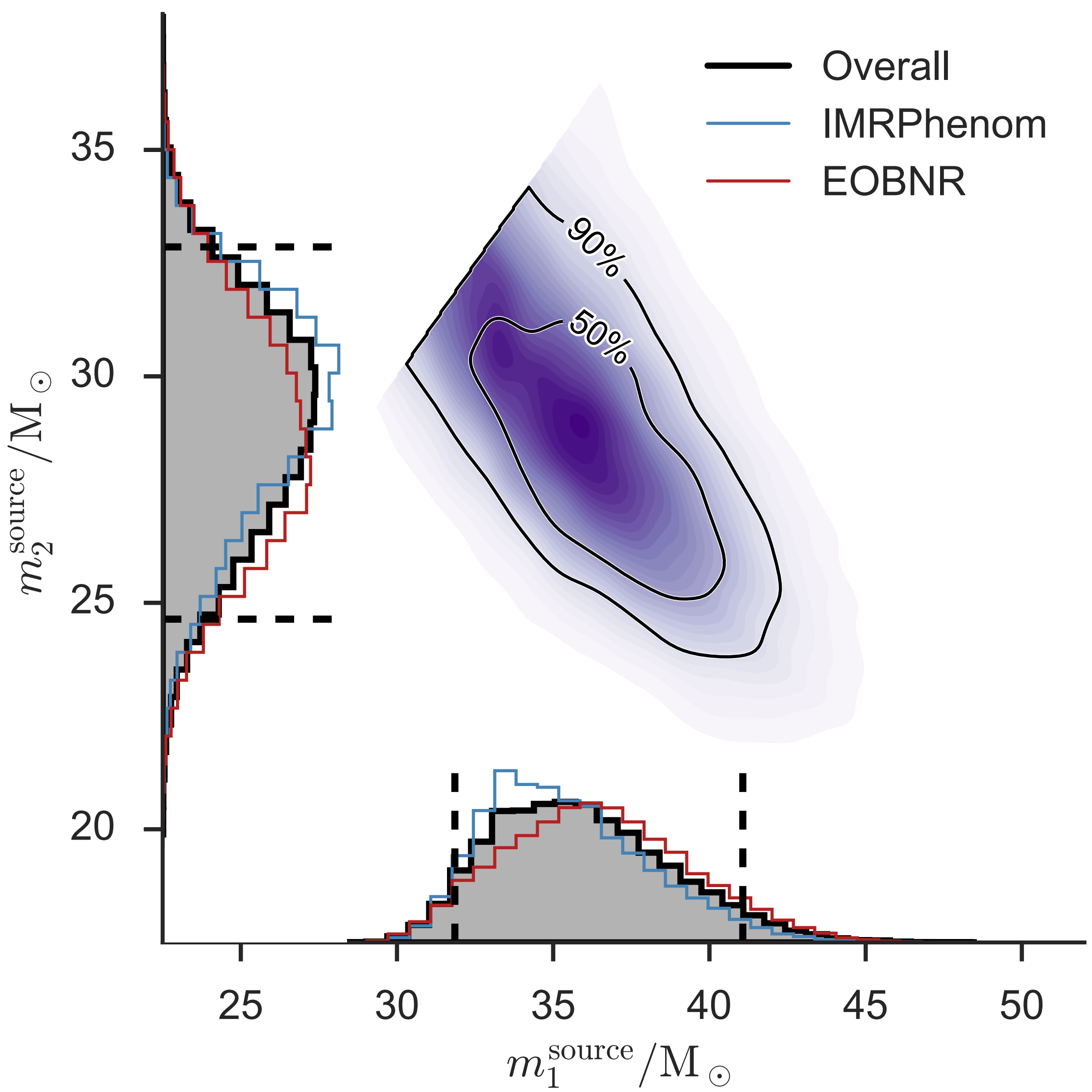
LIGO:

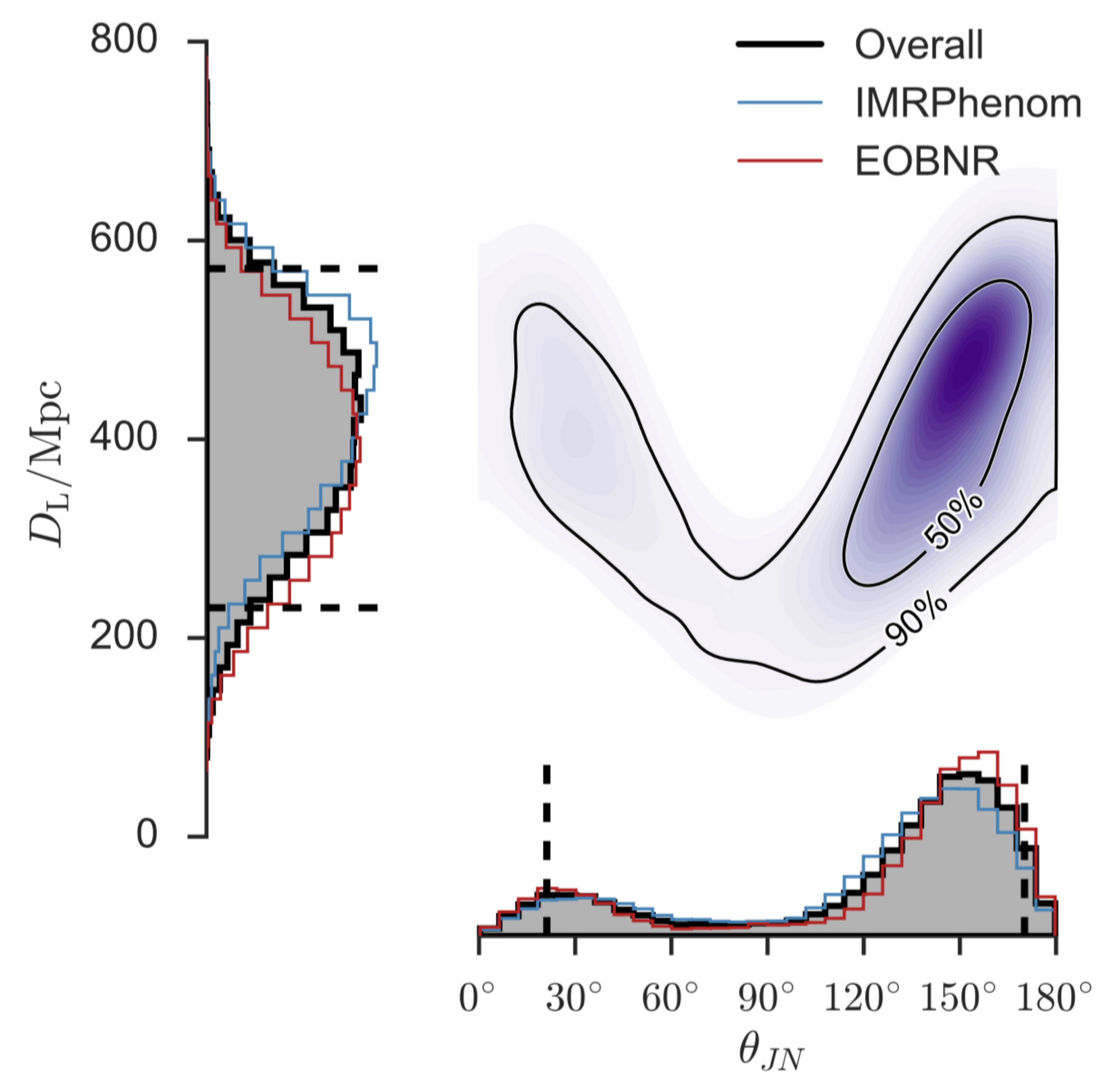
Laser Interferometer Gravitational-wave Observatory

What is the scale?
Zooming into an Hydrogen atom…



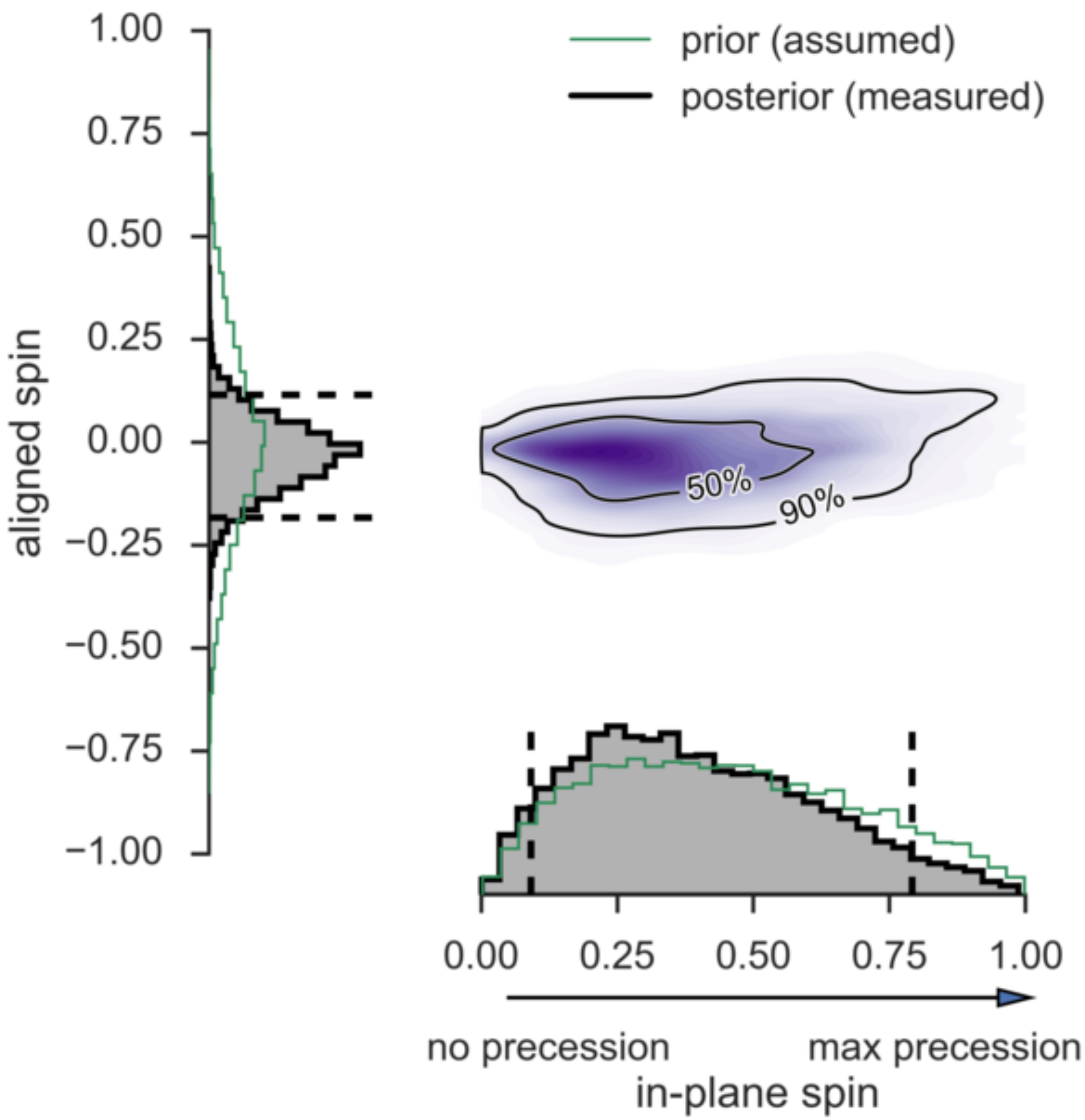
GW150914: Parameters





IMRPhenom = analytical + numerical hybrid waveforms, aligned spin (no precession)
EOBNR = effective one body approximation, calibrated with numerical relativity, aligned spin (no precession)

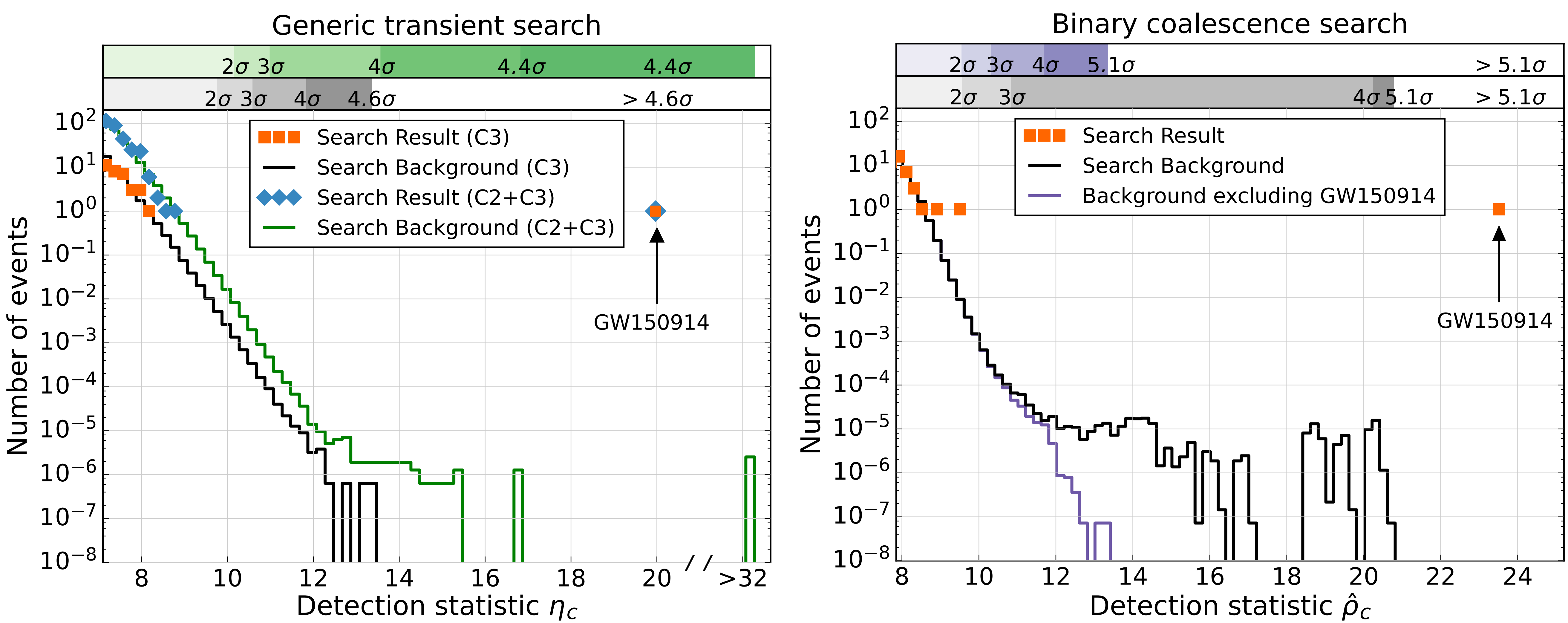
GW150914: Parameters



spins aligned
with orbital angular momentum
are constrained
to be small

in-plane spins
are unconstrained

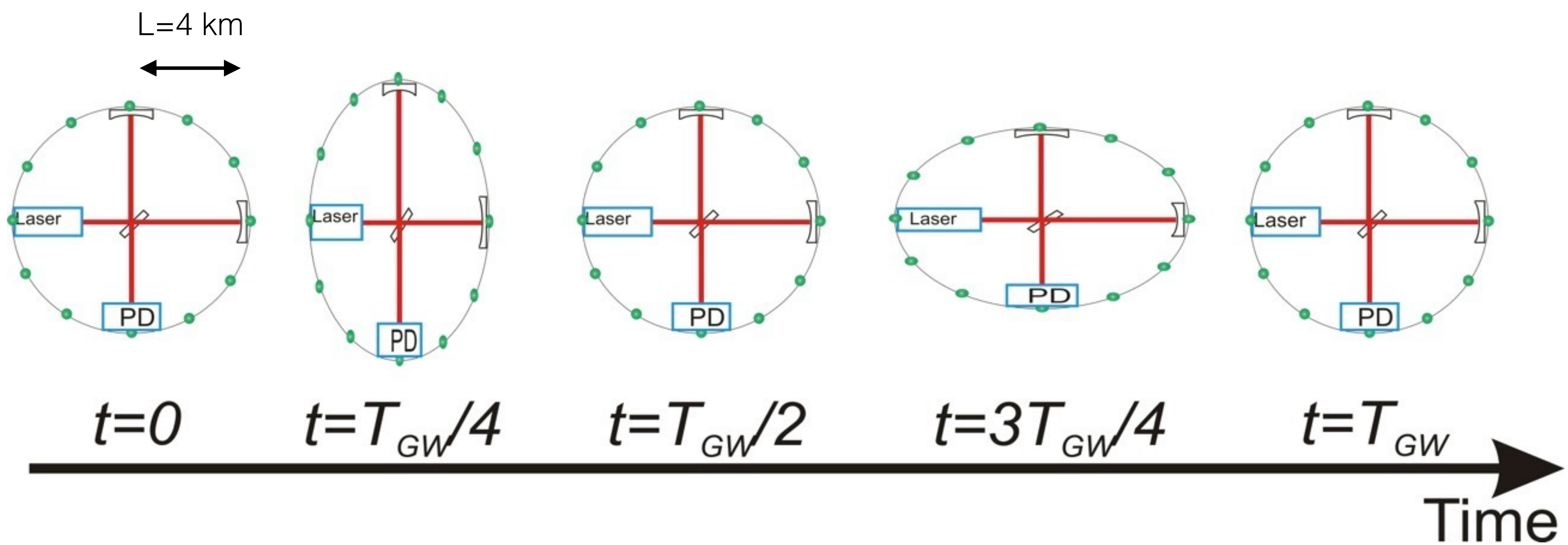
GW150914
Detection Confidence

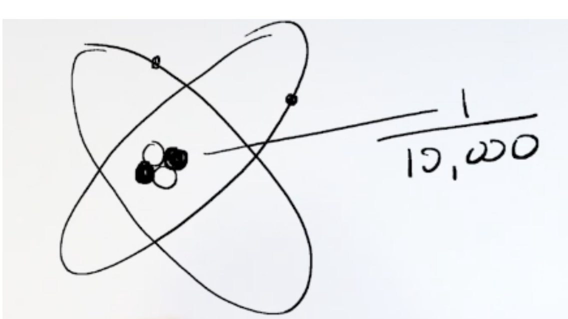


C3: frequency increases with time
C2: all other unknown morphology

How to Detect Gravitational Waves

Physically, gravitational waves are strains $\frac{\Delta L}{L}$

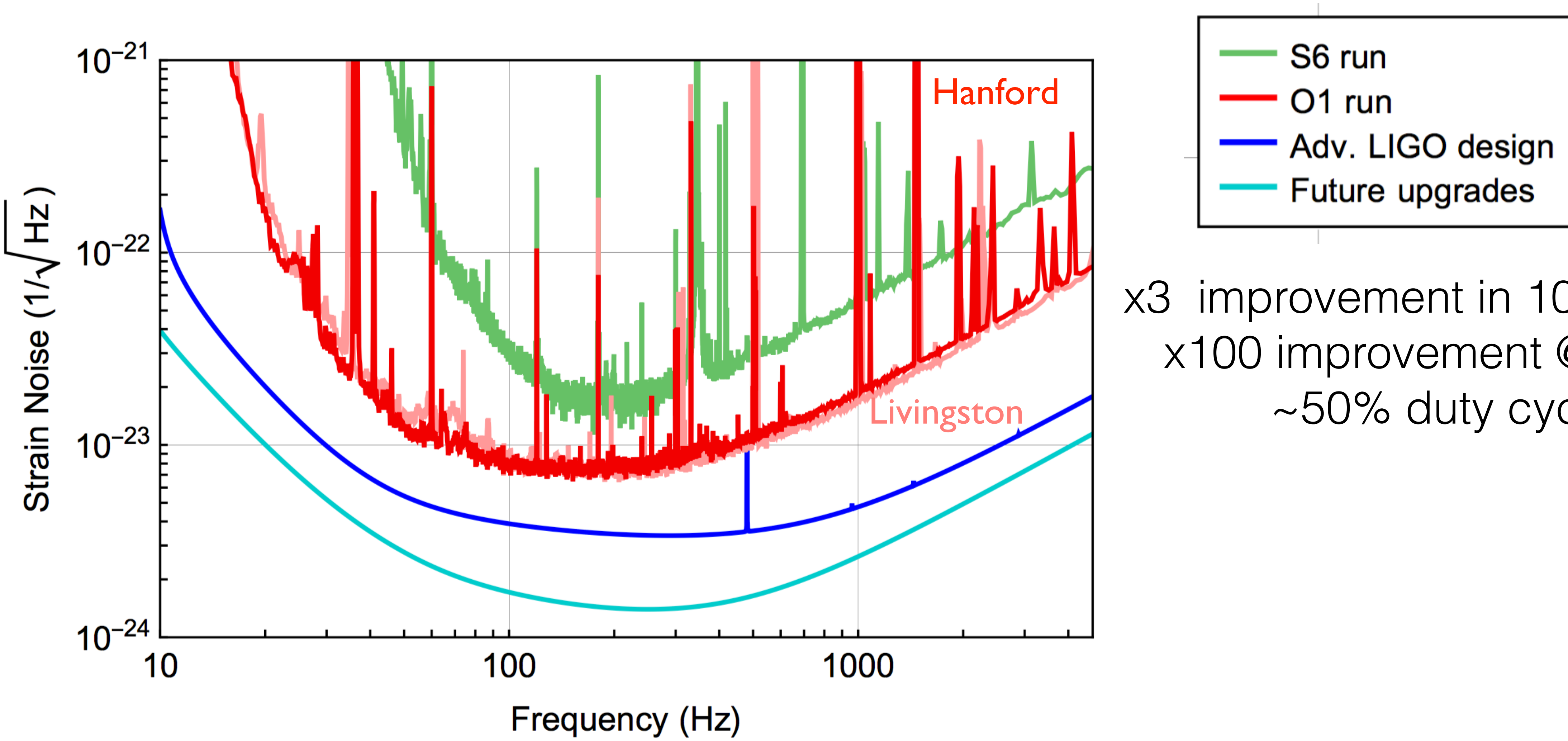




ΔL

Advanced LIGO First Science Run

(Sept 12, 2015 - Jan 12, 2016)



x3 improvement in 100-300 Hz

x100 improvement @ 50 Hz

~50% duty cycle

Discovery

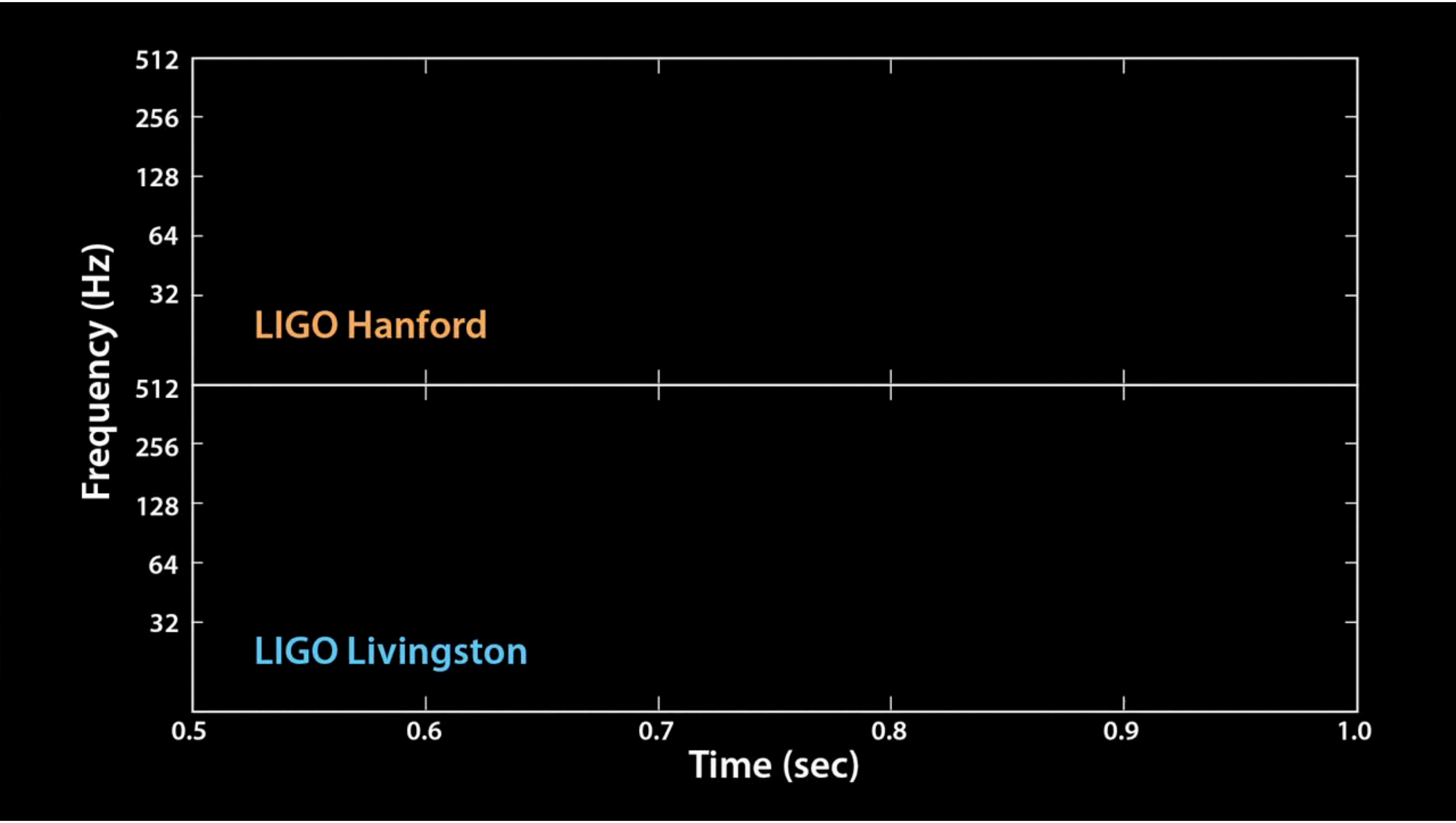
Observation of Gravitational Waves from a Binary Black Hole Merger

The LIGO Scientific Collaboration and The Virgo Collaboration

On September 14, 2015 at 09:50:45 UTC the two detectors of the Laser Interferometer Gravitational-wave Observatory (LIGO) simultaneously observed a transient gravitational-wave signal. The signal sweeps upwards in frequency from 35 Hz to 250 Hz with a peak gravitational-wave strain of 1.0×10^{-21}. It matches the waveform predicted by general relativity for the inspiral and merger of a pair of black holes and the ringdown of the resulting single black hole. The signal was observed with a matched filter signal-to-noise ratio of 24 and a false alarm rate estimated to be less than 1 event per 203 000 years, equivalent to a significance greater than $5.1\,\sigma$. The source lies at a luminosity distance of 410^{+160}_{-180} Mpc corresponding to a redshift $z = 0.09^{+0.03}_{-0.04}$. In the source frame, the initial black hole masses are $36^{+5}_{-4}\,\mathrm{M}_\odot$ and $29^{+4}_{-4}\,\mathrm{M}_\odot$, and the final black hole mass is $62^{+4}_{-4}\,\mathrm{M}_\odot$, with $3.0^{+0.5}_{-0.5}\,\mathrm{M}_\odot c^2$ radiated in gravitational waves. All uncertainties define 90% credible intervals. These observations demonstrate the existence of binary stellar-mass black hole systems. This is the first direct detection of gravitational waves and the first observation of a binary black hole merger.

PACS numbers: 04.80.Nn, 04.25.dg, 95.85.Sz, 97.80.-d

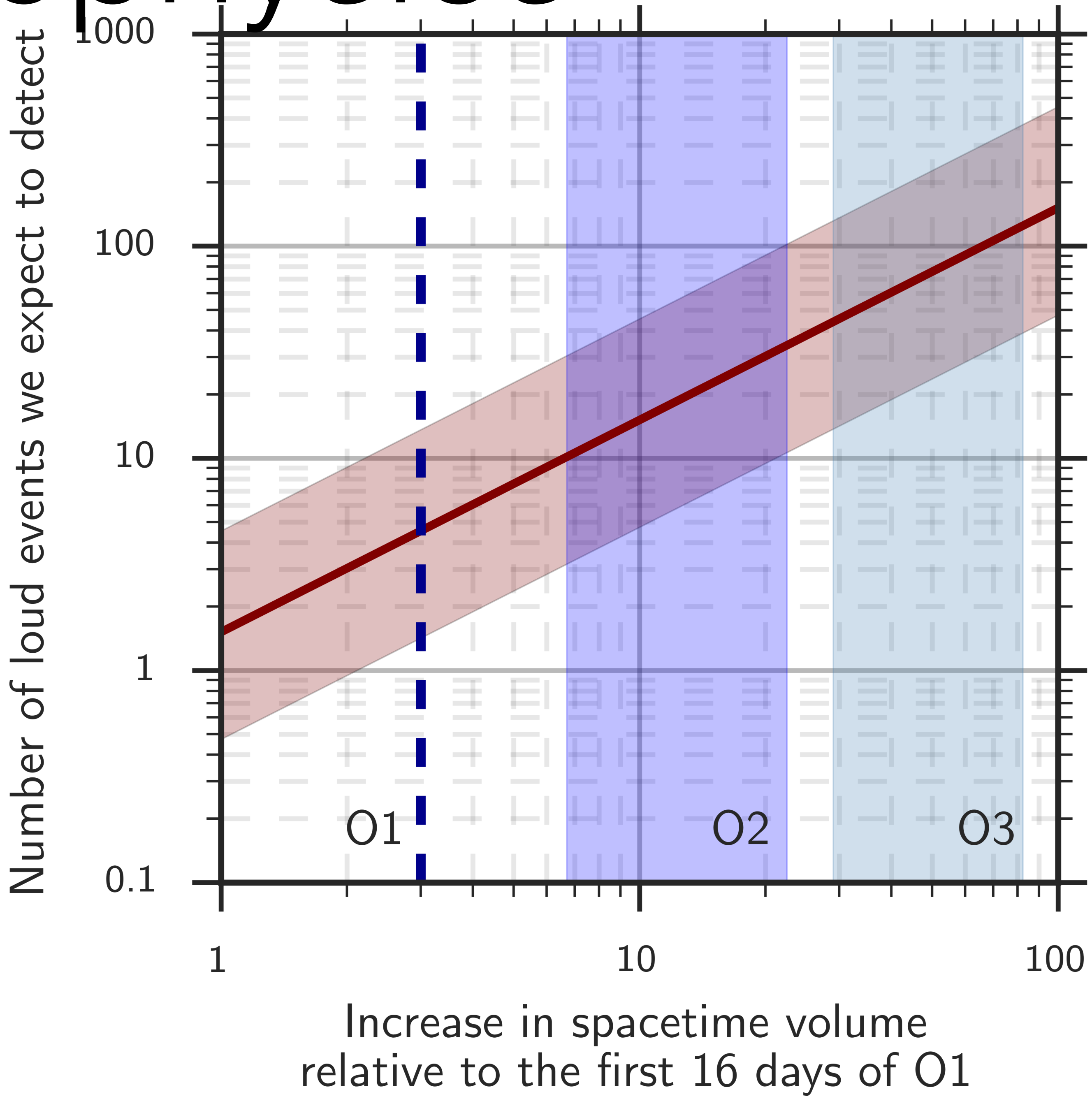
GW150914


512
256
128
64
32
LIGO Hanford
512
256
128
64
32
LIGO Livingston
Frequency (Hz)
0.5
0.6
0.7
0.8
0.9
1.0
Time (sec)

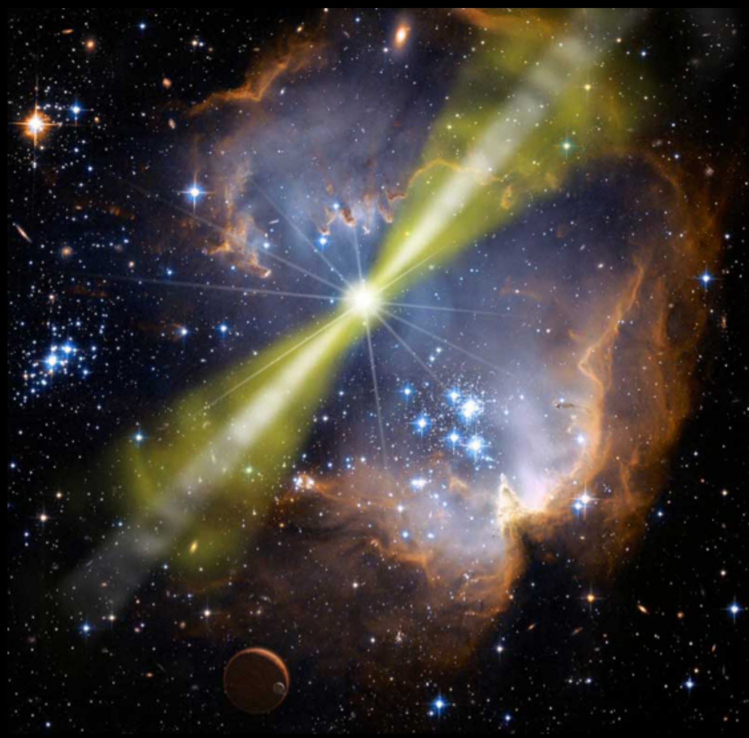
Gravitational-Wave Astrophysics

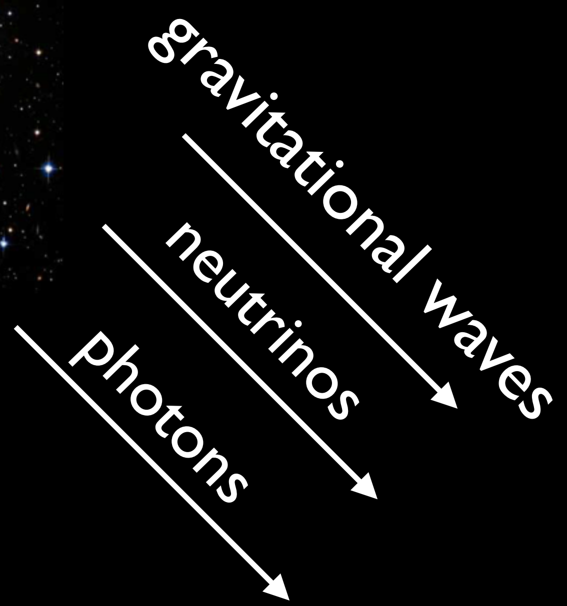
Rate from GW150914:
2-400/Gpc^3/yr

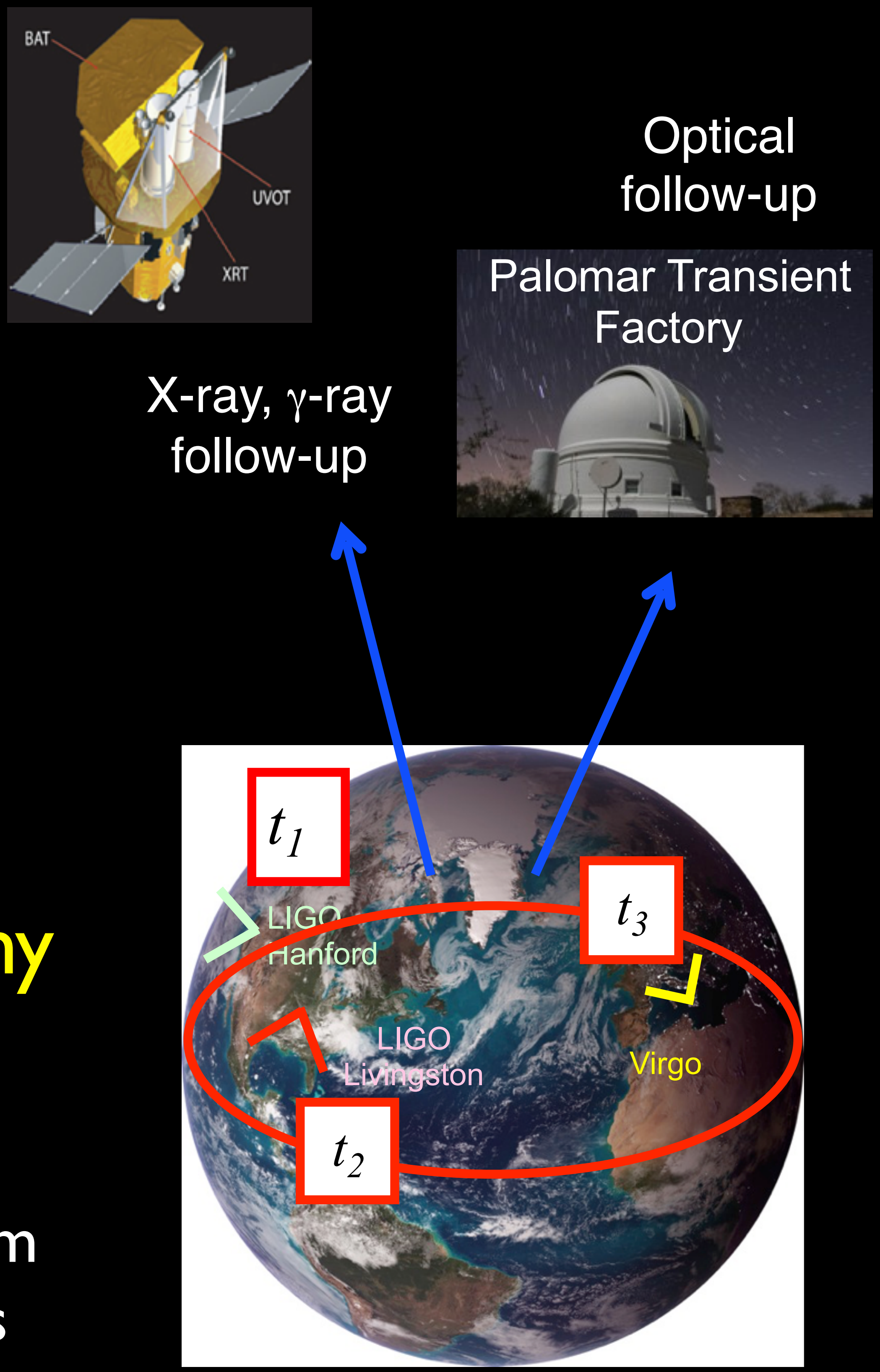
More BBH detections to come …
Reveal underlying BBH mass distribution
Quantitative model constraints
More reliable constraints



GW150914 analysis: 0.1 Gpc^3 yr (16 days calendar time)







Enabling multi-messenger astronomy with gravitational waves:

- ~60 Partners from 19 countries
- ~150 instruments covering the full spectrum from radio to very high-energy gamma-rays

The Journey of a Gravitational Wave