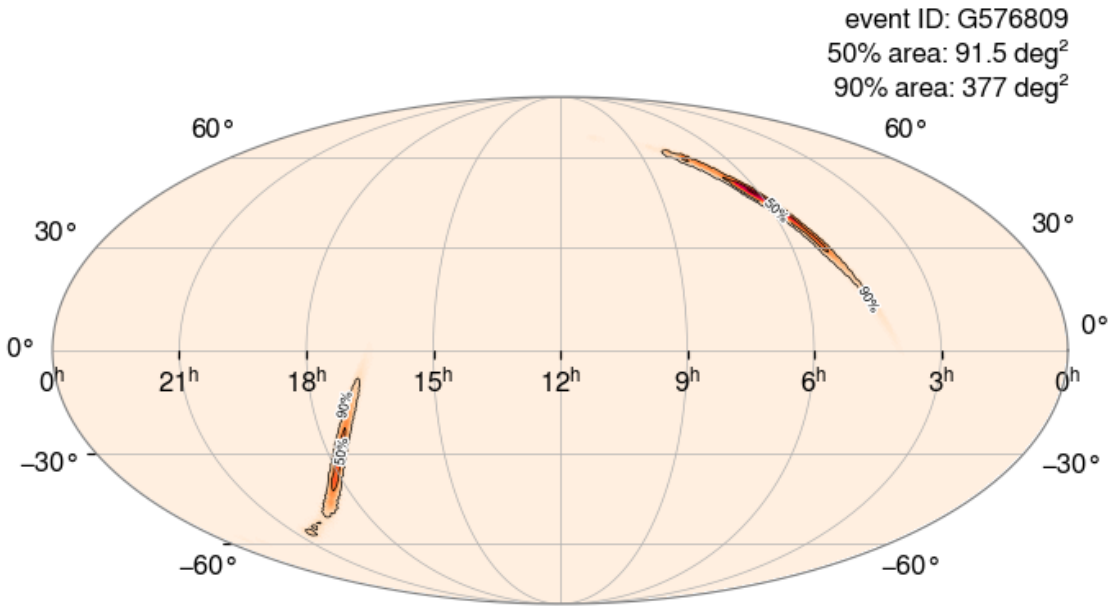
Gravitational-wave data policies: an LSC perspective

Stephen Fairhurst

Event alerts (https://gracedb.ligo.org/)

Superevent Information	
Superevent ID	S250628am
Category	Production
FAR (Hz)	1.683e-14
FAR (yr^{-1})	1 per 1.8824e+06 years
t_0 (GPS time)	1435170217.78
t_{end} (GPS time)	1435170218.82
Submitted ▾	2025-06-28 18:23:34 UTC
Links	Data

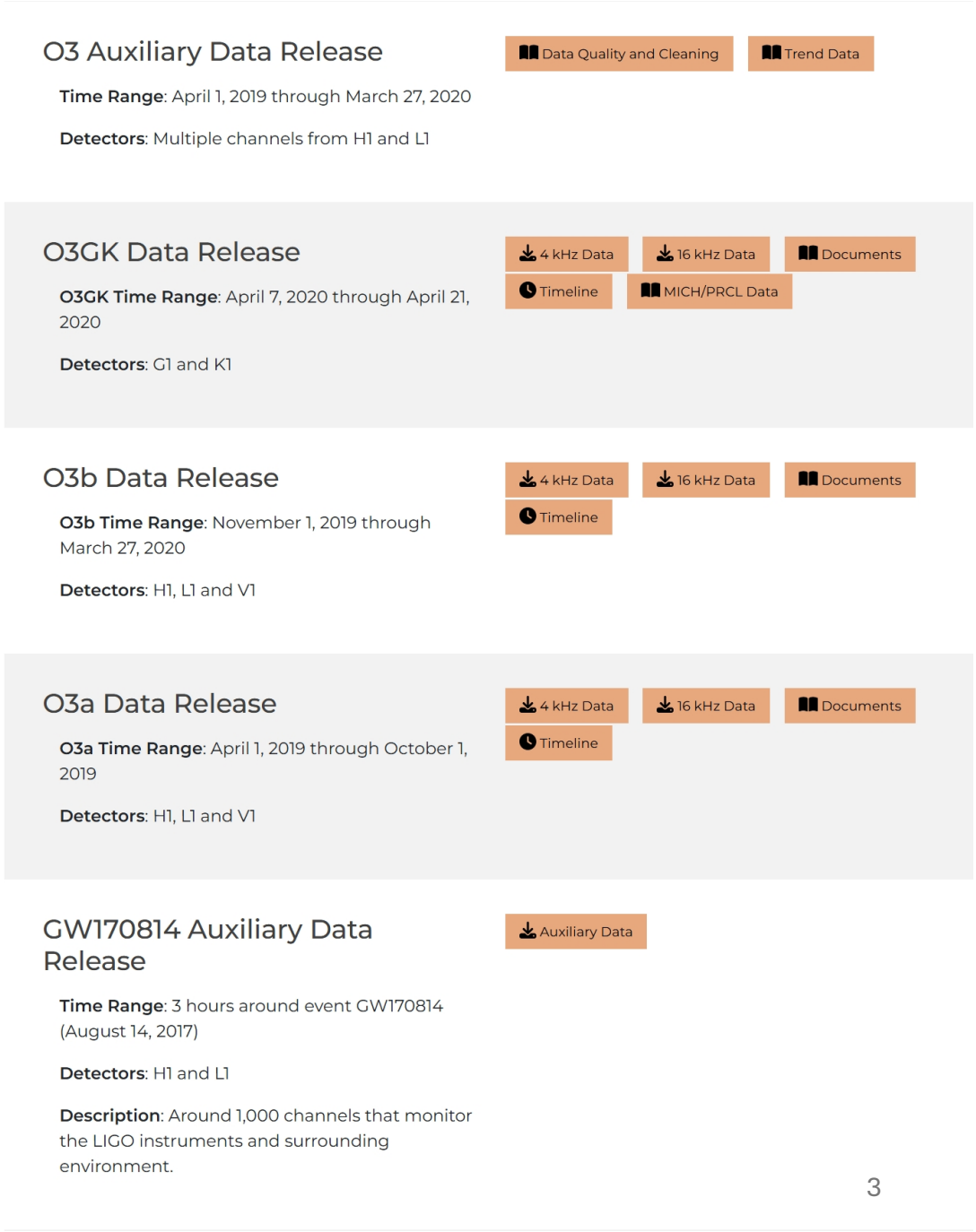


Recently added: a coarse mass estimate

{ "bin_edges": [0.1, 0.87, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.7, 1.9, 2.1, 2.3, 3, 5.5, 11, 22, 44, 88, 1000], "probabilities": [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.9999, 0, 0, 0, 0] }

Gravitational-wave data (https://gwosc.org/)

- Bulk data release from observing runs
- Auxiliary data from a subset of channels



Event Catalogs (https://gwosc.org/)

Name	Version	Release	GPS	Mass 1 ($M_\odot$)	Mass 2 ($M_\odot$)	Network SNR	Distance (Mpc)	χ_{eff}	Total Mass ($M_\odot$)
GW200322_091133	v1	GWTC-3-confident	1268903511.3	$^{+130}_{38\ -22}$	$^{+24.3}_{11.3\ -6.0}$	$^{+2.7}_{4.5\ -3.0}$	$^{+12500}_{3500\ -2200}$	$^{+0.54}_{0.27\ -0.58}$	$^{+132}_{50\ -22}$
GW200316_215756	v1	GWTC-3-confident	1268431094.1	$^{+10.2}_{13.1\ -2.9}$	$^{+2.0}_{7.8\ -2.9}$	$^{+0.4}_{10.3\ -0.7}$	$^{+480}_{1120\ -440}$	$^{+0.27}_{0.13\ -0.10}$	$^{+7.2}_{21.2\ -2.0}$
GW200311_115853	v1	GWTC-3-confident	1267963151.3	$^{+6.4}_{34.2\ -3.8}$	$^{+4.1}_{27.7\ -5.9}$	$^{+0.2}_{17.8\ -0.2}$	$^{+280}_{1170\ -400}$	$^{+0.16}_{-0.02\ -0.20}$	$^{+5.3}_{61.9\ -4.2}$
GW200308_173609	v1	GWTC-3-confident	1267724187.7	$^{+166}_{60\ -29}$	$^{+36}_{24\ -13}$	$^{+2.5}_{4.7\ -2.9}$	$^{+13900}_{7100\ -4400}$	$^{+0.58}_{0.16\ -0.49}$	$^{+169.0}_{92.0\ -48.0}$
GW200306_093714	v1	GWTC-3-confident	1267522652.1	$^{+17.1}_{28.3\ -7.7}$	$^{+6.5}_{14.8\ -6.4}$	$^{+0.4}_{7.8\ -0.6}$	$^{+1700}_{2100\ -1100}$	$^{+0.28}_{0.32\ -0.46}$	$^{+11.8}_{43.9\ -7.5}$
GW200302_015811	v1	GWTC-3-confident	1267149509.5	$^{+8.7}_{37.8\ -8.5}$	$^{+8.1}_{20.0\ -5.7}$	$^{+0.3}_{10.8\ -0.4}$	$^{+1020}_{1480\ -700}$	$^{+0.25}_{0.01\ -0.26}$	$^{+9.6}_{57.8\ -6.9}$
GW200225_060421	v1	GWTC-3-confident	1266645879.3	$^{+5.0}_{19.3\ -3.0}$	$^{+2.8}_{14.0\ -3.5}$	$^{+0.3}_{12.5\ -0.4}$	$^{+510}_{1150\ -530}$	$^{+0.17}_{-0.12\ -0.28}$	$^{+3.6}_{33.5\ -3.0}$
GW200224_222234	v1	GWTC-3-confident	1266618172.4	$^{+6.7}_{40.0\ -4.5}$	$^{+4.8}_{32.7\ -7.2}$	$^{+0.2}_{20.0\ -0.2}$	$^{+500}_{1710\ -650}$	$^{+0.15}_{0.10\ -0.16}$	$^{+7.2}_{72.3\ -5.3}$
GW200220_124850	v1	GWTC-3-confident	1266238148.1	$^{+14.1}_{38.9\ -8.6}$	$^{+9.2}_{27.9\ -9.0}$	$^{+0.3}_{8.5\ -0.5}$	$^{+2800}_{4000\ -2200}$	$^{+0.27}_{-0.07\ -0.33}$	$^{+17}_{67\ -12}$
GW200220_061928	v1	GWTC-3-confident	1266214786.7	$^{+40}_{87\ -23}$	$^{+26}_{61\ -25}$	$^{+0.4}_{7.2\ -0.7}$	$^{+4800}_{6000\ -3100}$	$^{+0.40}_{0.06\ -0.38}$	$^{+55}_{148\ -33}$
GW200219_094415	v1	GWTC-3-confident	1266140673.1	$^{+10.1}_{37.5\ -6.9}$	$^{+7.4}_{27.9\ -8.4}$	$^{+0.3}_{10.7\ -0.5}$	$^{+1700}_{3400\ -1500}$	$^{+0.23}_{-0.08\ -0.29}$	$^{+12.6}_{65.0\ -8.2}$
GW200216_220804	v1	GWTC-3-confident	1265926102.8	$^{+22}_{51\ -13}$	$^{+14}_{30\ -16}$	$^{+0.4}_{8.1\ -0.5}$	$^{+3000}_{3800\ -2000}$	$^{+0.34}_{0.10\ -0.36}$	$^{+20}_{81\ -14}$
GW200210_092254	v1	GWTC-3-confident	1265361792.9	$^{+7.5}_{24.1\ -4.6}$	$^{+0.47}_{2.83\ -0.42}$	$^{+0.5}_{8.4\ -0.7}$	$^{+430}_{940\ -340}$	$^{+0.22}_{0.02\ -0.21}$	$^{+7.1}_{27.0\ -4.3}$

And event (https://zenodo.org/records/5546663) and population parameters (https://zenodo.org/records/11254021)

Some Observations

- A single, vetted source of strain data, event alerts & catalogues, population properties and other derived data are widely used by the broader community
  - Data are used in both expected and unanticipated ways
  - Ensuring that internal and released data are identical helps reduce errors and review burden
- We are consistently evolving to more data/information being released and on shorter timescales
  - Leads to collaboration concerns that major results will be derived first by the outside community
  - Collaboration publications tend to become the definitive statements, even if not the first

Some Observations

- Data releases must be very well documented and explained, with clear example codes, tutorials, etc
  - Important to be able to update, and version control, released data
- Realistically, data release policies are decided (or at least highly influenced) by funding council requirements.
- Many of the data products require data from all operating instruments, e.g. localization, event catalogs
  - Requires a single, consistent data policy between all observatories

Data Management

- We often focus on what data is released and when, but there's much more to data management including
  - Storage, preservation and archiving
  - Documentation, metadata, training, licensing
  - Costs

References

- LIGO-M080072-v2: Providing Open Access of LIGO Data to the Broader Research Community
- LIGO-T1000414-v13: An Astrophysical Metric for LIGO Open Data Release
- LIGO-M1200055-v2: LSC AND VIRGO POLICY ON RELEASING GRAVITATIONAL WAVE TRIGGERS TO THE PUBLIC IN THE ADVANCED DETECTORS ERA
- LIGO-M1000066-v30: LIGO Data Management Plan