Providing a Cleaned and Characterized Dataset for Gravitational-Wave Astrophysics

Derek Davis, Syracuse University
Thomas J. Massinger, LIGO Laboratory, Caltech

DESCRIPTION OF PROJECT

Our focus project is the production of the cleaned and characterized data set for Advanced LIGO's second observing run that will be used in production astrophysical searches and released to the public through the GWOSC. The two main products in the O2 dataset will be cleaned h(t) for both Hanford and Livingston and a set of accompanying data quality flags. Our contributions have led to an increase in sensitivity of greater than 30% for gravitational wave analyses and facilitated a higher number of confident detections of gravitational waves from compact binary coalescences.

DATA QUALITY VETOES

We produced data quality (DQ) vetoes for all gravitational wave searches. Our focus was on CBC searches, but our products are also used in burst, continuous wave, and stochastic background searches. All searches use Category 1 (CAT1) DQ products, which designate times with severe data quality issues that should not be used in analyses. Transient searches also use Category 2 (CAT2) vetoes, which don't remove data from the input to analyses but indicate periods of excess noise. Development of these vetoes involved identifying instrumental noise sources and creating algorithms to systematically flag them. These instrumental investigations sometimes lead to commissioners resolving problems in the detector, which is the preferred outcome.

In O2, 1.7% of data was removed by CAT1 vetoes, while 0.4% was removed by CAT2 vetoes. Data quality vetoes were shown in O1 to increase sensitivity of CBC searches by up to 50% (Abbott et al. 2018a) and reduce the background of Burst searches. In cases where a veto is not developed, instrumental investigations resulted in expertise that allowed us to follow up and validate detection candidates in O2 based on knowledge of instrumental noise.

CLEANING

We subtracted linearly coupled noise from the entire O2 data set (Davis et al. 2018), resulting in a cleaned dataset for all of O2 for both interferometers. At both sites, calibration lines and contributions from 60 Hz power mains were subtracted out of the data. At LHO, broadband noise due to beam jitter, which was the dominant noise source at H1, was subtracted from the 10-1024 Hz range. This noise subtracted dataset will be the input to all O2 result papers going forward and will be released to the public via GWOSC.

To accomplish this large scale noise subtraction, we developed a noise subtraction pipeline from scratch using the methods of (Allen et al. 1999). This subtraction pipeline will, if necessary, be used to subtract generic linearly coupled noise in O3. While initial subtraction was done for individual candidate events (Driggers et al 2018), our pipeline was developed to run as a fully parallelized workflow via Condor and Pegasus. The increased efficiency and throughput of our pipeline allowed for the entire O2 dataset to be cleaned in weeks.

Subtraction of linearly coupled noise resulted in a 30% increase in network sensitivity as measured by an injection campaign by a production CBC search (Davis et al. 2018). This increase in sensitivity will apply to all analyses of LIGO data, resulting in the most sensitive gravitational wave searches to date. In terms of individual gravitational wave events, the use of the cleaned O2 dataset enabled GW170809 to be found by multiple pipelines and GW170818 (the loudest signal detected by Virgo to date) be identified as an astrophysical signal, in addition to increasing the significance of each of the previously discovered signals (Abbott et al. 2018b).

In addition to the detection of individual events, there are more broad astrophysical questions that can be answered. The full, updated foreground distribution of gravitational wave candidates produced with cleaned data will be used for establishing more confident BBH and BNS rate estimates in the O2 catalog. The use of this dataset also allows for long duration analyses to be performed on noise subtracted data for the first time, such as investigations of post merger signals after GW170817 (Abbott et al. 2018c). In terms of parameter estimation, the use of cleaned data on marginal events strengthens the efficacy of Bayesian coherence tests and provides secondary checks on the astrophysical results that used the initial noise subtracted data from Driggers et al.

BROADER IMPACT

Our work has improved the scientific output of the LSC by increasing both the sensitivity of our searches and the confidence in published detections. Within the LSC, the scope of our work required us to serve as a bridge between the instrument science and astrophysics working groups. This increased communication is crucial to the continued scientific success of the LSC. An optimal astrophysical search pipeline requires understanding of instrumental noise artifacts and the optimal detector requires understanding what types of noise are negatively impacting the astrophysical searches. In a broader sense, we have a responsibility to the public to provide a characterized and vetted data set. This knowledge will help support the entire astronomical community to effectively utilize LIGO data and will prevent computing and observational resources from being wasted on data that is unfit for astrophysical analysis.

INDIVIDUAL CONTRIBUTIONS

It is very difficult to detangle our individual contributions; we worked as a team on all of the aforementioned work. We both developed data quality vetoes during O2, developed the noise cleaning pipeline, and split the data down the middle when subtracting linearly coupled noise. We have both been vetting candidate events for the O2 catalog paper based on our knowledge of transient noise in the detectors.

CITATIONS

Abbott et al. 2018a, Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO's first observing run, Class. Quantum Grav., 35, 6
- 2018b, Catalog of Compact Binary Coalescences in the First and Second Observing Runs of Advanced LIGO and Virgo, LIGO-P1800307-v1
- 2018c, Search for gravitational waves from a long-lived remnant of the binary neutron star merger GW170817, arXiv:1810.02581

Allen et al. 1999, Automatic cross-talk removal from multi-channel data, arXiv:9909083

Davis et al. 2018, Improving the Sensitivity of Advanced LIGO Using Noise Subtraction, arXiv:1809.05348

Driggers et al. 2018, Improving astrophysical parameter estimation via offline noise subtraction for Advanced LIGO, arXiv:1806.00532