

Memo on LVK proprietary data procedures

June 26, 2024

The LIGO-Virgo(-KAGRA) data are released together according to a schedule agreed within the LVK collaboration and with the funding agencies. The data contain all the information to perform full analyses and, once it is public, external groups can be effectively competitive with the internal LVK efforts.

In addition to being responsible for operating the interferometers, collecting the data and ensuring their highest possible quality, the LVK hosts a vibrant data analysis community and a unique variety of expertise, delivering high profile scientific results out of the LVK data. These cover a wide set of areas in modern physics and they are distributed over a large range of timescales, from the ~ 90 seconds it takes to uncover and broadcast low-latency triggers to astronomers via GCN notices and circulars to the 1 year or more it can take to perform deep analyses and collect and interpret their results in collaboration-wide publications.

The data proprietary period that follows a data taking period allows the LVK Collaboration to ensure that the production of these publications follows high-quality standards. All LVK members with authorship rights sign all LVK publications, which, especially in the European system, helps sustain the attractiveness of working within the Collaboration (including on detector activities).

The data proprietary period for LVK data has been subject to negotiations within the Collaboration, with constraints from the funding agencies. As detailed in appendix A, there is a general trend to reduce the proprietary data period, relative to the effective time of observation. The goal set for O5 is to operate in the so-called “continuous observing mode”, where the interferometers collect data for “long” periods of time (O5 is supposed to last for about 3 years), and release data in chunks of 12 months, 12 months after the end of each chunk.

Within the LVK, this direction is mainly motivated by an argument of pressure received by the NSF. The 12+12 months mode is generally accepted within the collaboration, and scientists are working relentlessly to optimize the data treatment and analysis procedures in order to meet this target. With no intention to compromise on the quality of our results, the reduction of proprietary time is effectively reducing the variety of the LVK scientific production. As an example, already for O4, the data analysis working groups have been forced to make a choice, and determine which papers are “key” and need to be prioritized to be published by the data release deadline. While the 12+12 model seems manageable by the working groups searching for transient GW signals (regardless of whether they use signal models to analyze the data), it should be considered as a hard limit and the community is worried about the general trend of reduction of proprietary data.

The 12+12 model remains problematic for groups looking for persistent (quasi-)monochromatic GW signals or a persistent stochastic GW background. Notably, these groups are aiming for their first detection, as opposed to the transient search groups. All else being fixed, their

chances of uncovering a signal increase as data is accumulated, but so do the computational costs of the analyses.

Appendix A - history of data taking periods and proprietary data periods

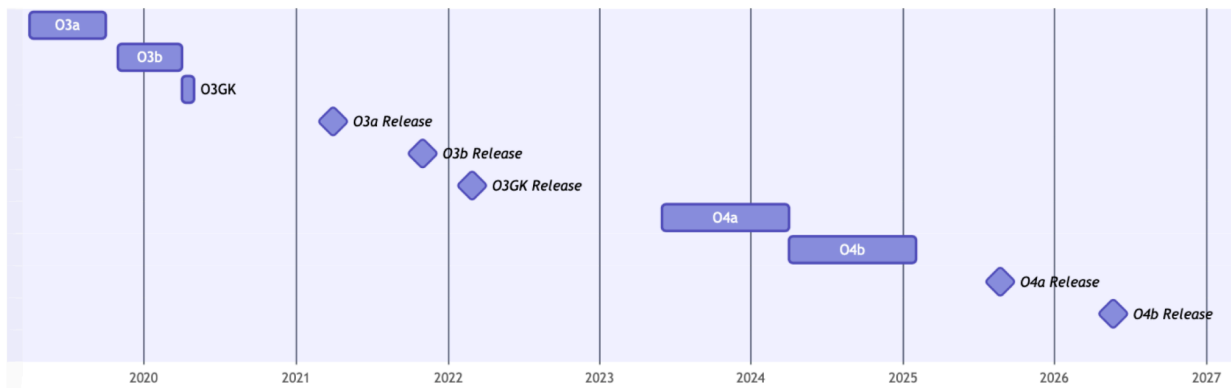
- O1 (2015-09-12 to 2016-01-19) – initially only data around 3 events was released with the associated publications and eventually the whole dataset was released 24 months after the end of the 4 months long run, in January 2018.
- O2 (2016-11-30 to 2017-08-25) – data released in February 2019, 18 months after the end of the 9 months long run.
- O3a (2019-04-01 to 2019-10-01) – data released on 2021-04-30, approximately 18 months after the end of the 6 months long observing block.
- O3b (2019-11-01 to 2020-03-27) – data released on 2021-11-07, approximately 18 months after the end of the 5 months long observing block, abruptly ended due to the COVID-19 pandemic.
- O3GK (2020-04-07 to 2020-04-21) – data released on 2022-03-03. This was an additional data release that included 2 weeks of data from GEO 600 and KAGRA.
- O4a (2023-05-24 - 2024-01-16) – data to be released on August 2025-08-23, approximately 19 months after the end of the 8 months long observing block.
- O4b (2024-04-10 - 2025-06-09) – data release under discussion, following the run plan extension from 2025-01-23 to 2025-06-09. The LVK proposal is to release this on 2026-09-30.

This data is summarized in the following table, with the O5 target for reference. The next to last column reports the proprietary period to released block duration ratio, a quantity which serves as a proxy for the pressure the data release places on the collaboration. Finally the last column is the median wait time normalized to the block duration: this quantifies how long non-LVK members wait for the data, compared to the amount of data received itself.

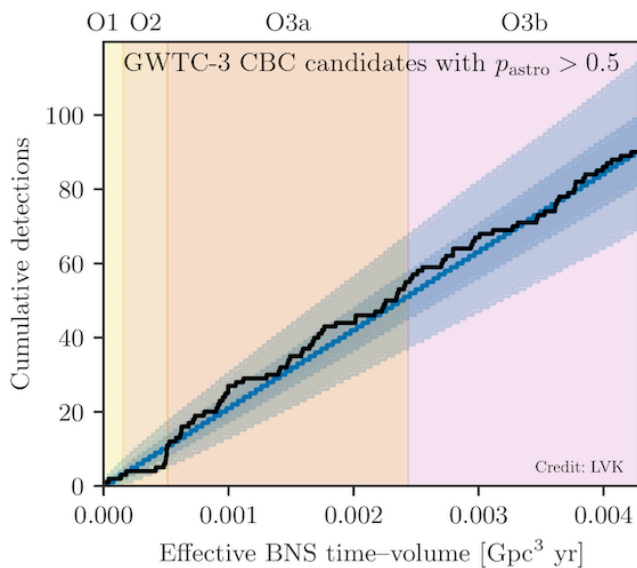
| Observing period | Run or block duration (months) | Proprietary period since end (months) | Proprietary period since start (months) | Block duration/ proprietary period since end | Median wait time / block duration |
|------------------|--------------------------------|---------------------------------------|---|--|-----------------------------------|
| O1 | 4 (LIGO) | 24 | 28 | 0.17 | 6.5 |
| O2 | 9 (LIGO), 1 (Virgo) | 18 | 27 | 0.33 | 2.5 |
| O3a | 6 (LIGO-Virgo) | 18 | 25 | 0.33 | 3.5 |

| | | | | | |
|-------------|-----------------|------------------------|----|------|-----|
| O3b | 5 (LIGO-Virgo) | 18 | 24 | 0.28 | 4.1 |
| O4a | 8 (LIGO) | 19 | 27 | 0.42 | 2.9 |
| O4b | 14 (LIGO-Virgo) | 18 (<i>proposed</i>) | 32 | 0.78 | 1.8 |
| O5 (target) | 12 | 12 | 24 | 1 | 1.5 |

Note that the effective observing time for O2 is about 3.9 months (LIGO) and 0.5 months (Virgo). The last column is calculated as $0.5 + \text{Proprietary period/block duration}$ and represents the median wait for data for scientists outside the LVK collaboration.



O3 blocks and data releases, and O4 planned blocks and data releases, prior to the decision to extend O4.



Number of confirmed GW signals observed during O1, O2, O3a and O3b (each observing run is displayed with a different background color) as a function of a proxy of the detectors sensitivity.