Incorporation of Statistical Data Quality Information into the GstLAL Search Analysis

Patrick Godwin,[1,2] Reed Essick,[3] Chad Hanna,[1,2,4,5] Kipp Cannon,[6] Sarah Caudill,[7] Chiwai Chan,[6] Jolien D. E. Creighton,[8] Heather Fong,[6,9] Erik Katsavounidis,[10] Ryan Magee,[1,2] Duncan Meacher,[8] Cody Messick,[1,2] Soichiro Morisaki,[11] Debnandini Mukherjee,[1,2] Hiroaki Ohta,[6] Alexander Pace,[1,2] Iris de Ruiter,[12,7] Surabhi Sachdev,[1,2,13] Leo Tsukada,[6,9] Takuya Tsutsui,[6] Koh Ueno,[6] Leslie Wade,[14] and Madeline Wade[14]

[1]*Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA*
[2]*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*
[3]*Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, Illinois, 60637, USA*
[4]*Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*
[5]*Institute for CyberScience, The Pennsylvania State University, University Park, PA 16802, USA*
[6]*RESCEU, The University of Tokyo, Tokyo, 113-0033, Japan*
[7]*Nikhef, Science Park, 1098 XG Amsterdam, Netherlands*
[8]*Leonard E. Parker Center for Gravitation, Cosmology, and Astrophysics, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA*
[9]*Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan*
[10]*LIGO, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[11]*Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan*
[12]*Astronomical Institute, Anton Pannekoek, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, Netherlands*
[13]*LIGO Laboratory, California Institute of Technology, MS 100-36, Pasadena, California 91125, USA*
[14]*Department of Physics, Hayes Hall, Kenyon College, Gambier, Ohio 43022, USA*

(Dated: May 15, 2020)

We present updates to GstLAL, a matched filter gravitational-wave search pipeline, in Advanced LIGO and Virgo's third observing run. We discuss the incorporation of statistical data quality information into GstLAL's multi-dimensional likelihood ratio ranking statistic and additional improvements to search for gravitational wave candidates found in only one detector. Statistical data quality information is provided by iDQ, a data quality pipeline that infers the presence of short-duration transient noise in gravitational-wave data using the interferometer's auxiliary state, which has operated in near real-time since before LIGO's first observing run in 2015. We look at the performance and impact on noise rejection by the inclusion of iDQ information in GstLAL's ranking statistic, and discuss GstLAL results in the GWTC-2 catalog, focusing on two case studies; GW190424A, a single-detector gravitational-wave event found by GstLAL and a period of time in Livingston impacted by a thunderstorm.

I. INTRODUCTION

During Advanced LIGO and Virgo's second observing run, the detection of a binary neutron star merger, GW170817 [1] and its joint detection of a short gamma ray burst [2, 3] observed by Fermi-GBM and INTEGRAL sparked a massive follow-up campaign with more than 70 telescopes and observatories participating [4]. GW170817 started a new era of multi-messenger astronomy and provided insight into compact binary mergers and their associated electromagnetic counterparts, prompting an avalanche of further study, from advancing our understanding of short Gamma-ray burst beaming profiles (e.g. [2, 5, 6]), kilonova light-curves and evolution (e.g. [7–10]), cosmology and the expansion rate of the universe [11, 12], the equation of state of dense nuclear matter (e.g. [13, 14]), and the mass distribution of compact objects [15] among countless others.

GW170817 was identified in low-latency as a single-detctor candidate in the LIGO Hanford detector by GstLAL, a matched filter gravitational-wave search for compact binary coalescences [16, 17]. However, at that time both the LIGO Livingston and Virgo detectors also were recording science-quality data. LIGO Livingston, in particular, was expected to have witnessed a similar signal to the Hanford detector, and it was quickly discovered that a non-Gaussian noise transient, or *glitch* [18], in the Livingston interferometer coincided with GW170817's inspiral track, thereby causing GstLAL to neglect data from that detector at that time. While the noise transient was from a familiar class of such glitches and was later modeled and subtracted from the detector data, enabling precision measurements of the gravitational-wave signal using Livingston data [19, 20], the presence of non-Gaussian noise initially complicated the detection process, highlighting the need for rapid, reliable data quality information.

iDQ, a statistical inference framework that generates probabilistic data quality information in low-latency [21], has operated throughout the advanced detector era. iDQ autonomously identified auxiliary witnesses to the type of non-Gaussian noise coincident with GW170817 and correctly labeled the subsecond interval containing the arti-

fact as extremely likely to contain a glitch in low-latency, less than 8 seconds after GstLAL reported the candidate. Unfortunately, at that time, iDQ's statistical data quality information was not incorporated into the GstLAL search. Nonetheless, iDQ's autonomous identification of the noise transient, along with further human vetting, allowed GW170817 to be announced to the broader astrophysical community in time to inform electromagnetic follow-up [22].

GW170817 points out the benefits of folding probabilistic data quality information into searches directly. This work describes the incorporation of statistical data quality information from iDQ within the GstLAL search during the third observing run, an important milestone in mitigating the impact of non-Gaussian noise within searches in low-latency.

In April 2019, the advanced LIGO [23] and Virgo [24] interferometers started their third observational run, O3, and with it, began sending out automated open public alerts from gravitational-wave candidates. The goal of low-latency gravitational-wave detection and alerting infrastructure [25] is to detect the onset of electromagnetic emission coming from compact binary coalesences within seconds of the merger [26]. This requires the need for rapid validation and follow-up of gravitational-wave candidates.

Non-Gaussian noise in LIGO and Virgo can produce short-duration noise transients in the gravitational-wave strain, $h(t)$, as seen in GW170817. The presence of glitches in gravitational-wave data limits the search sensitivity to gravitational-wave signals. Being able to identify periods of non-Gaussianities and limit their effect on searches is essential for detecting gravitational-wave signals. The increased sensitivity of gravitational-wave detectors in O3 has increased the rate at which gravitational-wave detections are made. However, the rate of glitches has also increased in both LIGO detectors during the first half of advanced LIGO and Virgo's third observing run, O3a [27], making the use of automated noise mitigation methods and rapid follow-up more critical.

The detector's auxiliary state in advanced LIGO [23] is monitored by the instrumental control system, which includes the optical configuration [28], seismic isolation [29], and many control systems which are responsible for keeping the interferometer in lock [30, 31]. Beyond this, many sensors monitor the detector's environment [32]. In total, $O(10^5)$ channels monitor the detector's auxiliary state in advanced LIGO. The noise sources that produce glitches in $h(t)$ are occasionally witnessed by at least one of the detector's many subsystems or in its surrounding environment [32], which can be detected in one or more auxiliary channels.

Understanding the physical mechanism in which glitches couple to $h(t)$ allows one to either identify and remove the source of the noise or identify times when glitches due to this coupling are present [33]. If a physical mechanism can not be observed, one has to rely on statistical couplings to infer the presence of a glitch, by identifying statistical correlations between $h(t)$ and auxliary channels. Indeed, given the large number of auxiliary channels, the vast majority of couplings to $h(t)$ remain unmeasured and statistical inference remains the best chance to systematically identify terrestrial noise sources. iDQ provides a framework to infer the presence of glitches in $h(t)$ by identifying statistical correlations between $h(t)$ and auxiliary channels, as well as providing a mechanism to discover new statistical couplings as they arise in low-latency. It has operated in near-real-time since before O1 in 2015, and was sped up for the latest observing run to produce probabilistic data quality information encoded as a set of streaming timeseries available concurrently with $h(t)$ in low-latency.

Historically, gravitational-wave searches handle times associated with poor data quality by removing the problematic data from being analyzed or by vetoing any candidates during this time [34]. These periods of poor data quality are flagged by various data quality vetoes based on their severity [18, 35, 36], and are either generated in low-latency from select known couplings or produced in an offline fashion after various data quality investigations. One benefit of using non-binary data products is the ability to continuously downrank periods associated with various non-Gaussianities. This allows searches to detect extremely confident astrophysical signals in the presence of excess non-Gaussian noise instead of forcing them to reject the entire stretch of data, regardless of whether there is a signal present, based on non-Gaussian noise. In this manner, iDQ can be used to fold in data quality information into a gravitational-wave search without vetoing an astrophysical event coincident with a glitch, as in the case with GW190424A, discussed in Sec. III B 1.

GstLAL is a matched filter gravitational-wave search pipeline aimed at detecting compact binary coalesences in near real-time, providing event significance estimates and point estimates for binary parameters. GstLAL operates in both low-latency for near real-time detection, as well as an offline mode to process gravitational-wave data with background statistics collected over the analysis period as well as additional information generated offline such as data quality vetoes and improvements in $h(t)$ calibration [37]. GstLAL identifies candidates by using a multi-dimensional likelihood ratio ranking statistic $\mathcal{L}$, described in [16, 17, 38–40], which folds in information about the candidate event's SNR in each detector, a multi-detector signal consistency test [38], time-averaged detector sensitivity, the signal population model, and information about the collected background.

In this work, we describe methods of handling transient noise in GstLAL in Sec. II, and discuss the inclusion of statistical data quality information via iDQ into a matched filter search, by including it as a term in GstLAL's ranking statistic in Sec. II B. We also discuss the performance and pipeline sensitivity from its inclusion and its impact on the GWTC-2 catalog in Sec. III.

In particular, we look at two case studies; GW190424A, a single-detector gravitational-wave event found by GstLAL which is found to be vetoed by a data quality flag and a period of non-Gaussianity caused by a nearby thunderstorm in LIGO-Livingston and identified by iDQ.

II. INCORPORATION OF STATISTICAL DATA QUALITY INFORMATION

As mentioned in Sec. I, gravitational-wave searches have historically handled periods of poor data quality by either vetoing candidates during this time at the event identification stage or by removing such segments from $h(t)$ in the data conditioning stage, through a procedure called gating [16, 41]. GstLAL handles times associated with poor data quality with several complimentary approaches; (1) signal-based vetoes using a ξ^2 test [16], (2) an auto-gating procedure based on loud excursions from $h(t)$ [16], and (3), gating times associated with poor data quality from a known list of data quality vetoes [36].

GstLAL filters the gravitational-wave strain, $h(t)$, against a bank of known compact binary template waveforms. Both the data and the template waveforms are whitened using a measured noise power spectral density (PSD) estimate. The template waveforms are whitened ahead of time using a reference PSD. The data is whitened on-the-fly during operation to track drifts in the PSD over a detector lock segment, using a median-mean PSD estimation approach [16]. This provides an estimation technique which is robust against short timescale fluctuations, including many glitch classes.

Different categories of data quality vetoes are treated based on their severity. Of interest in compact binary searches are category 1 (CAT1) vetoes, flagging exceptionally egregious times, and category 2 (CAT2) vetoes based on physical couplings. Times associated with CAT1 vetoes are not analyzed to avoid causing issues with PSD estimation. Gating based on loud $h(t)$ excursions and CAT2 data quality vetoes are treated similarly; the target data associated with a segment is padded on both sides and a windowing function is applied to it. During the first observational run, O1, a square window was used, replacing the data within the window with zeros. In the second observational run, O2, a Tukey window was used to smoothly transition between the gate and the whitened data. In addition, a trigger from an individual detector is vetoed if it coincides with a gate.

The procedure discussed in this paper replaces the treatment of vetoes in (3), and instead relies on iDQ to infer the presence of non-Gaussian noise in the data. Although Ref. [21] discusses many possible ways to do this, we focus on a relatively simple, yet effective, scheme. This is done by adding a term in GstLAL's multi-dimensional ranking statistic that includes information from iDQ to downrank candidate events that are in close proximity with a glitch. This term is targetted at single detector candidates rather than events found in coincidence, as discussed in II B, since many of the consistency checks found for coincident events are not present for single detector candidates. Among other data products, iDQ produces a log-likelihood estimate of non-Gaussian noise, $\log \mathcal{L}_{\rm iDQ}$. We do not use this statistic as-is, however. Instead, we renormalize $\log \mathcal{L}_{\rm iDQ}$ to better control its possible impact on the GstLAL search. This renormalized quantity, $\log \widehat{\mathcal{L}}_{\rm iDQ}$, is maximized over a ± 1 second window around the candidate event.

One benefit of this approach is that data quality information from external sources can be added in without vetoing gravitational-wave events which coincide with periods of poor data quality. Furthermore, iDQ data products are available in low-latency at the same time as $h(t)$, so this work can be extended to run in low-latency operation as well, as would be required for astrophysical events coincidence with non-Gaussian noise like GW170817.

A. iDQ Renormalization Procedure

iDQ produces data products that are available at the same time as gravitational-wave strain data in its low-latency operation, but due to its low-latency nature can be subject to data dropouts or sub-optimal calibration during periods of extreme non-Gaussianity. In order to mitigate such effects from the low-latency runs, $\log \mathcal{L}_{\rm iDQ}$ timeseries were collected across the time of interest and renormalized. This also allows us to better control the possible impact that iDQ information could have on the search. Since the period of interest is over a period of several months, we instead collect data over smaller periods of time, about a week of coincident data, and renormalized for each stretch of time.

For each timespan, the $\log \mathcal{L}_{\rm iDQ}$ timeseries is ingested and the raw timeseries sampled at 128 Hz is aggregated and maximized over 1 second windows. These aggregated values are used to calculate the transformation, defined as

$$\begin{aligned} \log \widehat{\mathcal{L}}_{\rm iDQ} &= \begin{cases} 0 & P < P_{min} \\ \log \frac{P}{100-P} & P_{min} \leq P \leq P_{max} \\ 15 & P \geq P_{max} \end{cases}, \\ P_{min} &= 50, \\ P_{max} &= \frac{100}{1+e^{-\max(\log \widehat{\mathcal{L}}_{\rm iDQ})}} = \frac{100}{1+e^{-15}}. \end{aligned} \tag{1}$$

where P is the percentile of $\log \widehat{\mathcal{L}}_{\rm iDQ}$ values over a given chunk.

P_{min} and P_{max} are chosen to restrict the dynamic range to $\log \widehat{\mathcal{L}}_{\rm iDQ} \in [0, 15]$, to enforce that candidates are only downranked as well as limit the dynamic range of $\log \widehat{\mathcal{L}}_{\rm iDQ}$. Here, we cap $\log \widehat{\mathcal{L}}_{\rm iDQ}$ to be 15, and values below the 50th percentile are mapped to $\log \widehat{\mathcal{L}}_{\rm iDQ} = 0$. Higher values of $\log \widehat{\mathcal{L}}_{\rm iDQ}$ suggest a higher degree of non-

Gaussianity in the data, so restricting $\log \widehat{\mathcal{L}}_{\rm iDQ}$ to be positive implies that candidate events can only be penalized from its inclusion into GstLAL. Restricting $\log \widehat{\mathcal{L}}_{\rm iDQ}$ to 15 scales the contribution from iDQ so that it is of a similar size as a confident event seen from GstLAL ($\log \mathcal{L} \approx 15$). The mapping between percentile and likelihood is shown in figure 1. After computing the transformation for each timespan, all timespans across the analysis time of interest are combined and stored as a single file to be ingested by GstLAL.

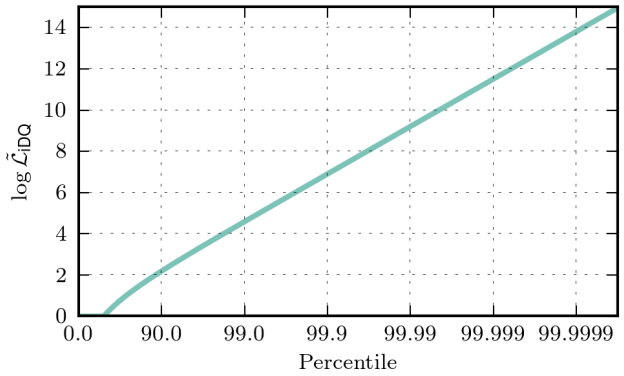

FIG. 1. Mapping between $\log \mathcal{L}_{\rm iDQ}$ and $\log \widehat{\mathcal{L}}_{\rm iDQ}$ based on percentile. The effects of limiting the dynamic range based on P_{min} and P_{max} are such that $\log \widehat{\mathcal{L}}_{\rm iDQ} \geq 0$ and $\log \widehat{\mathcal{L}}_{\rm iDQ} \to 15$ as $P \to 100$.

B. Modification of the Ranking Statistic

As mentioned in Sec. I the multi-dimensional likelihood ratio ranking statistic in GstLAL is responsible for ranking candidate events. The likelihood ratio is defined as

$$\mathcal{L} = \frac{P(\,\vec{D_H}, \vec{O}, \vec{\rho}, \vec{\xi^2}, \Delta\vec{\phi}, \Delta\vec{t}\,|\,\text{signal}\,)}{P(\,\vec{D_H}, \vec{O}, \vec{\rho}, \vec{\xi^2}, \Delta\vec{\phi}, \Delta\vec{t}\,|\,\text{noise}\,)}, \tag{2}$$

where each vector of parameters denotes detector-specific quantities.

With the set of detectors LIGO Hanford (H1), LIGO Livingston (L1), and Virgo (V1), a vector of parameters for X is given by $\vec{X} = \{X_{H1}, X_{L1}, X_{V1}\}$. $\vec{D_H}$ represents the horizon distance, and accounts for the detector sensitivity at the time of the event. $\vec{O}$ represents the detectors which were observing at the time of the detection. ρ is the detected signal-to-noise ratio (SNR), and ξ^2 is the signal-based-veto, which tests the goodness of fit to the template waveform. Finally, Δt and $\Delta\phi$ are the time and phase difference between two detectors. The full factorization of the likelihood ratio is described in [16], [17], [38], and [39].

For single detector candidates, however, signal consistency and coincidence tests are not present as with the multi-detector case. In this case, the main term in likelihood ratio takes on a simpler form, with the Δt and $\Delta\phi$ terms not present:

$$\mathcal{L} = \frac{P(\,D_H, O, \rho, \xi^2\,|\,\text{signal}\,)}{P(\,D_H, O, \rho, \xi^2\,|\,\text{noise}\,)} \left(\mathcal{L}_{\rm penalty}\widehat{\mathcal{L}}_{\rm iDQ}\right)^{-1}, \tag{3}$$

Due to the lack of consistency checks compared with coincident triggers, single detector triggers are more prone to spurious non-Gaussianities and so have two additional terms to mitigate these effects. The first is an emperically determined penalty, $\mathcal{L}_{\rm penalty}$, ensuring that only single detector triggers which can be cleanly distinguished from the background are considered to be significant. For this analysis, $\log \mathcal{L}_{\rm penalty} = 10$. The second term, $\widehat{\mathcal{L}}_{\rm iDQ}$, accounts for non-Gaussian noise and is informed directly by iDQ.

For the iDQ term, a ± 1 second window is applied around the coalescence time, where $\log \widehat{\mathcal{L}}_{\rm iDQ}$ is maximized over this window and applied to single detector candidates. This term downweighs the significance of the event candidate by the likelihood that there is a noise transient in the 2 second window.

III. RESULTS

Here, we focus on the impact of data quality information incorporated into the GstLAL search for GWTC-2 results. We will discuss general statements about the impact of iDQ on rejecting glitches as well as the bulk properties and impact on the pipeline. We provide two specific times in O3a as concrete examples. The first is a single-detector gravitational-wave event, GW190419A, found by GstLAL that was vetoed by a CAT2 flag. The second is a period of time of poor data quality due to a thunderstorm in Livingston, which was rejected by iDQ as likely being terrestrial in origin. This same time was also flagged in a data quality flag offline.

A. Glitch Rejection

First, we look at the efficiency of iDQ timeseries on glitch rejection. In Figure 2, we see the Receiver Operating Characteristic (ROC) curves for both H1 and L1, calculated as the fraction of glitches removed versus False Alarm Probability (FAP), defined as the fraction of time that would be falsely identified as containing glitches based on the detector's auxiliary state. Glitches were determined from times identified by the Stream-based Noise Acquisition and eXtraction (SNAX) pipeline [42] in $h(t)$ with an SNR at or above 8. Similarly, times determined to be clean were sampled over times where SNAX had no glitches at or above 6. This criterion was chosen

in the same manner as is used for iDQ during low-latency operation for O3.

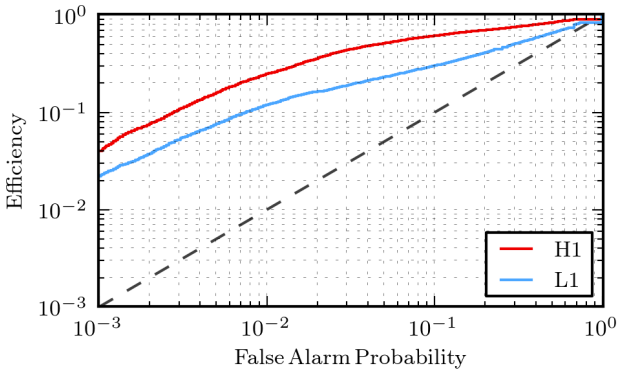

FIG. 2. Receiver Operating Characteristic (ROC) curves for $\log \widehat{\mathcal{L}}_{\rm iDQ}$ for H1 and L1. The dashed line corresponds with a classifier that randomly assigns whether a given time corresponds with a terrestrial noise or the absence of it. We see that iDQ is better able to identify non-Gaussian noise at H1 than at L1, likely due to the fact that each detector sees different noise sources and has slightly different arrays of auxiliary sensors.

It is expected that the performance of using the $\log \widehat{\mathcal{L}}_{\rm iDQ}$ timeseries would be diminished compared to the raw timeseries, due to washing away some of the knowledge used by iDQ in the calibration stage. However, since the goal of the transformation was to take a conservative approach and tackle only the most egregious times of poor data quality, this is acceptable. [1] At a 1% false alarm probability, we can remove 25% of glitches in H1 and 12% of glitches in L1 , and at a 0.1% false alarm rate, we can remove 4% of glitches in H1 and $1-2\%$ in L1. Note that this also incorporates times in which iDQ was not available due to data dropouts, and therefore a small fraction of glitches that would otherwise have been identified may have been missed. Offline iDQ runs would not suffer from such issues.

B. Search Impact

While the ROC curves demonstrate the recalibrated iDQ timeseries' ability to identify non-Gaussian noise efficiently, it is not clear that the identified noise actually limits the search's sensitivity. As such, we look at the impact in the GstLAL pipeline over GWTC-2 results. The sensitivity of the search pipeline is typically measured by performing injections with simulated gravitational-wave signals and measuring the number of injections that are recovered, providing a measure of the sensitive volume [34]. This quantity is multiplied by the analysis time to give Volume x Time (VT), taking into account any removed time from the analysis when applying vetoes. For an astrophysical population distributed uniformly in co-moving volume and source-frame time, the expected number of detections is directly proportional to VT. Any improvement in VT should directly correspond to more detected events. Figure 3 gives the change in search sensitivity with and without iDQ information incorporated into the analysis.

Improvements in VT are seen across all template waveforms in the intermediate FAR region at the few percent level, and appears to be more significant for templates fully encompassing the window where $\log \widehat{\mathcal{L}}_{\rm iDQ}$ is applied. A slight decrease in VT is seen at FAR $\approx 10^{-5}$ for shorter duration templates, which is not well-understood but should not cause recovered candidates found with high significance to be rejected. A decrease in VT is also seen for templates longer than $t = 2\,$s at higher FARs which may be due to only considering $\log \widehat{\mathcal{L}}_{\rm iDQ}$ immediately surrounding the merger time.

Looking at the candidate list, a new candidate event is found which was vetoed previously in a data quality flag. In addition, using iDQ information has downranked a potential candidate found previously by GstLAL and associated with a thunderstorm in LIGO-Livingston. This thunderstorm was flagged as being significant by the iDQ analysis. It was also added as a data quality flag independently by the Detector Characterization group for use in offline analyses.

1. *GW190424A*

This event was found by GstLAL as a single detector BBH in Livingston. However, the time surrounding this event was vetoed by an offline CAT2 flag. This veto in particular was created due to periodic glitching from a camera shutter [43]. An accelerometer near the Y-end test mass witnesses the source of the glitches and was used to create a veto for offline studies. This glitching coincided with a gravitational-wave signal and applying a veto would also veto the gravitational-wave event.

In Figure 4, one can see the inspiral track as well as excess noise from camera glitching about 0.2 seconds before. iDQ also has a peak around the same area and continuing on 1-2 seconds after the merger due to ringing in response to the initial impulse witnessed by the beam-tube accelerometer [43]. Since information provided by iDQ can only downrank a candidate event, GW190421A was not vetoed but merely downranked significantly due to the presence of the glitch. Despite that, the GstLAL analysis found this event as significant enough at the catalog threshold of 1 per 6 months.

[1] Characteristic performance for iDQ during O3 for $\log \mathcal{L}_{\rm iDQ}$ timeseries can be seen in [21].

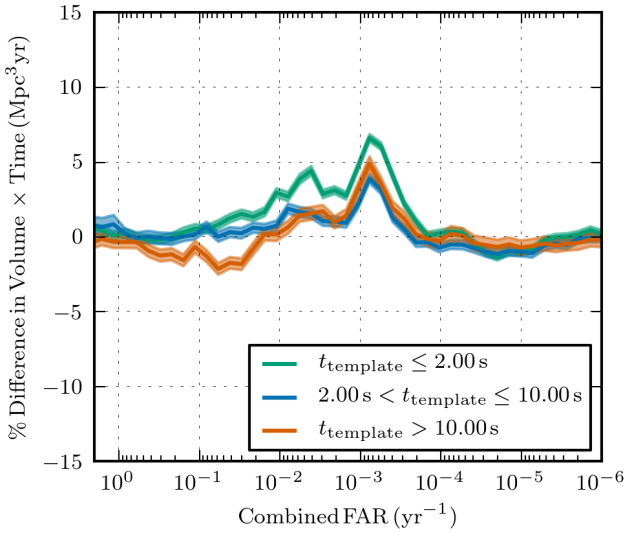

FIG. 3. Change in search sensitivity with and without iDQ information included as a term in the ranking statistic, binned by $t_{\rm template}$ corresponding to the template duration of the compact binary waveform. The shaded regions represent 90% credible intervals. $t = 2\,{\rm s}$ is the window in which $\log \widehat{\mathcal{L}}_{\rm iDQ}$ is maximized over and applied to single-detector event candidates. The high FAR limit is capped at the GWTC-2 FAR threshold at 1 per 6 months.

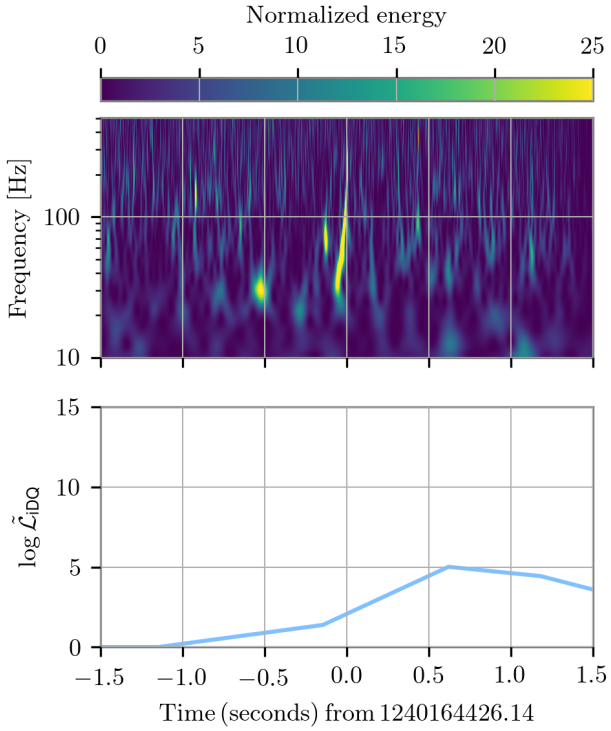

FIG. 4. Time surrounding GW190424A in LIGO-Livingston, shown in a time-frequency spectrogram [44] (above) and in the $\log \widehat{\mathcal{L}}_{\rm iDQ}$ timeseries (below). A clear inspiral track is visible in the time-frequency spectrogram, indicating a strong signal candidate. While the $\log \widehat{\mathcal{L}}_{\rm iDQ}$ term in GstLAL's ranking statistic heavily penalizes the timespan surrounding this event, other terms in the ranking statistic significantly points to the event being signal-like so that the candidate is recovered in an offline analysis.

2. *Thunderstorm Event*

Here, we look at a timespan of a few seconds of detector non-Gaussianity caused by a nearby thunderstorm in LIGO-Livingston during the third observing run. This time was flagged and captured in a data quality veto offline by the Detector Characterization group after it was found that the poor data quality in $h(t)$ in this stretch of time in L1 was caused by a thunderstorm. It was initially flagged by GstLAL as a potential single-detector candidate in an offline analysis without using vetoes or iDQ information. After incorporating iDQ information into the analysis, the candidate event was heavily downranked and no longer was considered to be a suitable candidate. In Figure 5, we see the iDQ timeseries for L1 around the time of the thunderstorm event and that iDQ flags the time as glitchy.

IV. CONCLUSION

We demonstrated a method of handling transient noise in GstLAL by including iDQ information as a term in GstLAL's ranking statistic, and saw the impact on GWTC-2 search catalog results by looking at single-detector gravitational-wave event, GW190424A and a period of non-Gaussianity caused by a thunderstorm. A clear benefit of using non-binary data quality information arises from downranking periods of non-Gaussianities and preventing a scenario where a gravitational-wave candidate in coincidence with a glitch could be vetoed. Note that while a gravitational-wave candidate can not be vetoed with this method, marginal candidates that fall below the threshold for consideration in being a candidate may still be lost after downranking the event. In addition, incorporating iDQ information available in near-real-time alongside $h(t)$ paves the way for autonomously folding in data quality information from the detector's auxiliary state for a wider set of glitch classes.

A few aspects could be further improved from the current implementation. First, since the analysis expects the iDQ timeseries to be available for all times ahead of time, this implementation currently doesn't work in the low-latency scenario. Moving forward, it would be ideal to allow the iDQ timeseries to be collected in low-latency during the filtering stage and fed into the ranking statistic directly while collecting event candidates. However, the incorporation of iDQ information in low-latency and subsequent event identification in the presence of glitches may still impact parameter estimation [19] and sky local-

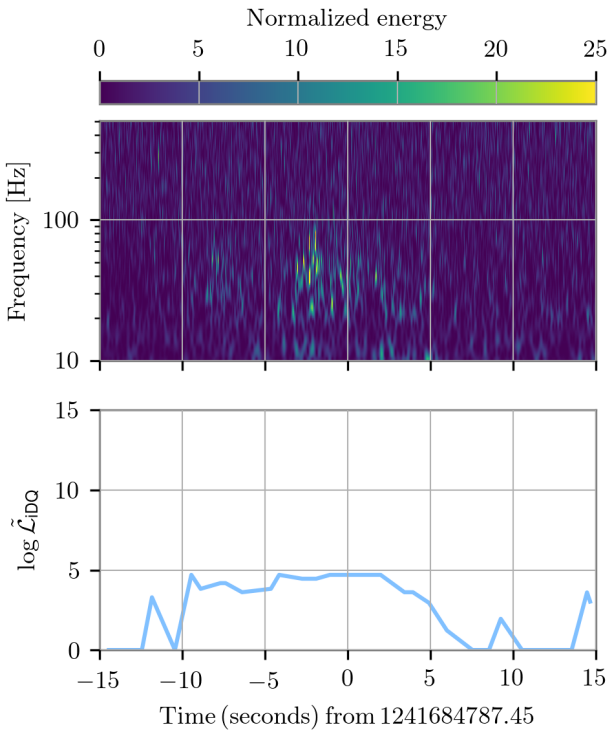

FIG. 5. Non-Gaussian noise in $h(t)$ caused by a thunderstorm in LIGO-Livingston, shown in a time-frequency spectrogram (above) and in the $\log \widehat{\mathcal{L}}_{\rm iDQ}$ timeseries (below). Excess noise is seen in the time-frequency spectrogram due to the thunderstorm, corresponding to periods of non-Gaussianity identified by iDQ.

ization estimates, and isn't addressed in this work. Allowing iDQ data products to be fed in asynchronously would also be beneficial for the offline case for the purposes of reranking candidates, since the filtering stage is costly compared to subsequent candidate rerankings.

Second, a more comprehensive study of applying iDQ information, including choosing an optimal window around the candidate event, whether to maximize or integrate over $\log \widehat{\mathcal{L}}_{\rm iDQ}$ across the window, and whether we can choose a less conservative approach and directly use $\log \mathcal{L}_{\rm iDQ}$ would be beneficial.

Additionally, any improvements to iDQ in performing its statistical inference may be reflected directly as improvements in VT in the GstLAL search. Many of the improvements within iDQ in O3 were targetted at improving the latency so that data products generated by iDQ could be available concurrently with calibrated $h(t)$. Moving forward, being able to provide iDQ a richer set of auxiliary information to perform its statistical inference such as time-frequency representations may improve iDQ's ability to detect non-Gaussian noise. This may also be combined with algorithmic improvements that would leverage a richer feature set.

Finally, the incorporation of iDQ was limited solely to single detector candidates, but it would be beneficial to apply this to coincident events as well. One complication from using $\log \widehat{\mathcal{L}}_{\rm iDQ}$ timeseries in coincident events arises due to not allowing candidates to be upranked from iDQ after the transformation is applied. Since on average there was a small contribution coming from each detector, this caused coincident events to be downranked a bit too strongly since we were effectively applying a penalty multiple times. This issue may end up being much less of a concern if iDQ information is also allowed to uprank during clean times.

We demonstrated an approach of folding in data quality information via iDQ into the GstLAL search without losing livetime and mitigating the possibility of vetoing gravitational-wave candidates. Since iDQ is available alongside $h(t)$, this gives a clear path to allow the inclusion of data quality information leveraging the detector's auxiliary state in low latency. This also moves us towards understanding how to optimally apply non-binary data quality information in a gravitational-wave search, which will be increasingly important as the detector network sensitivity increases in future observing runs.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation (NSF) through OAC-1841480, OAC-1642391, and PHY-1454389. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by NSF grants PHY-0757058 and PHY-0823459. Computations for this research were also performed on the Pennsylvania State University's Institute for Computational and Data Sciences Advanced Cyber-Infrastructure (ICDS-ACI). We thank Jolien Creighton, Keith Riles and Gabriele Vajente for their help in the review of the inclusion of iDQ information into GstLAL. R.E. is supported at the University of Chicago by the Kavli Institute for Cosmological Physics through an endowment from the Kavli Foundation and its founder Fred Kavli. J.C. is supported by NSF grant PHY-1912649. D.M. is supported by NSF grant PHY 14-54389, ACI grant 16-42391, and OAC grant 18-41480.

[1] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].

[2] B. Abbott *et al.* (LIGO Scientific, Virgo, Fermi-GBM, INTEGRAL), Astrophys. J. **848**, L13 (2017), arXiv:1710.05834 [astro-ph.HE].

[3] V. Savchenko *et al.*, Astrophys. J. **848**, L15 (2017), arXiv:1710.05449 [astro-ph.HE].

[4] B. Abbott *et al.* (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT), Astrophys. J. **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].

[5] K. P. Mooley, A. T. Deller, O. Gottlieb, E. Nakar, G. Hallinan, S. Bourke, D. A. Frail, A. Horesh, A. Corsi, and K. Hotokezaka, Nature **561**, 355 (2018).

[6] A. Farah, R. Essick, Z. Doctor, M. Fishbach, and D. E. Holz, (2019), arXiv:1912.04906 [astro-ph.HE].

[7] B. P. Abbott *et al.*, The Astrophysical Journal **850**, L39 (2017).

[8] T. Dietrich, M. W. Coughlin, P. T. Pang, M. Bulla, J. Heinzel, L. Issa, I. Tews, and S. Antier, (2020), arXiv:2002.11355 [astro-ph.HE].

[9] D. Coulter *et al.*, Science **358**, 1556 (2017), arXiv:1710.05452 [astro-ph.HE].

[10] C. D. Kilpatrick *et al.*, Science **358**, 1583 (2017), arXiv:1710.05434 [astro-ph.HE].

[11] B. Abbott *et al.* (LIGO Scientific, Virgo), (2019), arXiv:1908.06060 [astro-ph.CO].

[12] M. Fishbach *et al.* (LIGO Scientific, Virgo), Astrophys. J. **871**, L13 (2019), arXiv:1807.05667 [astro-ph.CO].

[13] B. P. Abbott *et al.* (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. Lett. **121**, 161101 (2018).

[14] R. Essick, P. Landry, and D. E. Holz, Phys. Rev. D **101**, 063007 (2020).

[15] B. P. Abbott *et al.*, The Astrophysical Journal **882**, L24 (2019).

[16] C. Messick *et al.*, Phys. Rev. **D95**, 042001 (2017), arXiv:1604.04324 [astro-ph.IM].

[17] S. Sachdev *et al.*, (2019), arXiv:1901.08580 [gr-qc].

[18] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Class. Quant. Grav. **33**, 134001 (2016), arXiv:1602.03844 [gr-qc].

[19] C. Pankow *et al.*, Phys. Rev. D **98**, 084016 (2018), arXiv:1808.03619 [gr-qc].

[20] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. X **9**, 011001 (2019).

[21] R. Essick, P. Godwin, C. Hanna, and E. Katsavounidis, In preparation (2020).

[22] Reed Essick (on behalf of the LIGO-Virgo Scientific Collaborations), https://gcn.gsfc.nasa.gov/gcn3/21509.gcn3 (2017).

[23] J. Aasi *et al.* (LIGO Scientific), Class. Quant. Grav. **32**, 074001 (2015), arXiv:1411.4547 [gr-qc].

[24] F. Acernese *et al.* (Virgo Collaboration), Class. Quant. Grav. **32**, 024001 (2015), arXiv:1408.3978 [gr-qc].

[25] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Astrophys. J. **875**, 161 (2019), arXiv:1901.03310 [astro-ph.HE].

[26] K. Cannon *et al.*, Astrophys. J. **748**, 136 (2012), arXiv:1107.2665 [astro-ph.IM].

[27] LVC, In preparation (2020).

[28] C. L. Mueller *et al.* (aLIGO), Rev. Sci. Instrum. **87**, 014502 (2016), arXiv:1601.05442 [physics.ins-det].

[29] F. Matichard *et al.*, Class. Quant. Grav. **32**, 185003 (2015), arXiv:1502.06300 [physics.ins-det].

[30] A. Staley, *Locking the Advanced LIGO Gravitational Wave Detector: with a focus on the Arm Length Stabilization Technique*, Ph.D. thesis, Columbia U. (2015).

[31] J. Graef Rollins, (2016), arXiv:1604.01456 [astro-ph.IM].

[32] A. Effler, R. M. S. Schofield, V. V. Frolov, G. González, K. Kawabe, J. R. Smith, J. Birch, and R. McCarthy, Class. Quant. Grav. **32**, 035017 (2015), arXiv:1409.5160 [astro-ph.IM].

[33] L. Nuttall *et al.*, Class. Quant. Grav. **32**, 245005 (2015), arXiv:1508.07316 [gr-qc].

[34] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Class. Quant. Grav. **35**, 065010 (2018), arXiv:1710.02185 [gr-qc].

[35] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), (2019), arXiv:1908.11170 [gr-qc].

[36] J. Aasi *et al.* (LIGO Scientific, VIRGO), Class. Quant. Grav. **32**, 115012 (2015), arXiv:1410.7764 [gr-qc].

[37] C. Cahillane *et al.* (LIGO Scientific), Phys. Rev. **D96**, 102001 (2017), arXiv:1708.03023 [astro-ph.IM].

[38] C. Hanna *et al.*, Phys. Rev. **D101**, 022003 (2020), arXiv:1901.02227 [gr-qc].

[39] K. Cannon, C. Hanna, and J. Peoples, (2015), arXiv:1504.04632 [astro-ph.IM].

[40] K. C. Cannon, Class. Quant. Grav. **25**, 105024 (2008).

[41] S. A. Usman *et al.*, Class. Quant. Grav. **33**, 215004 (2016), arXiv:1508.02357 [gr-qc].

[42] P. Godwin, "Low-latency statistical data quality in the era of multi-messenger astronomy," (2020).

[43] "aLIGO LLO Logbook," Accessed April 2020.

[44] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis, *8th Gravitational Wave Data Analysis Workshop (GWDAW 2003) Milwaukee, Wisconsin, December 17-20, 2003*, Class. Quant. Grav. **21**, S1809 (2004), arXiv:gr-qc/0412119 [gr-qc].