LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Technical Note **LIGO-T2200207–v2** 2022/06/09

Identifying Witnesses to LIGO Glitches Using Auxiliary Channels

Luis Gabriel C. Bariuan

California Institute of Technology
LIGO Project, MS 18-34
Pasadena, CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project, Room NW22-295
Cambridge, MA 02139
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

LIGO Hanford Observatory
Route 10, Mile Marker 2
Richland, WA 99352
Phone (509) 372-8106
Fax (509) 372-8137
E-mail: info@ligo.caltech.edu

LIGO Livingston Observatory
19100 LIGO Lane
Livingston, LA 70754
Phone (225) 686-3100
Fax (225) 686-7189
E-mail: info@ligo.caltech.edu

http://www.ligo.caltech.edu/

1 Introduction

The Laser Interferometer Gravitational Wave Observatory (LIGO) operates state-of-the-art ground-based gravitational wave (GW) detectors that made the first direct detection of GWs in 2015. Highly energetic astrophysical events generate GWs that propagate across spacetime. These *ripples* in spacetime are detected by LIGO due to the induced change in length of the interferometer arm. The resulting time-dependent data, strain $h(t)$, encodes the physical properties of the origin of GWs [1].

GWs reach the earth as an extremely faint signal inducing length changes smaller than a proton diameter in LIGO's 4 km long interferometer arms. To successfully detect such a small effect, the LIGO detectors are designed to be exquisitely sensitive. The extreme sensitivity makes these detectors vulnerable to terrestrial (environmental and instrumental) noise. Despite all the noise shielding, a particular type of noise which causes false alarms in GW search pipelines and adversely affects LIGO's sensitivity are short-lived, transient noise artifacts of largely unknown origins called "glitches". Glitches plague the main channel of the LIGO detectors which often results in false alarms [2]. Figure 1, adapted from [3], depicts representative examples of glitches across the 21 classes from the Gravity Spy Catalog [2].

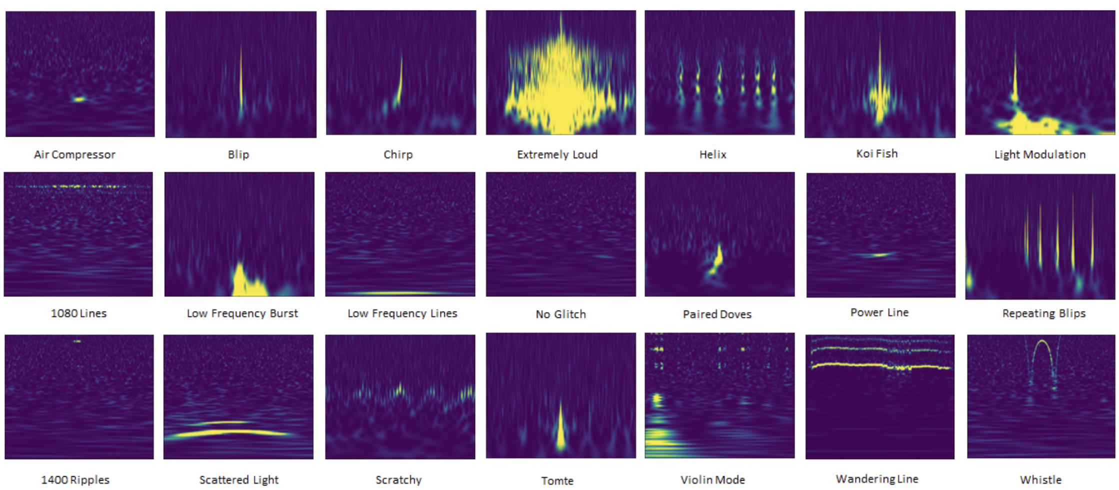

Figure 1: The figure, adapted from [3], depicts representative examples from 21 classes of glitches from the Gravity Spy Catalog [2] found in the strain data across the 3 observing runs of LIGO. The diversity in morphology across classes may hint towards various environmental and instrumental origins of these noise transients which result in false alarms in LIGO GW event search pipelines.

Along with the main channel, LIGO maintains a set of auxiliary channels to monitor the instrument and its environment using various sensors. These auxiliary channels may bear witness to the glitches. As a result, work has been done in associating noise transients within the main channel to recorded loud trigger events within the auxiliary channels to learn more about the origin of noise transients. By doing so, these witness auxiliary channels can then be used as veto generators to remove time segments in the main channel data that contain glitches [4].

Given the large number of auxiliary channels with an upper bound of $O(10^5)$ and a wide variety of glitches that may be present, machine learning methods were employed to find these witness auxiliary channels. Pipelines such as iDQ, a supervised low-latency glitch prediction method which uses algorithms like the Ordered Veto List algorithm (OVL) were developed to determine the probability of glitch occurrence within the main channel at time t given the data recorded in auxiliary channels [5, 4].

Recent work by [6], on the other hand, utilized unsupervised algorithms such as matrix and tensor factorizations to cluster triggers in a set of auxiliary channels as a method of finding valid channels that can serve as witnesses to the main channel glitches [6]. This method is unique from existing LIGO efforts in that, in addition to coming up with precise veto generators, it has the capacity to discover groups of auxiliary channels associated with certain glitches which may help domain experts generate hypotheses about the mechanism that causes those glitches, test those hypotheses, and potentially fix the glitches at their source.

In this work, our primary goal is to use the proposed method in [6] to explore different clustering models to find sets of channels that bear witness to glitches. Figure 2, top depicts a generic machine learning pipeline for reference and Figure 2, bottom depicts the end-to-end machine learning pipeline that we will create to find and validate channels which witness glitches. Our work may explore other features of the data associated with auxiliary channel triggers (an example of which was signal-to-noise ratio and peak frequency in [6]) and possibly morphologies of glitches to better understand which channels are candidate veto generators given a specific glitch type. We will use the rate of false and true positives as a function of the SNR threshold for triggers in the auxiliary channels to examine and select suitable veto generators. Finally, as a secondary and possibly long-term goal, we will examine characteristics of the sets of veto generators to determine which channels are associated with specific glitch classes which may then be communicated to domain experts to localize and potentially fix the glitch at their source.

2 Objectives

We present a general outline of our objectives for this project.

- **Design and implement clustering algorithms**

  In this work, we aim to explore different features of the data associated with auxiliary channels triggers (eg. signal-to-noise ratio and peak frequency in [6]) to devise clustering algorithms for identifying groups of channels that witness glitches.

- **Utilize clustering approaches to determine suitable veto generators.**

  Within this objective, we will compare the resulting number of false and true positives within the candidate witness channels as a function of the signal-to-noise ratio (SNR) threshold obtained from the candidate witnesses to determine if the channel can serve as a suitable veto generator.

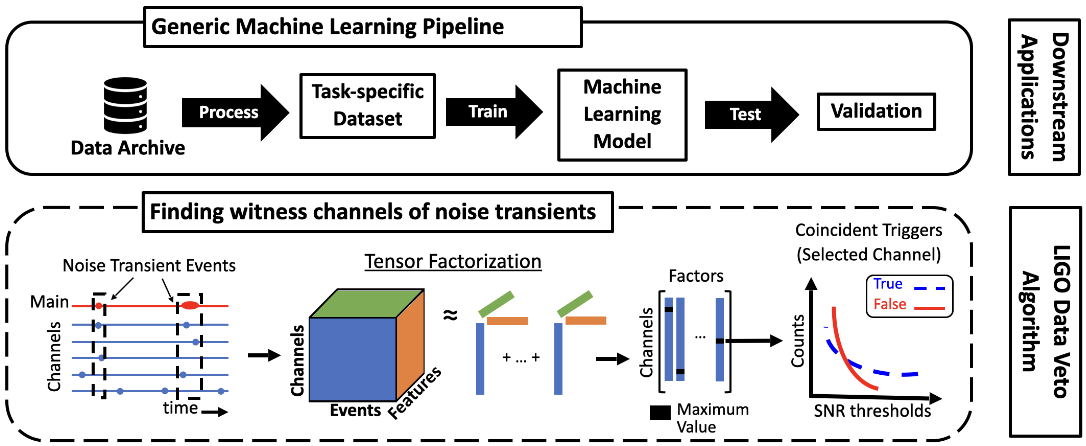

Figure 2: (*Top*) Diagrammatic representation of a generic machine learning pipeline, where a large dataset is processed to create a task-specific dataset. The smaller dataset is used to train a specified machine learning model. Finally, the model is validated through testing. (*Bottom*) Diagram of the proposed pipeline for finding witness channels of noise transients. Triggers within the main channel are associated with events in auxiliary channels. The triggers are clustered through algorithms such as tensor factorization to determine witness channels. The false and true positives within the witness channels are analyzed to determine if the candidate witness channels that bear witness to coincident triggers are suitable veto generators [3].

3 Approach

In this work, our goal is to obtain a set of auxiliary channels that serve as witnesses to glitches. These witness channels will serve to increase LIGO's existing toolkit to veto glitch-contaminated data.

Previous work developed tools such as procedural algorithms (eg. UPV and HVeto) and more recently, the iDQ, a low-latency supervised machine learning algorithm to veto glitches [7, 8, 4].

In this work, however, we propose to investigate extensions of recently-proposed matrix and tensor factorization based unsupervised machine learning techniques by experimenting with other parameters associated with auxiliary channel triggers to find different sets of veto generators [6]. We will investigate two particular directions:

- **Experimenting with different feature representations for the raw data.** Previous work [6], utilized matrix and tensor decomposition (which incorporated frequency as a tertiary tensor mode) to co-cluster triggers within the auxiliary channels. The triggers were described by the signal-to-noise of the raw data. In this work, we will experiment with other features of the auxiliary channels triggers like peak frequency, bandwith, duration, amplitude, etc to find different sets of witness channels that may serve as veto generators.
- **Incorporating domain knowledge and prior information to the clustering**

analysis. Prior information about individual channel behavior or channel to channel proximity/similarity) may aid in clustering glitches and creating models that maximize the number of useful channels extracted.

From this approach, our goal is to obtain a set of suitable auxiliary channels that may serve as veto generators.

4 Project Schedule

We present an outline of the projected project schedule.

- Before arrival: Complete background readings to gain familiarity with the approach and techniques utilized. Familiarization with coding toolkits by reading documentation.
- Week 1: Initial orientation to gain familiarity with the toolkits and manipulating LIGO datasets. Explore and discuss possible algorithms and clustering methods.
- Week 2-3: Explore an algorithm and clustering method. Tests results against previous work.
- Week 4: Optimize methods to create a pipeline for exploring other clustering methods.
- Week 4-6: Complete interim report. Experiment with using other features of the raw data during the clustering process.
- Week 7: If time permits, consider looking at vetoing glitches through morphologies by analyzing spectrograms.
- Week 8: Begin drafting final project summary, while possibly completing possibly unfinished tests from previous weeks. Exploration of extensions of the work to the morphologies may be continued, if time permits.
- Week 9-10: Continue to finish the final project summary drafts and presentation. Perform necessary revisions to complete the final report.

References

[1] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[2] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. K. Katsaggelos, S. L. Larson, T. K. Lee, C. Lintott, T. B. Littenberg, A. Lundgren, C. Østerlund, J. R. Smith, L. Trouille, and V. Kalogera. Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science. *Classical and Quantum Gravity*, 34(6):064003, March 2017.

[3] https://dcc.ligo.org/G2200883 v1.

[4] Reed Essick, Patrick Godwin, Chad Hanna, Lindy Blackburn, and Erik Katsavounidis. iDQ: Statistical Inference of Non-Gaussian Noise with Auxiliary Degrees of Freedom in Gravitational-Wave Detectors. *arXiv e-prints*, page arXiv:2005.12761, May 2020.

[5] R. Essick, L. Blackburn, and E. Katsavounidis. Optimizing vetoes for gravitational-wave transient searches. *Classical and Quantum Gravity*, 30(15):155010, August 2013.

[6] R. Gurav, B. Barish, G. Vajente, and E. Papalexakis. Unsupervised matrix and tensor factorization for LIGO glitch identification using auxiliary channels. *Association for the Advancement of Artificial Intelligence*, 2020.

[7] Tomoki Isogai. Used percentage veto for LIGO and virgo binary inspiral searches. *J. Phys. Conf. Ser.*, 243:012005, 2010.

[8] Joshua R. Smith, Thomas Abbott, Eiichi Hirose, Nicolas Leroy, Duncan MacLeod, Jessica McIver, Peter Saulson, and Peter Shawhan. A hierarchical method for vetoing noise transients in gravitational-wave detectors. *Classical and Quantum Gravity*, 28(23):235005, December 2011.