Discovering features in gravitational-wave data through detector characterization, citizen science and machine learning

S Soni[1], C P L Berry[2,3], S B Coughlin[2], M Harandi[4],
C B Jackson[5], K Crowston[4], C Østerlund[4], O Patane[6],
A K Katsaggelos[7], L Trouille[8,2], V-G Baranowski[9],
W F Domainko[9], K Kaminski[9], M A Lobato Rodriguez[9],
U Marciniak[9], P Nauta[9], G Niklasch[9], R R Rote[9], B Téglás[9],
C Unsworth[9] and C Zhang[9]

[1]Department of Physics, Louisiana State University, 202 Nicholson Hall Baton Rouge, LA 70803 USA

[2]Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Department of Physics and Astronomy, Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA

[3]SUPA, School of Physics and Astronomy, University of Glasgow, Kelvin Building, University Ave, Glasgow G12 8QQ, UK

[4]School of Information Studies, Syracuse University, 343 Hinds Hall. Syracuse, NY 13210, USA

[5]Information School, University of Wisconsin–Madison, Helen C White Hall, 600 N Park Street, Madison, WI 53706, USA

[6]Nicholas and Lee Begovich Center for Gravitational-Wave Physics and Astronomy (GWPAC), Department of Physics, California State University Fullerton, Fullerton, 800 North State College Blvd, CA 92831, USA

[7]Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

[8]Zooniverse, The Adler Planetarium, 1300 South Lake Shore Drive, Chicago, IL, 60605, USA

[9]Gravity Spy

E-mail: siddharthsoni22@gmail.com

Abstract. The observation of gravitational waves is hindered by the presence of transient noise (glitches). We study data from the third observing run of the Advanced LIGO detectors, and identify new glitch classes: ***Fast Scattering/Crown*** **and** ***Low-frequency Blips***. Using training sets assembled by monitoring of the state of the detector, and by citizen-science volunteers, we update the Gravity Spy machine-learning algorithm for glitch classification. We find that **Fast Scattering/Crown** is linked to ground motion at the detector sites is especially prevalent, and identify two subclasses linked to different types of ground motion. Reclassification of data based on the updated model finds that $\sim 27\%$ of all transient noise at LIGO Livingston belongs to **the Fast Scattering class**, **while** $\sim 8\%$ **belongs to the Low-frequency**

Blip class, making **them** the most frequent **and fourth most frequent** source**s** of transient noise at that site. Our results demonstrate both how glitch classification can reveal potential improvements to gravitational-wave detectors, and how, given an appropriate framework, citizen-science volunteers may make discoveries in large data sets.

1. Introduction

Advances in detector technology and in data collection drive discovery—access to new types of observations and more comprehensive data sets reveal previously unknown phenomena. A prime example is the operation of the Advanced LIGO [1] and Advanced Virgo [2] detectors which has enabled the first direct observations of gravitational waves [3]. Gravitational waves provide a means to study merging neutron star and black hole binaries, which would be almost impossible otherwise, revolutionising our understanding of these astrophysical objects [4–6]. Since the first gravitational-wave observation in 2015, the rate of detection has accelerated, reaching about one per week in the most recent observing run (O3) [7] **which ran from April 2019 until March 2020**. The wealth of data provided from ever-more sophisticated detectors enables scientific breakthroughs, but only if we can effectively use all of this data.

Gravitational-wave detectors produce many forms of data—not only observations of gravitational waves from astrophysical sources, but also data on the operational state of the detector and the status of their environments [8–10]. How to intelligently use big data sets is a difficult challenge: because of their size, it is not always feasible for experts to carefully examine entire data sets. One approach is crowd-sourcing analysis through citizen-science projects [11–13]; here volunteers perform activities such as image classification at a scale impossible for a small science team. Alternatively, analysis may be automated through use of machine-learning algorithms [14, 15]; these can be trained to reproduce complicated operations enabling computers to quickly analyze data. Both citizen-science and machine-learning approaches have been successfully applied to a wide range of problems. However, in both citizen-science and machine-learning approaches it can be difficult to discover new features in the data—citizen-science volunteers, being non-experts, may not recognize interesting new phenomena without guidance, while machine-learning algorithms can struggle with unexpected features in the data. We examine the problem of identifying new **transient** noise features (*glitches*) within data from the two LIGO gravitational-wave observatories in Livingston, Louisiana and Hanford, Washington.

LIGO data quality is adversely affected by glitches. These short-duration **bursts of** noise can mimic or overlap a true gravitational-wave signal, complicating the process of detecting **transient** signals and inferring the properties of the source of the gravitational waves [16–21]. Glitches can have environmental or instrumental origins [8, 22–24]. There are a wide range of different glitch types; Figure 1 shows a time–frequency spectrogram, an Omega-scan [25], of common glitch types. As the detectors undergo commissioning,
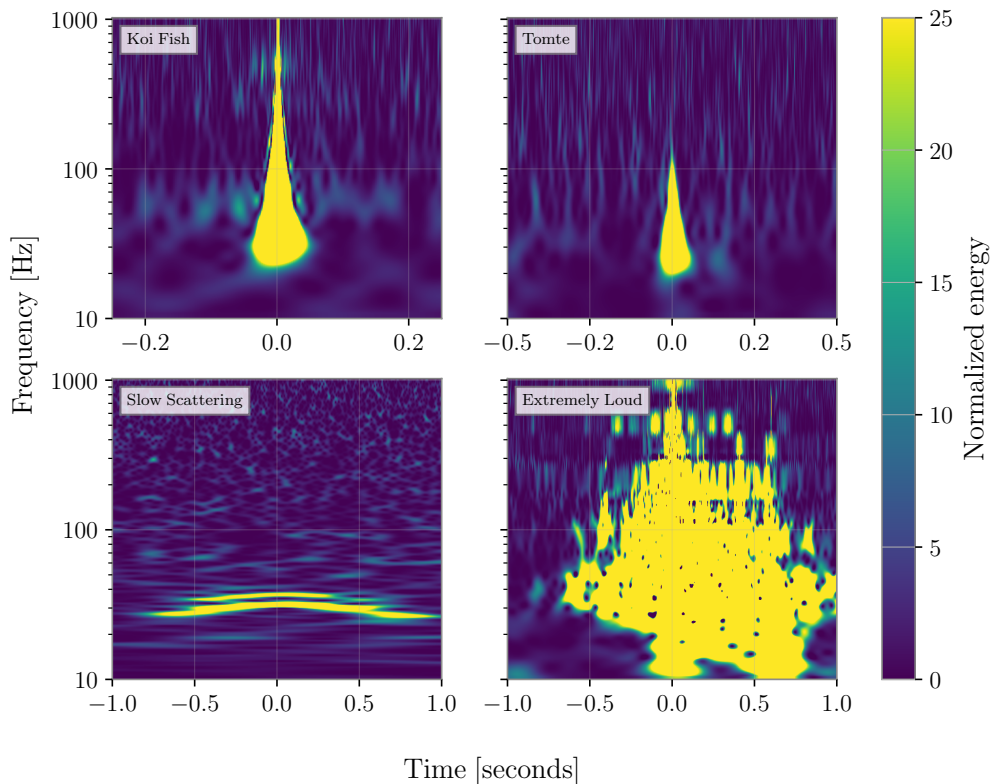
Figure 1: A selection of transient noise categories in gravitational-wave strain data. These spectrograms illustrate the variety of time–frequency morphologies of glitches in LIGO data. *Top left:* Koi Fish is short-duration broadband noise, usually with high SNR. *Top right:* Tomte is another short-duration noise category which occurred with a high rate at LIGO Livingston during O3. *Bottom left:* Slow Scattering (Scattered Light) appears as long-duration arches between 10 Hz and 120 Hz. *Bottom right:* Extremely Loud are very high SNR triggers that saturate the time–frequency plane and cause large drops in the astrophysical reach of the detector. While the rate of Slow Scattering goes up with ground motion close to the detector, the other transient noise categories have not shown any environmental coupling.

glitch classes can be eliminated as their sources are identified, but new classes may also arise (either because a change to a detector subsystem inadvertently introduces a glitch, or because improvements in detector sensitivity reveal a previously subdominant noise source). For instance, before the start of O3, a number of upgrades were applied to the detectors, chief among them an increase in the laser power and the addition of squeezed light [26]. These improvements increased the detectors' sensitivities by a factor of 2 in the most sensitive frequency band compared to O2 [7, 26]. However, this was accompanied by an increase in the glitch rate; at the Livingston Observatory the rate of glitches with signal-to-noise ratio (SNR) above 6.5 quadrupled [7, 24]. Identifying the origins of glitches is key to making the best use of gravitational-wave detector data.

Gravity Spy [27] is a project that combines the power of citizen science and machine

learning.‡ Using the Zooniverse platform, citizen-science volunteers are asked to sort glitches into a number of classes (including None of the Above and No Glitch). These classifications are used to build a training set used to train a convolutional neutral network (CNN) [28]. The volunteers are trained through a series of levels, starting with classifying glitches the machine-learning algorithm calculates as confidently belonging to a subset of easily distinguishable classes, and progressing to classifying glitches which received a low probability of belonging to a given class by the algorithm. The database of Gravity Spy classified glitches are routinely used in studies of the performance of the gravitational-wave detectors [29–34].

The glitch classes included in Gravity Spy were originally defined by LIGO experts based on their studies of the detectors, with additional classes suggested by volunteers during early testing [27, 35]. To aid in identifying new classes, Gravity Spy provides a search tool [36] (available to both volunteers and experts) that identifies glitches in the data set that are similar to a nominal example.§ The similarity search uses the machine-learning algorithm's feature space to identify morphologically similar glitches [36], where similarity is measured using the cosine distance between glitches in a feature space that is constructed by the CNN [37]. Here we study how Gravity Spy enabled the discovery of **two** new prevalent glitch classes—*Fast Scattering/Crown* and *Low-frequency Blips*—during the O3 run of Advanced LIGO.

Assessing LIGO data quality is a time-consuming process. As early as October–November 2018, commissioners at LIGO Livingston began to notice a new type of daytime noise that appeared as short-duration arches in time–frequency spectrograms; this came to attention of detector-characterization experts in March 2019, and the new glitch would eventually become known as Fast Scattering/Crown [38, 39]. Similarly, during O3 detector-characterization experts noticed a new population of short-duration, broadband noise transients that were similar to the established Blip class of glitches, but with a lower maximum frequency. This category came to be known as Low-frequency Blips [40]. We have identified a potential correlation between Fast Scattering/Crown glitches and instrument vibration, while for Low-frequency Blips (as for Blips [23]), the cause still remains a mystery. Both these types of transient noise were prevalent in the O3 data, impacting the data quality in different ways. Low-frequency Blips, due to their morphology, can be confused for short-duration gravitational-wave transients, such as signals from high-mass compact binaries hindering detection [24, 41]. Fast Scattering/Crown glitches transients on the other hand, negatively impacted the overall sensitivity of the detectors, and also overlapped with multiple detections, necessitating their excision from the data before the sources of the signals could be inferred [6, 7]. The output of our work is a classification of glitches from O3 that can be used in future

‡ Gravity Spy Zooniverse project gravityspy.org
§ Gravity Spy Tools gravityspytools.ciera.northwestern.edu/

studies of LIGO data.

Identification of new glitch classes is an important step in improving the LIGO detectors. We show how Gravity Spy has enabled the independent identification of a new glitch class via two complementary methods. The first is the traditional approach of carefully studying the state of the LIGO instruments and their environments to identify potential noise sources [9]. The second is through Gravity Spy citizen-science volunteers exploring the data from the detectors to identify new patterns. Despite the citizen-science volunteers not having access to auxiliary information about the instruments or environment, an analogous new glitch type was found by both methods, highlighting the potential of citizen-science volunteers for making discoveries based upon their own in-depth investigations. We show how the updated glitch classification is applied to O3 data, demonstrate how this new class is related to ground motion at the observatory sites, and discuss how future discoveries of new **transient** noise features could be facilitated.

The success of Gravity Spy in enabling the identification of the Fast Scattering/Crown glitch class may serve as a case study for how citizen-science volunteers can make discoveries if supplied with appropriate training, analysis tools and expert feedback. Given the increasing prevalence of large scientific data sets, such citizen-science investigations could lead to breakthroughs in many other fields.‖

2. Methods

In Sec. 2.1 we introduce how glitches are identified in LIGO data, and highlight the need to assemble training sets for new glitch classes. In Sec. 2.2 we provide examples of how new glitches are identified by LIGO detector-characterization studies, while in Sec. 2.3 we discuss the equivalent practices of the citizen-science volunteers. During O3, both approaches lead to the convergent identification of a new glitch class linked to ground-motion at the detector sites, named Fast Scattering by the LIGO experts and Crown by the volunteers; we concentrate on this class, while also discussing a second class identified by the experts, Low-frequency Blips.

2.1. Detector characterization & Gravity Spy

Detector characterization is an umbrella term for monitoring data quality, investigating the origin of noise in the detector, enhancing the detector performance by mitigating noise, and validating gravitational-wave signals [24]. Apart from Gravity Spy, several other software tools such as Hveto [42], gwpy [43], gwdetchar [44], Omicron [45, 46] and iDQ [47] are used in detector characterization **for studying transient noise properties**. The output from noise-analysis tools as well as the state of various

‖ Since 2007, the Zooniverse platform has launched over 200 projects across disciplines from astronomy to zoology, cancer research to climate science, and the arts and the humanities.

subsystems of the detector, is shown on a set of webpages known as LIGO Summary pages [48]. Detector characterization is essential to gravitational-wave astronomy.

Potential glitches in the LIGO data are found by event trigger generators that identify excess power in the data stream. All the **noise transients** analyzed in this paper were detected by the Omicron algorithm [45, 46]. The Omicron algorithm annotates each of these detected triggers with specific characteristics such as event time, peak frequency, central frequency and SNR. Other than these characteristics, the time–frequency structure of a glitch is the most important feature that separates it from other transient noise sources. The glitch morphology of the trigger can be visualized in a time–frequency spectrogram commonly known as Omega-scans [25, 44]. These Omega-scans are used frequently in data-quality studies to establish potential noise correlations between different parts of the detector [49]. Examples of the range of morphologies are shown in Figure 1; these examples are some of the most common glitch classes found in O3. Different sources of transient noise saturate different areas of the time–frequency plane. These distinguishing patterns can be used to categorize glitches.

2.1.1. Machine-learning algorithm & training Gravity Spy uses a CNN to classify glitches into categories [27, 28]. Given an image, a trained CNN architecture will recognize different aspects in the image and assign varying degrees of weights and biases to these predictive features. For Gravity Spy, each input image is a combination of four time–frequency spectrograms of 0.5 s, 1.0 s, 2.0 s and 4.0 s duration centered on the same transient [27, 28]. Some of the noise categories such as the Blip are sub-second in duration whereas others like Slow Scattering last 3–4 s. This concatenation of multiple duration images ensures that relevant features for each class are captured by the training set. For each image fed to the model, the algorithm will calculate a probability score (confidence) for each class, and the image is assigned the class with the highest associated confidence.

The CNN architecture of Gravity Spy [28] consists of multiple layers. There are four convolutional layers to extract features, each following by a max-pooling and a rectified linear unit (ReLU) activation layer. After these layers, there are a fully connected layer and a final softmax layer. The number of nodes in the final softmax layer is equal to the number of glitch classes, and the output score of the softmax layer is used a probability distribution across glitch classes [27]. The CNN is trained using a cross-entropy loss function.

The output of the softmax layer is given by:

$$o_c^i = \frac{\exp(w_c \cdot x)}{\sum_{c=1}^{C} \exp(w_c \cdot x)}, \tag{1}$$

where o_i^c represents the probability score that the image i belongs to the glitch category c, C is the total number of glitch classes ,and w represents the weight vector that connects the output of previous layer to the cth node in softmax layer. Weights as neural network parameter represents the

strength of connection between different nodes. While training, the CNN model optimizes the cross entropy loss function:

$$\mathcal{L} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_c^i \log o_c^i, \tag{2}$$

where y_c^i is an indicator function for true label of the image i that equals 1 if the image belongs to class c and equals 0 for all other classes, and N is the total number of training samples.

Key to the performance of the CNN is the training set. The Gravity Spy training set is composed of time–frequency spectrogram images **of different glitches that have been selected by experts and volunteers to exemplify the various classes**. The total number of training images for each category is a function of how frequently it appears in the data. Blip and Slow Scattering, for example, have more training data as they occur with a relatively higher rate: Blips occurred at a rate of about 2 per hour in each of the two LIGO detectors during the second observing run (O2) [24]. The previous training set contains a total of 7932 training images distributed among 22 classes [35]; the Gravity Spy model trained on this data set was used during the first two LIGO observing runs as well as during O3.¶

Here, we discuss updating the Gravity Spy training set to account for changes to the data quality in O3. Our new training set includes two new glitch classes: Fast Scattering/Crown and Low-frequency Blips. **Adding the new glitch classes enables us to use the machine-learning algorithm to identify examples of these glitches, so that they can be studied in detail.** Following O3, we retrained the CNN using the updated training sets, and reclassified the data set with this new model (**statistics for the most common glitch classes as identified by our old and new machine-learning models are given in Table** 1).

2.2. Overview of O3 data quality & expert investigations

During O3, Gravity Spy classified 195445 and 125256 Omicron triggers from LIGO Livingston and LIGO Hanford, respectively, into one of the 22 pre-O3 classes with confidence above 90%. **In comparison to LIGO Hanford, the Livingston Observatory experiences a higher rate of glitches due to its better sensitivity in roughly** 10–60 Hz **band and worse ground motion conditions near the site.** Noise due to light scattering dominated at both Livingston and Hanford with 40.3% and 46.6% of the triggers being classified as *Scattered Light* at the two sites respectively (see Table 1). This is not surprising as ground motion is a persistent source of noise at both **of** the observatories. During high ground motion, the vibrating hardware in the detector

---

¶ The old 22 glitch classes are: 1080 Lines, 1400 Ripples, Air Compressor, Blip, Chirp (simulated gravitational-wave signals), Extremely Loud, Helix, Koi Fish, Light Modulation, Low-frequency Burst, Low-frequency Line, Paired Dove, Power Line, Repeating Blips, Scattered Light (which we will rename to Slow Scattering), Scratchy, Tomte, Violin Mode Harmonic, Wandering Line, Whistle, No Glitch (for when no glitch is visible), and None of the Above (for glitches that do not match the other categories).

can reflect any stray light incident on it back into the main beam; the relative motion between the vibrating surface and the main optic's mirrors introduces a time-varying phase to this stray light which shows up as noise [**33**, **50**–**52**]. At Livingston, the glitch categories Tomte and None of the Above are the most dominant after Scattered Light while at Hanford, Low-frequency Burst and Extremely Loud are the second and third most frequent source of transient noise respectively. The frequency of different glitch classes varies both between the two observatories and between observing runs.

Two major clues pointed towards the need to recognize at least one new glitch category during O3:

(i) A high number of None of the Above at Livingston (around 12% of all the classified triggers; see Table 1).

(ii) There were triggers classified as Scattered Light with a short duration (0.25–0.3 s) glitch morphology compared to the prototypical scattering arch shown in the bottom left of Figure 1.

The SNR and duration distribution of the Livingston glitches classified as Scattered Light hinted towards the presence of another population, one with a shorter duration and lower SNR than the conventional Scattered Light glitches [33]. The time–frequency spectrograms of a number of None of the Above glitches appeared similar to the new population of Scattered Light. **The differences from the traditional Scattered Light suggest a separate physical origin for this new population.** A time comparison showed that days with a high number of None of the Above at Livingston were also the days with high ground motion, which is the main contributor of noise due to light scattering. These investigations pointed towards the presence of a new class of glitches with a coupling to ground motion similar to, but distinct from, the previously observed Scattered Light. This new glitch category came to be known within LIGO as *Fast Scattering* due to its short-duration arches, while the longer duration Scattered Light became known as *Slow Scattering*.

Fast Scattering glitches are subsecond-duration noise transients that affect LIGO data quality in the 20–60 Hz range. A comparison of Slow Scattering and Fast Scattering is shown in Figure 2. The rate of Fast Scattering **is** found to be strongly correlated with ground motion in 0.1–0.3 Hz and 1–6 Hz frequency bands [**53**]. Human activity, thunderstorms, logging and trains (in the case of Livingston) near the site are major causes of ground motion in this band. Fast Scattering is more common at Livingston than at Hanford due to differences in ground motion as well as the detector sensitivity [54].

In addition to Fast Scattering, another addition to the Gravity Spy classification set is the Low-frequency Blip. Blips are sub-second duration broadband glitches that show up frequently at both the LIGO detectors [23]; they are yet to show any environmental coupling. Due to their morphology, these transients are considered especially harmful to search for high-mass binaries [29, 55, 56]. Efforts to understand their origin have largely been unsuccessful though many sources have been ruled out [57–59]. During
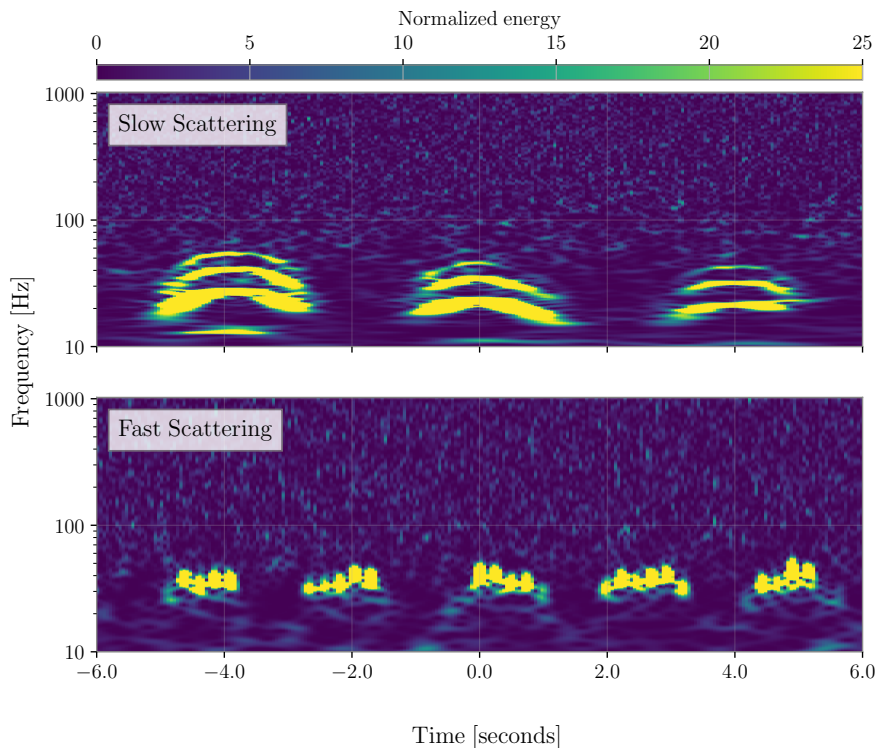
Figure 2: Comparison of glitches caused by light scattering. *Top:* Slow Scattering shows up as long duration arches and is correlated with an increase in ground motion in the earthquake (0.03–0.1 Hz) and microseismic (0.1–0.3 Hz) band. *Bottom:* Fast Scattering often consists of clusters of arches with a higher frequency, and is more common during an increase in anthropogenic band ground motion (1–6 Hz). The noise transients shown here are during the passing of a train close to the Livingston site when Fast Scattering triggers occur in clusters.

O3, we noticed many glitches with duration similar to Blips but with much lower peak frequencies were being classified as Blips or None of the Above. Blips usually have the frequency range $\sim$ 32–2048 Hz, while the majority of transients from the new population are in the band $\sim$ 32–128 Hz. Therefore, we created a new Gravity Spy category for these glitches.

2.2.1. Expert-selected training set To update the Gravity Spy model to include O3 glitches, training examples were needed. Our search for Fast Scattering began with finding days in O3 with high ground motion in the 1–6 Hz frequency band at Livingston. We concentrated on Livingston as this noise is more prevalent there. Next, for these days we randomly selected 400 triggers classified as Scattered Light by the old model with Q-value between 8 and 14. **The Quality factor or Q-value is defined as the ratio of the central frequency of the glitch to it's bandwidth:**

$$Q = \frac{f_0}{\sigma_f} \tag{3}$$

The Q-value intervals for Fast Scattering and Slow Scattering are disjoint due to the difference in glitch morphology. These 400 triggers formed the Fast Scattering training data. For Low-frequency Blips, we selected 630 glitches classified as Blips during O3 with peak frequency between 10 Hz and 50 Hz. Low-frequency Blip and Tomte glitches affect the LIGO data quality in a similar frequency band of ~ 30–120 Hz. To ensure that the machine-learning classifier does not get confused between these two classes, we also increased the number of Tomte in the training set. We found new examples of Tomte by randomly sampling 300 glitches from the O3 data, and added these to the training set. We also added 150 examples of Slow Scattering from Hanford, as some examples of Slow Scattering from Hanford were being misclassified as Extremely Loud by the old model. All the data added to the training set were previously classified with confidence above 0.9. As a final change, we removed the training data associated with the None of the Above category. None of the Above was primarily intended to enable our citizen-science volunteers to flag new glitch classes; having now established the major classes for the O3 set it was not needed (although it would be reintroduced for future observing runs). The new O3 training data set contains 9631 training examples distributed over 23 classes.

The Gravity Spy algorithm was trained on this data set for a total of 100 epochs, with a 87.5:12.5 split between training and validation set. At the end of the training, the training and validation set accuracies were 99.9% and 98.8% respectively.

2.3. Volunteer methods

Gravity Spy citizen-science volunteers have the ability to suggest new glitch classes. To better understand the volunteers' process for submitting a new glitch proposal, we conducted interviews with 9 Gravity Spy participants. Six one-hour interviews were carried out via video conference. Three other interviews were conducted via email. The interviews provided insights about work practices e.g., using computational tools, and interactions e.g., collaboration with other Gravity Spy participants when investigating and curating a possible new glitch. The Gravity Spy Zooniverse platform includes a forum where volunteers can discuss and ask questions of experts. In addition to the interviews, we extracted all comments from the Gravity Spy forum, and used these to supplement interviewees' stories, focusing on the collaboration, articulations of work and how they coordinate action through the Gravity Spy's discussion interfaces.

2.3.1. Citizen-science volunteer investigations To propose a new glitch class, volunteers must show that their proposed class is distinct from existing classes and abundant enough in the LIGO data to warrant attention. They do so by developing a new glitch proposal following a template outlined by our science team (see Figure 3). Glitch proposals describe the morphological characteristics of a glitch and link to collections of exemplary subjects. Proposals are vetted by moderators (science-team members or experienced volunteers) prior to submission to ensure they include all the necessary
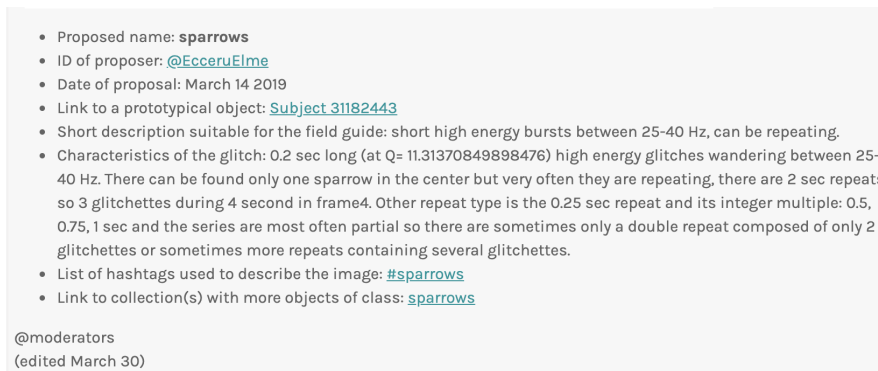
- Proposed name: **sparrows**
- ID of proposer: @EcceruElme
- Date of proposal: March 14 2019
- Link to a prototypical object: Subject 31182443
- Short description suitable for the field guide: short high energy bursts between 25-40 Hz, can be repeating.
- Characteristics of the glitch: 0.2 sec long (at Q= 11.31370849898476) high energy glitches wandering between 25-40 Hz. There can be found only one sparrow in the center but very often they are repeating, there are 2 sec repeats so 3 glitchettes during 4 second in frame4. Other repeat type is the 0.25 sec repeat and its integer multiple: 0.5, 0.75, 1 sec and the series are most often partial so there are sometimes only a double repeat composed of only 2 glitchettes or sometimes more repeats containing several glitchettes.
- List of hashtags used to describe the image: #sparrows
- Link to collection(s) with more objects of class: sparrows

@moderators
(edited March 30)

Figure 3: An example of a glitch proposal for a Sparrow class, which would eventually become the Crown class. **After collecting volunteer input,** moderators submit proposals and include: a link to an exemplary image, a short description of the glitch **class**, a link to the hashtag and link to collections containing ~ 75 examples of the glitch **class**.

information and to verify that the proposed class is novel. Significant work goes into developing these proposals including searching for unusual glitches while classifying; curating subjects using hashtags (#) in forum discussions; building collections; using the similarity search tool; describing glitch characteristics in discussion posts, and exploring the LIGO aLog.[+] Interviews and trace-data analysis suggest that volunteers rely on a range of search strategies. For instance, one volunteer reported that they relied on their photographic memory and hashtags to find potential new glitches, while another volunteer combined hashtag searches, the similarity search and collections. This work often continues after the submissions of a proposal as volunteers wait for science-team feedback.

Ideas for new glitch classes often emerge during volunteers' regular classification activity where they come across (None of the Above) glitches that do not fit any of the known classes. However, there are tens of thousands of None of the Above subjects in the Gravity Spy data, and volunteers apply a range of strategies to group these subjects into coherent classes. Figure 4 illustrates the use of hashtags, comments and collections.

Many participants will mark an unknown subject with #none_of_the_above and sometimes propose a name for this potential new class (see top post in Figure 4), which was part of the Sparrow proposal shown in Figure 3. Subsequently, other volunteers might comment on posts as we see in the second post in Figure 4, where an experienced participant describes the differences between three glitches under investigation (crown-sparrow, sparrows and serpents) pointing to specific morphological features of each glitch e.g., "sparrow serpent subset has 7 glitches during 4 seconds." Such comments may help newcomers learn which glitches are being curated and the characteristics of their shape [61].

[+] aLogs are public records of issues at the LIGO detector sites: Hanford aLog alog.ligo-wa.caltech.edu/aLOG/ and Livingston aLog alog.ligo-wa.caltech.edu/aLOG/.
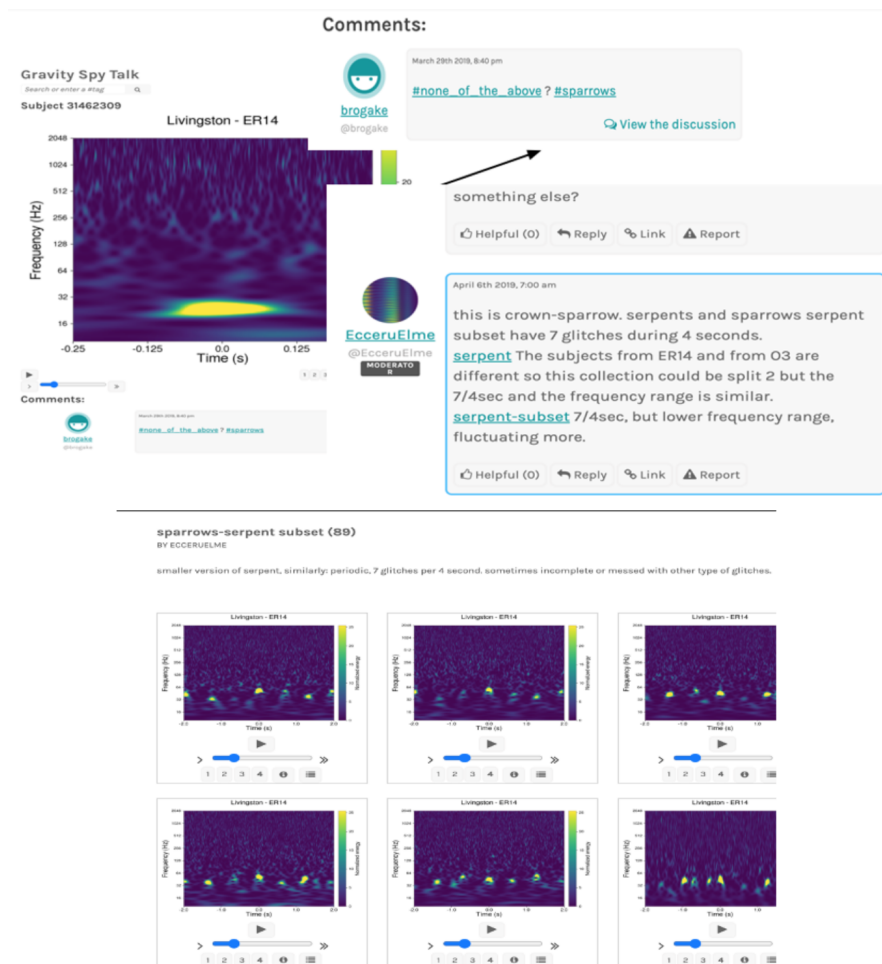
Figure 4: Instance of one volunteer curating glitches using hashtags and another articulating the differences between three glitch classes (*top*) and a collection of subjects which the collection creator labelled sparrow-serpent subset (*bottom*). There are 89 subjects that the volunteer believes matches this class. ER14 and O3 refer to when the data were taken: Engineering Run 14 and Observing Run 3, respectively. ER14 immediately preceded O3, and was a time to test the running of the detectors before making science-quality observations [60]; the behavior of the detectors should be similar in ER14 and O3.

The Zooniverse platform offers the ability to create collections of subjects, which can be public or private; some volunteers will build collections of related glitch images which they populate with subjects they find during their regular classification, by searching for particular hashtags or keywords, adding glitches from other collections, or using the similarity search. Participants often use hashtags to search for subjects with similar names. Yet, we do not find common policies about tag use and how to reconcile conflicting naming practice. It is not uncommon to find deliberations about whether a subject belongs to e.g., a #crown, #sparrow or #serpent class.

Advanced volunteers rely on external resources such as the similarity search to

curate subjects [62]. The similarity search tool can be used to display subjects which appear most similar to a prototypical glitch. This process supplements the manual and serendipitous search for similar subjects. The output of the similarity search includes auxiliary data e.g., peak frequency and duration which volunteers use to determine whether a subject fits within a collection. As evidenced in the glitch proposals, volunteers download and analyze auxiliary data to determine whether subjects returned from the search fit within the numeric bounds of subjects already within the collection. Additional outputs of the similarity search are also used to provide evidence of the glitches and temporal prevalence. Distributional charts (another output of similarity search) were posted and synthesized by a volunteer who noted, "(1) crown cluster started around mid-March 2019, (2) . . . the same problem was present already in mid-December 2018, (3) invisible cluster (1e+20 subjects, not on Gravity Spy) might be similar, and (4) between mid-February 2019 and beginning of July 2019 there seem to be two larger time clusters of crowns (couldn't feel the gap yet)".* Here the volunteer notes that Crown glitches become more common after the start of O3, that the machine-learning algorithm found them similar to pre-existing Scattered Light glitches, and that the prevalence of Crowns changes with time (following the prevalence of ground motion at the sites).

Finally, some volunteers reference aLogs in comments. Written by LIGO scientists, aLogs provide insights into the source of glitches. One volunteer composed an archaeology of aLog posts related to a potential new glitch named Crown.

> **Summary:** Shaker injections at End-Y produce noise similar to 'daytime noise' (noise that correlates with increased anthropogenic motion during daytime).
>
> **Characteristic shape:** As noted by Josh and Andy in alog 44803, the daytime noise that we see when anthropogenic ground motion is increased has a characteristic shape with both 'slow' and 'fast' arches. There are a few examples in the linked alog, and the first attached image file shows this same characteristic shape during a time when a train was passing.

These posts help to inform volunteers on potential details to look for in assembling collections, and also give an indication of changes in the state of the detectors which could impact the glitches present in the data. This aLog [38] points out the source of the Crown glitch at the LIGO detector site appearing in the Gravity Spy data and its characteristic shape—this volunteer-curated class is equivalent to the Fast Scattering glitch class.

After compiling evidence in a glitch proposal, volunteer suggestions for new glitch classes are reviewed by the LIGO detector-characterization experts. The experts check for possible origins of the glitch, and provide feedback after evaluating the proposal. In some cases, the proposed class may be a subset of an existing class or may be extremely

---

* Glitches uploaded to Zooniverse have a numerical identification; for glitches in the Gravity Spy database yet to be uploaded to Zooniverse, the similarity search returns an identification number 1e+20.

localized in time (for example, the proposed Falcon class only occurred in a two-hour period on 20 June 2017 when there was an unreplicated issue with the Hanford detector), and so a new class may not be justified. If a new prevalent glitch class is identified, the volunteer-curated collection becomes the seed of a training set for the CNN.

2.3.2. Volunteer-selected training set The Volunteer O3 training data set contains 9496 training examples distributed over 23 classes, with 265 Crown examples: to aid the comparison between the proposed Fast Scattering and Crown classes, we include the expert-curated Low-frequency Blip examples, the additional Tomte examples and remove the None of the Above class for this model too. As for the Expert training set, the machine-learning algorithm was trained on this data set for a total of 100 epochs, with a 87.5:12.5 split between training and validation set. At the end of the training, the training and validation set accuracies were 99.2% and 97.5% respectively.

3. Results

Following the methods discussed above, we have updated the Gravity Spy machine-learning algorithm using two new training sets. The first training set was assembled by LIGO detector characterization experts (Sec. 2.2) using knowledge of the detectors status; this included two new classes: Fast Scattering and Low-frequency Blip. The second training set was assembled by citizen-science volunteers (Sec. 2.3) following their investigations of spectrograms; this included the new class of Crown. We identify the proposed Fast Scattering and Crown classes as representing the same noise source caused by ground-motion at the detector site. Using the two versions of the retrained machine-learning algorithm, we analyzed the O3 data from Livingston and Hanford to classify the glitch populations, and compared the models derived from the Expert and Volunteer training sets.

3.1. Performance of the O3 models

Before reclassifying the O3 data with the new models, we need to check that the newly trained model successfully identifies the new classes (Fast Scattering and Low-frequency Blips for the Expert model) with high confidence levels. To ensure this, we reclassified 100 triggers labeled as Scattered Light by the previous model on days with high ground motion in the anthropogenic band and with Q-value between 8 and 14. The selected range of Q-values ensured that they were Fast Scattering, and the new Expert model classified all of them as such. Similarly, we reclassified 80 Blips with peak frequency between 10 Hz and 50 Hz, and the new model classified 79 of them as Low-frequency Blips and one as Tomte. We visually inspected all the reclassified triggers as a further measure of confidence. Correctly recognizing the newly added glitch categories is not enough: the new model should also not misclassify other glitch categories.
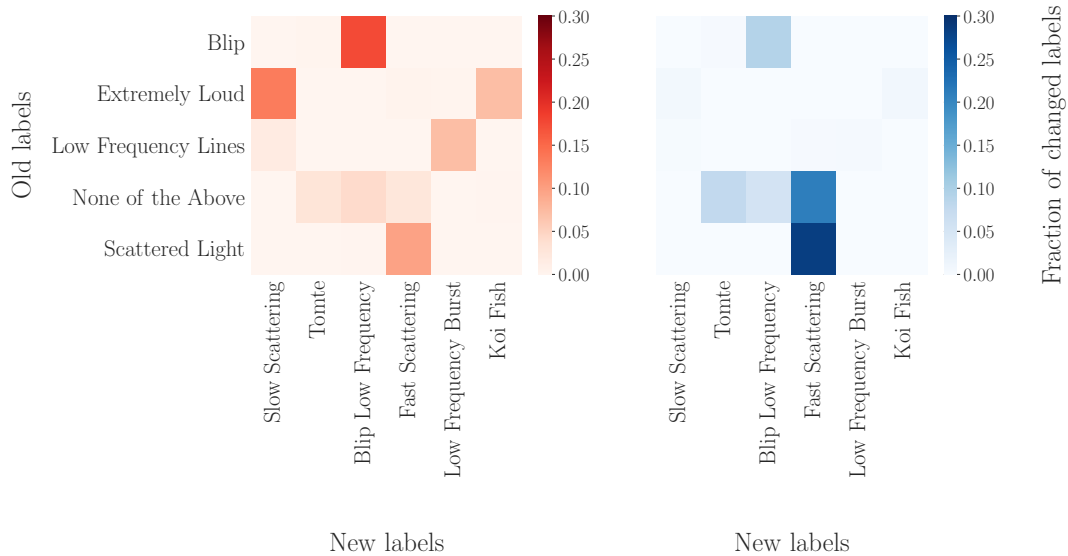
Figure 5: Heatmaps showing changes of labels following reclassification by the Expert model for Hanford (*left*) and Livingston (*right*) data. The change of labels was restricted to a handful of classes, and we only show the classes that contributed to at least 5% of the changes in labels on reclassification.

To probe more in-depth and to check the overall performance of this model, we reclassified 20% of O3 data at Hanford and Livingston with the new model to determine:

(i) Is there a large change in the confidence assigned to data by the new model compared to the old model?

(ii) What percentage of data is labelled with a new glitch category?

(iii) What is the overall movement of labels from the old classification to the new classification?

A comparison of the confidence distribution of the reclassified sample with the old classification did not show any significant difference for either of the detectors. At Livingston, the new model classified 42% of the triggers in the sample with a new label, while at Hanford, only 11% of the data was given a new glitch category. This is expected since Fast Scattering is more frequent at Livingston, and for reasons we do not yet understand, we have observed a higher rate of Low-frequency Blips there.

Figure 5 illustrates the main changes in glitch classification between the old and new Expert models. As expected, the majority of newly classified Fast Scattering at both detectors was labelled earlier as either Slow Scattering (previously named Scattered Light) or None of the Above. Similarly most of the Low-frequency Blips were previously classified as either Blips or None of the Above. At Hanford, some of the triggers previously classified as Extremely Loud are now labelled as Slow Scattering by the new model. Extremely Loud does not have a typical morphology similar to Slow Scattering, being more defined by its high SNR, but visual inspection of the reclassified examples confirm them as Slow Scattering. As shown in Figure 1, glitches

with a lot of saturation in the time–frequency plane between 10–1000 Hz are classified as Extremely Loud. A number of Koi Fish glitches occurring within a short duration from each other can saturate the time–frequency plane as well, and hence impersonate Extremely Loud glitches. Due to this thin barrier between the two glitch categories from a morphological point of view, the movement from Extremely Loud to Koi Fish at Hanford is not surprising. The results indicate that the updated model is classifying glitches as expected.

A summary of the most prevalent glitches, as classified by the old and new models is presented in Table 1.

3.2. Fast Scattering at LIGO Livingston

The primary motivation behind the retraining and reclassification was to generate a large data set for improved noise and statistical analysis, and this abundance of data provided us the resource to achieve this. We performed detailed investigations using the results of the Expert model; using this we found 55279 glitches labelled as Fast Scattering with confidence above 90% at Livingston. To study the origins of Fast Scattering, we focused on the month of February 2020 when 14536 triggers were classified as Fast Scattering with $\geq 90\%$ confidence. Before this reclassification, we were aware of the 4 Hz Fast Scattering that correlates strongly with ground motion in the anthropogenic 1–6 Hz band. In the reclassified data set, we also found a strong presence of a 2 Hz Fast Scattering population with a dependence on ground motion in 0.1–0.3 Hz microseismic band. Fast Scattering **glitches** associated with these two distinct types of ground motion are shown in Figure 6. Ocean waves in the Gulf of Mexico are the primary source of microseismic noise at Livingston **which** is usually associated with Slow Scattering [33]. The microseismic coupling **(and the reduction in Slow Scattering discussed next)** explains the overall high rate of Fast Scattering in February, as microseismic ground motion is higher at Livingston during winter.

The LIGO detectors use a four-level suspension used to reduce the effects of ground vibration [63]. This consists of two parallel chains: the main chain, which includes the test mass as its bottom layer, and the reaction chain (RC), which provides an isolated set of masses that can apply control signals to the main chain [1, 64]. During an increase in ground motion in the earthquake (0.03–0.1 Hz) and microseismic (0.1–0.3 Hz) bands, the relative motion between these two chains increases. This relative motion was observed to be highly correlated with an increase in the rate of Slow Scattering glitches. To try to mitigate the impact of ground motion, a technique known as RC tracking was implemented at the detectors in the first half of January 2020 [33]. **RC tracking reduces the relative motion between the two chains by sending a part of the control drive originally applied to the penultimate stage, to the top (R0) stage of the suspension.** This reduced the rate of Slow Scattering by a factor of ~ 100 for microseismic ground motion above 1000 nm s^{-1} at

Table 1: Most common glitch classes in LIGO Livingston (*top*) and LIGO Hanford (*bottom*) O3 data as labelled by the old and new (Expert and Volunteer) Gravity Spy models with confidence above 90%. Percentages indicate the fraction of all glitches analyzed from the given observatory. The Scattered Light class was renamed Slow Scattering in the new models. The Expert model introduces the new classes Fast Scattering while the Volunteer model introduces Crown. Both of the models include the new Low-frequency Blip class, and neither includes a None of the Above class.

	Livingston		
Rank	Old	Expert	Volunteer
1	Scattered Light 40.3%	Fast Scattering 27.1%	Slow Scattering 31.5%
2	Tomte 13.9%	Slow Scattering 23.2%	Crown 24.0%
3	None of the Above 12.2%	Tomte 19.3%	Tomte 18.2%
4	Blip 5.9%	Low-frequency Blip 7.7%	Low-frequency Blip 6.8%
5	No Glitch 5.2%	Extremely Loud 3.6%	Extremely Loud 3.6%
Total	195445	203469	216423
	Hanford		
Rank	Old	Expert	Volunteer
1	Scattered Light 46.6%	Slow Scattering 46.9%	Slow Scattering 52.9%
2	Low-frequency Burst 13.3%	Low-frequency Burst 15.9%	Low-frequency Burst 14.6%
3	Extremely Loud 11.0%	Extremely Loud 9.0%	Extremely Loud 9.0%
4	Koi Fish 7.2%	Koi Fish 7.0%	Koi Fish 6.8%
5	Blip 6.9%	Blip 4.9%	Blip 4.2%
Total	125256	121632	121113
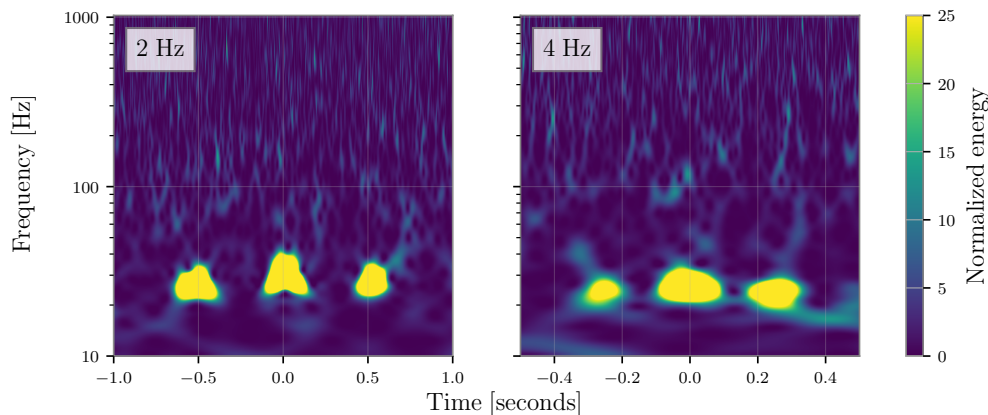
Figure 6: Spectrograms showing examples of the two subclasses of Fast Scattering. *Left:* The 2 Hz Fast Scattering is associated with an increase in microseismic ground motion (0.1–0.3 Hz) at the observatory site. *Right:* The rate of 4 Hz Fast Scattering goes up with an increase in the anthropogenic band (1–6 Hz) caused due to trains, thunderstorms or human activity near the instrument. Each subclass of Fast Scattering is made up of (a variable number) of individual bursts of noise which appear as blobs or arches in the spectrogram.

Hanford and 1500 $\mathrm{nm\,s^{-1}}$ at Livingston. The reduction in noise could have contributed to an increased visibility of 2 Hz Fast Scattering, which may have been previously hidden underneath (louder) Slow Scattering.

3.3. Comparing Expert and Volunteer

In order to analyze the similarities and differences between the Expert model's Fast Scattering and the Volunteer model's Crown, we ask:

 (i) What is the distribution of glitches classified as Fast Scattering and Crown glitches by our two models?

(ii) Is the rate of Fast Scattering and Crown impacted by similar physical conditions near the detector?

Due to an increase in microseismic as well as anthropogenic ground motion near Livingston, February 2020 registered a large number of glitches classified as Fast Scattering and Crown by our two models, and we used this sample to answer our questions.

To check the consistency of the Crown and Fast Scattering labels, we reclassified the glitches labelled as Fast Scattering with the Volunteer model and glitches labelled as Crown with the Expert model. This comparison is shown in Figure 7. The Expert model classifies almost all ($> 99\%$) the Crown glitches as Fast Scattering. The Volunteer model classifies the majority ($\sim 85\%$) of Fast Scattering glitches as Crown, but also classifies a significant minority ($\sim 14\%$) as Slow Scattering. We inspected the spectrograms of a
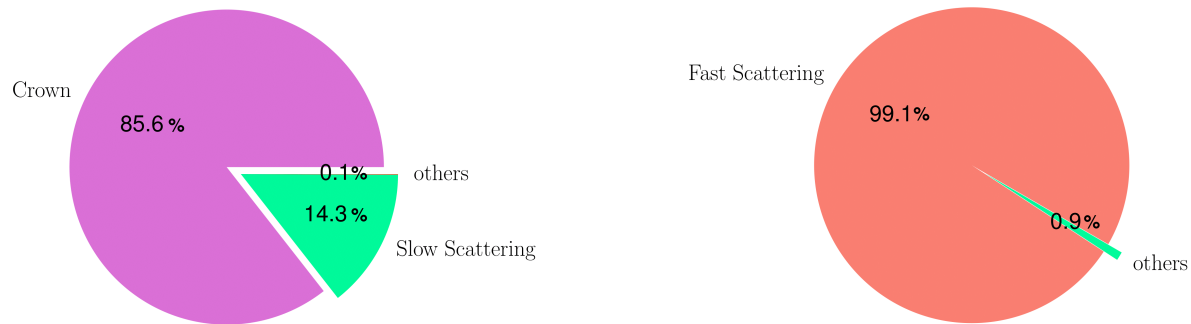
(a) Volunteer model classification of Fast Scattering data.

(b) Expert model classification of Crown data.

Figure 7: Consistency of the Fast Scattering and Crown classes. *Left*: Fast Scattering glitches classifications under the Volunteer model. *Right*: Crown glitches classifications under the Expert model. The glitches have a minimum SNR threshold of 7.5 and a minimum confidence of 0.9.

sample of these glitches labelled as Slow Scattering and confirmed they belong to the Fast Scattering/Crown category. This confusion originates from the features common to Fast Scattering/Crown and Slow Scattering which can make it hard for the CNN to distinguish them. It was this similarity which led to many of the Fast Scattering/Crown glitches being classified as Slow Scattering using the old model. Overall, both models do primarily identify the same types of glitch.

The differences between the Expert and Volunteer models can be explained due to the differences in the training sets. Figure 8 shows the classification of Fast Scattering glitches by the Volunteer model with different SNR thresholds. The fraction labelled as Slow Scattering increases with SNR. The Volunteer **model** classifies few high-SNR glitches as Crown. While the Expert training set contains 400 Fast Scattering examples randomly selected from the O3 data between 1 February 2020 and 1 March 2020, the majority of the Volunteer training set (221/265) comes from the Engineering Run (ER14) during March 2019, just before the beginning of O3 in April 2019. **Glitches from ER14 were uploaded to Zooniverse before those from O3, and so the first examples of Fast Scattering/Crown seen by volunteers were from this period.** We do not expect the behavior or the morphology of noise to change significantly from ER14 to O3, but we do find that there is a difference in the SNR of Fast Scattering/Crown glitches, mirroring the two training sets. **The variation in Fast Scattering/Crown properties between ER14 and O3, while small, shows the importance of verifying if changes in the detectors have affected the form of glitches in the data.**

Figure 9 shows the SNR probability density of the Fast Scattering and Crown training data. The Fast Scattering training set covers a broader range of SNR values,

(a) Volunteer model classification of Fast Scattering with SNR > 20.

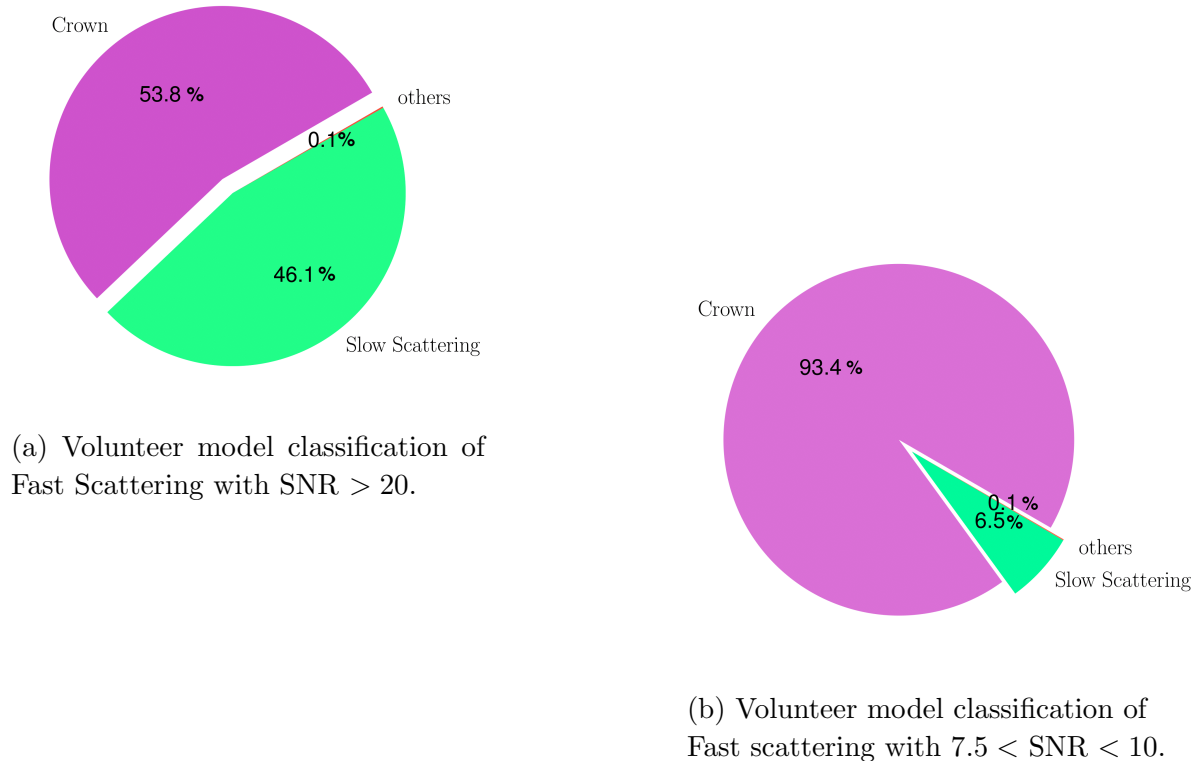(b) Volunteer model classification of Fast scattering with $7.5 < \mathrm{SNR} < 10$.

Figure 8: Consistency of the Fast Scattering and Crown classes for different SNR thresholds: Fast Scattering classifications using the Volunteer model. The fraction of Fast Scattering glitches labelled as Slow Scattering instead of Crown increases with SNR.

with a symmetric 90% range of 8.6–23.1. In comparison, the Crown training data has a narrower SNR distribution with a 90% range of 7.6–12.2. The Expert training set covers a larger range of SNR values, and this difference explains the disparity between the Expert and Volunteer model evident in Figure 7 and Figure 8.

The LIGO sites have a network of sensors to record their surrounding environments [8, 9]. **Ground motion (along the X, Y and Z directions) is measured by reference seismometers located at the two end stations and at the corner station of the LIGO interferometer [1, 9]. The output from these seismometers is then bandpassed into multiple frequency bands between** 0.03 Hz **and** 30 Hz**.** To verify that Fast Scattering and Crown glitches have similar physical origins, we examine their rate as antropogenic ground motion varies. Several times a day, trains on a railway track near the Y end of the Livingston detector cause anthropogenic-band ground motion. Figure 10 shows a period of increased ground motion in the antropogenic band (3–5 Hz) caused by a train on 17 February 2020 coincided with an increase in the rate of Fast Scattering and Crown glitches. The only significant difference between the two classes is observed at high SNR, as a consequence of the lack of high-SNR Crown glitches in the Volunteer training set. The correlation
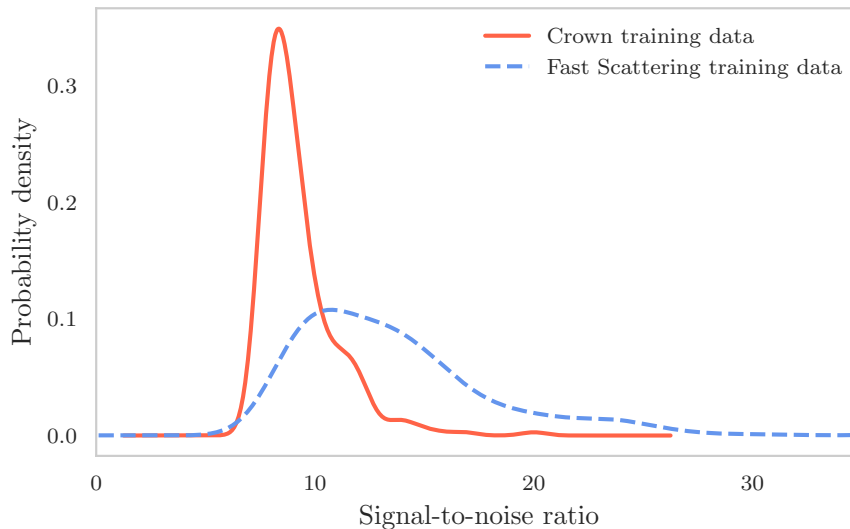
Figure 9: SNR distribution of Fast Scattering and Crown glitches in the Expert and Volunteer training sets, respectively. The Crown training set has a narrower SNR distribution and peaks at a smaller value in comparison to the Fast Scattering training set.

with ground motion shows that the noise recognized as Crown and Fast Scattering are affected by similar environmental factors. Correlating specific behaviors of transient noise such as its rate to environmental conditions is one of the most crucial steps in noise characterization.

4. Discussion & conclusions

We have demonstrated how Gravity Spy can be used by both detector-characterization experts and citizen-science volunteers to efficiently identify new classes of glitch in the wealth of data provided by the LIGO gravitational-wave detectors. Examining the O3 data set, we have found two new families of glitches: Fast Scattering/Crown and Low-frequency Blip.

We observe a significant difference between the glitch populations in LIGO Livingston and LIGO Hanford data. Using our Expert/Volunteer Gravity Spy machine-learning algorithm, 55211/51887 glitches are classified as Fast Scattering/Crown with 90% confidence in O3 LIGO Livingston data, while only 1286/581 are classified as such in LIGO Hanford data. This highlights that while both observatories share the same base design [1], they differ in exact configuration and in their physical environments. Therefore, in future studies, it may be beneficial to tailor separate machine-learning algorithms to classify glitches in each detector. Since we expect that glitch classification will still require recognising similar features in time–frequency spectrograms, these observatory-specific algorithms could be built upon the existing Gravity Spy algorithm as an example of transfer learning [36, 65–67]. Isolating new glitch classes enables
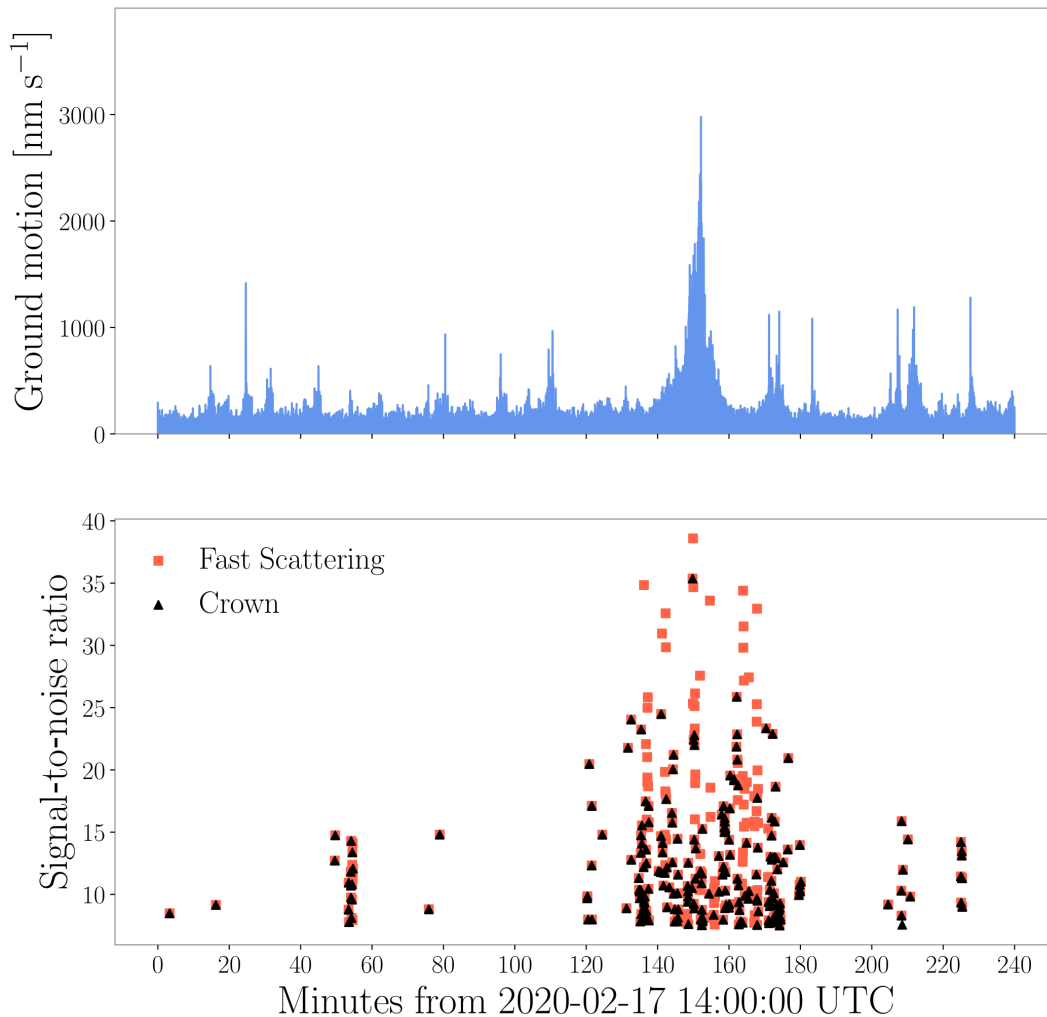
Figure 10: *Top:* Ground motion recorded at the Y end station of the LIGO Livingston Observatory, bandpassed between 3 Hz and 5 Hz. Trains on a track near this station regularly increase the ground motion in this anthropogenic band. *Bottom:* SNR versus time plot of Fast scattering and Crown. This shows an increase in the rate of both Fast Scattering and Crown glitches with an increase in the ground motion.

progress in improving gravitational-wave detector data quality, and studying the glitch populations of different observatories individually may enable us to more precisely trace the origins of glitches.

We have shown that Fast Scattering/Crown glitches are related to ground motion at the detector sites. Through examining glitches classified as Fast Scattering, we have discovered 4 Hz and 2 Hz subclasses which have different dependencies on ground motion, which is a significant step towards improved noise characterization. Physical environment monitoring tests at the end and corner stations at both detector sites revealed high quality-factor resonances at frequencies close to 4 Hz believed to be responsible for the 4 Hz Fast Scattering observed during O3 [68–71]. Scientists at

both Livingston and Hanford have planned to damp this motion and re-evaluate the impact of ground motion on the detectors to see if data quality can be improved. As glitches caused by light scattering were a significant data-quality issue in O3, impacting the analyses of multiple detections [7], mitigation of their causes is a priority ahead of the up-coming fourth observing run (O4) [24].

The global detector network will continue to be enhanced to increase sensitivity to gravitational waves. The A+ upgrades of the LIGO detectors include higher laser power, new test mass mirrors with lower thermal noise, balanced homodyne readout and frequency-dependent squeezing [72]. Some of these upgrades will be completed before O4, increasing the binary neutron star inspiral ranges (an average distance a standard source can be detected [73]) from $\sim$ 110–130 Mpc [7, 26] to $\sim$ 160–190 Mpc [60]. Other improvements to the detectors are planned to address data-quality issues: new baffles will be installed to reduce stray light [26], which should decrease the amount of scattered light, and the adoption of new seismic-isolation control systems should improve the noise coupling from ground motion [33, 74]. However, the process of detector calibration, upgrades and installing new hardware contribute to a change in the state of the system and may introduce new sources of transient noise. A gravitational-wave detector with enhanced sensitivity, and thus a lowered noise floor, may also witness new sources of noise that were not a cause of concern earlier. We may also observe the same transient noise but in a different frequency band [75]. Therefore, we expect that the identification of new classes of glitch and their causes will remain a necessary task.

The similarity of the Fast Scattering and Crown classes identified independently by detector-characterization experts and citizen-science volunteers has provided a unique opportunity to study how the two groups work. We have shown that the crowd-sourcing potential of citizen-science projects does not have to be limited to simple classification tasks. Non-experts can explore complicated data sets and make impactful discoveries. We have found that volunteers engage in in-depth investigations, similar to professional scientists. They use a variety of tools in combination, and are capable of producing sophisticated analyses of their own design. While the Fast Scattering and Crown results are not identical, this is not surprising given that the volunteers lacked access to additional auxiliary information on the state of the detectors, such as the presence of ground motion around the time of the glitches, **and that the volunteers created their training set using glitches from primarily ER14 rather than O3**. Our results suggest that providing volunteers with more data, as well as the tools to explore it, may lead to even greater scope for discovery.

The benefit of making data accessible to non-experts is that it can then be examined by many more people. Gravity Spy volunteers, like other citizen-science volunteers [76, 77], are largely self-motivated, and like to work individually; however, expert feedback is appreciated and beneficial to their work. Therefore, we have found that it essential to provide (i) guidance to volunteers so that they can train themselves in the domain at their own pace (as through the Gravity Spy levels); (ii) tools to explore the data (such as the Gravity Spy similarity search [36]), and (iii) the ability

to interact with experts (as provided through evaluation of glitch proposals). These recommendations overlap with best practices advised for citizen-science projects to produce reliable results [78]. Previous studies have shown that citizen-science volunteers can produce high accuracy results, performing classifications comparable to professional scientists [79–81], and even detect novel features in data [82–84]; here, we have shown volunteers can go beyond this, and identify new classes analogously to experts. When provided with suitable resources, citizen-science volunteers can make significant scientific discoveries based upon complex data sets. As science progresses to produce more data than can be examined by small science teams, the potential of citizen-science volunteers may enable even greater scientific return.

Acknowledgments

References

[1] Aasi J *et al.* (LIGO Scientific) 2015 *Class. Quant. Grav.* **32** 074001 (*Preprint* 1411.4547)

[2] Acernese F *et al.* (Virgo) 2015 *Class. Quant. Grav.* **32** 024001 (*Preprint* 1408.3978)

[3] Abbott B *et al.* (LIGO Scientific, Virgo) 2016 *Phys. Rev. Lett.* **116** 061102 (*Preprint* 1602.03837)

[4] Abbott B *et al.* (LIGO Scientific, Virgo) 2016 *Astrophys. J. Lett.* **818** L22 (*Preprint* 1602.03846)

[5] Abbott R *et al.* (LIGO Scientific, Virgo) 2021 *Astrophys. J. Lett.* **913** L7 (*Preprint* 2010.14533)

[6] Abbott R *et al.* (LIGO Scientific, KAGRA, Virgo) 2021 *Astrophys. J. Lett.* **915** L5 (*Preprint* 2106.15163)

[7] Abbott R *et al.* (LIGO Scientific, Virgo) 2021 *Phys. Rev. X* **11** 021053 (*Preprint* 2010.14527)

[8] Abbott B *et al.* (LIGO Scientific, Virgo) 2016 *Class. Quant. Grav.* **33** 134001 (*Preprint* 1602.03844)

[9] Nguyen P *et al.* (LIGO Instrument Science) 2021 (*Preprint* 2101.09935)

[10] Fiori I *et al.* 2020 *Galaxies* **8** 82

[11] Méndez B J H 2008 *Astronom. Soc. Pacific Conf. Ser.* **389** 219

[12] Silvertown J 2009 *Trends Ecol. Evol.* **24** 467–471

[13] Bonney R, Shirk J L, Phillips T B, Wiggins A, Ballard H L, Miller-Rushing A J and Parrish J K 2014 *Science* **343** 1436–1437

[14] L'Heureux A, Grolinger K, Elyamany H F and Capretz M A M 2017 *IEEE Access* **5** 7776–7797

[15] Cuoco E *et al.* 2021 *Mach. Learn. Sci. Tech.* **2** 011002 (*Preprint* 2005.03745)

[16] Dal Canton T, Bhagwat S, Dhurandhar S V and Lundgren A 2014 *Class. Quant. Grav.* **31** 015016 (*Preprint* 1304.0008)

[17] Abbott B P *et al.* (LIGO Scientific, Virgo) 2018 *Class. Quant. Grav.* **35** 065010 (*Preprint* 1710.02185)

[18] Pankow C *et al.* 2018 *Phys. Rev. D* **98** 084016 (*Preprint* 1808.03619)

[19] Powell J 2018 *Class. Quant. Grav.* **35** 155017 (*Preprint* 1803.11346)

[20] Abbott B P *et al.* (LIGO Scientific, Virgo) 2020 *Class. Quant. Grav.* **37** 055002 (*Preprint* 1908.11170)

[21] Chatziioannou K, Cornish N, Wijngaarden M and Littenberg T B 2021 *Phys. Rev. D* **103** 044013 (*Preprint* 2101.01200)

[22] Nuttall L K 2018 *Phil. Trans. Roy. Soc. Lond. A* **376** 20170286 (*Preprint* 1804.07592)

[23] Cabero M *et al.* 2019 *Class. Quant. Grav.* **36** 15 (*Preprint* 1901.05093)

[24] Davis D *et al.* (LIGO Instrument Science) 2021 *Class. Quant. Grav.* **38** 135014 (*Preprint* 2101.11673)

[25] Chatterji S, Blackburn L, Martin G and Katsavounidis E 2004 *Class. Quant. Grav.* **21** S1809–S1818 (*Preprint* gr-qc/0412119)

[26] Buikema A *et al.* (LIGO Instrument Science) 2020 *Phys. Rev. D* **102** 062003 (*Preprint* 2008.01301)

[27] Zevin M *et al.* 2017 *Class. Quant. Grav.* **34** 064003 (*Preprint* 1611.04596)

[28] Bahaadini S, Rohani N, Coughlin S, Zevin M, Kalogera V and Katsaggelos A K 2017 Deep Multi-view Models for Glitch Classification *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* pp 2931–2935 (*Preprint* 1705.00034)

[29] Davis D, White L V and Saulson P R 2020 *Class. Quant. Grav.* **37** 145001 (*Preprint* 2002.09429)

[30] Torres-Forné A, Cuoco E, Font J A and Marquina A 2020 *Phys. Rev. D* **102** 023011 (*Preprint* 2002.11668)

[31] Cabero M, Mahabal A and McIver J 2020 *Astrophys. J. Lett.* **904** L9 (*Preprint* 2010.11829)

[32] Jadhav S, Mukund N, Gadre B, Mitra S and Abraham S 2020 (*Preprint* 2010.08584)

[33] Soni S *et al.* (LIGO Instrument Science) 2020 *Class. Quant. Grav.* **38** 025016 (*Preprint* 2007.14876)

[34] Krastev P G, Gill K, Villar V A and Berger E 2021 *Phys. Lett. B* **815** 136161 (*Preprint* 2012.13101)

[35] Bahaadini S, Noroozi V, Rohani N, Coughlin S, Zevin M, Smith J R, Kalogera V and Katsaggelos A 2018 *Info. Sci.* **444** 172–186

[36] Coughlin S *et al.* 2019 *Phys. Rev. D* **99** 082002 (*Preprint* 1903.04058)

[37] Bahaadini S, Noroozi V, Rohani N, Coughlin S, Zevin M and Katsaggelos A K 2018 DIRECT: Deep Discriminative Embedding for Clustering of LIGO Data *2018 25th IEEE International Conference on Image Processing (ICIP)* pp 748–752 (*Preprint* 1805.02296)

[38] Smith J 2019 aLIGO LLO Logbook 44803

[39] Austin C 2019 Daytime Scatter at LLO Tech. Rep. G1900753 LSC URL https://dcc.ligo.org/LIGO-G1900753/

[40] Soni S and Coughlin S 2020 Training a new GravitySpy model Tech. Rep. G2000844 LSC URL https://dcc.ligo.org/LIGO-G2000844/public

[41] Abbott R *et al.* (LIGO Scientific, Virgo, KAGRA) 2021 (*Preprint* 2107.03701)

[42] Macleod D, Urban A L, Smith J and Massinger T 2019 gwdetchar/hveto: 1.0.1 URL https://doi.org/10.5281/zenodo.3532131

[43] Macleod D, Urban A L, Coughlin S, Massinger T, Pitkin M, rngeorge, paulaltin, Areeda J, Singer L, Quintero E and Leinweber K 2020 gwpy/gwpy: 2.0.1 URL https://doi.org/10.5281/zenodo.3973364

[44] Urban A L, Macleod D, Massinger T, Bidler J, Smith J, Macedo A, Soni S, Coughlin S, Leman K, Davis D and Lundgren A 2019 gwdetchar/gwdetchar: 1.0.2 URL https://doi.org/10.5281/zenodo.3592169

[45] Robinet F 2015 Omicron: An Algorithm to Detect and Characterize Transient Noise in Gravitational-Wave Detectors Tech. Rep. VIR-0545C-14 Virgo URL https://tds.ego-gw.it/ql/?c=10651

[46] Robinet F, Arnaud N, Leroy N, Lundgren A, Macleod D and McIver J 2020 *SoftwareX* **12** 100620 (*Preprint* 2007.11374)

[47] Essick R, Godwin P, Hanna C, Blackburn L and Katsavounidis E 2020 (*Preprint* 2005.12761)

[48] Macleod D, Urban A L, Isi M, Massinger T, paulaltin, Pitkin M and Nitz A 2020 gwpy/gwsumm: 1.0.3 URL https://doi.org/10.5281/zenodo.3765457

[49] Soni S 2020 aLIGO LLO Logbook 52071

[50] Ottaway D J, Fritschel P and Waldman S J 2012 *Opt. Express* **20**

[51] Accadia T *et al.* 2010 *Class. Quant. Grav.* **27** 194011

[52] Valdes G, O'Reilly B and Diaz M 2017 *Class. Quant. Grav.* **34** 235009

[53] Soni S, Effler A, Frolov V and Schofield R 2020 Fast scattering noise at LIGO and DetChar Noise sprint Tech. Rep. G2001639 LSC URL https://dcc.ligo.org/LIGO-G2001639/public

[54] Patron A 2020 aLIGO LLO Logbook 53678

[55] Abbott B P *et al.* (LIGO Scientific, Virgo) 2019 *Phys. Rev. X* **9** 031040 (*Preprint* 1811.12907)

[56] Abbott B P *et al.* (LIGO Scientific, Virgo) 2019 *Phys. Rev. D* **100** 064064 (*Preprint* 1906.08000)

[57] Davis D 2018 aLIGO LHO Logbook 41263

[58] Helmling-Cornell A 2019 aLIGO LHO Logbook 50946

[59] Helmling-Cornell A 2020 Blip Glitches in LHO Tech. Rep. G2001007 LSC URL https://dcc.ligo.org/LIGO-G2001007/public

[60] Abbott B *et al.* (KAGRA, LIGO Scientific, Virgo) 2020 *Living Rev. Rel.* **23** 3 (*Preprint* 1304.0670)

[61] Jackson C B, Østerlund C, Mugar G, Hassman K D and Crowston K 2015 Motivations for Sustained Participation in Crowdsourcing: Case Studies of Citizen Science on the Role of Talk *2015 48th Hawaii International Conference on System Sciences* pp 1624–1634

[62] Oesterlund C S, Mugar G, Jackson C, Hassman K D and Crowston K 2018 *Academy of Management Proceedings* **376** 20170286

[63] Matichard F *et al.* 2015 *Class. Quant. Grav.* **32** 185003 (*Preprint* 1502.06300)

[64] Aston S M *et al.* 2012 *Class. Quant. Grav.* **29** 235004

[65] Shao L, Zhu F and Li X 2015 *IEEE Transactions on Neural Networks and Learning Systems* **26** 1019–1034

[66] Tan C, Sun F, Kong T, Zhang W, Yang C and Liu C 2018 A survey on deep transfer learning *Artificial Neural Networks and Machine Learning – ICANN 2018* ed Kůrková V, Manolopoulos Y, Hammer B, Iliadis L and Maglogiannis I (Cham: Springer International Publishing) pp 270–279 (*Preprint* 1808.01974)

[67] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H and He Q 2021 *Proceedings of the IEEE* **109** 43–76 (*Preprint* 1911.02685)

[68] Effler A 2020 aLIGO LLO Logbook 53185

[69] Effler A 2020 aLIGO LLO Logbook 53025

[70] Effler A 2020 aLIGO LLO Logbook 53057

[71] Effler A 2020 aLIGO LLO Logbook 53364

[72] Miller J, Barsotti L, Vitale S, Fritschel P, Evans M and Sigg D 2015 *Phys. Rev. D* **91** 062005 (*Preprint* 1410.5882)

[73] Chen H Y, Holz D E, Miller J, Evans M, Vitale S and Creighton J 2021 *Class. Quant. Grav.* **38** 055010 (*Preprint* 1709.08079)

[74] Schwartz E *et al.* (LIGO Instrument Science) 2020 *Class. Quant. Grav.* **37** 235007 (*Preprint* 2007.12847)

[75] Soni S 2020 aLIGO LLO Logbook 54383

[76] Raddick M J, Bracey G, Gay P L, Lintott C J, Cardamone C, Murray P, Schawinski K, Szalay A S and Vandenberg J 2013 *Astron. Educ. Rev.* **12** 010106 (*Preprint* 1303.6886)

[77] Jones M G, Childers G, Andre T, Corin E N and Hite R 2018 *Int. J. Sci. Educ. B* **8** 287–306

[78] Freitag A Meyer R and Whiteman L 2016 *Citizen Science: Theory and Practice* **1** 2

[79] Lintott C J *et al.* 2008 *Mon. Not. Roy. Astron. Soc.* **389** 1179–1189 (*Preprint* 0804.4483)

[80] Aceves-Bueno E, Adeleye A S, Feraud M, Huang Y, Tao M, Yang Y and Anderson S E 2017 *Bulletin of the Ecological Society of America* **98** 278–290

[81] McCarthy M S, Stephens C, Dieguez P, Samuni L, Després-Einspenner M L, Harder B, Landsmann A, Lynn L K, Maldonado N, Ročkaiová Z, Widness J, Wittig R M, Boesch C, Kühl H S and Arandjelovic M 2021 *Ecology and Evolution* **11** 1598–1608

[82] Lintott C *et al.* 2009 *Mon. Not. Roy. Astron. Soc.* **399** 129–140 (*Preprint* 0906. 5304)

[83] Schwamb M E *et al.* 2013 *Astrophys. J.* **768** 127 (*Preprint* 1210.3612)

[84] Debes J H, Thévenot M, Kuchner M J, Burgasser A J, Schneider A C, Meisner A M, Gagné J, Faherty J K, Rees J M, Allen M and et al 2019 *Astrophys. J. Lett.* **872** L25 (*Preprint* 1902.07073)