1-2023

Glitch subtraction from gravitational wave data using adaptive spline fitting

Soumya D. Mohanty

Mohammad A. T. Chowdhury

Glitch subtraction from gravitational wave data using adaptive spline fitting

Soumya D. Mohanty, Mohammad A. T. Chowdhury

Department of Physics and Astronomy, University of Texas Rio Grande Valley, One West University Blvd., Brownsville, Texas 78520, USA

E-mail: `soumya.mohanty@utrgv.edu`

E-mail: mohammadabuthaher.chowdhury01@utrgv.edu

Abstract. Transient signals of instrumental and environmental origins ("glitches") in gravitational wave data elevate the false alarm rate of searches for astrophysical signals and reduce their sensitivity. Glitches that directly overlap astrophysical signals hinder their detection and worsen parameter estimation errors. As the fraction of data occupied by detectable astrophysical signals will be higher in next generation detectors, such problematic overlaps could become more frequent. These adverse effects of glitches can be mitigated by estimating and subtracting them out from the data, but their unpredictable waveforms and large morphological diversity pose a challenge. Subtraction of glitches using data from auxiliary sensors as predictors works but not for the majority of cases. Thus, there is a need for nonparametric glitch mitigation methods that do not require auxiliary data, work for a large variety of glitches, and have minimal effect on astrophysical signals in the case of overlaps. In order to cope with the high rate of glitches, it is also desirable that such methods be computationally fast. We show that adaptive spline fitting, in which the placement of free knots is optimized to estimate both smooth and non-smooth curves in noisy data, offers a promising approach to satisfying these requirements for broadband short-duration glitches, the type that appear quite frequently. The method is demonstrated on glitches drawn from three distinct classes in the Gravity Spy database as well as on the glitch that overlapped the double neutron star signal GW170817. The impact of glitch subtraction on the GW170817 signal, or those like it injected into the data, is seen to be negligible.

1. Introduction

In a fairly short time since the first direct detection of a gravitational wave (GW) signal (GW150914) in 2015 [1] by the twin LIGO [2] detectors, GW astronomy has emerged as an information-rich field that will revolutionize our understanding of compact objects such as black holes and neutron stars. By now, the network of LIGO and Virgo [3] detectors has reported 90 confirmed detections of GW signals from compact binary coalescences (CBCs) across the first observing run (O1) [4] to the third (O3) [5]. The majority of these are binary black hole (BBH) mergers but the haul also includes a double neutron star (DNS) system (GW170817) [6].

The rate of detectable GW signals will grow as more detectors, namely KAGRA [7] and LIGO-India [8], join the network and increase its distance reach for GW sources. Design studies are already underway for the successors to the current generation of GW detectors [9, 10, 11] with the goal of achieving an order of magnitude improvement in sensitivity across the current operational frequency band. In addition, next-generation detectors will seek to expand the operational range to lower frequencies ($\approx$ 1 Hz), thereby increasing the duration of in-band GW signals across the board: for example, a DNS signal starting at $\approx$ 10 Hz will last for days compared to the $\approx$ 1 min for GW170817. Thus, future detectors will not only see a higher rate but also longer signals, raising the prospect [12] that there will be no data segment free of detectable GW signals.

The false alarm rate, hence the sensitivity, of searches for CBCs as well as generic short duration GW signals, or *bursts*, is dominated [13] by transient non-GW signals of instrumental or environmental origins, commonly called *glitches*. This is because glitches that populate the same frequency band as CBC or burst signals and happen to be transient in duration can falsely trigger the respective search pipelines. A glitch has a particularly adverse effect if it overlaps with a GW signal, as happened in the case of GW170817 [6], and causes the glitch rejection step of a search pipeline to also discard the signal. Even a non-overlapping glitch can severely degrade parameter estimation if it is close enough to a GW signal [14]. In the third observing run of the LIGO and Virgo detectors, $\approx 20\%$ of detected GW signals overlapped with glitches [15] due to the high rate of the latter. For future detectors, the frequency of chance overlaps will be enhanced by the higher rate of detectable GW signals as well as, for CBC signals, their longer durations.

Glitches have dissimilar and unpredictable waveforms but many of the observed ones tend to fall into distinct morphological classes. This has motivated the investigation of automated glitch classification using machine learning where a range of different methods have been proposed, such as Support Vector Machine [16], t-Sne [17], random forests [16], S-means [18], and Deep Convolutional Neural Networks [19]. The Gravity Spy [20] project uses a citizen science approach to engage the lay public in labeling glitches by visual inspection of their constant Q-transform [21, 22] images. This has created a high quality training dataset for machine learning methods. By now, there exist more than 20 named glitch classes in the Gravity Spy database, collected over multiple observing runs of the LIGO detectors [17].

Several different approaches have been developed to mitigate the adverse effects of glitches on GW searches. GW search pipelines typically compute secondary functionals, called *vetoes*, of the data besides the primary detection statistic that help in distinguishing genuine GW signals from glitches. A well-known example is the Chi-square [23] veto used in CBC search pipelines. For LIGO-Virgo data, a set of Data Quality flags have been developed that use information from a large number of auxiliary sensors to quantify the safety of analyzing a given segment of GW strain data [24]. For glitches that overlap a GW signal, the gating [25] method excises the rectangular time-frequency block, or just the time interval, containing an identified glitch from the data.

Cross-channel regression using data from auxiliary sensors [26, 27, 28, 29] has been used to reduce excess broadband noise and a few types of glitches [30].

A relatively recent approach is that of estimating the waveform of a glitch from the data time series itself and subtracting it out. Glitch subtraction was of critical importance in the case of GW170817 and has been shown to be an important requirement in reducing bias in the estimation of GW signal parameters [31]. The GW170817 glitch subtraction was carried out using the multi-detector BayesWave pipeline [32, 33], which has also been used for other types of glitches [15]. Another method, Glitschen [34], follows the approach of constructing parametrized waveform models for identified glitch classes using principal component analysis of training sets. A strong motivation for developing glitch estimation and subtraction methods is that one could, in principle, preprocess the data to clean out every sufficiently loud glitch of a known type and make glitch rejection in all downstream GW searches safer.

In this paper, we present a method for the estimation and subtraction of broadband, short-duration glitches that have appeared frequently in the observation runs of the LIGO detectors. The method is computationally cheap, works with single-detector data, does not require a training set of pre-identified glitches, and is not predicated on auxiliary sensor data. The core component of the method is SHAPES (Swarm Heuristics based Adaptive and Penalized Estimation of Splines), an adaptive spline curve fitting algorithm introduced in [35]‡. SHAPES uses splines with free placement of knots to fit both smooth and non-smooth curves in noisy data. In particular, point discontinuities in the curve or its derivatives (up to some order) can be accommodated in the fit by allowing knots to merge. The ability to handle both sharp and slow changes in a curve is a built-in form of multiresolution analysis in SHAPES and a critical requirement for effective estimation of broadband glitches. We examine the performance of our glitch subtraction method on the GW170817 glitch in LIGO-Livingston data and instances of glitches from three morphologically distinct classes, namely, *Blip*, *Koi Fish*, and *Tomte*, in the Gravity Spy database. In each of the latter three cases, we inject a DNS signal overlapping with the glitch to mimic the case of GW170817. We find that the impact of glitch subtraction on the signals, real or injected, is negligible.

The rest of the paper is organized as follows. Sec. 2 reviews SHAPES with the goal of providing a self-contained description of the algorithm that is pertinent to this paper. Further details, such as the motivation and justification for certain features of the algorithm, can be found in [35]. Sec. 3 describes the dataset used in this paper and the details of how SHAPES is used for glitch subtraction. Sec. 4 presents the results. Our conclusions and discussion of future work are presented in Sec. 5.

‡ The SHAPES code is available from the Github repository `mohanty-sd/SHAPES.git`.

2. Adaptive spline fitting: the SHAPES algorithm

SHAPES is derived under the following models for the noisy data, $\overline{y}$, and the signal $\overline{s}(\theta)$.

$$\overline{y} = \overline{s}(\theta) + \overline{\epsilon} \, , \tag{1}$$

where $\overline{y}$, $\overline{s}$, and $\overline{\epsilon}$ are row vectors with N elements, $y_i = y(t_i)$ and $s_i(\theta) = s(t_i;\theta)$, $i = 0, 1, \ldots, N-1$, are samples taken at $t_i = i/f_s$ with f_s being the sampling frequency, and θ denotes the set of signal parameters that need to be estimated from the data. The noise samples, ϵ_i, are drawn independently from the zero mean and unit variance normal (Gaussian) probability density function $N(0,1)$. This assumption, namely, that of a white Gaussian noise process does not entail a loss of generalization since GW data can always be whitened using the estimated noise power spectral density (PSD).

The signal $s(t;\theta)$ is assumed to be a spline of polynomial order k and, as such, can be represented by a linear combination of B-spline functions [36],

$$s\left(t;\theta = \{\overline{\alpha}, \overline{\tau}\}\right) \; = \; \sum_{j=0}^{P-k-1} \alpha_j B_{j,k}(t;\overline{\tau}) \, , \tag{2}$$

where $\overline{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{P-k-1})$, and $\overline{\tau} = (\tau_0, \tau_1, \ldots, \tau_{P-1})$, $\tau_{i+1} \geq \tau_i$, is a sequence of P *knots* that marks the end points of the contiguous intervals containing the polynomial pieces of the spline. Note that knots are allowed to be equal, leading to knots with multiplicity higher than one. Repeating knots create discontinuity in either the value of a B-spline function or its derivatives (up to order $k-2$). This allows the $s(t;\theta)$ in Eq. 2 to model signals with point discontinuities in value or derivatives. In the rest of the paper, we will set $k = 4$, making $s(t;\theta)$ a cubic spline.

The best fit spline parameters, $\widehat{\alpha}$ and $\widehat{\tau}$, are the ones that minimize a penalized least-squares function,

$$L_\lambda(\overline{\alpha}, \overline{\tau}) = L(\overline{\alpha}, \overline{\tau}) + \lambda R(\overline{\alpha}) \, , \tag{3}$$

$$L(\overline{\alpha}, \overline{\tau}) \; = \; \sum_{i=0}^{N-1} \left(y_i - s_i(\overline{\alpha}, \overline{\tau})\right)^2 \, , \tag{4}$$

where the penalty term,

$$R(\overline{\alpha}) = \sum_{j=0}^{P-k-1} \alpha_j^2 \, , \tag{5}$$

is found to be useful in the suppression of spurious clustering of the knots. These clusters are observed when the method tries to minimize $L_\lambda(\overline{\alpha}, \overline{\tau})$ by fitting out outlier data points arising from the noise alone. The strength of the penalty is controlled by the gain factor λ, with higher values of λ leading to smoother estimates.

The optimization of $L_\lambda(\overline{\alpha}, \overline{\tau})$ over the non-linear parameters $\overline{\tau}$ has been a long-standing computational barrier [37, 38, 39, 40] for adaptive spline fitting. At the same time, the benefits of optimizing the placement of knots have also been demonstrated extensively [38, 41]. It was shown in [42], and independently in [43], that Particle Swarm Optimization (PSO) [44, 45], a widely used nature-inspired metaheuristic for global optimization, has good performance on the free knot placement problem. Moreover, being a continuous optimization method, PSO can explore all arrangements of knots, including the ones where knots are sufficiently close to be merged into a single knot of higher multiplicity. This allows the fitting of functions that have a mix of smooth and non-smooth parts.

There are many variations [46] among the algorithms that fall under the umbrella of the PSO metaheuristic but they all share the following common features. (i) The function to be optimized, called the *fitness function*, is sampled at multiple locations, called *particles*, that move iteratively to explore the domain, called the *search space*, over which the the global optimum of the fitness is to be found. The set of particles is called a *swarm*. (ii) The location of each particle is updated following a dynamical rule that incorporates randomness. The rule typically uses the best location found by a particle in its history, called its *personal best*, and the best location found by the particles in its *neighborhood*, called its *local best*. Here, the fitness value at a location defines how good it is: for a minimization problem, the lower the fitness, the better the location. (iii) Each particle explores the search space independently but is constantly attracted towards the personal and local bests. This leads to a form of communication between the particles that speeds up convergence to a promising region, followed by refinement of the solution until the iterations are terminated.

The best location among all the particles at termination is the final solution found by the swarm for the global optimum. While there is no guarantee that the final solution is the true global optimum, the probability of successful convergence can be boosted exponentially by running multiple independent runs of PSO and picking the one with the best final solution. Most of the parameters involved in the PSO algorithm, such as the number of particles or the weights attached to the attractive forces, have very robust values across a wide variety of benchmark optimization problems [47] and rarely need to be changed. In our experience, there are typically only two quantities that need tuning: the number of iterations, $N_{\rm iter}$, to termination and the hyper-parameter $N_{\rm runs}$, the number of independent PSO runs. In this paper, we fix $N_{\rm iter} = 2000$ and $N_{\rm runs} = 8$ throughout. The number of particles is always set to 40 and the settings for the remaining parameters, as well as the definition of the neighborhood used for the local best, are described in [35].

The description above was for the case where the number of knots, P, is fixed. The complete SHAPES algorithm incorporates model selection using the Akaike Information Criterion (AIC) [48], where the optimum number of knots minimizes,

$$\mathrm{AIC} = 4P + L_\lambda(\widehat{\alpha}, \widehat{\tau}) \; . \tag{6}$$

While, given sufficient computing resources, model selection could be performed over all values of P until the minimum value of AIC is found, practical considerations dictate that the set of knot numbers used be a finite and small one. In this paper, for example, we use knot numbers in the set starting at 5 and ending at 60 in increments of 5. It is important to note that this restriction of knot numbers is not a fundamental limitation but a technical one meant to manage the computational burden of model selection. Thus, the only significant free parameter that needs to be set by the user in the current version of SHAPES is λ.

Since SHAPES assumes that the noise in the data is white, GW strain data must be whitened prior to glitch estimation and subtraction. The data conditioning steps involved are as follows (in sequential order). (a) Suppression of the seismic noise below 10 Hz, (b) robust estimation of the power spectral density (PSD) noise floor, (c) whitening of the noise floor using the estimated PSD [49], and (d) automated identification of high-power narrowband noise features ("lines") and their suppression using notch filters. These steps are common to all GW search pipelines, so they do not need to be elaborated further here.

3. Demonstration data

The glitches considered in this paper for demonstrating the performance of SHAPES are listed in Table 1. The corresponding GW strain data files can be located and downloaded from the Gravitational Wave Open Science Center (GWOSC) [50] using the information provided in this table. We have used the standard 4096 sec long GWOSC data files sampled at 4 kHz.

The GW170817 glitch presents a particularly interesting example of the deleterious effect of glitches on GW searches. The GW signal appeared in both LIGO-Hanford (H1) and LIGO-Livingston (L1) with a combined network signal to noise ratio (SNR) of 32.4. Such a strong signal would have been detected easily in coincidence across L1 and H1 by the GW search pipelines in operation at the time. However, a coincident detection was prevented by a large overlapping glitch in L1 causing the release of only an unusual single-detector GW detection alert to the astronomical community. About 11 hours elapsed between the initial alert and the release of the skymap localizing GW170817, a process that included the subtraction of the glitch using BayesWave.

In addition to the GW170817 glitch, we have taken three representative glitches from the Blip, Koi Fish, and Tomte, classes in the Gravity Spy [20] database [51]. These glitches were selected by taking the loudest 5 events, in terms of their signal-to-noise ratio (SNR) as given in the Gravity Spy database, for each class and then picking the first one in this list for which the corresponding GWOSC file had 100% science data that was also reasonably stationary. As can be seen from Table 1, this results in the selected glitches spanning a wide range in SNR.

After conditioning the data, we use the start time of a glitch, recorded in Table 1, to locate the glitch. Starting from the peak of the glitch, the data time series is scanned visually in both directions to identify a segment, containing the glitch, that tapers off at both its boundaries to the general noise level of the conditioned data.

To mimic the case of GW170817 and to study the effect of glitch subtraction on an overlapping GW signal, we injected a whitened restricted-2PN circularized binary inspiral signal with equal 1.4 $M_{\odot}$ components in the conditioned data. The SNR (in

Glitch Name	GPS start (sec)	SNR	Detector	run
GW170817 glitch	1187008880	–	L1	O2
Blip	1182397347	109.1	H1	O2
Koi Fish	1169847108	608.1	H1	O2
Tomte	1173086299	19.6	H1	O2

Table 1. Glitches considered in this paper along with their GPS start times, SNRs, the detectors in which they appeared, and the observation runs. For the Blip, Koi Fish, and Tomte glitches, the start times are taken from the Gravity Spy database. To the best of our knowledge, there is no SNR available in the literature for the GW170817 glitch.

white noise with unit variance) of the injected signal is set at 37.3, which is an ad hoc factor of $\sqrt{2}$ higher than the observed SNR of 26.4 of GW170817 in L1 [6]. The enhancement in SNR allows clearer visibility of the signal in time-frequency images while also posing a stronger challenge to SHAPES in terms how well it ignores the GW signal when estimating a glitch. The segment containing the glitch, taken from the conditioned data with the injected signal, is passed to SHAPES for estimation of the glitch waveform followed by its subtraction.

4. Results

In common with other papers on glitch estimation and subtraction, we present our results in the form of constant Q-transform (CQT) time-frequency images and time series plots. These are obtained by taking projections of the data on a set of windowed sinusoids. The width of the window decreases with an increase in the carrier frequency, f_c, such that $Q = f_c/\Delta f$, where Δf is the -3 dB bandwidth of the Fourier transform of the window, remains constant. We use the CQT code provided in the `librosa` [52] Python package for audio processing. For each glitch, we show CQTs of the conditioned data with injected signal and the residual after subtraction of the glitch estimate.

Fig. 1 shows the data segments that were processed using SHAPES and the corresponding estimated glitch waveforms. Except for GW170817, each segment was processed as a whole to obtain the glitch estimate. In the case of GW170817, SHAPES was applied independently to three separate but contiguous time intervals to estimate the complete glitch. This was necessitated by the presence of extended wings, preceding and trailing the core broadband (and rapidly varying) part in the middle, that dominate the conditioned data for ≈ 0.5 sec on each side. Applying SHAPES to the complete segment would have required using a very large number of knots (> 60), making it unnecessarily expensive computationally given that splitting the segment achieves a good solution.

As mentioned in Sec. 2, the penalty gain λ controls the smoothness of the estimate and is a user-specified parameter of the SHAPES algorithm. Typically, when a glitch is loud and has a complex shape, $\lambda = 0.01$ allows SHAPES to provide a better fit. For

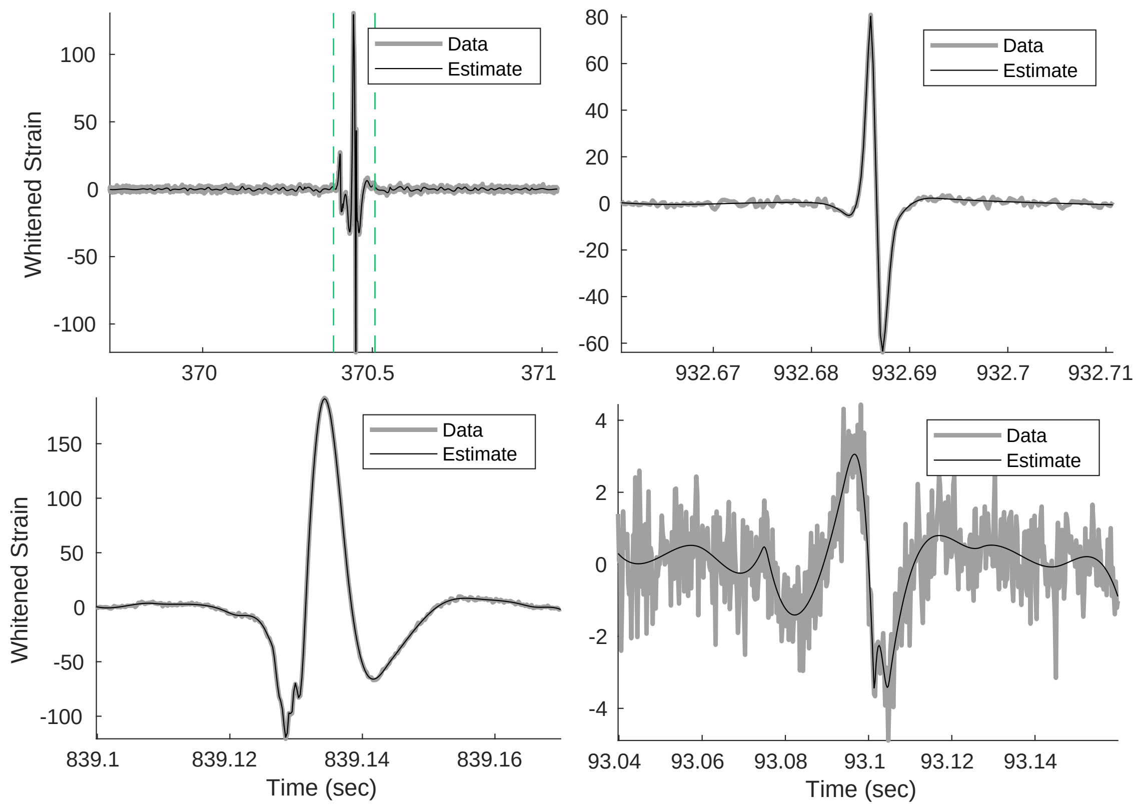

Figure 1. The conditioned strain data and the glitch waveform estimated by SHAPES for each of the glitches considered in this paper. Top row: GW170817 (left) and Blip (right). Bottom row: Koi Fish (left) and Tomte (right). The X-axis in each plot shows the time (sec) since the start of the open data file containing the glitch as provided by GWOSC. For GW170817, the dashed vertical lines demarcate the three adjacent segments that were analyzed separately.

low SNR and simple glitch waveforms, or if the data is just plain white noise, $\lambda = 0.1$ does an adequate job. In general, estimates from SHAPES are not sensitive to small variations of λ around these values because the model selection is able to compensate for a lower value of λ by selecting a higher knot number and vice versa. Without much fine tuning, we found that the values of λ listed in Table 2 work well for the glitches studied in this paper. We have also listed in this table the number of knots for the best fit models selected by the AIC.

Fig. 2 to Fig. 5 show the CQTs of the conditioned data and residuals after glitch subtraction for the glitches in the sequence GW170817, Blip, Koi Fish, and Tomte, respectively. In all cases, we see that the subtraction of the glitch does not affect the overlapping GW signal (real or injected) in any significant way. Some overfitting to the data, seen as very small CQT values, is visible in the residual for the GW170817 glitch at frequencies below ≈ 32 Hz but this band has no overlap with the signal. The overfitted parts are the two wings of the GW170817 glitch mentioned earlier. The CQTs of the residuals for the Blip and Tomte glitches show near perfect removal of the glitch. (For

Glitch Name	Penalty gain (λ)	Number of knots
GW170817 glitch	0.1, 0.01, 0.1	60,40,50
Blip	0.01	15
Koi Fish	0.01	30
Tomte	0.1	15

Table 2. The penalty gain λ used for the glitches and the number of knots in the best fit model. For the GW170817 glitch, there are three segments with the middle one containing the principal glitch and adjacent ones containing the wings. The penalty gains and best fit model are listed for all three segments in sequential order from left to right.

Tomte, the coalescence time of the GW signal was kept further away from the glitch in order to create an overlap between the signal track and the glitch.) The residual for Koi Fish shows effective removal of the glitch with the exception of a transient and low frequency narrowband component. This leftover component does not overlap with the signal.

The principal computational cost in SHAPES is the global optimization of the fitness function in Eq. 3. The time taken by the current `MATLAB` [53] code for a single PSO run on a segment with ≈ 300 samples and knot numbers $P \in [10, 60]$ (in steps of 5) is < 10 min on an Intel Xeon E5 processor (clock rate 3 GHz). The runtime increases with the number of knots used, mainly due to an increase in the number of B-spline functions that need to be computed. With a code currently under construction in the C language, and implementation of further hardware acceleration (e.g., using Graphics Processing Units), the runtime is expected to decrease substantially. We also note that the segments containing glitches can be processed in parallel since SHAPES is a purely time-domain method. Hence, the computational cost will scale slower than linearly with the number of glitches when analyzing data containing multiple glitches.

5. Discussion and Conclusions

We have presented a new approach to glitch subtraction using an adaptive spline fitting method called SHAPES. The method was demonstrated on the GW170817 glitch as well as other representative short duration and broadband glitches. In a single detector and in the absence of strong prior information about the signal, it is not possible to distinguish a GW signal from a glitch in the part where they overlap. Hence, it is expected that the signal power will be removed in that part along with the glitch when the latter is estimated and subtracted out. Nonetheless, as far as the DNS signal used in this paper is concerned, we observe very little impact on the signal across a wide range of glitch SNRs. While this conclusion will be quantified in future studies using a much larger number of glitches, it is clear that SHAPES is effective at addressing glitch subtraction.

SHAPES is not well adapted to fitting highly oscillatory waveforms since these

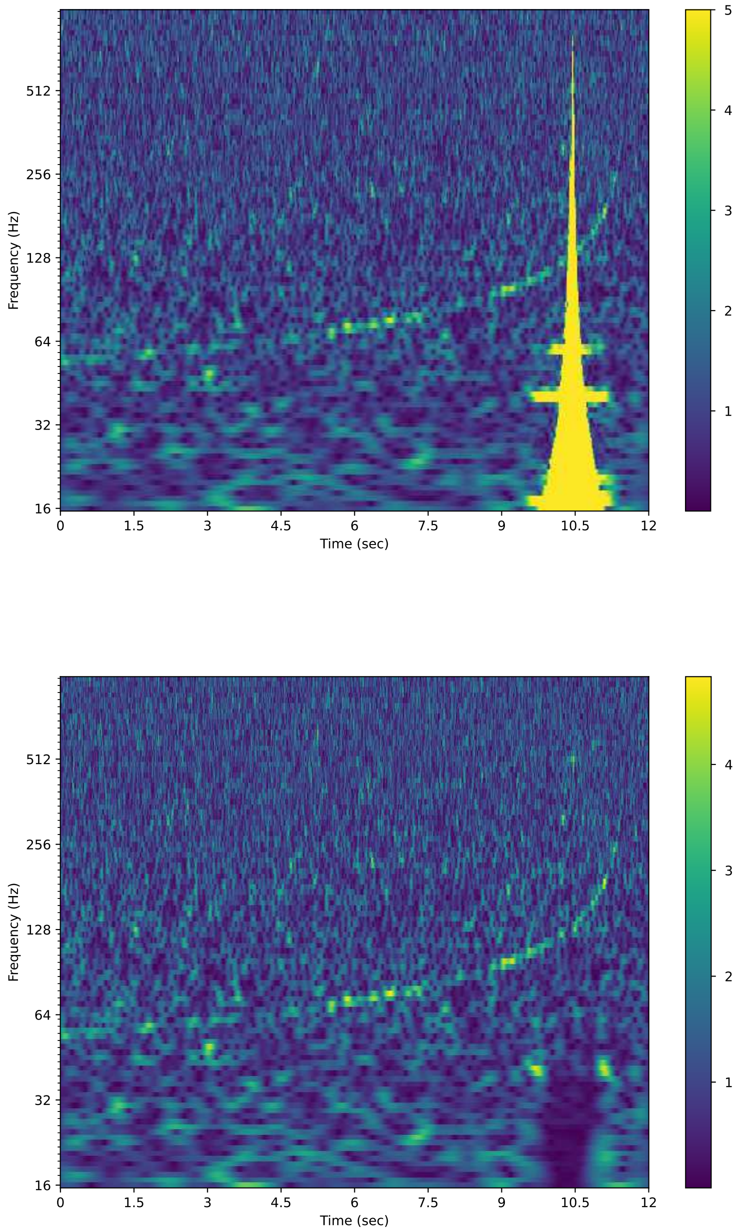

Figure 2. Subtraction of the GW170817 Glitch. The top and bottom panels show the CQT of the data and residual, respectively. The glitch is the vertical feature at $\approx$ 10.5 sec. In order to show both the glitch and the signal in the same image, a threshold has been applied to the CQT as indicated by the maximum value in the colorbar of the top panel.

are are not represented well by splines without using an inordinate number of knots. Therefore, the direct use of SHAPES for glitches in the Gravity Spy database such as whistlers or wandering lines is not viable. However, chirp signals such as these could be estimated using the method proposed in [54, 55], where adaptive splines figure indirectly

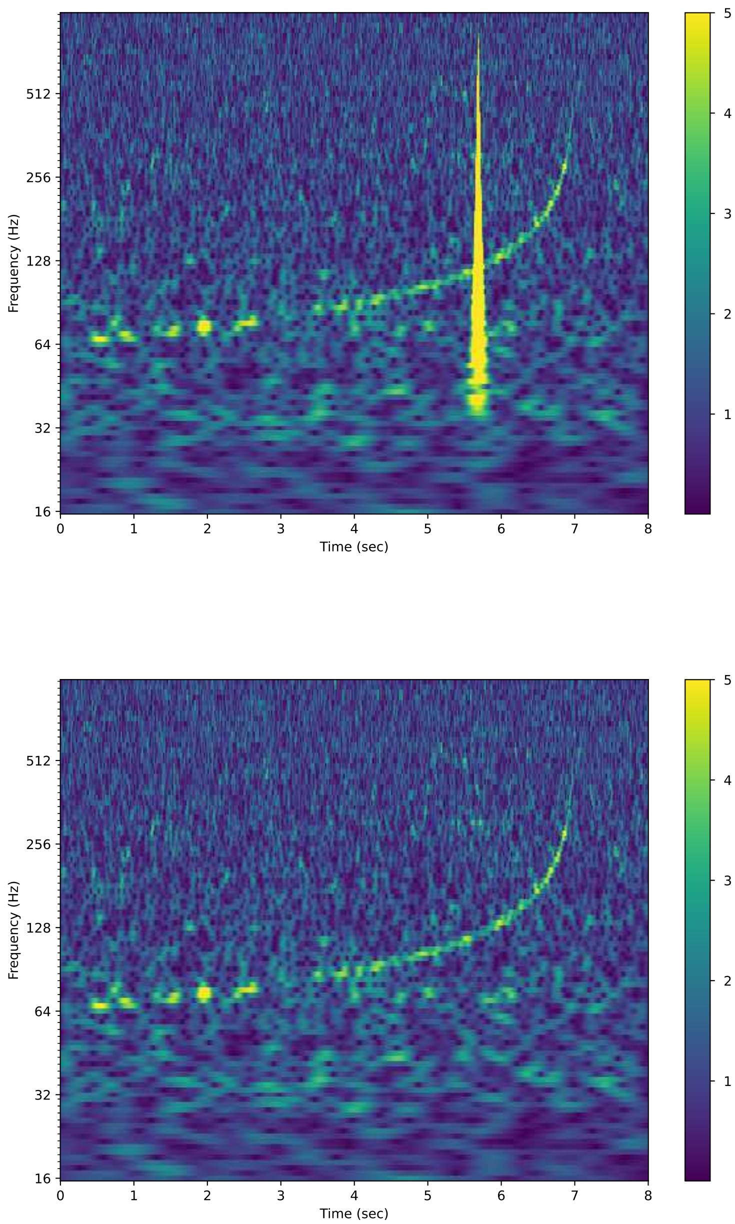

Figure 3. Subtraction of the Blip Glitch. The top and bottom panels show the CQT of the data and residual, respectively. The glitch is the vertical feature at ≈ 6 sec. In order to show both the glitch and the signal in the same image, a threshold has been applied to the CQT as indicated by the maximum value in the colorbar of the top panel.

in a non-linear signal model. This is an interesting direction that will be pursued in future work.

Other current limitations of SHAPES, which are technical in nature, are that the penalty gain parameter λ as well as the segment length to be processed must be specified

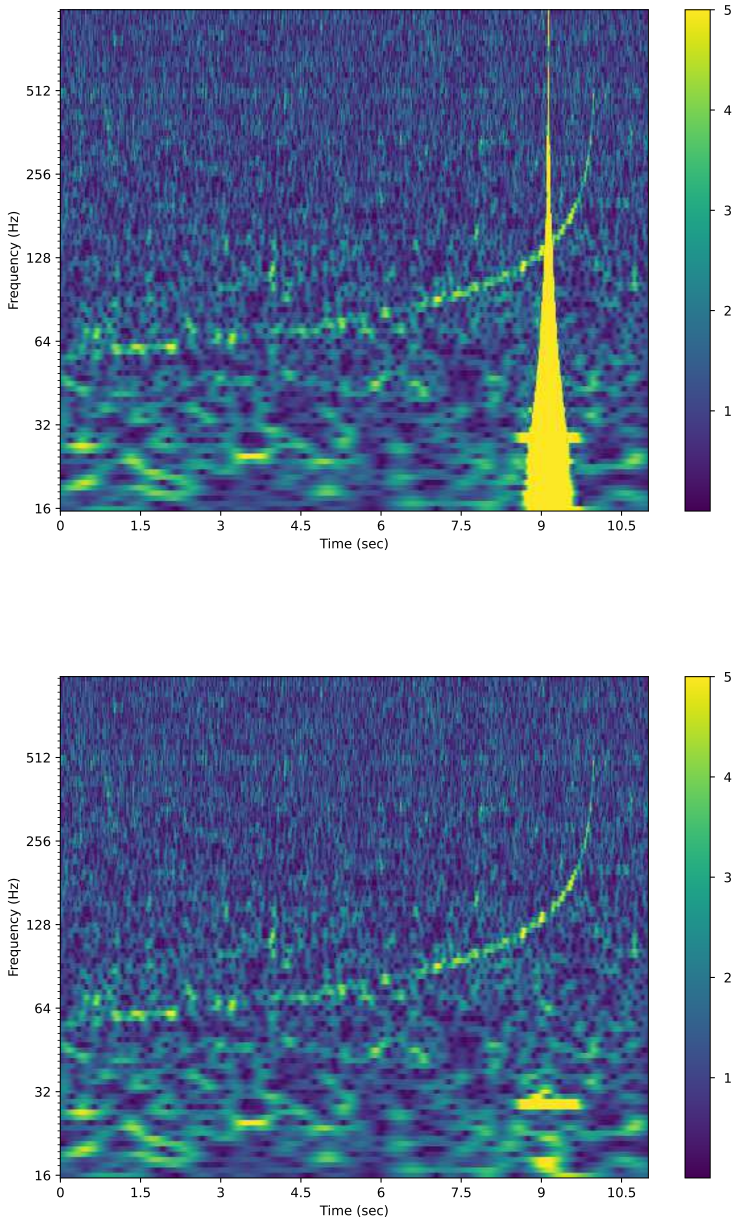

Figure 4. Subtraction of the Koi Fish glitch. The top and bottom panels show the CQT of the data and residual, respectively. The glitch is the vertical feature at $\approx$ 9.0 sec. In order to show both the glitch and the signal in the same image, a threshold has been applied to the CQT as indicated by the maximum value in the colorbar of the top panel.

by the user. The choice of the latter, along with the nature of the data, influences the number of knots used in the fit and led to the necessity of breaking up the data for the GW170817 glitch into three ad hoc parts. Work is in progress to address both of these limitations.

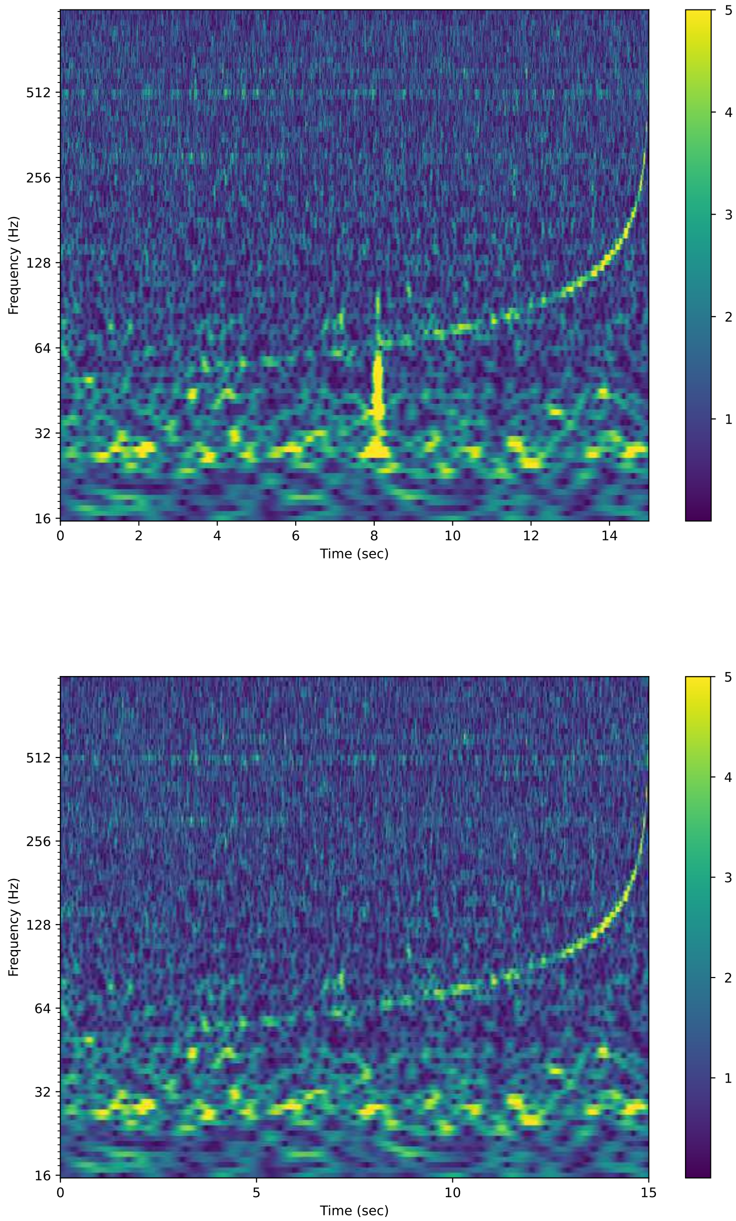

Figure 5. Subtraction of the Tomte Glitch. The top and bottom panels show the CQT of the data and residual, respectively. The glitch is the vertical feature at ≈ 8.0 sec. In order to show both the glitch and the signal in the same image, a threshold has been applied to the CQT as indicated by the maximum value in the colorbar of the top panel.

Our results show that SHAPES is a promising addition to the toolbox of glitch subtraction methods that will become increasingly important as GW detectors become more sensitive. SHAPES is computationally inexpensive, taking on the order of a few minutes for each glitch, and will be made much faster by planned code improvements.

This could allow, in principle, the subtraction of a large number of broadband glitches of known types as part of data conditioning and provide significantly cleaner data for any type of GW search.

Acknowledgments

S.D.M is supported by U.S. National Science Foundation (NSF) grant PHY-2207935 and partially supported by the U.S. Department of Defense grant W911NF2110169. MATC acknowledges support from the Presidential Graduate Research Award at the University of Texas Rio Grande Valley. We acknowledge the Texas Advanced Computing Center (TACC) at the University of Texas at Austin (www.tacc.utexas.edu) for providing high performance computing resources.

References

[1] B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116:061102, Feb 2016.

[2] P. Fritschel. Second generation instruments for the Laser Interferometer Gravitational Wave Observatory (LIGO). In M. Cruise and P. Saulson, editors, *Gravitational-Wave Detection*, volume 4856 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 282–291, March 2003.

[3] J. Degallaix, T. Accadia, F. Acernese, M. Agathos, A. Allocca, et al. Advanced Virgo Status. In G. Auger, P. Binétruy, and E. Plagnol, editors, *9th LISA Symposium*, volume 467 of *Astronomical Society of the Pacific Conference Series*, page 151, January 2013.

[4] BP Abbott, R Abbott, TD Abbott, S Abraham, F Acernese, K Ackley, C Adams, RX Adhikari, VB Adya, C Affeldt, et al. Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs. *Physical Review X*, 9(3):031040, 2019.

[5] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, R. Abbott, et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. *arXiv e-prints*, page arXiv:2111.03606, 2021.

[6] B. P. Abbott et al. GW170817: Observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119:161101, Oct 2017.

[7] K. Somiya. Detector configuration of KAGRA-the Japanese cryogenic gravitational-wave detector. *Classical and Quantum Gravity*, 29(12):124007, June 2012.

[8] C. S. Unnikrishnan. Indigo and ligo-india: Scope and plans for gravitational wave research and precision metrology in india. *International Journal of Modern Physics D*, 22(01):1341010, 2013.

[9] M Punturo et al. The Einstein Telescope: a third-generation gravitational wave observatory. *Classical and Quantum Gravity*, 27(19):194002, 2010.

[10] Sheila Dwyer, Daniel Sigg, Stefan W. Ballmer, Lisa Barsotti, Nergis Mavalvala, and Matthew Evans. Gravitational wave detector with cosmological reach. *Physical Review D*, 91:082001, Apr 2015.

[11] Benjamin P Abbott, R Abbott, TD Abbott, MR Abernathy, K Ackley, C Adams, P Addesso, RX Adhikari, VB Adya, C Affeldt, et al. Exploring the sensitivity of next generation gravitational wave detectors. *Classical and Quantum Gravity*, 34(4):044001, 2017.

[12] Tania Regimbau, Thomas Dent, Walter Del Pozzo, Stefanos Giampanis, Tjonnie G. F. Li, Craig Robinson, Chris Van Den Broeck, Duncan Meacher, Carl Rodriguez, B. S. Sathyaprakash, and

Katarzyna Wójcik. Mock data challenge for the einstein gravitational-wave telescope. *Phys. Rev. D*, 86:122001, 2012.

[13] B P Abbott et al. Effects of data quality vetoes on a search for compact binary coalescences in advanced LIGO's first observing run. *Classical and Quantum Gravity*, 35:065010, 2018.

[14] Jade Powell. Parameter estimation and model selection of gravitational wave signals contaminated by transient detector noise glitches. *Classical and Quantum Gravity*, 35(15):155017, 2018.

[15] Derek Davis, TB Littenberg, IM Romero-Shaw, M Millhouse, J McIver, F Di Renzo, and G Ashton. Subtracting glitches from gravitational-wave detector data during the third observing run. *arXiv preprint arXiv:2207.03429*, 2022.

[16] Rahul Biswas, Lindy Blackburn, Junwei Cao, Reed Essick, Kari Alison Hodge, Erotokritos Katsavounidis, Kyungmin Kim, Young-Min Kim, Eric-Olivier Le Bigot, Chang-Hwan Lee, John J. Oh, Sang Hoon Oh, Edwin J. Son, Ye Tao, Ruslan Vaulin, and Xiaoge Wang. Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data. *Phys. Rev. D*, 88:062003, Sep 2013.

[17] S. Bahaadini, V. Noroozi, N. Rohani, S. Coughlin, M. Zevin, J.R. Smith, V. Kalogera, and A. Katsaggelos. Machine learning for gravity spy: Glitch classification and dataset. *Information Sciences*, 444:172–186, 2018.

[18] S Mukherjee, R Obaid, and B Matkarimov. Classification of glitch waveforms in gravitational wave detector characterization. *Journal of Physics: Conference Series*, 243:012006, aug 2010.

[19] Daniel George, Hongyu Shen, and E. A. Huerta. Classification and unsupervised clustering of ligo data with deep transfer learning. *Phys. Rev. D*, 97:101501, 2018.

[20] Michael Zevin, Scott Coughlin, Sara Bahaadini, Emre Besler, Neda Rohani, Sarah Allen, Miriam Cabero, Kevin Crowston, Aggelos K Katsaggelos, Shane L Larson, et al. Gravity spy: integrating advanced ligo detector characterization, machine learning, and citizen science. *Classical and Quantum Gravity*, 34(6):064003, 2017.

[21] Judith C Brown. Calculation of a constant Q spectral transform. *The Journal of the Acoustical Society of America*, 89(1):425–434, 1991.

[22] S Chatterji, L Blackburn, G Martin, and E Katsavounidis. Multiresolution techniques for the detection of gravitational-wave bursts. *Classical and Quantum Gravity*, 21(20):S1809, 2004.

[23] Bruce Allen. χ^2 time-frequency discriminator for gravitational wave detection. *Physical Review D*, 71:062001, Mar 2005.

[24] BP Abbott et al. Characterization of transient noise in advanced ligo relevant to gravitational wave signal GW150914. *Classical and Quantum Gravity*, 33(13):134001, 2016.

[25] S. A. Usman et al. An improved pipeline to search for gravitational waves from compact binary coalescence. *Classical and quantum gravity*, 33:215004, 2016.

[26] B. Allen, W. Hua, and A. Ottewill. Automatic cross-talk removal from multi-channel data. *ArXiv General Relativity and Quantum Cosmology e-prints*, September 1999.

[27] Jennifer C Driggers, Matthew Evans, Keenan Pepper, and Rana Adhikari. Active noise cancellation in a suspended interferometer. *Review of Scientific Instruments*, 83(2):024501, 2012.

[28] V Tiwari, M Drago, V Frolov, S Klimenko, G Mitselmakher, V Necula, G Prodi, V Re, F Salemi, G Vedovato, and I Yakushin. Regression of environmental noise in ligo data. *Classical and Quantum Gravity*, 32(16):165014, 2015.

[29] G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepańczyk, and S. Vitale. Machine-learning nonstationary noise out of gravitational-wave detectors. *Phys. Rev. D*, 101:042003, Feb 2020.

[30] J. C. Driggers et al. Improving astrophysical parameter estimation via offline noise subtraction for Advanced LIGO. *Phys. Rev. D*, 99:042001, Feb 2019.

[31] Chris Pankow, Katerina Chatziioannou, Eve A. Chase, Tyson B. Littenberg, Matthew Evans, Jessica McIver, Neil J. Cornish, Carl-Johan Haster, Jonah Kanner, Vivien Raymond, Salvatore Vitale, and Aaron Zimmerman. Mitigation of the instrumental noise transient in gravitational-

wave data surrounding gw170817. *Phys. Rev. D*, 98:084016, Oct 2018.

[32] Neil J Cornish and Tyson B Littenberg. Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches. *Classical and Quantum Gravity*, 32(13):135012, 2015.

[33] Katerina Chatziioannou, Neil Cornish, Marcella Wijngaarden, and Tyson B. Littenberg. Modeling compact binary signals and instrumental glitches in gravitational wave data. *Phys. Rev. D*, 103:044013, Feb 2021.

[34] JD Merritt, Ben Farr, Rachel Hur, Bruce Edelman, and Zoheyr Doctor. Transient glitch mitigation in advanced ligo data. *Physical Review D*, 104(10):102004, 2021.

[35] Soumya D Mohanty and Ethan Fahnestock. Adaptive spline fitting with particle swarm optimization. *Computational Statistics*, pages 1–37, 2020.

[36] Carl de Boor. On calculating with b-splines. *Journal of Approximation Theory*, 6(1):50 – 62, 1972.

[37] Svante Wold. Spline functions in data analysis. *Technometrics*, 16(1):1–11, 1974.

[38] Hermann G Burchard. Splines (with optimal knots) are better. *Applicable Analysis*, 3(4):309–319, 1974.

[39] David LB Jupp. Approximation to data by splines with free knots. *SIAM Journal on Numerical Analysis*, 15(2):328–343, 1978.

[40] Zhen Luo and Grace Wahba. Hybrid adaptive splines. *Journal of the American Statistical Association*, 92(437):107–116, 1997.

[41] Satoshi Miyata and Xiaotong Shen. Adaptive free-knot splines. *Journal of Computational and Graphical Statistics*, 12(1):197–213, 2003.

[42] Akemi Gálvez and Andrés Iglesias. Efficient particle swarm optimization approach for data fitting with free knot b-splines. *Computer-Aided Design*, 43(12):1683–1692, 2011.

[43] Soumya D. Mohanty. Particle swarm optimization and regression analysis I. *Astronomical Review*, 7(2):29–35, 2012.

[44] J. Kennedy and R. C. Eberhart. Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks: Perth, WA, Australia*, volume 4, page 1942. IEEE, 1995.

[45] Soumya D. Mohanty. *Swarm Intelligence Methods for Statistical Regression.* Chapman and Hall/CRC, 2018.

[46] Andries P Engelbrecht. *Fundamentals of computational swarm intelligence*, volume 1. Wiley Chichester, 2005.

[47] Daniel Bratton and James Kennedy. Defining a standard for particle swarm optimization. In *Swarm Intelligence Symposium, 2007. SIS 2007. IEEE*, pages 120–127. IEEE, 2007.

[48] Hirotogu Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer, 1998.

[49] S Mukherjee. Median-based noise floor tracker (MNFT): robust estimation of noise floor drifts in interferometric data. *Classical and Quantum Gravity*, 20(17):S925, 2003.

[50] Michele Vallisneri, Jonah Kanner, Roy Williams, Alan Weinstein, and Branson Stephens. The ligo open science center. In *Journal of Physics: Conference Series*, volume 610, page 012021. IOP Publishing, 2015.

[51] Scott Coughlin, Michael Zevin, Sara Bahaadini, Neda Rohani, Sara Allen, Christopher Berry, Kevin Crowston, Mabi Harandi, Corey Jackson, Vicky Kalogera, Aggelos Katsaggelos, Vahid Noroozi, Carsten Osterlund, Oli Patane, Joshua Smith, Siddharth Soni, and Laura Trouille. Gravity Spy Machine Learning Classifications of LIGO Glitches from Observing Runs O1, O2, O3a, and O3b, Zenodo v1.0.0 https://doi.org/10.5281/zenodo.5649212, November 2021.

[52] Brian McFee, Colin Raffel, Dawen Liang, Daniel P Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. librosa: Audio and music signal analysis in python. In *Proceedings of the 14th python in science conference*, volume 8, pages 18–25, 2015.

[53] Matlab Release2020b. *The MathWorks, Inc., Natick, Massachusetts, United States.;* `http://www.mathworks.com/`.

[54] Soumya D. Mohanty. Spline based search method for unmodeled transient gravitational wave chirps. *Physical Review D*, 96:102008, Nov 2017.

[55] S. D. Mohanty. Detection and estimation of unmodeled chirps. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2643–2647, Sep. 2018.