# Identification of noise artifacts in searches for long-duration gravitational-wave transients

**Tanner Prestegard**[1]**, Eric Thrane**[1]**, Nelson L Christensen**[2]**, Michael W Coughlin**[2]**, Ben Hubbert**[2]**, Shivaraj Kandhasamy**[1]**, Evan MacAyeal**[2] **and Vuk Mandic**[1]

[1] School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
[2] Physics and Astronomy, Carleton College, Northfield, MN 55057, USA

E-mail: ethrane@physics.umn.edu

## Abstract

We present an algorithm for the identification of transient noise artifacts (glitches) in cross-correlation searches for long gravitational-wave (GW) transients lasting seconds to weeks. The algorithm utilizes the auto-power in each detector as a discriminator between well-behaved stationary noise (possibly including a GW signal) and non-stationary noise transients. We test the algorithm with both Monte Carlo noise and time-shifted data from the LIGO S5 science run and find that it removes a significant fraction of glitches while keeping the vast majority (99.6%) of the data. We show that this cleaned data can be used to observe GW signals at a significantly lower amplitude than can otherwise be achieved. Using an accretion disk instability signal model, we estimate that the algorithm is accidentally triggered at a rate of less than $10^{-5}\%$ by realistic signals, and less than 3% even for exceptionally loud signals. We conclude that the algorithm is a safe and effective method for cleaning the cross-correlation data used in searches for long GW transients.

PACS number: 95.55.Ym

(Some figures may appear in colour only in the online journal)

## 1. Introduction

Our aim is to detect long-lasting gravitational-wave (GW) transients (lasting seconds to weeks) in the presence of 'glitches': non-stationary noise artifacts that contaminate the otherwise approximately Gaussian strain noise in GW interferometers. We focus our attention on the cross-correlation method of [1], though, it may be possible to extend this formalism to other search algorithms as well—a topic of ongoing research. Possible sources of long GW transients include convection in proto-neutron stars [2–7], rotational instabilities associated with nascent

neutron stars [2, 8–13], instabilities in the disks of accreting systems [14–17], neutron star glitches [18], soft gamma repeaters/anomalous x-ray binaries [19–25] and dynamically formed black hole binaries [26–29].

Glitches can arise from environmental contamination such as mechanical vibrations, electromagnetic disturbances, circuit breaker trips, power shorts and asymmetric photodiode response [30]. While some glitches can be identified and removed by comparing GW strain channels with environmental and sub-system monitoring channels, many remain after the first stages of data cleaning (see, e.g., [30–37]). These remaining glitches require special attention for two reasons. First, a high glitch rate can diminish the sensitivity of a search by raising the threshold required for an event to be statistically significant[3]. Indeed, below we shall show a realistic example wherein the required signal power for a $p = 0.1\%$ false alarm probability event drops two-fold when we use our algorithm to remove glitchy segments from GW data. This level of improvement is not achievable with the application of existing data-quality flags. Second, robust glitch identification methods can improve our confidence in a GW candidate if it does not resemble non-stationary noise.

We describe an algorithm to check the consistency of the auto-power from two terrestrial GW detectors to identify glitches in searches using the cross-power statistic described in [1]. (Throughout, we use the expressions 'auto-power' and 'cross-power' instead of 'power spectrum', which can refer to either.) We demonstrate the ability of the algorithm to improve the sensitivity of targeted searches by cleaning real interferometer data to a level approaching optimally well-behaved Gaussian noise.

This work builds on [33], which described how environmental monitoring channels can be used to identify *long-lasting noise* transients. However, it differs because first, *we utilize only GW strain channels*, and second, because we are interested in recovering *long-lasting GW signals* in the presence of what are sometimes very *short bursts of noise*. It also differs from [38, 39] and other consistency-check algorithms that the authors are aware of because we are not checking the consistency of GW triggers, but rather we are checking the consistency of *data segments*—many of which will together constitute a GW trigger. This is born of necessity from our focus on long transients. We shall see in section 4 that by flagging individual segments as glitchy, we are able in principle to observe a GW event temporarily disturbed by non-stationary noise.

To illustrate our glitch identification algorithm, we use Monte Carlo (MC) and time-shifted (TS) data from the 4 km LIGO H1 and L1 interferometers [40] in Hanford, WA and Livingston, LA, respectively. Time-shifting one strain time series with respect to another by an amount greater than the GW travel time between interferometers removes astrophysical signals while preserving non-Gaussian noise artifacts that are otherwise difficult to simulate. Our MC assumes Gaussian noise with an initial LIGO design sensitivity, and our TS data are from the 5 November 2005–30 September 2007 S5 science run (see, e.g., [41, 42]). During S5, the LIGO interferometers achieved a strain sensitivity of $\approx 3 \times 10^{-23}$ Hz$^{-1/2}$ in the most sensitive band around 100–200 Hz. We utilize a few days of accumulated data from GPS = 816 065 659–819 039 020. By comparing how the glitch identification algorithm performs for MC and TS results, we can measure how close we can get to ideal Gaussian noise by cleaning non-Gaussian noise. While we use the LIGO H1 and L1 detectors for illustrative purposes,

---

[3] The astute reader may wonder how the present concern about glitches should be squared with the finding in [1] that the signal-to-noise ratio distributions for time-shifted (TS) and Monte Carlo (MC) 'are in qualitative agreement'. Do we really need to worry about glitches in searches for long GW transients? The answer is yes. The results presented in [1] compared the standard deviation and approximate shape of distributions of *pixel* SNR for MC and time-shift data. While this comparison showed that the distributions are similar, our present analysis focuses on the high-SNR tail of the distribution of *clusters* of pixels. Since glitches tend to produce clusters of pixels of non-Gaussian noise, their importance is magnified when we study the distribution of cluster SNR.

we expect that these techniques can be extended to additional pairs of detectors including interferometers such as Virgo [43–46], LCGT [47, 48] and GEO [49–52].

The outline for the rest of this paper is as follows. In section 2 we summarize the cross power-based analysis framework from [1]. In section 3 we develop an autopower difference statistic that can be used to evaluate whether the autopower in a pair of detectors is consistent with noise plus a GW signal. We analyze the behavior of this statistic for stationary noise, signals, and glitches. In section 4 we present a glitch identification algorithm based on the autopower difference statistic and demonstrate its ability to clean TS LIGO data. In section 5 we introduce an accretion disk instability (ADI) waveform, which we use in section 6 to investigate the safeness of our algorithm, i.e., the probability that it falsely identifies a signal as glitch-like. In section 7 we investigate the complementarity of our algorithm to data quality flags based on instrumental and environmental noise artifacts. Section 8 contains concluding remarks.

## 2. Formalism

Our starting point is [1], which is described in greater detail in the appendix. We use the cross-correlation of two or more spatially separated interferometers to construct a statistic $\hat{Y}(t; f)$, which is an unbiased estimator for the GW power $H(t; f)$ between times $t$ and $t + \delta t$ in some frequency bin between $f$ and $f + \delta f$. $H(t; f)$ is defined in terms of the GW field Fourier coefficients, $\tilde{h}_A(f)$ (see the appendix):

$$H(t; f) = \text{Tr}\,[H_{AA'}(t; f)] = \text{Tr}\left[\frac{2}{\mathcal{N}} \langle \tilde{h}_A(t; f)\tilde{h}_{A'}(t; f)\rangle\right] \tag{1}$$

Here the brackets $\langle \ldots \rangle$ denote the expectation value of the enclosed quantity. The semicolon emphasizes that $t$ refers to the beginning of a data segment of length $\delta t$ and not to the many sampling times associated with each segment. It is important that the noise in the two interferometers is uncorrelated, which is easily achieved for spatially separated interferometers.

The set of $\hat{Y}(t; f)$ can be represented as an $ft$-map (spectrogram). The same is true of $\hat{\sigma}(t; f)$, an estimator for the uncertainty associated with $\hat{Y}(t; f)$. GW candidates are identified as clusters of high SNR $\equiv \hat{Y}/\hat{\sigma}$ pixels [1]. The significance of a cluster $\Gamma$ can be estimated by calculating the total SNR for the entire cluster, denoted SNR($\Gamma$), and comparing it to the distribution of SNR($\Gamma$) obtained with TS data [1].

For sufficiently long signals, the effect of non-stationary noise is averaged away and SNR($\Gamma$) becomes Gaussian distributed by the central limit theorem. This limiting case is the stochastic radiometer—a technique for mapping the GW sky with two or more spatially separated interferometers [53–55]. Here, however, we study (relatively) shorter time scales where glitches play a role in our ability to determine the significance of an event. The question we aim to investigate in the rest of the paper is: how can we discriminate between large values of SNR($\Gamma$) due to a GW signal and large values due to glitches?

Our glitch identification algorithm will utilize cross-power $\hat{C}_{IJ}(t; f)$ and auto-power $\hat{P}_I(t; f)$, which are related to $H(t; f)$ by the 'pair efficiency' $\epsilon_{IJ}(t; \hat{\Omega})$:

$$\langle \hat{C}_{IJ}(t; f)\rangle \equiv \epsilon_{IJ}(t; \hat{\Omega}, \vec{\alpha})H(t; f)\,e^{-2\pi i f \tau_{IJ}} \tag{2}$$

$$\langle \hat{P}_I(t; f)\rangle \equiv \epsilon_{II}(t; \hat{\Omega}, \vec{\alpha})H(t; f) + N_I(t; f). \tag{3}$$

Here $\tau_{IJ}$ is the direction-dependent time delay between detector $I$ and detector $J$ and $N_I(t; f)$ is the noise power in detector $I$. For additional details, including an expression for $\epsilon$, see the appendix. It is also useful to define $\hat{P}'_I(t; f)$, the power in the $2n$ segments neighboring $t$:

$$\hat{P}'_I(t_0; f) \equiv \frac{1}{2n} \left[ \sum_{t=t_0-n\delta t}^{t=t_0+n\delta t} \hat{P}_I(t; f) \right] - \frac{1}{2n} \hat{P}_I(t_0; f). \tag{4}$$

In this analysis we use $n = 4$ neighboring segments on each side.

## 3. An auto-power difference statistic

Since the noise and the signal are uncorrelated, the expectation value of $\hat{P}_I(t; f)$ is given by equation (3). If we assume that $N_I(t; f)$ can be estimated by looking at neighboring segments of noise, (i.e., the noise is stationary), then we can construct an estimator for the observed auto-power in detector $I$ due to GWs:

$$\frac{\hat{P}_I(t; f) - \hat{P}'_I(t; f)}{\epsilon_{II}}. \tag{5}$$

We assume that there is no (or comparatively little) signal present in the same frequency bin during these neighboring times[4] so that

$$\langle \hat{P}'_I(t; f) \rangle \approx N_I(t; f). \tag{6}$$

Similarly, an estimator for the GW auto-power in detector $J$ is

$$\frac{\hat{P}_J(t; f) - \hat{P}'_J(t; f)}{\epsilon_{JJ}}. \tag{7}$$

We now construct an estimator, which represents the GW auto-power difference between detectors $I$ and $J$:

$$\hat{\Xi}(t; f) \equiv \frac{\hat{P}_I(t; f) - \hat{P}'_I(t; f)}{\epsilon_{II}} - \frac{\hat{P}_J(t; f) - \hat{P}'_J(t; f)}{\epsilon_{JJ}}. \tag{8}$$

By construction, we expect that $\langle \hat{\Xi}(t; f) \rangle = 0$ for well-behaved noise plus a signal that is well modeled by the pair efficiencies $\epsilon_{II}, \epsilon_{JJ}$. We note that $|\hat{\Xi}(t; f)|$ is invariant under $I \leftrightarrow J$.

It is desirable to normalize $\hat{\Xi}(t; f)$ such that the new quantity is unitless with a near-unity variance. The variance of $\hat{\Xi}(t; f)$ is given by

$$\sigma_\Xi^2(t; f) = \frac{\langle \hat{P}_I(t; f) \rangle^2 + \langle \hat{P}'_I(t; f) \rangle^2}{\epsilon_{II}^2} + \frac{\langle \hat{P}_J(t; f) \rangle^2 + \langle \hat{P}'_J(t; f) \rangle^2}{\epsilon_{JJ}^2} - \frac{2 \epsilon_{IJ}^2}{\epsilon_{II} \epsilon_{JJ}} |\langle \hat{Y}(t; f) \rangle|^2. \tag{9}$$

This motivates a normalization factor denoted $\hat{\sigma}_\Xi^2(t; f)$, which we choose to be

$$\hat{\sigma}_\Xi^2(t; f) \equiv \frac{\hat{P}_I^2(t; f) + \hat{P}_I'^2(t; f)}{\epsilon_{II}^2} + \frac{\hat{P}_J^2(t; f) + \hat{P}_J'^2(t; f)}{\epsilon_{JJ}^2} - \frac{2 \epsilon_{IJ}^2}{\epsilon_{II} \epsilon_{JJ}} |\hat{Y}(t; f)|^2. \tag{10}$$

We shall see below that this normalization provides an effective means of creating a unitless signal-to-noise ratio $\mathrm{SNR}_\Xi \equiv \hat{\Xi}/\hat{\sigma}_\Xi$, which we can use to determine if the auto-power in two interferometers is consistent with a GW signal plus well-behaved (stationary) noise. The $(t; f)$ dependence of $\mathrm{SNR}_\Xi(t; f)$ is implicit. Note that $\mathrm{SNR}_\Xi$ is not equivalent to the cross-correlation statistic $\mathrm{SNR} \equiv \hat{Y}/\sigma_Y$.

---

[4] This approximation works best for narrowband signals whose frequency varies significantly with time, as is the case for the examples shown here (see, e.g., figure 1). When the approximation is poor, e.g., for a monochromatic signal, then $\langle \hat{P}'_I(t; f) \rangle$ may include a significant GW component as well, though, $\hat{\Xi}(t; f)$ (defined in equation (8)) will behave much the same way as it is still the case that $\langle \hat{\Xi}(t; f) \rangle = 0$ by construction.
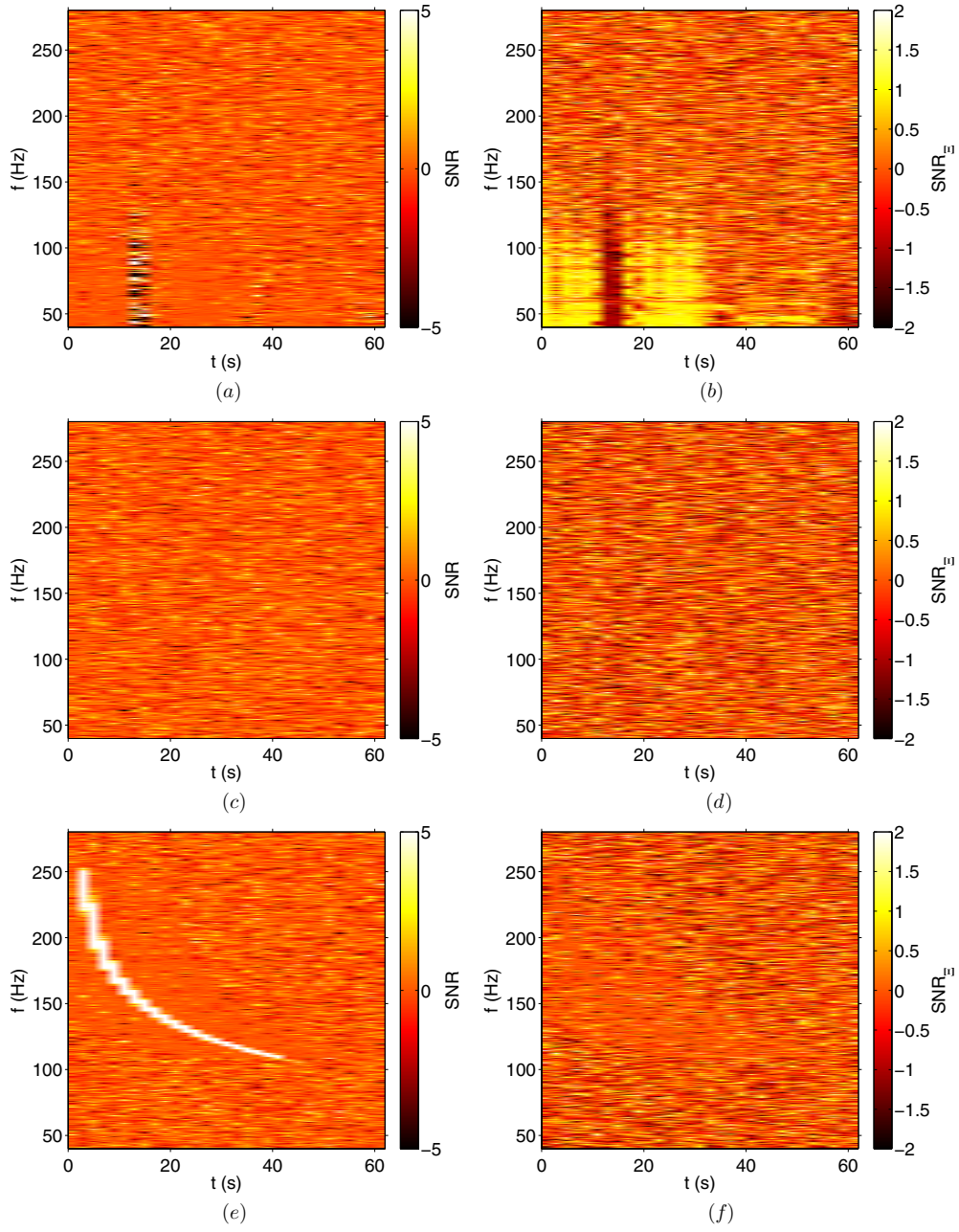
**Figure 1.** *Ft*-maps of TS LIGO S5 data. The left-hand column shows cross-power SNR while the right-hand column shows $\mathrm{SNR}_\Xi$ for the same data. Top row: a likely glitch. Middle row: nearly stationary noise. Bottom row: stationary noise plus a simulated circularly polarized ADI waveform ($d = 5$ Mpc) (see [15–17]).

By considering equations (8) and (9), it is apparent that the qualitative behavior of $\mathrm{SNR}_\Xi$ is different for signals $\langle \hat{Y}(t; f) \rangle > 0$ and glitches $P_I(t; f) \gg P_J(t; f)$. Loud glitches in detectors $I, J$ cause $\mathrm{SNR}_\Xi \approx \pm 1$ surrounded by $\mathrm{SNR}_\Xi \approx \mp 1$ (see, e.g., figure 1, top row).

Neighboring segments are affected due to our noise estimation technique, which averages adjacent segments in time (see the appendix). Loud GW signals, on the other hand, cause $\text{SNR}_\Xi \approx 0$ surrounded by $\text{SNR}_\Xi \approx 0$ with larger fluctuations in the neighboring segments. This qualitative description of the $\text{SNR}_\Xi$ in the presence of a GW signal is demonstrated in figure 2 as well as the bottom row of $ft$-maps in figure 1.

## 4. Glitch identification

Having introduced the auto-power difference statistic $\text{SNR}_\Xi$, we now present an $\text{SNR}_\Xi$-based algorithm to identify glitches. We use data collected from the LIGO S5 science run. Our network consists of the two 4 km LIGO interferometers mentioned in section 1. The data are time-shifted by a duration greater than the GW travel time between H1 and L1 in order to wash out the presence of astrophysical signals. To begin, we utilize $ft$-maps with 4 s × 0.25 Hz pixels.

In figure 2 we show the distribution of $\text{SNR}_\Xi$ for well-behaved noise (top), glitchy noise (middle) and a simulated ADI GW signal (see section 5) injected on top of Gaussian noise (bottom). 'Well-behaved' means that there are no obvious high-level glitches visible in an $ft$-map of SNR, which is to say that the data approximate stationary Gaussian noise. As examples of glitchy noise, we utilize data from two extreme glitches; one from H1 and one from L1. As stated in section 3, we observe that glitches cause an excess of pixels near $|\text{SNR}_\Xi| = 1$. However, if we simply flag segments with $|\text{SNR}_\Xi| \approx 1$, we will throw out more data than necessary because segments neighboring a glitch also exhibit $|\text{SNR}_\Xi| \approx 1$.

To discriminate between the glitch segment and its neighbors, we define an additional metric, the *auto-power stationarity ratio*:

$$R_I(t) = \frac{1}{N_f} \sum_f \frac{\hat{P}_I(t; f)}{\hat{P}'_I(t; f)}. \tag{11}$$

Here $N_f$ is the number of frequency bins. We expect segments with a glitch to have $R_I(t) \gtrsim 1$ whereas neighboring segments should have $R_I(t) < 1$. (Of course, GW signals can also lead to $R_I(t) \gtrsim 1$ so it is necessary to use $R$ in conjunction with $\text{SNR}_\Xi$ in order to separate glitches from GW events.) Glitches are unlikely to occur in two interferometers at the same time.

Now we are ready to devise our glitch likely flag. A data segment (or equivalently, an $ft$-map column) is identified as *glitch-like* if either of the following criteria are satisfied:

$$\geqslant 2.7\% \text{ of pixels have } 0.95 < \text{SNR}_\Xi < 1.05 \text{ and } R_I(t) > 2 \text{ and } R_J(t) \leqslant 2. \tag{12a}$$

$$\geqslant 2.7\% \text{ of pixels have } -1.05 < \text{SNR}_\Xi < -0.95 \text{ and } R_J(t) > 2 \text{ and } R_I(t) \leqslant 2. \tag{12b}$$

These parameters are chosen primarily to optimize the efficiency of our algorithm at rejecting glitches, though some fine tuning is necessary to ensure the safety of a particular signal model. In this case, the parameters are adjusted for the ADI model (see figure 8), but we shall see that they are also effective for a very different signal model (based on accretion disk fragmentation) in section 5. Before we continue, it will be useful to define $\mathcal{F}$ as the ratio of the number of pixels at some time $t$ satisfying the criterion $0.95 < |\text{SNR}_\Xi| < 1.05$ to the total number of pixels at time $t$. Note that $\mathcal{F}$, by definition, must take on discrete values.

In order for this to be an effective flag, it must not only identify glitch-like structures in the data, but it should also have a low *false glitch rate*. We define the false glitch rate as the fraction of Gaussian-noise segments (containing no glitches) flagged as glitch-like per unit time. Using simulated Gaussian data, we estimate a false glitch rate of $\lesssim 1 \times 10^{-3}$ day$^{-1}$. This false glitch rate is calculated for a frequency range between 100–250 Hz consisting of
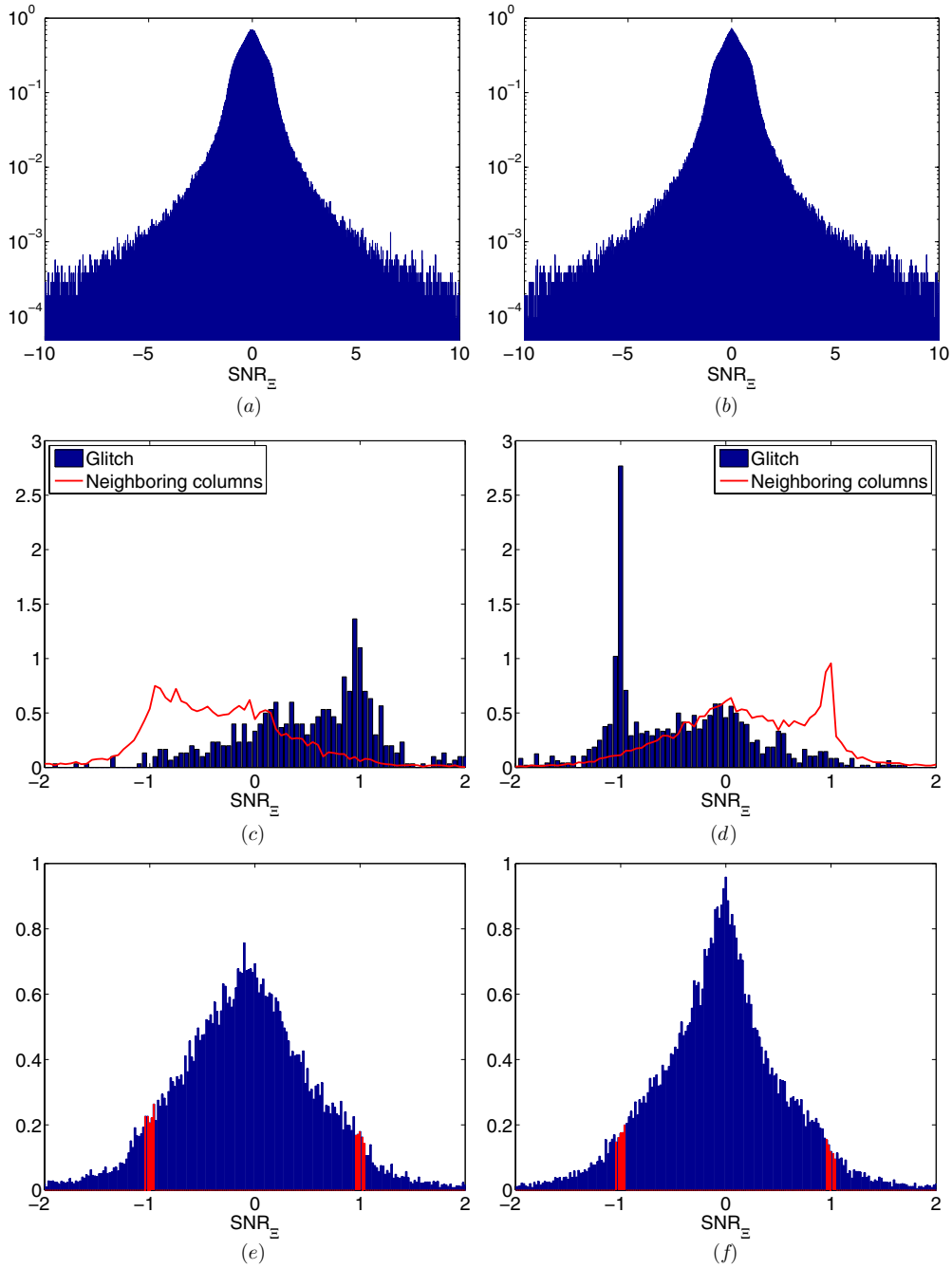
**Figure 2.** Histograms of $\mathrm{SNR}_\Xi$ $ft$-map pixels. Top: 2200 s of Gaussian MC noise (left) and well-behaved (nearly-Gaussian) TS LIGO S5 data (right). Middle: two examples of glitches in H1 (left) and L1 (right) each consisting of 2 s of data. These data segments were chosen to illustrate examples of strong glitches. Bottom: 40 s-long circularly-polarized ADI injections recovered with an unpolarized filter at 30 Mpc (left) and at 5 Mpc (right). The red bars indicate $0.95 < |\mathrm{SNR}_\Xi| < 1.05$. The $x$-axis range differs between rows due to the different amount of data being analyzed in each case.
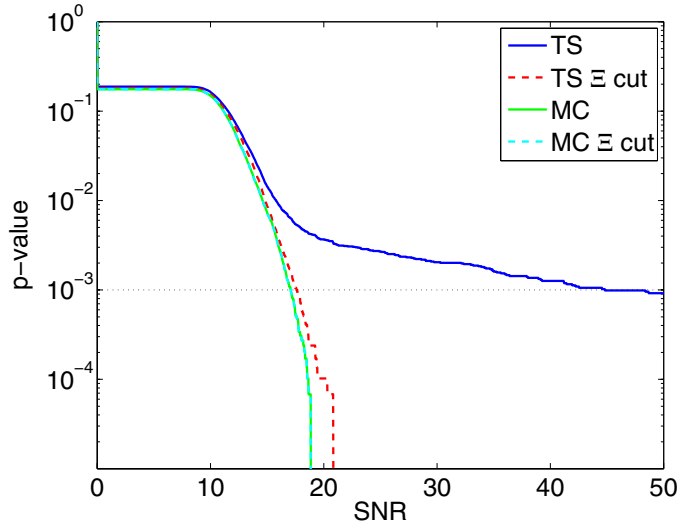
**Figure 3.** Plot of *p*-value versus SNR for a density-based search algorithm [56] applied to TS LIGO S5 data (TS) and Gaussian MC data, with and without our $SNR_\Xi$-based glitch cut applied, for 4 s × 0.25 Hz pixels in a frequency band of 100–250 Hz. The $SNR_\Xi$-based glitch cut improves the sensitivity at $p = 0.1\%$ (marked with a black dotted line) by more than a factor of two. The asymptotic *p*-value at low SNR is the probability that any above-threshold cluster is identified.

150 pixels, a range suitable for the ADI model that we will use to test this algorithm (see section 5).

　　To determine the effectiveness of our flag, we perform a background study comparing TS data (containing stationary noise and glitches) with MC (stationary noise) with and without flagged data removed. An effective flag eliminates high-SNR events from the tail of the distribution, thereby creating better agreement between TS and MC data. We utilize a density-based search algorithm [56] to analyze 12 s × 150 Hz *ft*-maps with 4 s × 0.25 Hz pixels[5]. We focus on a frequency range of 100–250 Hz in order to study the ADI signal discussed in section 5 (see figure 1, bottom row). In figure 3, we plot *p*-value (false alarm probability) versus SNR for MC and TS data with and without the glitch-likely flag. The results indicate a significant improvement in the agreement between TS and MC data with the application of the flag. The required SNR for a $p = 0.1\%$ event is reduced more than twofold through the use of the glitch identification flag.

　　Having demonstrated the efficacy of our glitch identification algorithm for the case of 4 s × 0.25 Hz pixels in the 100–250 Hz band chosen to study the ADI model (see section 5), we now consider a few other cases. An exhaustive exploration of the domain of utility for this algorithm is beyond our present scope. Rather, we aim to highlight both the promise and the limitations of this technique by considering a few more special cases. In the top-left panel of figure 4, we plot *p*-value versus SNR for 1 *s* × 1 Hz pixels in the same 100–250 Hz frequency band used in figure 3. Based on the agreement between TS and MC data, we conclude that even relatively short segments of data can be effectively cleaned with the glitch identification algorithm so as to achieve good agreement between MC and TS data.

---

[5] The algorithm is a modified version of `BurstCluster` by Peter Kalmus and Rubab Khan created for the LIGO Flare Pipeline (see [57, 58]).
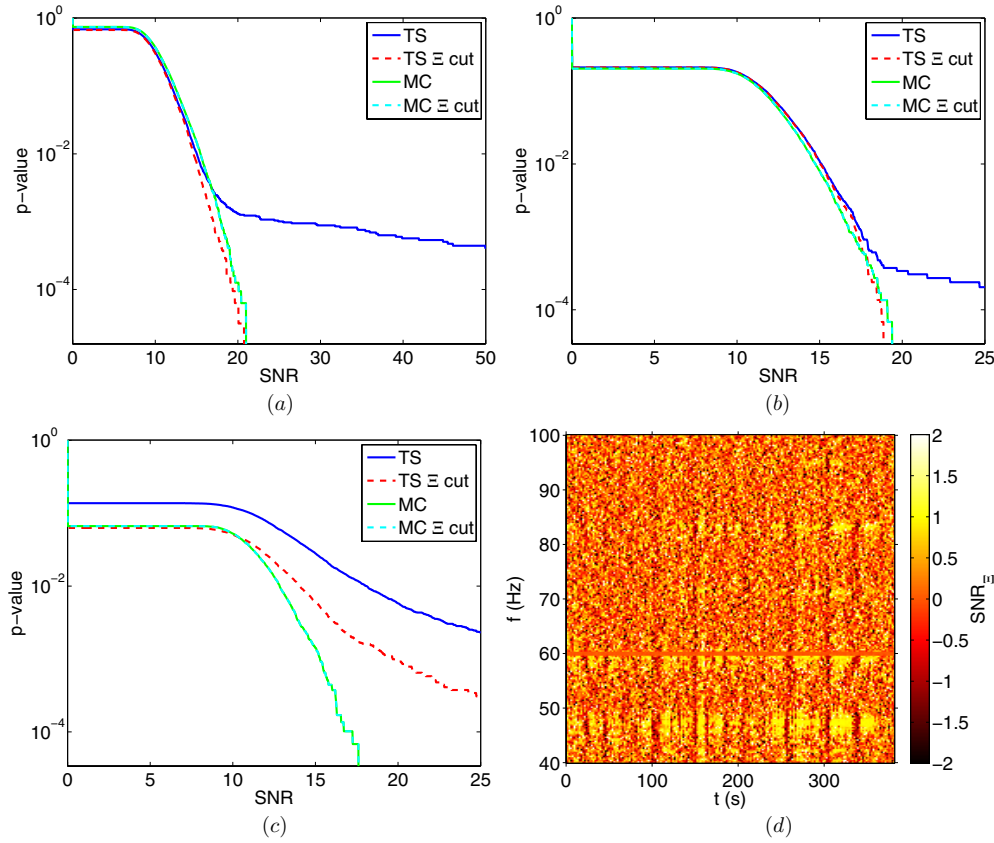
**Figure 4.** Top-left: plot of *p*-value versus SNR for the 100–250 Hz band using 1 s × 1 Hz
pixels (the asymptotic *p*-value at low SNR differs from the 4 s × 0.25 Hz case since we tune the
clustering algorithm differently for different segment durations). The relatively good agreement
between MC and TS data suggests that even short $\mathcal{O}(1\,\text{s})$ segments of cross-correlated data can be
effectively cleaned with our glitch identification flag. Top-right: plot of *p*-value versus SNR for the
375–525 Hz band using 4 s × 0.25 Hz pixels. This higher frequency band exhibits good agreement
between MC and TS data due to the nearly stationary noise associated with higher frequencies.
Bottom-left: plot of *p*-value versus SNR for the 40–100 Hz band using 4 s × 0.25 Hz pixels. While
the cut dramatically improves the agreement between MC and TS data, significant disagreement
remains, possibly due to non-stationary noise associated with this band. Bottom-right: an *ft*-map
of $\text{SNR}_\Xi$ for TS LIGO S5 data demonstrating the non-stationary noise sometimes associated with
low frequencies. The 60 Hz line is masked.

In the top-right plot, we show the case of 4 s × 0.25 Hz pixels in a higher frequency
band: 375–525 Hz. Again we observe good agreement between TS and MC data, though, this
is not surprising since this higher frequency band is typically dominated by nearly stationary
noise. On the bottom-left, we show the case of 4 s × 0.25 Hz pixels in a lower frequency
band: 40–100 Hz. While the agreement between TS and MC data is improved with the
glitch identification algorithm, significant disagreement remains due to non-stationary noise,
which is more common at lower frequencies. An $\text{SNR}_\Xi$ *ft*-map from a period of noisy low-
frequency data is included in the bottom-right panel, which indicates that this effect may be
due to quasi-continuous broadband noise rather than infrequent glitches. It is possible that the
inclusion of additional vetoes utilizing physical environmental monitors such as microphones
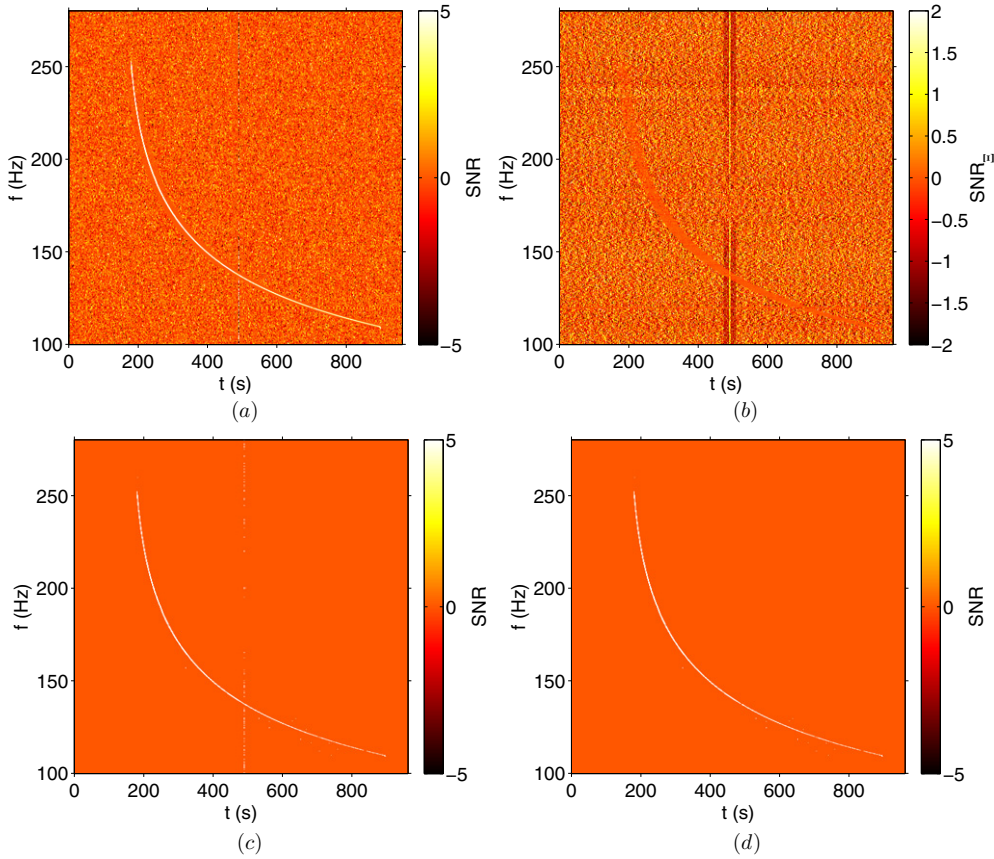
**Figure 5.** *Ft*-maps of an ADI software injection in TS data containing a glitch (near $t \approx 490$ s). Top-left is SNR and top-right is $\text{SNR}_\Xi$. The bottom plots show the recovered track without (left) and with (right) the glitch identification algorithm. The glitch identification algorithm excludes the glitch (visible as a vertical column of bright pixels) from the reconstructed track.

and seismometers may help achieve better agreement between TS data in this band and MC noise.

Finally, in figure 5, we demonstrate how the glitch identification algorithm can be used to improve the accuracy of a long reconstructed signal by removing one or more glitchy segments. Motivated by models of long GW transients, which may last for hundreds of seconds or longer (e.g. [59, 60]), we consider a 700 s long ADI waveform. We inject the waveform into TS Gaussian noise during a period with a known glitch (visible as a vertical column around $t \approx 490$ s). Using the density-based clustering algorithm described above, we recover the track without (bottom-left) and with (bottom-right) the glitch identification algorithm applied. The glitch identification algorithm correctly identifies the glitch, which is therefore excluded from the reconstructed event. This demonstrates not only that the glitch identification algorithm improves the accuracy of a reconstructed track, but also that it is in theory possible to observe a GW event disrupted by a glitch. While this possibility is discussed in [33], this is (to our knowledge) the first time that a method has been proposed for removing pieces of glitchy data from the middle of a GW trigger using only strain data.
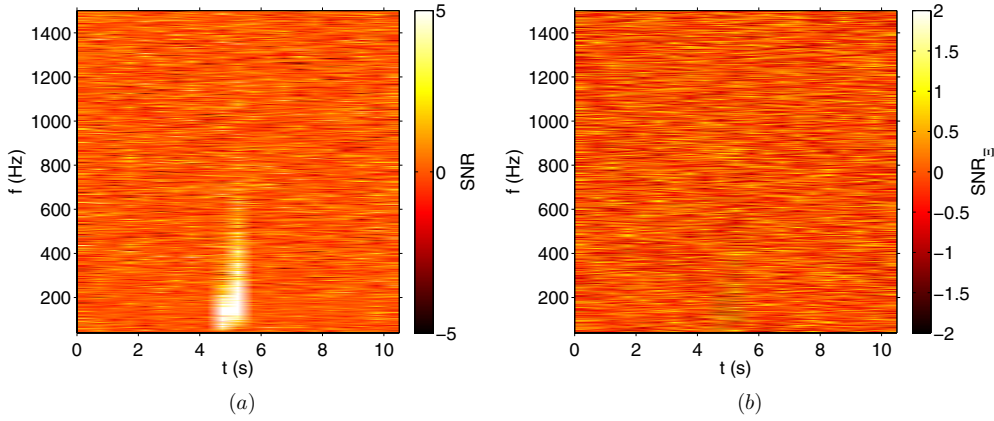
**Figure 6.** *Ft*-maps of SNR (left) and $\mathrm{SNR}_\Xi$ (right) for a $\approx 1$ s accretion disk fragmentation waveform injected into stationary noise ($d = 1$ Mpc).

## 5. Toy model waveforms

In order to demonstrate our glitch identification algorithm, we utilize a toy model [61] for a narrowband signal from an accretion disk instability, which can take place during the collapsar death of a star and may therefore be associated with long gamma-ray bursts (see also [15, 16]). In this 'suspended accretion' model, a spinning black hole (with mass $M$ and parameterized by a dimensionless spin parameter $a^\star$) is surrounded by a torus (mass $m$). The spinning black hole drives magneto-hydrodynamical turbulence in the torus, which causes it to form clumps with mass given by $\epsilon m$. These clumps emit elliptically polarized narrowband gravitational radiation for a duration of $\mathcal{O}(10\text{–}100$ s) as the central black hole transfers its angular momentum to the clumps. This toy model provides a useful test of our algorithm because we expect many sources of long GW transients to be both narrowband and elliptically polarized [1]. We use $M = 10\,M_\odot$, $m = 1.5\,M_\odot$, $\epsilon = 0.1$ and $a^\star = 0.95$ to create the $\approx 40$ s waveform used here. A spectrogram of this waveform can be seen in the bottom-left panel of figure 1. For more details see [61].

In addition to the ADI model, we also consider an accretion disk fragmentation model from [61]. In this model, an accretion disk associated with a long gamma-ray burst forms clumps through helium photo-disintegration [14]. These clumps inspiral into the remnant black hole, creating a chirp-like GW signal.

The fragmentation model can be tuned to produce shorter burst-like signals. Burst-like signals present an extra challenge to the glitch identification algorithm because, like a glitch, the power is typically concentrated in a single $\mathcal{O}(1\ s)$-wide *ft*-map column (though we still expect the auto-power between two interferometers to be consistent for a well-constructed filter). While we are primarily concerned here with *long* transients, we use a short $\approx 1$ s fragmentation waveform in section 6 in order to study the performance of the glitch identification flag in this limiting case. We shall see that, while the flag performs best for long transients, the false dismissal rate is low even for short signals unless the signal is unrealistically loud. The fragmentation waveforms from [61] are parameterized by the mass of the central black hole $M$, the torus scale height $\eta$, the torus viscosity $\alpha$ and the initial radius $r_0$ (in units of black hole mass). We use $M = 10\,M_\odot$, $\eta = 0.8$, $\alpha = 0.1$ and $r_0 = 200$ to create the $\approx 1$ s waveform here. *Ft*-maps of this fragmentation waveform are shown in figure 6.
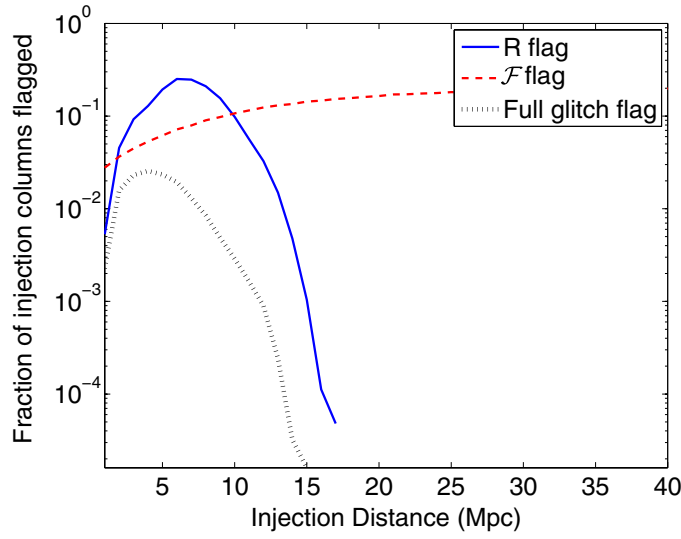
**Figure 7.** Safety study for simulated ADI signals. The *x*-axis is the distance to the source. For each distance, we average over 312 directions and 20 noise realizations. The *y*-axis is the fraction of segments satisfying the *R* criteria (solid blue), the $SNR_\Xi$ criteria (dashed red) and satisfying both in such a way as to be flagged as glitch-like (dotted black). Note that the fraction of segments flagged as glitch-like decreases at closer distances (corresponding to louder signals) because both detectors exceed the threshold on $R(t)$, which prevents the signal from being flagged as glitch-like (see equations (12*a*) and (12*b*)).

## 6. Safety

A critical aspect of any glitch identification algorithm is its *safeness*: the probability that it falsely identifies a segment associated with a GW signal as glitch-like. To test the safeness of our glitch flag, we apply it to ADI injections in Gaussian simulated noise at different sky locations. Many long transient signals (including the ADI model considered here) are expected to be elliptically polarized [1]. In practice, however, it is possible to search for such signals with an unpolarized filter since the two-detector statistic $\hat{Y}(t; f)$ is largely unaffected by polarization details, if the signal is not so long that the polarization degeneracy is resolved by the rotation of the Earth. In this analysis we use circularly polarized waveforms, a plausible model for many elliptically polarized sources with electromagnetic triggers, which tend to be observed head-on [62].

In figure 7 we present the results of a safety study in which we perform MC injections of ADI signals on top of Gaussian simulated noise. We use 312 uniformly distributed sky directions with 20 noise realizations for each direction. To be flagged as glitch-like, a segment must satisfy our requirements on $R(t)$ and $SNR_\Xi$ (see equations (12*a*) and (12*b*)). For each injection we record the fraction of segments satisfying the requirements on $R(t)$ alone (blue), $SNR_\Xi$ alone (red) and segments meeting both criteria which are therefore identified as glitch-like (black). Note that our ADI signal spans $\approx$ 39 data segments. For marginally detectable signals (a $d$ = 38 Mpc signal can be recovered with $p$ = 0.1%), the fraction of flagged segments is negligible. For a very loud signal at $d$ = 5 Mpc (see the lower-left-hand plot in figure 1), the fraction of flagged segments becomes 3%. We conclude that for realistic (marginally detectable signals), the proposed glitch identification flag leads to an acceptably
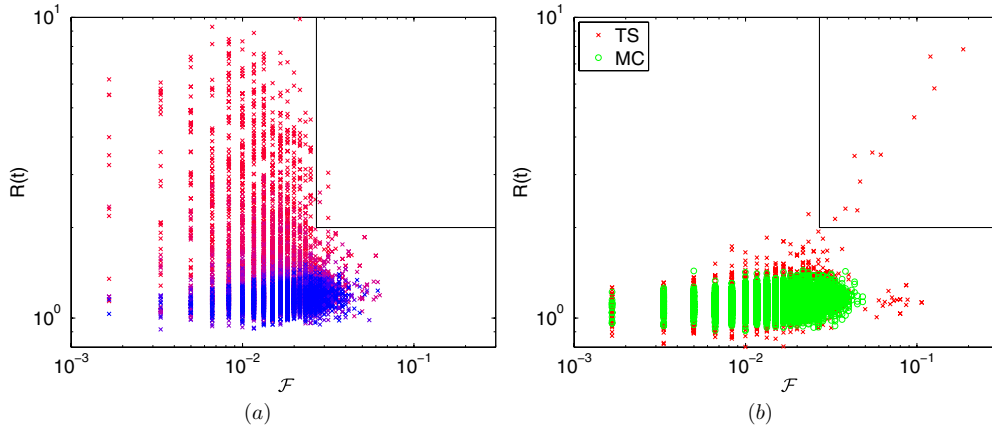
**Figure 8.** Scatter plots of injected ADI signals (left) and TS data (right red ×'s) and MC noise (right green ○'s) in the plane of our glitch identification parameters $R(t)$ and $\mathcal{F}$. Injection distances range from 5–40 Mpc with smaller distances corresponding to redder data markers. The glitch identification thresholds for each parameter are represented by black lines and points in the upper right quadrant are flagged as glitch-like. Note that $\mathcal{F}$ takes on discrete values as discussed below equation (12a).

small false dismissal rate. In order to further reduce the false dismissal rate for very high-SNR signals, one could design a less aggressive auto-power cut for triggers with extremely high $SNR_\Xi$, but this is beyond our present scope.

We also consider the case of the short $t \approx 1$ s accretion disk fragmentation signal described in section 5. In order to test the glitch rejection algorithm on this short signal, we inject the waveform on top of MC noise. We vary the distance of the injection and perform many trials at each distance, averaging over sky location. For a very loud $d = 1$ Mpc signal, the false dismissal probability is high: 21%. However, we find that false dismissal probability is $< 1\%$ for signals at $d > 2.4$ Mpc. While our clustering algorithm is not designed for signals that are vertical $ft$-map columns, we can estimate our sensitivity to short signals by summing all the pixels in the brightest column in order to calculate a total SNR for that segment [1]. For signals at $d = 2.4$ Mpc, the total SNR $\approx 12$ on average. For a quasi-normally distributed quantity like total SNR, this corresponds to an extremely small $p$-value. We conclude that even for very short signals, the false dismissal probability is small for signals with realistic values of SNR, though, unrealistically high values of SNR have a significant probability of being flagged as glitch-like.

Having discussed both the efficacy of the algorithm flagging glitches as well as its safeness *not* flagging segments associated with GW signals, it is interesting to consider the parameter space of the cut. In figure 8, we show scatter plots of injected ADI signals (left) and noise (right) in the plane of our glitch identification parameters $R(t)$ and $\mathcal{F}$. The horizontal and vertical lines indicate the glitch-likely thresholds. Data markers in the upper right-hand quadrant satisfying both cuts are flagged as glitch-like. The left-hand plot includes eight different ADI injection distances ranging from $d = 5$ Mpc to 40 Mpc; redder data markers correspond to smaller distances. We consider injections from 50 random directions at each distance, each of which is associated with 20 time segments, giving a total of $20 \times 50 \times 8 = 8000$ data markers. The right-hand plot includes 8000 data markers for S5 LIGO TS data (red ×'s) and 8000 data markers for MC Gaussian noise (green ○'s). Our cut is chosen to exclude the 'glitch tail' of the
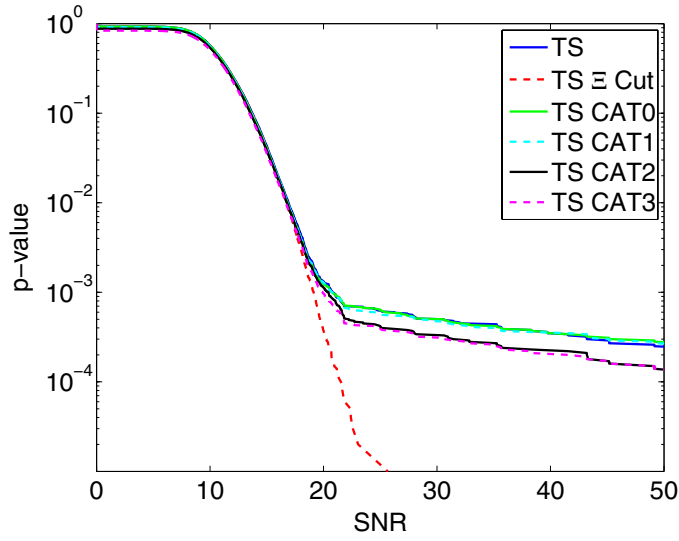
**Figure 9.** A plot of *p*-value versus SNR for 1 s $\times$ 1 Hz resolution TS data (TS) with no flags applied (solid blue), with the $SNR_\Xi$-based flag applied (dashed red), and with the data quality flags from [30, 37] applied in succession (CAT0 representing no flags applied, CAT1 representing the application of the category 1 flags, CAT2 representing categories 1 and 2 flags, etc). The data are parsed into 12 s $\times$ 150 Hz *ft*-maps.

red time shift distribution extending up and to the right while preserving most of the injected signals. Different signal models and different noise environments may require different cuts than the ones presented here.

## 7. Comparison with other data-quality flags

As noted above, numerous methods have been devised in order to determine when the strain channel is contaminated or corrupted by environmental or subsystem noise (see, e.g., [30–37]). A natural question, therefore, is: to what extent does the glitch identification flag developed here provide information complementary to existing data-quality flags? During the S5 science run, LIGO data quality flags were classified in terms of numbered categories 1–4 [30, 37]. These four categories describe different levels of severity: category 1, which includes data that will not be analyzed as it is corrupted or contaminated by known and identified processes; category 2, where the data is analyzed but various vetoes [30–37] will be applied only in post-processing; category 3, which are advisory flags used for detection confidence; and category 4, which are advisory flags used to exert caution in case of a detection candidate. Comprehensive descriptions of the S5 data quality flags are fully described elsewhere [30, 37].

The numbering is meant to convey the usability of the data, with category 1 flags representing the most contaminated data. In figure 9, we plot *p*-value versus SNR for TS data with no flag applied (solid blue), with $SNR_\Xi$-based flag applied (dashed red), and with various data quality flag categories applied in succession (no flags, category 1 applied, then categories 1 and 2 applied, etc). We find that the $SNR_\Xi$-based flag removes a significant number of glitches that are not already identified by category-numbered flags. It is evident that the two types of flags are complementary—our glitch identification flag finds inconsistencies

in autopower between detectors while the category-numbered flags identify and characterize specific instrumental and environmental fluctuations.

## 8. Conclusions

There is strong motivation for searches for long unmodeled GW transients, but searches utilizing an excess cross-power statistic [1] must contend with glitches, which hamper sensitivity. We introduce an auto-power consistency algorithm for identifying glitch-like data segments in searches for long GW transients and we study its behavior in various regimes: well-behaved noise, glitchy noise and potentially detectable GW signals. We find that it is effective at identifying glitches with minimal losses in data and live-time, thereby improving sensitivity. Yet it is safe in the sense that it does not flag GW signals at a high rate. Finally, we note that the glitch identification algorithm presented here may be useful for searches for short-duration transients. This is an area of ongoing research.

## Acknowledgments

## Appendix. Additional formalism

We consider the general form of a metric perturbation from a point source in the transverse-traceless gauge $h_{ab}(t, \vec{x})$. It can be written in terms of GW field Fourier coefficients, $\tilde{h}_A(f)$:

$$h_{ab}(t, \vec{x}) = \sum_A \int_{-\infty}^{\infty} df \, e_{ab}^A(\hat{\Omega}) \, \tilde{h}_A(f) \, e^{2\pi i f(t + \hat{\Omega} \cdot \vec{x}/c)}. \tag{A.1}$$

Here $t$ is time, $\vec{x}$ is the position vector, $\{e_{ab}^A\}$ are the GW polarization tensors, $\hat{\Omega}$ is the direction to the source and $A$ runs over $+$ and $\times$ polarizations. The dependence of $h_{ab}(t, \vec{x})$ on $\hat{\Omega}$ is implicit. The GW strain power between times $t$ and $t + \delta t$ in some frequency band between $f$ and $f + \delta f$ is

$$H_{AA'}(t; f) = \frac{2}{\mathcal{N}} \langle \tilde{h}_A(t; f) \tilde{h}_{A'}(t; f) \rangle. \tag{A.2}$$

The factor of two comes from the fact that we consider the single-sided power spectrum and $\mathcal{N}$ is a normalization factor arising from the use of a discrete Fourier transform. The semicolon emphasizes that $t$ refers to the beginning a data segment of length $\delta t$ and not to the many sampling times associated with each segment. Following [1], we define $H(t; f)$ so as to be invariant under change of polarization basis:

$$H(t; f) = \text{Tr} \left[ H_{AA'}(t; f) \right]. \tag{A.3}$$

The metric perturbation in equation (A.1) induces a strain in detector $I$ given by

$$\tilde{h}_I(t; f) = \sum_A \tilde{h}_A(t; f, \hat{\Omega}) \, e^{2\pi i f(t + \hat{\Omega} \cdot \vec{x}_I/c)} F_I^A(t; \hat{\Omega}). \tag{A.4}$$

Here $F_I^A(t; \hat{\Omega})$ is the antenna factor for detector $I$ (see [63]) and $\vec{x}_I$ is its position vector. The measured strain in detector $I$ is given by the sum of $\tilde{h}_I(t; f)$ with a noise term $\tilde{n}_I(t; f)$:

$$\tilde{s}_I(t; f) = \tilde{h}_I(t; f) + \tilde{n}_I(t; f). \tag{A.5}$$

15

We assume that the noise in two interferometers is uncorrelated, which is easily achieved for spatially separated interferometers.

In [1] it was shown that one can construct an unbiased estimator for $H(t; f)$ using the cross-power $\hat{C}_{IJ}(f)$ created from two spatially separated interferometers $I$ and $J$. This estimator is given by

$$\hat{Y}(t; f) = \text{Re}[Q_{IJ}(t; f, \hat{\Omega}, \vec{\alpha}) \hat{C}_{IJ}(t; f)] = \frac{2}{\mathcal{N}} \text{Re}[Q_{IJ}(t; f, \hat{\Omega}, \vec{\alpha}) \tilde{s}_I^{\star}(t; f)\tilde{s}_J(t; f)]. \qquad (A.6)$$

Here $Q_{IJ}(t; f, \hat{\Omega}, \vec{\alpha})$ is a filter function which takes into account the phase delay from the spatial separation of the interferometers as well as the detection efficiency of interferometers $I$ and $J$. It also depends on $\vec{\alpha}$, which is a set of parameters that characterizes the expected form of $H_{AA'}(f)$ such as the polarization of the source. We can write the filter function as

$$Q_{IJ}(t; f, \hat{\Omega}, \vec{\alpha}) = \frac{1}{\epsilon_{IJ}(t; \hat{\Omega}, \vec{\alpha})} e^{2\pi i f \hat{\Omega} \cdot \Delta \vec{x}_{IJ}/c}. \qquad (A.7)$$

Here $\Delta \vec{x}_{IJ} \equiv \vec{x}_I - \vec{x}_J$ is the difference in position vectors for detectors $I$ and $J$. $\epsilon_{IJ}(t; \hat{\Omega}, \vec{\alpha})$ is the 'pair efficiency', which is defined in terms of the expectation value of interferometer cross- and auto-powers:

$$\langle \hat{C}_{IJ}(t; f) \rangle \equiv \epsilon_{IJ}(t; \hat{\Omega}, \vec{\alpha})H(t; f) e^{-2\pi i f \hat{\Omega} \cdot \Delta \vec{x}_{IJ}/c} \qquad (A.8)$$

$$\langle \hat{P}_I(t; f) \rangle \equiv \epsilon_{II}(t; \hat{\Omega}, \vec{\alpha})H(t; f) + N_I(t; f), \qquad (A.9)$$

where $N_I(t; f) \equiv (2/\mathcal{N})|\tilde{n}_I(t; f)|^2$ and $H(t; f)$ is defined in equation (A.3). The pair efficiency for an unpolarized source is

$$\epsilon_{IJ}(t; \hat{\Omega}, \text{unpolarized}) = \frac{1}{2} \sum_A F_I^A(t; \hat{\Omega})F_J^A(t; \hat{\Omega}). \qquad (A.10)$$

Hereafter we abbreviate $\epsilon_{IJ}(t; \hat{\Omega}, \vec{\alpha})$ as simply $\epsilon_{IJ}$.

Through our definition of $Q_{IJ}(t; f, \hat{\Omega}, \vec{\alpha})$, we implicitly assume that the direction of the source $\hat{\Omega}$ is known. In order to estimate how well we must know $\hat{\Omega}$, we consider how large the error in $\hat{\Omega}$ (denoted $\delta\theta$) must be before we lose too much signal power. If we demand that we measure a fraction of at least $R$ of the total possible power, then we can tolerate angular errors $\delta\theta$ of

$$\delta\theta \lesssim \cos^{-1}(R) \left( \frac{c}{2\pi f |\Delta \vec{x}_{IJ}|} \right). \qquad (A.11)$$

For the Hanford–Livingston pair, this implies that we can tolerate angular errors of $\delta\theta \lesssim 0.8°$ up to 500 Hz with $R = 90\%$. For comparison, we note that the Swift experiment has an angular resolution of $\approx 0.25°$ [64]. For the remainder of the paper, we consider a single search direction. For triggers with large error regions on the sky, one can iterate over a grid of points inside a search cone, but this is a trivial extra step.

Since by assumption the noise in detectors $I$ and $J$ is uncorrelated, it follows that

$$\langle \hat{Y}(t; f) \rangle = H(t; f). \qquad (A.12)$$

An estimator for the variance of $\hat{Y}(t; f)$ is given by [1]:

$$\hat{\sigma}(t; f)^2 = \frac{1}{2}|Q_{IJ}(t; f, \hat{\Omega}, \vec{\alpha})|^2 \hat{P}_I'(t; f)\hat{P}_J'(t; f). \qquad (A.13)$$

Here $\hat{P}_I'(t; f)$ and $\hat{P}_J'(t; f)$ are the auto-powers measured in detectors $I$ and $J$, respectively. The prime denotes that they are calculated using $2n$ neighboring segments in order to obtain an estimate of the noise associated with the segment beginning at $t$:

$$\hat{P}_I'(t_0; f) \equiv \frac{1}{2n} \left[ \sum_{t=t_0-n\delta t}^{t=t_0+n\delta t} \hat{P}_I(t; f) \right] - \frac{1}{2n} \hat{P}_I(t_0; f). \qquad (A.14)$$

16

# References

[1] Thrane E *et al* 2011 *Phys. Rev.* D **83** 083004
[2] Ott C D 2009 *Class. Quantum Grav.* **26** 063001
[3] Dessart L, Burrows A, Livne E and Ott C D 2006 *Astrophys. J.* **645** 534
[4] Müller E, Rampp M, Buras R, Janka H T and Shoemaker D H 2004 *Astrophys. J.* **603** 221
[5] Keil W, Janka H T and Müller E 1996 *Astrophys. J. Lett.* **473** 111
[6] Miralles J A, Pons J A and Urpin V A 2000 *Astrophys. J.* **543** 1001
[7] Miralles J A, Pons J A and Urpin V 2004 *Astron. Astrophys.* **420** 245–9
[8] Corsi A and Mészáros P 2009 *Astrophys. J.* **702** 1171
[9] Piro A L and Ott C D 2011 *Astrophys. J.* **736** 108
[10] Ott C D, Dimmelmeier H, Marek A, Janka H T, Hawke I, Zink B and Schnetter E 2007 *Phys. Rev. Lett.* **98** 261101
[11] Ou S, Tohline J E and Lindblom L 2004 *Astrophys. J.* **617** 490
[12] Scheidegger S, Käppeli R, Whitehouse S C, Fischer T and Liebendörfer M 2010 *Astron. Astrophys.* **514** A51
[13] Lai D and Shapiro S L 1995 *Astrophys. J.* **442** 259
[14] Piro A L and Pfahl E 2007 *Astrophys. J.* **658** 1173
[15] van Putten M 2002 *Astrophys. J. Lett.* **575** 71–4
[16] van Putten M 2001 *Phys. Rev. Lett.* **87** 091101
[17] van Putten M 2008 *Astrophys. J. Lett.* **684** 91
[18] Andersson N, Comer G L and Langlois D 2002 *Phys. Rev.* D **66** 104002
[19] Glampedakis K, Samuelsson L and Andersson N 2006 *Mon. Not. R. Astron. Soc. Lett.* **371** 74
[20] Samuelsson L and Andersson N 2007 *Mon. Not. R. Astron. Soc.* **374** 256
[21] Levin Y 2006 *Mon. Not. R. Astron. Soc. Lett.* **368** 35
[22] Sotani H, Kokkotas K D and Stergioulas N 2008 *Mon. Not. R. Astron. Soc. Lett.* **385** 5
[23] Horvath J E 2005 *Mod. Phys. Lett.* A **20** 2799
[24] de Freitas Pacheco J A 1998 *Astron. Astrophys.* **336** 397
[25] Ioka K 2001 *Mon. Not. R. Astron. Soc.* **327** 639
[26] Vaishnav B, Hinder I, Shoemaker D and Herrmann F 2009 *Class. Quantum Grav.* **26** 204008
[27] Levin J and Contreras H 2011 *Class. Quantum Grav.* **28** 175001
[28] O'Leary R M, Kocsis B and Loeb A 2009 *Mon. Not. R. Astron. Soc.* **395** 2127
[29] Kocsis B and Levin J 2011 in preparation (arXiv:1109.4170)
[30] Blackburn L *et al* 2008 *Class. Quantum Grav.* **25** 184004
[31] Smith J R *et al* 2011 *Class. Quantum Grav.* **28** 235005
[32] Ajith P *et al* 2006 *Class. Quantum Grav.* **23** 5825
[33] Coughlin M (for the LIGO Scientific Collaboration and the Virgo Collaboration) 2011 arXiv:1108.1521
[34] Isogai T (for the LIGO Scientific Collaboration and the Virgo Collaboration) 2010 *J. Phys.: Conf. Ser.* **243** 012005
[35] Ballinger T (for the LIGO Scientific Collaboration and the Virgo Collaboration) 2009 *Class. Quantum Grav.* **26** 204003
[36] Christensen N (for the LIGO Scientific Collaboration and the Virgo Collaboration) 2010 *Class. Quantum Grav.* **27** 194010
[37] Slutsky J *et al* 2010 *Class. Quantum Grav.* **27** 165023
[38] Cannon K C 2008 *Class. Quantum Grav.* **24** 105024
[39] Hild S *et al* 2007 *Class. Quantum Grav.* **24** 3783
[40] Abbott B *et al* (LIGO Scientific Collaboration) 2009 *Rep. Prog. Phys.* **72** 076901
[41] Abbott B *et al* (LIGO Scientific Collaboration and the Virgo Collaboration) 2009 *Nature* **460** 990
[42] Abbott B *et al* (LIGO Scientific Collaboration) 2008 *Astrophys. J. Lett.* **683** 45
[43] Acernese F (for the Virgo Collaboration) 2006 *Class. Quantum Grav.* **23** S63
[44] Acernese F *et al* 2008 *Class. Quantum Grav.* **25** 184001
[45] Accadia T *et al* 2011 *Class. Quantum Grav.* **28** 114002
[46] https://wwwcascina.virgo.infn.it/advirgo/
[47] Kuroda K (for the LCGT Collaboration) 2010 *Class. Quantum Grav.* **27** 084004
[48] http://gw.icrr.u-tokyo.ac.jp/lcgt/
[49] The LIGO Scientific Collaboration 2011 *Nature Phys.* **7** 962–5
[50] Willke B *et al* 2006 *Class. Quantum Grav.* **23** S207–14
[51] Grote H (for the LIGO Scientific Collaboration) 2008 *Class. Quantum Grav.* **25** 114043

[52] Grote H (for the LIGO Scientific Collaboration) 2010 *Class. Quantum Grav.* **27** 084003
[53] Abbott B *et al* (LIGO Scientific Collaboration) 2007 *Phys. Rev.* D **76** 082003
[54] Ballmer S 2006 LIGO interferometer operating at design sensitivity with application to gravitational radiometry *PhD Thesis* Massachusetts Institute of Technology
[55] Abadie J *et al* (LIGO Scientific Collaboration and the Virgo Collaboration) 2011 *Phys. Rev. Lett.* **107** 271102
[56] Khan R and Chatterji S 2009 *Class. Quantum Grav.* **26** 155009
[57] Abbott B *et al* (LIGO Scientific Collaboration) 2009 *Astrophys. J. Lett.* **701** 68–74
[58] Abbott B *et al* (LIGO Scientific Collaboration) 2008 *Phys. Rev. Lett.* **101** 211102
[59] Corsi A and Mészáros P 2009 *Class. Quantum Grav.* **26** 204016
[60] Stella L *et al* 2005 *Astrophys. J. Lett.* **634** L165
[61] Santamaría L and Ott C D 2011 *LIGO DCC* T1100093 https://dcc.ligo.org/cgi-bin/DocDB/ShowDocument?docid=38606
[62] Kobayashi S and Mészáros P 2003 *Astrophys. J. Lett.* **585** L89
[63] Allen B and Romano J D 1999 *Phys. Rev.* D **59** 102001
[64] Gehrels N *et al* 2004 *Astrophys. J.* **611** 1005