First search for long-duration transient gravitational waves after glitches in the Vela and Crab pulsars

DOI: 10.1103/PhysRevD.100.064058

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Physical Review D (Particles, Fields, Gravitation and Cosmology)

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
First search for long-duration transient gravitational waves after glitches in the Vela and Crab pulsars

David Keitel,1,2 David Woan,2 Matthew Pitkin,2 Courtney Schumacher,3 Brynley Pearlstone,2 Keith Riles,4 Andrew G. Lyne,5 Jim Palfreyman,6 Benjamin Stappers,5 and Patrick Weltevrede5

1 University of Portsmouth, Institute of Cosmology and Gravitation, Portsmouth PO1 3FX, United Kingdom
2 University of Glasgow, School of Physics and Astronomy, Kelvin Building, Glasgow G12 8QX, Scotland, United Kingdom
3 Southern Methodist University, Dallas TX 75205, United States of America
4 University of Michigan, Ann Arbor, MI 48109, USA
5 Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK
6 Department of Physical Sciences, University of Tasmania, Private Bag 37, Hobart, Tasmania 7001, Australia

(Dated: September 9, 2019, revision 6a3c73ab – LIGO-P1900183-v5)

Gravitational waves (GWs) can offer a novel window into the structure and dynamics of neutron stars. Here we present the first search for long-duration quasi-monochromatic GW transients triggered by pulsar glitches. We focus on two glitches observed in radio timing of the Vela pulsar (PSR J0835–4510) on 12 December 2016 and the Crab pulsar (PSR J0534+2200) on 27 March 2017, during the Advanced LIGO second observing run (O2). We assume the GW frequency lies within a narrow band around twice the spin frequency as known from radio observations. Using the fully-coherent transient-enabled F-statistic method, we search for transients of up to four months in length. We find no credible GW candidates for either target, and through simulated signal injections we set 90% upper limits on (constant) GW strain as a function of transient duration. For the larger Vela glitch, we come close to beating an indirect upper limit for when the total energy liberated in the glitch would be emitted as GWs, thus demonstrating that similar post-glitch searches at improved detector sensitivity can soon yield physical constraints on glitch models.

I. INTRODUCTION

Neutron stars (NSs) provide a rich astrophysical laboratory for nuclear physics at extreme densities. Gravitational waves (GWs) can contribute to probing NS structure and dynamics through observations of binary mergers like GW170817 [1], but also by searching for signals from individual, rapidly spinning objects [2]. One of the most prominent dynamic features of individual NSs is the presence of timing glitches (sudden spin-up events) in pulsars [3]. Glitches are considered promising probes of the NS interior [4, 5] and possible emitters of detectable GW signals [7–11].

Despite 50 years of glitch observations and a lot of productive model development [12], the mechanism (or mechanisms) behind glitches are not well understood. This is where detecting transient GW signals during or following a glitch could yield valuable insights, as they would directly measure changes in the quadrupole moment of the NS.

So far, the only dedicated analysis of LIGO-Virgo [13] data that touched on this topic was a search for short (∼ O(s)) signals from a 2006 glitch of the Vela pulsar in initial LIGO data [14]. Targeted searches for continuous wave (CW) signals from known pulsars (most recently [16–19]) have also taken observed glitches into account as breaks in the pulsar ephemerides, but these always focus on persistent signals lasting for the full observation times before/after the glitch. Intermediate-duration transient searches have also been performed on magnetar bursts (most recently in [20] and to look for a post-merger remnant of GW170817 [21, 22], but for those targets the emission mechanisms and parameter space are very different than for pulsar glitches.

Here we perform a dedicated transient analysis of GW data from the Advanced LIGO (aLIGO) second observing run (O2, December 2016 to August 2017), searching for transient GW signals lasting hours to months after two glitches observed from the Vela pulsar (PSR J0835–4510) on 12 December 2016 (MJD 57839 ± 0.000029 [23–25]) and the Crab pulsar (PSR J0534+2200) on 27 March 2017 (MJD 57839.92 ± 0.06) [26, 27]. Due to their relatively young ages and close distances, these were the first two pulsars for which the indirect spin-down upper limit on CW emission (see e.g. Sec. 2.3 of [28]) was beaten [29, 30]. They are not the only known pulsars to have glitched during O2, and targeting a larger sample will be interesting in the future. Sec. 2.2 of [18] gives an overview of known pulsars accessible to LIGO-Virgo searches, though not all are known to glitch; and many known glitching pulsars listed in the standard public catalogues [27, 31] are unfortunately not within the aLIGO sensitivity band. The search presented in this paper is intended as a pilot instance of this new type of transient analysis, and the Crab and Vela pulsars were natural choices as the highest-priority targets.

To this end, we for the first time apply the tran-
sient \( F \)-statistic method introduced by Prix, Giampa- 

is & Messenger \[9\] to actual LIGO data. It focuses on 

long-duration (hours to months) transient GW 

signals that are quasi-monochromatic, i.e. narrowly 

localized in frequency (near twice the pulsar’s rota-

tion frequency) at each point during the observation 

time and slowly evolving in both frequency and am-

plitude. The data is split into Short Fourier Transforms 

(SFTs) and then, for a bank of templates from a sim-

ple frequency-evolution and transient-window model, a 

maximum-likelihood matched-filter statistic is computed. 

The method does not assume a particular glitch 

model, but only that the GW emission resembles 

these simple phenomenological templates. 

We present the data set used for this study in Sec. \[II\] 

general aspects of the signal model and \( F \)-statistic 

method in Sec. \[III\] and the specific search setup used in 

Sec. \[IV\]. Sec. \[V\] summarizes the search results and pro-

vides upper limits (ULs) on GW strain under the 

transient parameters \( \mathcal{T} \) consist of the window shape, sig-

nal start time \( t_0 \) and a duration parameter \( \tau \). The 

CW part depends on a set of phase evolution param-

ters \( \lambda = \{ \alpha, \delta, f, \dot{f}, \ldots \} \) (sky position, frequency, and 

frequency derivatives or ‘spindowns’) and on four am-

plitude parameters \( \mathcal{A} = \{ h_0, \cos \iota, \psi, \phi_0 \} \), which our de-

tection statistic (the \( F \)-statistic first introduced in \[33\]) 

analytically maximizes over. \( h_0 \): dimensionless am-

plitude, \( \iota \) and \( \psi \): orientation and polarization of the source, 

\( \phi_0 \): GW phase at reference time \( T_{\text{ref}} \). The signal am-

plitude for a deformed NS at distance \( d \), emitting GWs 

at \( f = 2f_{\text{spin}} \) in the dominant \( l = m = 2 \) mode from a 

quadrupolar ellipticity \( \epsilon \), with principal moment of iner-

cia \( I \), is given by 

\[
h(t, \lambda, \mathcal{A}, \mathcal{T}) = \varpi(t, t_0, \tau) h(t, \lambda, \mathcal{A}) \tag{1} \]

\[
h_0 = \frac{4\pi^2 G \epsilon I}{c^4 d^2 f^2}. \tag{2}
\]

Hence, the simplest interpretation of our transient signal 

model would be a temporary increase in the quadrupo-

lar deformation after the glitch, with the product \( \epsilon I \) 

falling off again as determined by the transient window 

function with timescale \( \tau \). This would be an intu-

itive behaviour within the ‘starquake’ model of pulsar 

glitches \[44,46\], where the crust is vio-

lently deformed. In the more popular class of 

glitch models based on a rotation lag between the 

bulk of the NS and an interior superfluid com-

ponent, the interpretation becomes more compli-

cated, and could involve processes such as crustal 

heating \[47\], non-axisymmetric oscillations \[8\] or 

post-glitch excitation of Ekman flows \[7,8,11\]. 

These options are reviewed in more detail in 

\[9,12\]. For our search, the mechanism does not 

initially matter as long as a bank of templates 

from the simple signal model provides a sufficient 

fit to the real signal. The connection to physical 

models can be made later, based on the signal 

durations and amplitudes observed, or from the 

obtained upper limits (ULs). 

For each target pulsar, \( \alpha \) and \( \delta \) are fixed. We 

place a grid in \((f,\dot{f})\) space with fixed spacings.
The search method, implemented in the program lalapps_ComputeFstatistic, then loops over this grid, and for each $\lambda = (f, \dot{f})$ pair it computes the transient $F$-statistic map

$$F_{mn}(\lambda) = F(\lambda, t_{0n}, \tau_n)$$

by computing partial sums, corresponding to all \{(t_{0n}, \tau_n)\} combinations, of the per-SFT 'atomic' ingredients of the $F$-statistic. This corresponds to demodulating the detector data over the whole observation time once for each $\lambda$ template, taking into account the time-varying detector response, and then simple arithmetic operations for the transient aspect.

In this search, we use a simple 'rectangular window', i.e. $h_0 = \text{const.}$ for $t \in [t_0, t_0 + \tau]$ and $h_0 = 0$ outside. The maximum loss compared to more realistic exponential-decay windows has already been estimated as acceptable in [9]. A search with exponential windows would also be computationally much more costly; see [50] for a detailed discussion. Also see Sec. IV and appendix C for more details on this point.

To obtain detection candidates, we are interested in peaks over the full $(\lambda, T) = (f, \dot{f}, t_0, \tau)$ space. To reduce the number of strong $F$-statistic outliers due to single-detector noise disturbances, we also apply the line-robust statistic $B_{\text{SGL}}$ from [51] at every point, and for each $\lambda$ we only store the values of $\max F_{mn}(\lambda)$ over the subset where $B_{\text{SGL}}(\lambda, t_0, \tau)$ is above some threshold. In the end, candidates are identified from the $F$-ordered results, as the $F$-statistic is directly related to the signal-to-noise ratio (SNR) and its analytically known distribution will allow us to set a detection threshold without additional pure-noise Monte Carlos. $B_{\text{SGL}}$ is thus used here simply as an intermediate veto step, not as a full replacement detection statistic.

**IV. SEARCH SETUP AND PARAMETER SPACE COVERED**

The key parameters of the two target pulsars and their glitches during O2 are summarized in Table I. For each glitch, we search for transients starting in a $\Delta t_0 = 1$ day window centred around the nominal glitch epoch and with durations $\tau$ up to 121 days. Hence we search a data set of $T_{\text{obs}} \approx 4$ months, chosen for two reasons:

(i) With the whole of O2 lasting about 9 months, for much longer $\tau$ we would no longer get sufficient benefits from the transient $F$-statistic over the results of full-O2 CW analyses [18][19]. From Eq. (62) in [9], the mismatch (relative loss in squared SNR) for observing a signal of true length $\tau_s$ with a rectangular template window of length $\tau$ is $m \approx 1 - (\tau - \tau_s)^2/(\tau \tau_s)$. So for a maximum $\tau_s = 4$ months and a full-O2 CW search’s $\tau = 9$ months, we expect $m \approx 30\%$ corresponding to still about a factor of 2 gain in SNR for a transient with our maximum $\tau$ compared to the full-O2 CW search. For longer $\tau_s$, the gain would be correspondingly smaller.

(ii) $T_{\text{obs}} \approx 4$ months also matches the duration of the first aLIGO observing run (O1), so that, at the longest $\tau$, our search becomes comparable in setup to the O1 narrow-band search [17] and we can draw some direct comparisons in the following.

Similar to [17][19], we allow for some mismatch between the true GW frequency $f$ and its nominal value (twice the radio-observed $f_{\text{spin}}$), constructing a rectangular search grid in $(f, \dot{f})$ space. We choose resolutions in GW frequency and spin-down of $df = 1/\Delta f_{\text{obs}} \approx 9.57 \times 10^{-8}$ Hz and $d\dot{f} = 1/\Delta \dot{f}_{\text{obs}} \approx 9.15 \times 10^{-15}$ Hz s$^{-1}$, and cover a frequency range of 0.1 Hz and a spindown range of $11df \approx 0.01 \times 10^{-13}$ Hz s$^{-1}$ both centred on a point $(f, \dot{f})$ corresponding to twice the values from the pulsar’s radio ephemerides.

The ephemerides were obtained from observations at the University of Tasmania’s Mount Pleasant Radio Observatory for Vela and at Jodrell Bank (UK) for the Crab, using the TEMPO2 software [55]. Both were originally fitted with the goal of minimal residuals over the whole respective data ranges of the CW searches in [18]. Vela, in addition to glitch recovery, has a lot of timing noise [56, 57] and micro-glitches [58]; hence, to minimize overall residuals, the post-glitch ephemeris for O2 was fitted down to the twelfth derivative (without an exponential recovery term). Thus, and due to the intrinsic rapid evolution of the frequency and its derivatives after a glitch, the $f$ in Table I should not be compared directly with the long-time value reported by [59]. Similarly, the Crab pulsar also has significant timing noise and glitch recovery complicating its spin-down [60]. For its O2 ephemeris, no explicit glitch model was used, but again twelve spin-down derivatives were included to minimize overall residuals.

For both targets, we use the $(f, \dot{f})$ grid as described above, a single fixed value for $f$ and, set higher-order terms to zero, which is valid in the sense that we are accumulating less than a bin of mismatch in $f$ over our $T_{\text{obs}}$ from the second derivative and higher. In addition, while both ephemerides have reference times far from the glitches, the ranges covered in $f$ and $\dot{f}$ mean that either we resolve the glitch step with multiple templates (Vela), or the glitch step is itself smaller than our search resolution (Crab), so that extrapolation around the glitch epoch is safe for the purpose of this search.

The overall frequency evolution template count is $N_f = N_f N_{\dot{f}} \approx 1.5 \times 10^7$. Comparing with [17], these choices mean we cover slightly wider ranges in $f$ and $\dot{f}$ for Vela while for the Crab we cover the same $f$ and a narrower $\dot{f}$ range by a factor 15.

For each $(f, \dot{f})$ parameter space point, transient parameters of $t_0 \in [T_{\text{glitch}} - 0.5 \text{ days}, T_{\text{glitch}} + 0.5 \text{ days}]$
TABLE I. Search targets and their key parameters. Much higher-precision sky location ($\alpha$, $\delta$) and frequency evolution ($f$, $\dot{f}$, $\ddot{f}$) from the radio ephemerides are used in the search, with a factor 2 to convert from rotation to GW emission rates, and ranges of 0.1 Hz in $f$ and $\approx 10^{-13}$ Hz s$^{-1}$ in $\dot{f}$ are explored around the nominal values. For the Vela pulsar, all parameters are referenced to MJD 58000, while for the Crab pulsar the position epoch is MJD $\approx$ 53254 and the frequency epoch is MJD $\approx$ 57185.127. The glitch times $T_{\text{glitch}}$ (in GPS seconds) and glitch sizes $\Delta f/\dot{f}$ are taken from [27].

<table>
<thead>
<tr>
<th>target</th>
<th>$d$ [pc]</th>
<th>$\alpha$ [rad]</th>
<th>$\delta$ [rad]</th>
<th>$f$ [Hz]</th>
<th>$\dot{f}$ [Hz s$^{-1}$]</th>
<th>$\ddot{f}$ [Hz s$^{-2}$]</th>
<th>$T_{\text{glitch}}$ [s]</th>
<th>$\Delta f/\dot{f}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vela (J0835–4510)</td>
<td>287 [52]</td>
<td>2.2486</td>
<td>−0.7885</td>
<td>22.3722</td>
<td>$-3.12 \cdot 10^{-11}$</td>
<td>1.16 $\cdot 10^{-19}$</td>
<td>1165577920</td>
<td>1.43 $\cdot 10^{-6}$</td>
</tr>
<tr>
<td>Crab (J0534+2200)</td>
<td>2000 [53, 54]</td>
<td>1.4597</td>
<td>0.3842</td>
<td>59.3295</td>
<td>$-7.39 \cdot 10^{-10}$</td>
<td>1.72 $\cdot 10^{-20}$</td>
<td>1174687506</td>
<td>2.14 $\cdot 10^{-9}$</td>
</tr>
</tbody>
</table>

and $\tau \in [0.5 \text{days}, T_{\text{obs}}]$ are analyzed with a resolution $dt_0 = d\tau = T_{\text{SFT}} = 1800$ s, yielding $N_\tau = N_t N_\tau = 48 \times 5784 \approx 2.8 \times 10^5$ grid points.

The lower limit of $\tau \geq 0.5$ days is chosen empirically to avoid spurious outliers when statistics fluctuate too much over a small number of SFTs. In principle, the transient-$F$-statistic method can be extended to much shorter $\tau$ but this would require additional copies of the data set with shorter $T_{\text{SFT}}$. This could be done in the future, or a more adaptive multi-timescale approach could be pursued; but for this pilot search we limit ourselves to the standard data set with $T_{\text{SFT}} = 1800$ s and correspondingly do not explore very short transients.

The line-robust $B_{\text{SGL}}$ statistic we use as a veto has a free tuning parameter $\mathcal{F}^{(0)}_s$ determining how strongly it is allowed to deviate from the standard $F$-statistic [51]. Since in this search we are dealing with narrow frequency bands which are already known to contain some disturbances, we choose a relatively low value $\mathcal{F}^{(0)}_s = 10$ corresponding to a statistic that would be slightly suboptimal in purely Gaussian noise but is stricter in suppressing lines. We then set a rather lenient threshold of $B_{\text{SGL}} > -10$ to only cut out very strong single-detector artifacts that might have passed our line cleaning procedure (see appendix A) without severely affecting the distribution of $F$-statistic values.

V. RESULTS

A. Search results: no significant candidates

Full results for the max $2F(f, \dot{f})$ statistic (maximized over $t_0$ and $\tau$), projected onto the frequency axis, are shown in Fig. 1 for both the 20161212 Vela glitch and the 20170327 Crab glitch.

To determine whether the loudest of these per-template results constitute promising detection candidates, we can consider the well-known [e.g. 51] statistical properties of the $F$-statistic: as $2F$ in pure Gaussian noise follows a $\chi^2$ distribution (with 4 degrees of freedom), the loudest value $2F^*$ from $N$ independent trials is distributed as

$$p(2F^*; N) = N \chi^2_4(2F^*) \chi^2_4(2F < 2F^*)^{N-1},$$ (4)

where $\chi^2_4(2F^*)$ is the probability distribution function evaluated at $2F^*$ and $\chi^2_4(2F < 2F^*)$ is the cumulative distribution function integrated up to $2F^*$. For this search, we need to consider the loudest overall outlier $2F^* = \max_{f, \dot{f}}(\max_{t_0, \tau} 2F)$ for each target. Since the $(t_0, \tau)$ templates at each $(f, \dot{f})$ point re-use the same data atoms many times, the total effective number of templates is much lower than $N_t \cdot N_\tau$. We find acceptable fits to the max $2F$ distribution (over all $(f, \dot{f})$) for $N^\text{eff} \approx 50$, and then use a total $N^\text{eff} \approx 50 \cdot N_\lambda$ to obtain (numerically from Eq. 4) an expectation value $E[2F^*] \approx 48$ with standard deviation $\sigma[2F^*] \approx 2.7$.

The loudest candidate from the Vela search has $2F^* \approx 47.19$, and from the Crab search the loudest outlier is at $2F^* \approx 47.18$. These do not even cross a nominal $E[2F^*] \approx 48$ threshold, which might be considered a low-level criterion for further follow-up study. And even though the precise threshold to put could be shifted somewhat when revisiting the assumptions in effective template counting, any truly promising candidate would need to lie a few standard deviations above expectation. For comparison, levels of $E$, $E + \sigma$ and $E + 2\sigma$ are indicated by horizontal lines in Fig. 1.

Hence, while there is some evident substructure in the search results that could be further investigated by follow-ups with different $T_{\text{SFT}}$, gridless MCMC [62] or through more detailed data quality studies, we conclude that the loudest candidates from both targets are so weak that no such effort is warranted at this point. We proceed next to set upper limits (ULs) on GW strain from our targets based on the absence of a detection.

B. Upper Limits

To obtain ULs on the emitted strain from/after the two targeted glitches, still under the simplified assumption of a ‘rectangular’ transient, we perform software injections of simulated signals into the same data sets as used for the original searches, and determine the required scale of $h_0$ at which 90% of signals are recovered above the nominal $E[2F^*] \approx 48$ threshold. Physically, for a fixed overall energy budget of the glitch we would expect lower $h_0$ for longer $\tau$. At the same time, our search sensitivity also improves roughly with $\sqrt{\tau}$. Hence, we present ULs on $h_0$ as a function of $\tau$, randomizing over all other parameters. Details on the procedure are given in appendix B.
FIG. 1. Search results for transients following the 20161212 Vela glitch (left panel) and 20170327 Crab glitch (right panel) with the detection statistic $\max 2F(f, \dot{f})$ maximized over transient parameters $\{t_0, \tau\}$ and projected onto the $f$ axis. Horizontal lines correspond to expected maximum outliers (and also $+1$ and $+2$ standard deviations) for an estimated effective template count (per target) of $5.75 \cdot 10^5$.

FIG. 2. ULs on GW strain for transients following the 20161212 Vela glitch (left panel) and 20170327 Crab glitch (right panel) as a function of the duration $\tau$ of rectangular-windowed signals. The black points and dotted lines represent our ULs at 90% confidence from simulated signal injections randomized over all other parameters. Error bars are estimated from a sigmoid fit to $p_{\text{det}}(h_0)$ curves, with a small contribution from calibration uncertainty (see appendix B). For comparison, we also show the glitch excess energy indirect ULs from Eq. 5 (solid red lines), a semi-analytic sensitivity estimate for the search (dashed blue lines) and the ULs for persistent CWs previously obtained in [17] (magenta hexagons, full O1 data) and [19] (cyan diamonds, O2 data starting 20170104 for Vela and after 20170327 for the Crab).

and the results are shown in Fig. 2.

In the same figures, we also compare the measured ULs against a sensitivity estimate for our search based on the octapps [63] CW sensitivity calculator (see [64] for a detailed description). To obtain this curve, we consider CW-like searches over a grid in $T_{\text{obs}} = \tau$ and feed the estimation code with cumulative duty factors and harmonic-mean averages of the detector power spectral density (PSD) over each of these durations, as well as an estimated average mismatch of 0.2 for our $(f, \dot{f})$ grid and a $2F$ threshold of 48 as above. The agreement is remarkably close.

We also find that at the longest signal duration probed ($\tau = 120$ d) our ULs approach those from the LIGO-Virgo narrow-band searches for CWs from the same targets (see Table VI of [17] and Table IV of [19]).

VI. DISCUSSION

In the absence of significant GW detection candidates after the two glitches in the Vela and Crab pulsars during the aLIGO O2 run, we have set ULs on the emitted GW strain as a function of signal duration $\tau$. To check how physically constraining these ULs are, we consider an indirect transient energy UL in analogy to the well-known spin-down limit from the CW case (see e.g. Sec. 2.3 of [28]). Here we briefly summarize the derivation of this indirect UL from Sec. II.C of [9].

According to the basic two-fluid model of NS glitches (see e.g. [65]), in an intra-glitch period the bulk of the NS gradually spins down, while an interior superfluid component retains most of its initial angular momentum. Then assuming
(i) a glitch transfers the whole angular momentum difference previously built up between the two components, (ii) the moments of inertia \( I_c \) (superfluid) and \( I_b \) (bulk) do not change during the glitch, and (iii) \( I_c \ll I_b \), then the excess superfluid energy liberated in this transfer is 
\[
E_{\text{glitch}} \approx 4\pi^2 I_b f \Delta \nu_{\text{s-c}},
\]
where \( \Delta \nu_{\text{s-c}} \) is the lag (built-up difference in \( f \)) between the two components right before the glitch (\( \Delta \nu \) in the notation of [28]).

If we further assume that (iv) all of \( E_{\text{glitch}} \) is emitted in GWs, and rewrite in terms of the observed relative frequency change \( \Delta f/f \) at the glitch and total moment of inertia \( I \approx I_c \), then the total emitted GW energy is independent of the signal duration \( \tau \), while the corresponding GW strain as a function of \( \tau \) is
\[
h_0 = \frac{1}{d} \sqrt{\frac{5G I \Delta f}{2c^5 \tau f^3}}.
\] (5)

As per [28], qualitatively similar ULs still hold for alternative glitch models such as crust-cracking starquakes [46], where instead the moment of inertia of the crust would change at the glitch.

This indirect UL is shown for comparison with our empirical ULs in Fig. 2. For this, we assume a fiducial value of \( I = 10^{38} \text{ kg m}^2 \) and distances of 287 pc for the Vela pulsar [62] and 2 kpc for the Crab pulsar [27].

With a frequency change of \( \Delta f/f = 1.431 \times 10^{-6} \), the Vela glitch was much larger than the Crab glitch with \( \Delta f/f = 2.14 \times 10^{-9} \) (both values from [27]). [66] recently suggested a larger initial frequency overshoot in the Vela glitch, which however has already decayed at the timescales probed in this search). Hence, though aLIGO sensitivity is better at the higher of the two target frequencies, for the Crab glitch our O2 search was still far away from the indirect energy UL; while for Vela we got very close to beating it at the shortest and longest \( \tau \). Overall, the search sensitivity was mostly limited by the significant time variation of detector sensitivities and duty factors. For example, the slow improvement in Vela ULs at intermediate durations (10–30 days) is largely due to the winter holiday break in O2 observing which made cumulative duty factors drop over this period.

Since the search presented in this paper is the first practical application of the transient-\( F \)-statistic method from [9] to GW detector data, we have made several choices to simplify the search setup and post-processing, but which can be improved over in future applications. Notably, the simple rectangular transient window allows for efficient computation, but while it does also recover most of the SNR for different signal shapes (see [9] and appendix C), a more general and sensitive analysis will be possible when including e.g. exponentially decaying window functions. (A natural choice considering the observed exponential recoveries in pulsar frequency after most glitches, see e.g. [67]). Since exponential templates are much more costly, using graphics processing units (GPUs) for the search would be highly beneficial [60].

The search setup can be improved in several other ways, including multi-timescale approaches: i.e. the use of input data at several different SFT durations to cover shorter transients; and varying the metric-based [68] frequency and spin-down resolutions as a function of \( \tau \).

To deal with outliers from either detector noise features or actual GW signals, MCMC-based follow-up [62] is a promising technique. If necessary, the detector noise can also be studied in much more detail than was required for this pilot search, e.g. using correlations with auxiliary channels to veto instrumental lines [42]. We also used only an ad-hoc veto version of the line-robust statistic from [51], while a semi-coherent transient-aware version [69] or a customized coherent version would offer the potential for more robust suppression of single-detector instrumental artifacts.

In addition to these improvements in the analysis method, the improved sensitivity and increased number of detectors in O3 (which has started in April 2019) and beyond [70] will be the strongest driver in bringing search results for future pulsar glitches into the physically constraining regime. Since the search method is computationally cheap (already on regular CPUs for rectangular windows, and when using GPUs this stays true also for exponential windows), with some more automation of the analysis pipeline it should be possible to target not only the highest-value objects such as the Vela and the Crab, but more of the large population of glitching pulsars [3].

ACKNOWLEDGMENTS

We thank members of the LIGO-Virgo Continuous Wave working group for many fruitful discussions; Reinhard Prix and Chris Messenger for initial advice on the transient-\( F \)-statistic method; Pep Covas Vidal, Evan Goetz, Ansel Neunzert and the rest of the LIGO detector characterization group for their invaluable work on data quality studies; Greg Ashton for detailed comments on the manuscript; and Paul Hopkins and Stuart Anderson for technical support. For part of this project, DK was funded under the EU Horizon2020 framework through the Marie Sklodowska-Curie grant agreement 704094 GRANITE. GW and MP are funded through the UK Science & Technology Facilities Council (STFC) grant ST/N005422/1. CS was supported through the Arcadia University Study Abroad programme. This research has made use of data obtained from the Gravitational Wave Open Science Center, a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes. The authors are grateful for computational resources provided by the LIGO Laboratory and Cardiff University and supported by National Sci-
TABLE II. Single-detector lines identified through outliers in the normalized SFT power and subsequently removed from the input SFTs for the $F$-statistic analyses. The more common cases where single bins in a single SFT exceed the threshold are not used for any cleaning. The width is given in terms of the number of bins $N_{\text{bins}}$ at resolution $1/T_{\text{SFT}} = 1/1800$ s. The duration is quoted as between the first and last SFT with $P^X(f_k) \geq 10$, not necessarily meaning that the line is persistently visible during this range, and not excluding that it is still present at a weaker level before and after. For simplicity and safety, the full set of SFTs is cleaned at these frequencies, not just the listed duration. Where the frequency matches up with a harmonic of a known comb of disturbances, the corresponding comb spacing is listed.

<table>
<thead>
<tr>
<th>detector</th>
<th>$f_{\text{start}}$</th>
<th>$N_{\text{bins}}$</th>
<th>duration</th>
<th>comb [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>22.2222</td>
<td>3</td>
<td>20161202–20170508</td>
<td>11.1111</td>
</tr>
<tr>
<td>H1</td>
<td>22.2500</td>
<td>1</td>
<td>20170204–20170418</td>
<td>0.9999862</td>
</tr>
<tr>
<td>L1</td>
<td>22.3600</td>
<td>3</td>
<td>20161214</td>
<td>unidentified</td>
</tr>
<tr>
<td>L1</td>
<td>22.4156</td>
<td>3</td>
<td>20161210–20170104</td>
<td>2.24154</td>
</tr>
<tr>
<td>L1</td>
<td>22.4706</td>
<td>1</td>
<td>20161213</td>
<td>unidentified</td>
</tr>
<tr>
<td>H1</td>
<td>59.2733</td>
<td>3</td>
<td>on/off through O2</td>
<td>0.9878881</td>
</tr>
</tbody>
</table>

Appendix A: Data cleaning

To identify single-detector noise artifacts in the form of narrow disturbances (spectral lines) within the analysis bands for each pulsar, we have considered a standard quantity for such data quality studies, the normalized SFT power [34]:

$$P^X(f_k) = \frac{2}{N_{\text{SFT}} T_{\text{SFT}}} \frac{|\tilde{x}_n^X(f_k)|^2}{S_n^{X^H}(f_k)}. \quad (A1)$$

Here, $\tilde{x}_n^X(f)$ is the data in bin $f_k$ of the $n$th SFT for detector $X$ and $S_n^{X^H}(f_k)$ is a running-median noise PSD estimate. This is computed for the bands in question using the lalapps_ComputePSD tool [11]. We then identified as (possibly transient) single-detector line features those frequency bins (or small number of adjacent bins) where at least 5 SFTs (not necessarily consecutive) crossed a threshold $P^X(f_k) > 10$. These are listed in Table I. For safety, we would not have vetoed any features common to both detectors – as could be produced by an (extremely loud) astrophysical signal – through this procedure. No such coincident features were found in the two bands.

Comparing with the detailed detector characterization approach of line hunting in [42], we find that the three easily identified H1 artifacts match up with harmonics of known instrumental frequency combs. On the other hand, the three L1 artifacts are all very short and limited to the first month of O2 when an improperly connected ethernet cable induced electronic crosstalk in the interferometer controls system, a period which has not been used in the LIGO-Virgo flagship CW searches [15, 19, 23]. These artifacts do not match any lines or combs listed in [42]. We have not investigated in any more detail whether these additional narrow and transient disturbances are clearly correlated with auxiliary channels, but since they are clearly limited to a single detector, they can still be considered safe for removal from the input data.

Using this list, we have cleaned the input SFTs of the affected detector by replacing the listed bins with samples drawn from a Gaussian distribution with variance matching the surrounding PSD estimate, using lalapps_SFTclean.

After this removal, no outliers with multi-detector max $2F \geq 48$ or single-detector max $2F \geq 52$ are found by the search. So while some of the remaining substructures visible in Fig. 1 are likely due to unidentified narrow instrumental disturbances, possibly including the December 2016 L1 and March–April 2017 H1 issues, none of these are strong enough to lead to significant outliers.

Appendix B: Details on UL procedure

To obtain the GW strain ULs as a function of signal length $\tau$ as presented in Sec. V B we perform software injections of simulated signals into the same input SFTs as used for the main search. At each sample value of $\tau$, we use the lalapps_MakeFakeData_v5 program [21] to simulate a set of signals with varying $h_0$, uniformly distributed over the whole search range in $\{f, f, t_0\}$ and also randomized over the remaining amplitude parameters $\{\cos \iota, \psi, \phi_0\}$. For each injection, a reduced parameter space covering $10^{-4}$ Hz in frequency (with the same $df$ as before) and only 1 spin-down bin is re-analyzed. We count an injection as recovered if it produces max $2F \geq 48$, as a candidate above this nominal threshold would have been considered for further follow-up if found in the main search.

For each $\tau$, the result is then an efficiency curve of detection probability $p_{\text{det}}$ against injected $h_0$. This can be fit with a sigmoid

$$p_{\text{det}}(h_0) = \frac{1}{1 + \exp[-a(h_0 - b)]} \quad (B1)$$

(see scipy.curve_fit) and evaluated at $p_{\text{det}} = 0.9$ to estimate a $h_{1000}$ UL at 90% confidence. The error on $h_{1000}$ is obtained from error propagation of the uncertainties in the fit coefficients $(a, b)$. We have run a relatively small set of UL simulations: 50 injections each at 10–20 $h_0$ steps per $\tau$ value, leading to ~10% uncertainty in $h_{1000}$ as evaluated from the fit. An additional contribution comes from calibration uncertainty in the measured strain at the detector [33, 36]. According to [37], for the 20–100 Hz band during O2 the amplitude uncertainties are 1.6% for H1 and 3.9% for L1.
signals turning on at some time \( t_0 \) and off again at \( t_0 + \tau \) with fixed amplitude, hence referred to as rectangular-windowed signals. The detection method as introduced by \[9\] however can also deal with more general transient window functions, and as an explicit example the implementations in LALSuite \[41\] and pyFstat \[50\] also include transients with exponentially decaying amplitude. Since pulsar rotation frequencies after glitches often show transients with exponentially decaying amplitude. Hence, one cannot directly compare against false-alarm probability \( p_{\text{fa}} \) which point the LALSuite implementation cuts off the signal draws. They found moderate losses with a worst case of \( \lesssim 10\% \) at relatively high \( p_{\text{fa}} \).

As a specific test with more direct relation to the analysis in this paper, let us revisit the UL injection-and-recovery procedure with exponential-window injections. Direct comparison of recovering these injections with both rectangular and exponential search windows is too expensive for the present purpose; instead, we have simply tested the recovery of exponential injections with rectangular search windows — though only for the 20161212 Vela glitch as an example target. Results are shown in Fig. 3 showing that there is indeed some loss in sensitivity from transient window mismatch, and that a dedicated exponential-window search (likely on GPUs \[50\]) can be valuable in the future. But for the moment, this test also demonstrates that our simplified pilot search also had sensitivity to exponentially decaying signals, with slightly higher ULs as per Fig. 3.

It is worth noting that, following the conventions of \[9\], the length of an exponential signal with \( h_0(t) \propto \exp[-(t - t_0)/\tau] \) is longer than that of a rectangular window with fixed length \( \tau \). the initial amplitude falls to 37\% after 1\( \tau \), 14\% after 2\( \tau \) and 5\% after 3\( \tau \), at which point the LALSuite implementation cuts off the window. Hence, one cannot directly compare \( p_{\text{det}} \) at a fixed \( \tau \) for injections of both window types and purely attribute the difference to window mismatch.

### Appendix C: Sensitivity compared with exponentially decaying transients

In this search we have only considered the simplest model for quasi-monochromatic transients, i.e. CW-like signals turning on at some time \( t_0 \) and off again at \( t_0 + \tau \) with fixed amplitude, hence referred to as rectangular-windowed signals. The detection method as introduced by \[9\] however can also deal with more general transient window functions, and as an explicit example the implementations in LALSuite \[41\] and pyFstat \[50\] also include transients with exponentially decaying amplitude. Since pulsar rotation frequencies after glitches often show an exponential recovery profile \[97\], exponential transient windows could be a more realistic option. Here we briefly investigate the effect of trying to recover exponentially-decaying signals with rectangular search windows.

This question was previously addressed in Sec. V.B of \[9\] through looking at detection probability \( p_{\text{det}} \) against false-alarm probability \( p_{\text{fa}} \) for synthetic noise and signal draws. They found moderate losses with a worst case of \( \lesssim 10\% \) at relatively high \( p_{\text{fa}} \).

FIG. 3. Injection recovery results for the Vela search, comparing transient injections with rectangular window function (black dotted line, same as in Fig. 2) and exponential window function (orange dashed line), both recovered with rectangular search windows.


[69] D. Keitel, Robust semicoherent searches for continuous gravitational waves with noise and signal models including hours to days long transients, Phys. Rev. D 93, 084024 (2016) arXiv:1509.02398 [gr-qc]

