Parkes Transient Events. I. Database of Single Pulses, Initial Results, and Missing Fast Radio Bursts

S.-B. Zhang\textsuperscript{1,2,3,4}, G. Hobbs\textsuperscript{3}, C. J. Russell\textsuperscript{5}, L. Toomey\textsuperscript{3}, S. Dai\textsuperscript{3} \textsuperscript{1}, J. Dempsey\textsuperscript{6}, R. N. Manchester\textsuperscript{3} \textsuperscript{1}, S. Johnston\textsuperscript{3}, L. Staveley-Smith\textsuperscript{\textsuperscript{1,7}}, X.-F. Wu\textsuperscript{1}, D. Li\textsuperscript{2,3,8,9}, Y.-Y. Yang\textsuperscript{9,10}, S.-Q. Wang\textsuperscript{2,3,11}, H. Qiu\textsuperscript{\textsuperscript{3,12}}, R. Luo\textsuperscript{3}, C. Wang\textsuperscript{13}, C. Zhang\textsuperscript{2,3,8,13}, L. Zhang\textsuperscript{2,3,8}, and R. Mandon\textsuperscript{3,14}

1 Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, People’s Republic of China; sbzhang@pmo.ac.cn
2 University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China
3 CSIRO Astronomy and Space Science, P.O. Box 76, Epping, NSW 1710, Australia
4 International Centre for Radio Astronomy Research, University of Western Australia, Crawley, WA 6009, Australia
5 CSIRO Scientific Computing Services, Australian Technology Park, Locked Bag 9013, Alexandria, NSW 1435, Australia
6 CSIRO Information Management & Technology, Box 225, Dickson, ACT 2602, Australia
7 ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
8 National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing 100101, People’s Republic of China
9 FAST Collaboration, CAS Key Laboratory of FAST, NAOC, Chinese Academy of Sciences, Beijing 100101, People’s Republic of China
10 School of Physics and Electronic Science, Guizhou Education University, Guiyang 550018, People’s Republic of China
11 Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, People’s Republic of China
12 Sydney Institute for Astronomy, School of Physics, University of Sydney, Sydney, NSW 2006, Australia
13 CSIRO Data61, Sydney, NSW 2015, Australia
14 Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Mail H30, P.O. Box 218, Hawthorn, VIC 3122, Australia

Received 2020 February 3; revised 2020 March 31; accepted 2020 April 1; published 2020 July 13

Abstract

A large number of observations from the Parkes 64 m diameter radio telescope, recorded with high time resolution, are publicly available. We have reprocessed all of the observations obtained during the first four years (from 1997–2001) of the Parkes Multibeam Receiver system in order to identify transient events, and have built a database that records the 568,736,756 pulse candidates generated during this search. We have discovered a new fast radio burst (FRB), FRB 010305, with a dispersion measure (DM) of 350 ± 5 cm\textsuperscript{-3} pc and explored why so few FRBs have been discovered in data prior to 2001. After accounting for the dispersion smearing across the channel bandwidth and the sky regions surveyed, the number of FRBs is found to be consistent with model predictions. We also present five single pulse candidates from unknown sources, but with Galactic DMs. We extract a diverse range of sources from the database, which can be used, for example, as a training set of data for new software being developed to search for FRBs in the presence of radio frequency interference.

Unified Astronomy Thesaurus concepts: Radio transient sources (2008); Radio bursts (1339); Astronomy databases (83)

1. Introduction

The Parkes 64 m diameter radio telescope has been used to discover more than half of the known pulsars through surveys such as those described by Manchester et al. (2001) and Hobbs et al. (2004), as well as the first rotating radio transients (RRATs; McLaughlin et al. 2006) and fast radio bursts (FRBs; Lorimer et al. 2007; Thornton et al. 2013). With updated receivers and signal processing equipment such as the Multi-beam (Staveley-Smith et al. 1996) and the Ultra Wideband Low (Hobbs et al. 2020) Receivers and their corresponding backend instrumentation, the telescope systems have remained versatile and continue to make new discoveries.

The majority of the Parkes high time resolution data sets are publicly available from CSIRO’s data archive\textsuperscript{15} (Hobbs et al. 2011). This archive allows new algorithms to be tested and then applied to the large data volumes (recent examples of new discoveries found in the archive include Pan et al. 2016; Zhang et al. 2018, 2019). The archive contains more than 100 observing projects, with each observing semester for each project stored as a data collection. Approximately 600 such data collections are now available for public access. This archive provides a very long data set (~29 yr) with stable observing systems.

It is likely that astronomical sources remain to be found in the archive. We have therefore started a project in which we will search all the archival data in a self-consistent manner for transient signals and create a database of all the single pulses detected. We have chosen to start by processing the search mode data sets available in the data archive that were recorded using the Multibeam Receiver between 1997 and 2001. The year 2001 was chosen as the initial cutoff date as this enables us to study the apparent lack of FRB discoveries before this date.\textsuperscript{16}

FRBs are bright, single pulses of millisecond duration and expected to have extragalactic origins. Almost 100 FRBs have now been published\textsuperscript{17} (Petrop et al. 2016). The Canadian Hydrogen Intensity Mapping Experiment (CHIME) has the highest rate of new FRB discoveries (The CHIME/FRB Collaboration et al. 2019a, 2019b) and the Australian

\textsuperscript{15} CSIRO Data Access Portal (DAP) https://data.csiro.au.

\textsuperscript{16} The data on the DAP are accessed through specific data collections, which include all observations for a specific project during one observing semester. The last data set used in this paper corresponds to the 2001 May observing semester.

\textsuperscript{17} http://www.frbcat.org
of HTRU observations was estimated by the archive of Parkes observing schedules that the HTRU survey commenced. The HTRU survey data sets are not currently available on the data archive and therefore not included in the dotted bars. The number of HTRU observations was estimated by the archive of Parkes observing schedules (https://www.parkes.atnf.csiro.au/observing/schedules/prev_schedules.php) and the allocated time (i.e., 7092 hr) for the HTRU survey. Note that this value is ∼39% larger than the expected survey length of ∼5100 hr (Keith et al. 2010).

Square Kilometre Array Pathfinder has demonstrated the ability to localize the sources within their host galaxies (Bannister et al. 2019). However, telescopes with high sensitivity such as Parkes continue to make FRB discoveries (e.g., Osłowski et al. 2019) and extend the distribution of FRBs on occurred time, dispersion measure (DM) and expected luminosity etc.

In Figure 1 we show the event times for FRB detections with Parkes overlaid on the epochs of the use of the Multibeam Receiver. From 1997–2010, only five FRBs were discovered, and four of them occurred in the first half of 2001 (Lorimer et al. 2007; Keane et al. 2012; Burke-Spolaor & Bannister 2014; Zhang et al. 2019) and one in 2009 (Champion et al. 2016). From Figure 1 it is clear that there are two intervals (from 1997–2001 and from 2002–2009) in the FRB detections where no FRBs have been detected.

Between 2001 and 2009, the Parkes telescope was involved in large-scale pulsar timing experiments and spectral-line and continuum surveys. The major pulsar survey during this time was the Bates et al. (2011) survey that made use of a multibeam methanol receiver. We only have a significant number of 20 cm multibeam observations for the first apparent gap in the number of detected FRBs, i.e., from 1997 to mid-2001. The data archive contains 38,190 hr of on-sky integration time across all beams using the Multibeam Receiver during this period; see the yellow dotted bars in Figure 1. We would expect events to be evenly distributed throughout the span of the data, but note that these early observations had lower frequency resolution than more recent Parkes observations. This limits the maximum DM and minimum width for detectable FRBs; the implications of this are described in Section 5. The choice of DM ranges and trials in previous analyses, methods of radio frequency interference (RFI) mitigation and signal-to-noise ratio (S/N) estimation could also have led to FRBs being missed (Keane et al. 2019).

In this paper we describe the details of the observations and data reduction in Section 2 before explaining the structure of the database in Section 3. The properties of the new discoveries are presented in Section 4 and discussed in Section 5. We conclude in Section 6. The database and associated software are separately available for public download from CSIRO as described in Appendix A. Appendix A also includes instructions for use of the database.

2. Observation and Data Reduction

The data sets used in the database were all obtained with the primary goal of discovering pulsars. The channelized and polarization-summed signals were one-bit sampled and recorded using an analog filter bank system. Each project semester has been saved as a data collection, named with the project code identifier and the semester period. The observing projects that have contributed observations include: P268, the Galactic Plane Parkes Multibeam Survey (e.g., Manchester et al. 2001; Hobbs et al. 2004; Lorimer et al. 2006); P269, a deep survey of the Large and Small Magellanic Clouds (Crawford et al. 2001; Manchester et al. 2006); P309, a survey of intermediate Galactic latitudes (Edwards et al. 2001); and P360 and P366 which were high-latitude pulsar surveys (Burgay et al. 2006; Jacoby et al. 2009). All the observations used the 20 cm Multibeam Receiver on the Parkes 64 m diameter radio telescope from 1997 May to 2001 August and together led to the discovery of more than 800 new pulsars. There are 21 collections in our current analysis. The central frequency, bandwidth, and number of channels are 1374 MHz, 288 MHz, and 96 respectively. We list the sampling times and integration time of these observations in Table 1.

The data were processed using the pulsar searching software package PRESTO18 (Ransom 2001) on CSIRO’s high performance computer facilities. Strong narrowband and short-duration broadband RFI were identified and marked using the PRESTO routine RFIFIND. We used a 1 s integration time for our RFI identification and the default cutoff to reject time-domain and frequency-domain interference in our pipeline.19 We also recorded the single pulse candidates at zero DM (without any

---

18 http://www.cv.nrao.edu/~sransom/presto/. Note that an update to PRESTO was required in order to handle the archived 1 bit data files from Parkes correctly.

19 Some bright signals, e.g., the pulses of Vela pulsar whose flux density at 1400 MHz is 1050 mJy (Jankowski et al. 2018), are sometimes also marked as RFI by this pipeline.
RFI mitigation) to enable long-term RFI studies with our database.

The data sets were dedispersed at DM values that were determined by the DDPLAN.PY algorithm with the option $R$ ("acceptable time resolution") set to be 0.5 ms and based on the central frequency, bandwidth, number of channels, and sampling time. The DM range was from 1 to 5000 cm$^{-3}$ pc. Data were then dedispersed at each of the trial DMs using the PREPDATA routine with RFI removal based on the mask file produced by RKIFIND. To avoid missing bright burst events, we used the option NOCLIP which disabled autoclipping of data during all the processing steps and the option B for the SINGLE_PULSE_SEARCH.PY routine which disabled the check for bad-blocks. This is an extremely parallel computational challenge so a Perl script, NPROC, was used to accelerate the single pulse searching steps by distributing individual computing jobs across multiple CPUs.

Single pulse candidates with $S/N$ larger than seven were identified using the SINGLE_PULSE_SEARCH.PY routine for each dedispersed time series and for different boxcar filtering parameters with filter widths of 1, 2, 3, 4, 6, 9, 14, 20, 30, 45, 70, 100, 150, 220, and 300 samples. We use the definition of $\sigma$ as presented by PRESTO as our $S/N$ value. A signal with $S/N$ above seven is unlikely to have been generated by receiver noise.

### 3. Database of the Single Pulses

We have created an SQLite\textsuperscript{20 }database to store the transient events. The database contains nine tables that store the data and three indices, which are used to speed up searches of the database. The database schema and description of the parameters are listed in Table 2 and graphically displayed in Figure 2.

| Collection Name | Number of Files | Integration Time (hr) | $t_{\text{samp}}$ (\mu s) | Detected Pulsars | Detected FRBs |
|-----------------|-----------------|-----------------------|---------------------------|-----------------|--------------|
| P268 – 1997AUGT | 4615            | 2668                  | 250                       | 35              |              |
| P268 – 1998JAN  | 4446            | 2594                  | 250                       | 58              |              |
| P268 – 1998MAY  | 5525            | 3128                  | 250                       | 71              |              |
| P268 – 1998SEP  | 3926            | 2287                  | 250                       | 49              |              |
| P309 – 1998SEP  | 20,959          | 1514                  | 125                       | 57              |              |
| P268 – 1999JAN  | 1638            | 956                   | 250                       | 19              |              |
| P309 – 1999JAN  | 23,940          | 1714                  | 125                       | 49              |              |
| P268 – 1999MAY  | 3198            | 1866                  | 250                       | 34              |              |
| P309 – 1999MAY  | 25,803          | 1875                  | 125                       | 48              |              |
| P269 – 1999SEP  | 3120            | 1813                  | 250                       | 34              |              |
| P269 – 2000JAN  | 2405            | 1403                  | 250                       | 19              |              |
| P269 – 2000MAY  | 3068            | 1782                  | 250                       | 19              |              |
| P269 – 2000MAY  | 2082            | 2444                  | 1000                      | 7               | FRB 010724; FRB 010312 |
| P268 – 2000OCT  | 1859            | 1084                  | 250                       | 18              |              |
| P269 – 2000OCT  | 951             | 1276                  | 1000                      | 2               |              |
| P268 – 2001JAN  | 1131            | 660                   | 250                       | 20              |              |
| P269 – 2001JAN  | 2258            | 2520                  | 1000                      | 7               | FRB 010724; FRB 010312 |
| P360 – 2001JAN  | 17,718          | 1277                  | 125                       | 15              | FRB 010125   |
| P268 – 2001MAY  | 1820            | 1062                  | 250                       | 13              | FRB 010621   |
| P360 – 2001MAY  | 13,885          | 1018                  | 125                       | 10              |              |
| P366 – 2001MAY  | 44,122          | 3249                  | 125                       | 19              | FRB 010305   |

Note. Here we also list the number of detections of known pulsars and FRBs during our processing. The DOIs of the collections are listed in Appendix B. Note that, for an unknown reason, some data collections contained observations outside of the observing semester specified. We include all the available data here and in our processing.

### Table 1

Data Collections used in This Study, the Number of Data Files, On-sky Integration Time and the Sample Time of Each Collection

\textsuperscript{20} https://www.sqlite.org/
The huge volume of candidates implies that it would be impossible to inspect them all by eye. Each candidate has the arrival time at the highest observing frequency.

To aid in inspecting these candidates, we have grouped different candidates showing common features. For instance, candidates with adjacent DMs and overlapping start and end times, often derive from the same wide-profile signal. We have therefore

### Table 2

| Name               | Type   | Description                                                                 |
|--------------------|--------|-----------------------------------------------------------------------------|
| collection         | text   | Unit of the data sets, normally contains the observations of one project semester |
| collectionID       | integer| The primary key                                                              |
| timeStamp          | date   | Time stamp of insertion                                                      |
| project            | text   | Observing project code identifier                                            |
| semester           | text   | Semester for the observing project                                           |
| description        | text   | Primary goal of the observation                                              |
| doi                | text   | Digital object identifier linked to the data sets                            |
| telescope          | text   | Telescope used to take the observation                                        |
| pi                 | text   | Principal investigator of the observing project                               |
| access             | text   | Data access status: available or embargo                                       |
| observation        | text   | One complete observing tracking                                              |
| observationID      | integer| The primary key                                                              |
| collectionID       | integer| Identifier for the collection, link to the table of collection                |
| timeStamp          | date   | Time stamp of insertion                                                      |
| timeStartMJD       | real   | Start MJD of the observation in UTC time                                     |
| frontend           | text   | Receiver used to acquire the data                                            |
| backend            | text   | Signal processor used for observation                                         |
| polnNum            | integer| Number of polarizations                                                      |
| sampleTime         | real   | Sample interval for SEARCH-mode data (s)                                     |
| freqC              | real   | Center frequency (MHz)                                                       |
| bandwidth          | real   | Observation bandwidth (MHz)                                                  |
| channelNum         | integer| Number of frequency channels                                                 |
| obs_length         | real   | The full duration of the observation (s)                                     |
| file               | integer| The primary key                                                              |
| fileID             | integer| Identifier for the observation, link to the table of observation              |
| observationID      | integer| The primary key                                                              |
| fileID             | integer| Identifier for the observation, link to the table of file                    |
| timeBegin          | real   | Time of the segment begin                                                    |
| timeEnd            | real   | Time of the segment end                                                      |
| type               | text   | Type of the events in the segment                                            |
| pulsar             | integer| Pulsars’ parameters provided by the latest ATNF pulsar catalog               |
| pulsarID           | integer| The primary key                                                              |
| timeStamp          | date   | Time stamp of insertion                                                      |
| jname              | text   | Pulsar name based on J2000 coordinates                                       |
| raj                | real   | R.A. in J2000 coordinates of the pointing center of the beam (hh:mm:ss,ssss) |
| rijd_s             | real   | R.A. in J2000 coordinates of the pointing center of the beam (deg)          |
| decl2000_s         | text   | Decl. in J2000 coordinates of the pointing center of the beam (dd:mm:ss,ssss) |
| decj               | real   | Decl. in J2000 coordinates of the pointing center of the beam (deg)          |
| azimuthAng         | real   | Azimuth angle (deg)                                                          |
| zenithAng          | real   | Zenith angle (deg)                                                           |
| beamNum            | integer| Beam number for multibeam systems (1 = central beam)                        |
| successProcess     | text   | Here will note the details if the file process failed                        |
| HPBW_d             | real   | Half power beamwidth of the beam (deg)                                       |
| label              | text   | Special file will be labeled                                                 |
| software           | text   | Software used to obtain the candidates                                       |
| softwareID         | integer| The primary key                                                              |
| timeBegin          | date   | Time stamp of insertion                                                      |
| name               | text   | Name of the software                                                         |
| version            | text   | Version of the software                                                      |
| homepage           | text   | Homepage of the software                                                     |
| repository         | text   | URL to software/code repository                                              |
| notes              | text   | Special changes when using the software                                       |
| pipeline           | text   | Pipeline built to process the data and obtain candidates.                    |
| pipelineID         | integer| The primary key                                                              |
| softwareID         | integer| Identifier for the software, link to the table of software                   |
| timeBegin          | date   | Time stamp of insertion                                                      |
| snrThreshold       | real   | S/N threshold to get the candidates                                          |

The huge volume of candidates implies that it would be impossible to inspect them all by eye. Each candidate has the arrival time at the highest observing frequency (the start time of the event) and at the lowest frequency (the end time). To aid in inspecting these candidates, we have grouped different candidates showing common features. For instance, candidates with adjacent DMs and overlapping start and end times, often derive from the same wide-profile signal. We have therefore...
grouped all the candidates that lie within the start and end time of a given event. The table FILESEGMEN records the fileID and the start and end time for these groupings.21

We also provide a table (PULSAR) of known pulsars based on the latest (i.e., version 1.62) ATNF Pulsar Catalogue22 (Manchester et al. 2005). Where possible (i.e., where the beam position is correct and single pulses have been detected) we have linked specific observations to known pulsars using the PSRFILELINK table in the database.

21 The database only contains the information for the candidates, but we plan to also provide the actual segments of raw data in a later version. Some examples are shown in the additional support directory of db_fileSeg mentioned in Appendix A.2

22 http://www.atnf.csiro.au/research/pulsar/psrcat

4. Results

We initially investigated all the files that contained candidates with $S/N > 8$ in the DM-time plane which were plotted by the usage plot_TimeDM described in Appendix A.3. Any candidate that was seen as a isolated burst and has DM larger than $D_{MMW}$ was selected for further analysis. Some files were significantly affected by RFI and therefore it was impossible to view every candidate with this $S/N$ threshold. We therefore also inspected the groupings of candidates. For these groups we only viewed the candidate with the highest $S/N$ present within that group. The $S/N$ of these candidates were slightly increased using a finer grid of DM trials than in the normal processing. We used DM values in a range $\pm 5 \text{cm}^{-3} \text{pc}$ centered at the candidate’s DM and with a
DM step of 0.1 cm$^{-3}$ pc. Eventually we obtained the DMs with the highest S/N.

We re-detected all four of the published FRBs. As listed in Table 1, they are FRB 010125 at a DM of 786.5 cm$^{-3}$ pc in beam 5 with S/N of 17.9, FRB 010312 at a DM of 1163 cm$^{-3}$ pc in beam 7 with S/N of 11.0, FRB 010621 at a DM of 749 cm$^{-3}$ pc in beam 10 with S/N of 15.8, and FRB 010724 at a DM of 373 cm$^{-3}$ pc in three beams (beams 6, 7, and 13) with S/N of 32.0, 15.0 and 24.1.

A new FRB (which, following tradition, we label as FRB 010305) with a DM of 350 ± 5 cm$^{-3}$ pc and S/N = 10.2 was detected in our search. Figure 3 shows the burst in the frequency–time plane and its integrated pulse profile after being dedispersed at the optimal DM value. The signal is much stronger in the lower part of the observing band, which is similar to FRBs such as FRB 010312 (Zhang et al. 2019), FRB 110214 (Petroff et al. 2019), and FRB 171019 (Shannon et al. 2018). Two strong narrowband RFI signatures in the higher band are clear from the figure, but they do not affect our confidence in the FRB detection. We are confident that this is a real FRB as our database contains no other unexpected event with a DM value larger than the Galactic DM contribution (with an S/N value between eight and ten).

We list the properties of FRB 010305 in Table 3. This FRB, the second-earliest known, was only detected in a single beam (Beam 3) of the Multibeam Receiver, and no clear pulse candidate or RFI occurred around the time of the burst in the remaining 12 beams. Detection in a single beam cannot provide a precise position or fluence. The coordinates listed are simply the pointing position of the beam. The burst has a width of 9 ± 2 ms at its 50% power point. In the lower half of Table 3, we provide various inferred properties of FRB 010305. The peak flux density was obtained from the single pulse radiometer equation (Cordes & McLaughlin 2003) and the S/N measurement. The YMW16 electron density model (Yao et al. 2017), which assumes $H_0 = 67.3$ km s$^{-1}$ Mpc$^{-1}$ (The Planck Collaboration et al. 2014) and the local intergalactic medium baryon density $n_{b_{\text{IGM}}} = 0.16$ m$^{-3}$ (Katz 2016), indicates a cosmic distance of 1.2 Gpc with the assumption of a host galaxy DM of 100 cm$^{-3}$ pc. In order to search for the possibility of the FRB repeating (Spitler et al. 2016), we identified 3.24 hr of observations whose beam pointing positions were within $1^\circ.0$ of the position of the beam in which FRB 010305 was detected. No convincing candidates were identified in these other observations.

Our search for new FRBs led to the detection of numerous single, dispersed pulse events with DM values likely to put the source within our Galaxy. Events that are detected in all the multibeam beams simultaneously are commonly referred to as “perytons” and they have been identified as being generated when a microwave oven door is opened prematurely (Petroff et al. 2015). We have detected 22 peryton events in total. Their details can be obtained by the tools presented in Section 5.1 and Appendix A.3.

Most of the candidates in the database are from locally generated interference or from single pulses of known pulsars. We have manually identified 1084 observation files that contain 385 unique pulsars with single pulses of S/N > 8 based on the positions, DMs and estimated periods; the mean observation time for each of these pulsars is ∼55 minutes. A detailed analysis of these pulses and RFI detections will be presented elsewhere. Examples are provided in Section 5.1.

Some dispersed events have only been detected in a single beam, have DMs indicating a Galactic source, and the beam pointing directions are not in the direction of known radio

### Table 3

| Event data UTC | 2001 Mar 5 |
| Event time UTC, $\nu_{1.374\text{\ GHz}}$ | 12:29:16:02 |
| Event time local (AEDT), $\nu_{1.374\text{\ GHz}}$ | 23:29:16:02 |
| Pointing R.A. (J2000) | 04:57:19.5 |
| Pointing decl. (J2000) | -52:36:24.668 |
| Galactic longitude | 260°06 |
| Galactic latitude | -38°34 |
| Beam 3 FWHM | 141 |
| DM (cm$^{-3}$ pc) | 350 ± 5 |
| Observed width (ms) | 9 ± 1.5 |
| S/N | 10.2 |

#### Inferred Properties

| Peak flux density (Jy) | 0.42 |
| Fluence (Jy ms) | 3.78 |
| DM$_{\text{MW,YMW16}}$ (cm$^{-3}$ pc) | 36 |
| Redshift$_{\text{YMW16}}, z$ | 0.3$^a$ |
| Distance$_{\text{YMW16}}$ (Gpc) | 1.2$^a$ |

**Notes.** The S/N was calculated after removing the RFI based on the mask file produced by RR FINF. The width was obtained by fitting the integrated pulse profile at its 50% power point.

$^a$ The DM of host galaxy was assumed to be 100 cm$^{-3}$ pc and the calculation used the YMW16 model (Yao et al. 2017).
pulsars. Five such events, those with S/N values above 10, are shown in Figure 4. The candidate label includes SPC (denoting “single pulse candidate”) and the detection date. We list the best-fitting DM, S/N, detecting beam, the beam pointing position in R.A. and decl. (and converted to Galactic coordinates), observed width at 50% power point, inferred peak flux density, fluence, DM$_{MW,YMW16}$, and Distance$_{YMW16}$ of these SPCs in Table 4. We note that SPC 991113 may be a bright pulse from the RRAT J1739−2521 with a DM of 186.4 cm$^{-3}$pc as their positions are close and the uncertainty for the DM of this SPC is relatively large (∼26 cm$^{-3}$pc). However, the single pulses from that RRAT are not expected to be so bright, nor so broad (Cui et al. 2017). No convincing candidate was obtained from the periodicity search of the files containing these candidates.

These single pulse candidates have only been detected in a single beam of the receiver suggesting that they are unlikely to be from terrestrial signals. There are no known pulsars or RRATs in the beam pointing directions and therefore these candidates are likely to be new RRATs, or giant pulses from currently unknown pulsars. No more convincing candidates were identified in the same sky directions of them from the data archive. We will reobserve these sky regions with the Parkes telescope, but note that all these sky regions could also be observed with the more sensitive MeerKAT (Jonas & MeerKAT Team 2016) telescope and SPC 000621 or SPC 010208 with the FAST telescope (Jiang et al. 2019). However, we note that the positions of these candidates are only poorly determined to date.

5. Discussion

Our primary goal was to process the first four years of Multibeam data from the Parkes telescope in order to understand why very few FRBs were discovered during that time. Between 1997 and 2001 June we now know of five FRBs
(including the new event presented in this paper). However, they all (including our new event) occurred in the first half of 2001.

The expected detection rate for a given set of observations depends upon (1) the DM range searched and (2) the channel bandwidth of the observations. The data sets between 1997 and mid-2001 contain four independent observing projects. These observing projects have different Galactic sky coverage, but they have the same channel bandwidth and we have been consistent in the DM range that was searched. Each of these four projects have successfully detected at least one FRB, as listed in Table 1.

Using the FRB event rate predicted by Champion et al. (2016), approximately 12 FRBs would have been detected in a HTRU-style survey (Keith et al. 2010) assuming the same integration time as our observations. However, the more modern surveys use signal processors with channel widths of only 0.39 MHz compared with 3 MHz for the data processed in this paper. In Figure 1 we present the start date (using a vertical line) of the HTRU survey, which made use of the narrower frequency channels. The use of wider channel widths implies that the observations are less sensitive to high-DM FRB events. For instance, the DM smearing across a 3 MHz channel is ~9 ms at 1.4 GHz for an FRB event with a DM of 1000 cm$^{-3}$ pc. Considering the DMs and observed pulse widths for the 28 FRBs that have been discovered using the Parkes telescope, we find that ~14 would have been missed (or difficult to detect because of significantly lower S/N values) in the earlier data. The event rate of detectable FRBs in the earlier data is therefore significantly lower than in the more recent data. After accounting for our wider channel bandwidths we expect ~5 FRB events in the earlier data (compared with five that we have detected). Note that our one-bit sampled data sets and the sampling time (for some of the data sets) of 1 ms can also affect the detection rate of our search. Our detections are therefore consistent with the FRB event prediction.

Using the definition in Bhandari et al. (2018) of three regions delineated in Galactic latitude of $|b| \leq 19^\circ$, $19^\circ < |b| \leq 42^\circ$, and $42^\circ < |b|$, we list the on-sky integration time in these three latitude ranges for our search in Table 5. Four of the five FRBs reported in our data sets occurred in the intermediate-latitude region from 2001 January to August and the remaining FRB (FRB 010621) occurred in the low-latitude region in the same time range.

Is it reasonable to detect five FRB events within six months, but no event in the previous 3.5 yr? Our search contained 28,690 hr of observations prior to 2001 January and 9500 hr of observations from 2001 January to August (see Table 5 for more details). Assuming that the occurrence of FRBs has a

### Table 4

| Entity | SPC 991113 | SPC 000115 | SPC 000621 | SPC 001122 | SPC 010208 |
|---|---|---|---|---|---|
| Event date UTC | 1999 Nov 13 | 2000 Jan 15 | 2000 Jun 21 | 2000 Nov 22 | 2001 Feb 08 |
| Event time UTC | 04:12:20.92 | 16:28:15.11 | 15:53:14.32 | 20:35:50.82 | 23:17:50.59 |
| Event time local (AEDT) | 15:12:20.92 | 03:28:15.11 | 01:53:14.32 | 07:35:50.82 | 09:17:50.59 |
| Pointing R.A. (J2000) | 17:39:49.6 | 13:28:55.8 | 16:05:35.7 | 08:08:10.6 | 19:05:49.0 |
| Pointing decl. (J2000) | −25:13:16.2 | −58:54:05.9 | −45:45:05.2 | −32:18:11.0 | −01:26:42.1 |
| Galactic longitude (°) | 2.48 | 307.77 | 334.59 | 249.96 | 33.28 |
| Galactic latitude (°) | 3.05 | 3.62 | 4.85 | 0.20 | −3.86 |
| Detected beam number | 12 | 12 | 3 | 7 | 13 |
| DM (cm$^{-3}$ pc$^{-1}$) | 203 | 213 | 65.1 | 135.5 | 102.6 |
| DM$_{MW, YMW16}$ (cm$^{-3}$ pc$^{-1}$) | 407.7 | 357.1 | 294.4 | 549.5 | 330.9 |
| Distance$_{YMW16}$ (kpc) | 4.8 | 6.1 | 1.9 | 0.44 | 3.9 |
| Observed width (ms) | 150 | 30 | 30 | 12.5 | 20 |
| S/N | 12.4 | 19.4 | 11.4 | 11.6 | 10.2 |
| Peak flux density (Jy) | 0.08 | 0.28 | 0.17 | 0.26 | 0.18 |

**Note.** The width was obtained by fitting the integrated pulse profile at its 50% power point.

### Table 5

| Observational Period | Galactic Latitude | Total (hr) |
|---|---|---|
| $0^\circ \leq |b| \leq 19^\circ$ | $19^\circ < |b| \leq 42^\circ$ | $42^\circ < |b| \leq 90^\circ$ | |
| Prior to 2001 Jan | 24,922 | 2217 | 1551 | 28,690 |
| 2001 Jan–Aug | 5016 | 3582 | 902 | 9500 |
| Complete | 29,938 | 5799 | 2453 | 38,190 |

**Note.** Four of the five FRBs reported in our data set occurred in the intermediate-latitude region from 2001 January to August and the remaining FRB (FRB 010621) occurred in the low-latitude region in the same time range.
Poisson distribution, the probability of observing \( k \) events in a given interval is:

\[
P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!},
\]

where \( \lambda \) is the mean number of events per interval. Therefore the probability of detecting no event in our 28,690 hr observations prior to 2001 January (\( \lambda = 3.756 \)) is \( \sim 0.0234 \), corresponding to a confidence level of 0.976, i.e., close to 2\( \sigma \) level of significance that the FRBs are unevenly distributed in time (Gehrels 1986).

5.1. Using the Database to Produce a Training Data Set

Our database contains 568,736,756 single pulse candidates, too many to view by eye. It is therefore likely that large-scale searches of the database will be carried out using machine learning or matched-filtering algorithms. Many such algorithms rely on data sets that can be used to train or optimize the algorithms and the labeled candidates in our database provides such a data set.

Known pulsars are labeled directly within the database. Currently we have labeled 385 pulsars for which single pulses are detectable. These pulsars have a range of properties with pulse period from 0.016 to 1.98 s and DMs between 2.64 and 1172.0 cm\(^{-3}\) pc. Descriptions of how to identify the relevant data files and segments are given in Appendix A.3.

In Appendix A.3 we also show how a user can extract the five FRB events and the single pulse candidates that have been described in this paper. We have also identified 10,614 segments of data that exhibit a range of RFI signatures (including 22 perytons) and 100 segments in which weak (\( S/N \sim 7 \)) burst events were recorded. Finally, we list 14,017 files in which no pulsed candidates were identified and therefore represent “clean” data.

6. Conclusion

As we have reprocessed all the data sets with a self-consistent search, the gap of the FRBs’ detection is unlikely caused by the events not being identified during previous searches. The paucity of FRBs in the early data sets can be explained by the large DM smearing and that the uneven distribution on time is only a fluctuation of \( \sim 2\sigma \).

Pulsars, RRATs, and FRBs continue to be discovered in both new and archival data sets. In the near future the data rates from the next generation of telescopes will be so high that the raw data files will unlikely be able to be preserved. Real-time algorithms will therefore process the data and top-ranked candidates will be presented to the astronomers. It is likely that a database of the pulsed candidates will be produced, and we have explored such a database in this paper. We note that our database is currently 41 GiB, which should be compared with 5.7 TiB for the raw data.

We have demonstrated how such databases can be used to discover new sources. Only \( \sim 199 \) hr of the total 38,190 hr (i.e., \( \sim 0.52\% \)) of the data sets contain candidates. The candidates are dominated by single pulses from known pulsars or RFI signals. We have chosen to also keep the candidate caused by RFI events in our database as we wish to carry out a long-term study of impulsive RFI at the Parkes observatory, but methods to identify the RFI (such as multibeam information; or using machine-learning algorithms trained to detect RFI) could significantly reduce the size of the database.

This is the first version of the database of the single pulses from Parkes telescope as we only searched a part of the archival observations. More data sets of the data archive are being processed and will be included in the next version. The database is publicly available. We hope it will be a useful resource for both the pulsar and FRB communities.

The Parkes radio telescope is part of the Australia Telescope National Facility which is funded by the Australian Government for operation as a National Facility managed by CSIRO. This paper includes archived data obtained through the CSIRO Data Access Portal (http://data.csiro.au). This project was supported by resources and expertise provided by CSIRO IMT Scientific Computing. This work was supported by a China Scholarship Council (CSC) Joint PhD Training Program grant, the National Natural Science Foundation of China (grant No. 11725314) and the Guizhou Provincial Science and Technology Foundation (grant No. [2020] 1Y019). Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

Appendix A

A.1. Downloading and Using the Database

The database is available from:

https://data.csiro.au/dap/landingpage?
pid=csiro:42640.

When connecting to the DAP, because of the size of the file, you will request access via an email address. A temporary password will be sent to your email to access the file directly. You need to request access via WebDAV.

In the email you will be able to click through the final location. You can click through to the file in most contemporary web browsers and 64-bit operating systems. Again, because of the size of the file (and the idiosyncrasies of different web browsers and operating systems) it may be easier to download with a scripting tool like wget. For example:

```
wget -m -np -nH -user = USER-password = PASSWORD https://webdav-bm.data.csiro.au/dap_prd_000042640v001/
```

replacing USER and PASSWORD with the user and password fields in the email.

A.2. Statistics

Currently:
1. it is distributed as a tar-ball (database_file.tar.gz);
2. the archive is \( \sim 9.3\) GiB;
3. when extracted the SQLite database:
   (a) requires \( \sim 41\) GiB for the single database file (singlePulseDAP.db);
   (b) has two (2) additional support directories:
      i. db_fileSeg containing 6944 file segments,
      requiring 573 MiB of disk space;
A.3. Usage

The database can be accessed using SQLITE directly. We also provide six (6) software tools to manage the database:

1. plot_TimeDM to plot the DM—time plane of one file/observation;
2. pfits_read_lbitExtraction to plot the frequency—time plane of one file segment using the extracted files;
3. cone_check_PSR.py to provide pulsars close to the beam pointing position;
4. cone_check_SPC.py to do a cone search for the candidates in the database;
5. searchFRB.py will output the filenames, or the fileIDs and fileSegmentIDs of those containing FRB-like phenomena;
6. get_training_data.py can output the information of different kinds of data set:
   (a) FRB;
   (b) perytons;
   (c) pulsar (DM and flux ranges can be specified).

To compile and execute the code:

1. Download the source code from the git repository:
   
   https://bitbucket.csiro.au/scm/spdb/spc-usertools.git

   with a web browser or using an appropriate git tool. On the command line this might look like:

   git clone https://bitbucket.csiro.au/scm/spdb/spc-usertools.git

2. run make to compile the C code, the dependencies are:
   SQLITE3, CFTSIO and PGPLOT;
3. Python code should run within a Python 3 environment;
4. use the \texttt{--h} command option for the supplied programs to print the help page;
5. for instance, to get the training data for pulsars with DM from 100–110 cm$^{-3}$ pc and flux at 1400 MHz from 1.0–2.0 mJy:

   
   get_training_data -t pulsar-dm1 100 -dmh 110 -fl 1 -fh 2

   generating the collection and filename information, for example:

   
   \begin{verbatim}
   Jname, filename, collection
   J1059-5742 PM0005_00881.sf 1997AUGT-P268
   J1059-5742 PM0005_01201.sf 1997AUGT-P268
   J1059-5742 PM0007_02251.sf 1997AUGT-P268
   J1059-5742 PM0016_02081.sf 1997AUGT-P268
   \end{verbatim}

Appendix B

The DOIs of the data collections used in this study are presented in Table B1.

| Collection Name | DOI |
|-----------------|-----|
| P268–1997AUGT   | doi:10.4225/08/583746ac2c4d6e |
| P268–1998JANT    | doi:10.4225/08/5808b59a36ecbf |
| P268–1998MAYT    | doi:10.4225/08/5850b6f1644170 |
| P268–1998SEPT    | doi:10.4225/08/587b1b111b1be |
| P309–1998SEPT    | doi:10.4225/08/577DD57C0F09F0 |
| P268–1999JANT    | doi:10.4225/08/587b1bd63e6eb |
| P309–1999JANT    | doi:10.4225/08/577F5A89BCB36 |
| P268–1999MAYT    | doi:10.4225/08/578838214A3A5 |
| P268–1999SEPT    | doi:10.4225/08/5884f09c4ed4cd |
| P268–2000JANT    | doi:10.4225/08/58919875e1975 |
| P268–2000MAYT    | doi:10.4225/08/58919875e1975 |
| P268–2000CTT     | doi:10.4225/08/58a0f51a47ab |
| P268–2001JANT    | doi:10.4225/08/58919875e1975 |
| P268–2001MAYT    | doi:10.4225/08/58919875e1975 |
| P360–2001JANT    | doi:10.4225/08/58919875e1975 |
| P360–2001MAYT    | doi:10.4225/08/58919875e1975 |
| P366–2001MAYT    | doi:10.4225/08/598c2d9103f0c |

**ORCID iDs**

S.-B. Zhang @ https://orcid.org/0000-0003-2366-219X
S. Dai @ https://orcid.org/0000-0002-9618-2499
R. N. Manchester @ https://orcid.org/0000-0001-9445-5732
L. Staveley-Smith @ https://orcid.org/0000-0002-8057-0294
X.-F. Wu @ https://orcid.org/0000-0002-6299-1263
D. Li @ https://orcid.org/0000-0003-3010-7661
S.-Q. Wang @ https://orcid.org/0000-0003-4498-6070
H. Qiu @ https://orcid.org/0000-0002-9586-7904
L. Zhang @ https://orcid.org/0000-0001-8539-4237

**References**

Bannister, K. W., Deller, A. T., Phillips, C., et al. 2019, Sci, 365, 565
Bates, S. D., Johnston, S., Lorimer, D. R., et al. 2011, MNRAS, 411, 1575
Bhandari, S., Keane, E. F., Barr, E. D., et al. 2018, MNRAS, 475, 1427
Burgay, M., Joshi, B. C., Deller, A. T., et al. 2019, MNRAS, 486, 19
Champion, D. J., Petroff, E., Kramer, M., et al. 2016, MNRAS, 460, L30
Cordes, J. M., & McLaughlin, M. A. 2003, ApJ, 596, 1142
Crawford, F., Kaspi, V. M., Manchester, R. N., et al. 2001, ApJ, 553, 367
Cui, B.-Y., Boyles, J., McLaughlin, M. A., et al. 2017, ApJ, 840, 5
Edwards, R. T., Bailes, M., van Straten, W., et al. 2001, MNRAS, 326, 358
Gehrels, N. 1986, ApJ, 303, 336
Hobbs, G., Faulkner, A., Stairs, I. H., et al. 2004, MNRAS, 352, 1439
Hobbs, G., Manchester, R. N., Dunning, A., et al. 2020, PASA, 17, 37
Hobbs, G., Miller, D., Manchester, R. N., et al. 2011, PASA, 28, 202
Jacoby, B. A., Bailes, M., Ord, S. M., et al. 2009, ApJ, 699, 2009
Jonas, J. & MeerKAT Team 2016, in Proc. MeerKAT Science: On the Pathway to the SKA (Trieste: PoS), 1
Katz, J. I. 2016, ApJ, 818, 19
Keith, M. J., Johnston, A., van Straten, W., et al. 2010, MNRAS, 409, 619
Lorimer, D. R., Bailes, M., McLaughlin, M. A., et al. 2007, Sci, 318, 777
Lorimer, D. R., Faulkner, A. J., Lyne, A. G., et al. 2006, MNRAS, 372, 777
Macquart, J.-P., & Johnston, S. 2015, MNRAS, 451, 3278
Manchester, R. N., Fan, G., Lyne, A. G., et al. 2006, ApJ, 649, 235
