Astronomy and Computing 45 (2023) 100754

Contents lists available at ScienceDirect

Astronomy and Computing

journal homepage: www.elsevier.com/locate/ascom

Full length article

Processing system for coherent dedispersion of pulsar radio emission

I.A. Girin, S.F. Likhachev, A.S. Andrianov, M.S. Burgin, M.V. Popov, A.G. Rudnitskiy *, V.A. Soglasnov, V.A. Zuga

Astro Space Center, Lebedev Physical Institute, Russian Academy of Sciences, Profsoyuznaya str. 84/32, Moscow, 117997, Russian Federation

ARTICLE INFO

Article history:
Received 3 May 2023
Accepted 22 August 2023
Available online 1 September 2023

MSC:
85-04
85-08

Keywords:
Radio interferometry
Data processing
Data pipeline
Pulsars
Radio astronomy

ABSTRACT

Our study provides pulsar researchers with the possibility to utilize VLBI data for looking at pulsar radio emission with the extreme time resolution typical for broad band VLBI recorders. Pulsars emit micropulses and giant pulses shorter than a microsecond. The short pulse durations indicate high brightness temperatures. This constrains the physical nature of the pulsar radio emission mechanism, whose theory is not yet completely understood.

Our paper describes a system for converting VLBI observation data using coherent dedispersion and compensation algorithms for two-bit signal sampling. Coherent dedispersion is the key to processing pulsar observations to obtain the highest temporal resolution. Correction for signal sampling makes it possible to eliminate parasitic effects that interfere with the analysis of pulsar diffraction patterns. A pipeline has been established that uses the developed converter and the Astro Space Center software correlator. It allows us to reprocess all Radioastron pulsar observations and conduct a search for giant pulses, which requires the highest temporal resolution possible.

1. Introduction

The Radioastron project comprised the 10-meter space radio telescope (SRT) that together with ground VLBI antennas formed a space-ground interferometer with a maximum baseline projection of up to 380000 km and a record angular resolution of about 8 μas (Baan et al., 2022). The SRT was launched on the 18th of July 2011, and successfully operated till January 2019.

Radioastron archived all the original raw baseband data. Such an approach allows re-correlation of original data if new scientific problems arise or improved methods of data reduction and interpretation are developed. At the end of this mission, the total raw data volume was approximately 3500 TB. More details about the archive and observations database can be found in Shatskaya et al. (2020a,b).

Pulsar observations were a significant part of the Radioastron scientific program (Kardashev et al., 2013, 2017). They were conducted at 324 MHz (P-band) or 1668 MHz (L-band) and in some cases at both frequencies simultaneously. Radioastron's intermediate frequency (IF) bandwidth was 16 MHz. The P-band receiver supported one sub-band, while 1668 MHz had two 16 MHz sub-bands. In total, the data of 25 pulsars were accumulated during the Radioastron operation, which included 98 observations having a total duration of 250 h. Usually, at least two large ground-based radio telescopes participate in space-ground VLBI sessions, such as GBT, Arecibo, or the WSRT aperture synthesis system.

As compared to single dish observations of pulsars that are usually performed using temporal resolution, Δt, of the order of 1 μs, interferometric observations are conducted with lower values of Δt. In particular, for the Radioastron pulsar observations $\Delta t = 62$ ns. High temporal resolution of the recorded signal combined with high sensitivity of the ground-based telescopes participating in the observations and their ability to simultaneously measure the flux density in two polarization channels permits, in principle, to study phenomena that involve rapid variations of intensity and/or polarization. An example of such a phenomenon is the longitudinal dependence on polarization of giant pulses from the Crab pulsar. An analysis of that dependence was carried out by Main et al. (2021) and later by Lin et al. (2023), who detected the difference between the locations of emitting regions where pulses and interpulses originate.

Thus, the complete dispersion compensation provided by the coherent method, is absolutely necessary for studying individual pulse structure. This is crucial for understanding the mechanism of pulsar radiation. Some pulsars have a pulse microstructure containing many short bursts, of which only about a dozen have been identified so far. As mentioned above, of particular importance was the study of the structure of giant pulses in the Crab pulsar and the millisecond PSR J1939+2134. These pulses contain unresolved nanoshot peaks. The energy density of radio emission in the region of nanoshot generation is monstrous.

* Corresponding author.
E-mail address: arud@asc.rssi.ru (A.G. Rudnitskiy).

https://doi.org/10.1016/j.ascom.2023.100754
2213-1337/

According to the theory, at such densities, radiation interaction with the medium changes radically. Nanoshot peaks are the only phenomena where the required energy density is achieved. The development of a theory based on the observed properties of nanoshots (so far there are only a few of them) is also interesting for physics in general, and not only for understanding the nature of pulsar radiation (Soglasnov et al., 2004; Soglasnov, 2007; Hankins and Eilek, 2007).

High temporal resolution of interferometric data is achieved partly through the low number of bits in each individual readout of the observed signal. Radioastron observations were performed using one- and two-bit digitization for space and ground telescopes respectively. Precision of observations with a low number of digitizing levels depends critically on quantization thresholds. Radioastron data processing routines assume thresholds are set close to optimal. This is described by Thompson et al. (2017) in Chapter 8, Section 8.3.1. A brief summary of two-bit level quantization can be found in Lawson et al. (1965), and fuller account is given by Van Vleck and Middleton (1966).

In processing data, it is usually assumed that the quantization thresholds are kept at their optimal values. The assumption usually holds for objects with slowly varying flux density, where necessary adjustments are performed by an automatic gain control (AGC) system. However, in observing pulsars the AGC is not used as it cannot function properly when the received signal is rapidly variable. Consequently, digitizing may be performed with quantization thresholds that deviate significantly from the optimal level. Reduction of the data using the standard approach introduces additional errors in the final results. To minimize that error, the algorithms for processing the digitized signal should be generalized to arbitrary quantization thresholds. This problem was considered by Jenet and Anderson (1998).

An example of how using a standard procedure for processing pulsar observations leads to erroneous conclusions was presented by Popov et al. (2023). In that paper it was shown that the dynamic spectra of the interstellar scintillations of the PSR J1239+2453 obtained using standard data processing algorithms exhibit frequency shifts that depend on the longitude and, if real, could be interpreted as a signature of the "interstellar interferometer" effect. However, the shift almost disappears when coherent dedispersion and proper correction for actual quantization threshold values is applied.

In this paper we describe the software for processing pulsar observations performed by the Radioastron. The newly developed preprocessor of raw observational data and the modified Astro Space (ASC) software correlator allow one to perform coherent dedispersion and correctly account for digitization with non-optimal quantization thresholds.

The paper is organized as follows: in Section 2 we briefly discuss dedispersion methods, and the numerical method used for digitizing correction is outlined in Section 3. Technical details of the implementation are described in Section 4. We present the results of application of the described methods to Radioastron observations of the PSR J1239+2453 in Section 5 and conclude in Section 6.

2. Correction of distortions caused by the dispersion

Due to the dispersion of radio waves, emission propagating through the ionized interstellar plasma arrives at the observer with a delay depending on the signal frequency. A short quasi-monochromatic impulse emitted at the low-frequency edge of the observing band will be received later than a synchronously emitted impulse at the high-frequency edge. As described in Lorimer and Kramer (2012), the relative delay, τ_{d}, is given by

$$\tau_{\mathrm{d}} = 8.3 \cdot 10^6 \cdot DM \cdot \frac{B}{f_c^3}, \tag{1}$$

where f_c is a central frequency, B is the bandwidth and DM is a dispersion measure expressed in pc cm^{-3}:

$$DM = \int_0^d n_e dl \tag{2}$$

is defined as the integral of the electron column density along the line of sight to the pulsar, and the dispersion constant D (MHz^2 pc^{-1} cm^3 s):

$$D = \frac{e^2}{2\pi m_e c} = 4.148808 \pm 0.000003 \times 10^3 \tag{3}$$

As a result of relative delay, the variability in integrated flux on time scales shorter than τ_{d} is temporally smeared.

The easiest way to deal with dispersion smearing is to split the receiver's band into several channels, hardware or software ("filter bank","digital spectrometer"), and sum the channel outputs with a time shift equal to the dispersion delay for the frequency of each channel (Hankins and Rickett, 1975; Lorimer and Kramer, 2012). This is a method of incoherent, or post-detector dedispersion. It is widespread because it requires very modest resources, and it is easy to implement in realtime. However, it cannot completely remove smearing, but only reduces it. With an increase in the number of channels, to compensate for delay, their width narrows. There comes a time when the time resolution limitation due to the uncertainty principle is more substantial than the smearing one.

Coherent, or pre-detector dedispersion, allows the signal structure to be restored completely by removing dispersion delay and smearing. Time resolution is limited by receiver bandwidth. An analysis of the signal's Fourier spectrum is its essence. The refractive index of a cold plasma differs from unity by an amount proportional to ν_L^2/ν^2, where ν_L is the Langmuir frequency, ν is the wave frequency, $\nu_L \ll \nu$ (out of resonance). A sine wave propagating through a medium with a refractive index other than one acquires a phase shift proportional to the integral of the refractive index along the line of sight. The dispersion measure, DM, can be easily calculated using Eq. (1)–(4) in Section 2. Therefore, all that is needed to completely eliminate dispersion is to shift each harmonic of the Fourier spectrum of the signal in phase to an opposite value. This will result in the restored signal spectrum.

By using coherent dedispersion, the smearing inside the receiver band can be completely removed. The result can be attributed either to the center of the band or to the edges of the band. In some tasks, such as synchronous observations at several frequencies, timing, it is necessary to consider the total dispersion delay. The result is attributed to an infinitely high frequency.

First of all, the full dispersion compensation provided by the coherent method is absolutely necessary for studying individual pulse structure. This is very critical for understanding the mechanism of pulsar radiation. It was found that some pulsars have a microstructure — a lot of short bursts, and now only about a dozen are known. Especially significant was the study of the structure of giant pulses, the pulsar in the Crab and the millisecond pulsar PSR J1939+2134. In these pulses, there are unresolved peaks — nanoshots. For the Crab pulsar coherent dedispersion made it possible to register pulses with an unprecedented peak flux density of several 10^6 Jy (Hankins and Eilek, 2007).

The coherent dedispersion method is based on the fact that the effect of dispersion on the signal received from a pulsar can be modeled as a linear filtering operation. The original signal can be recovered from the received signal by inverse filtering. The effect is best described in the frequency domain. The dispersion causes the phase shift $\phi(f)$, of a Fourier component of the original signal corresponding to frequency f:

$$\phi(f) = \frac{2\pi D f^2}{f_0^2(f_0 + f)}, \tag{4}$$

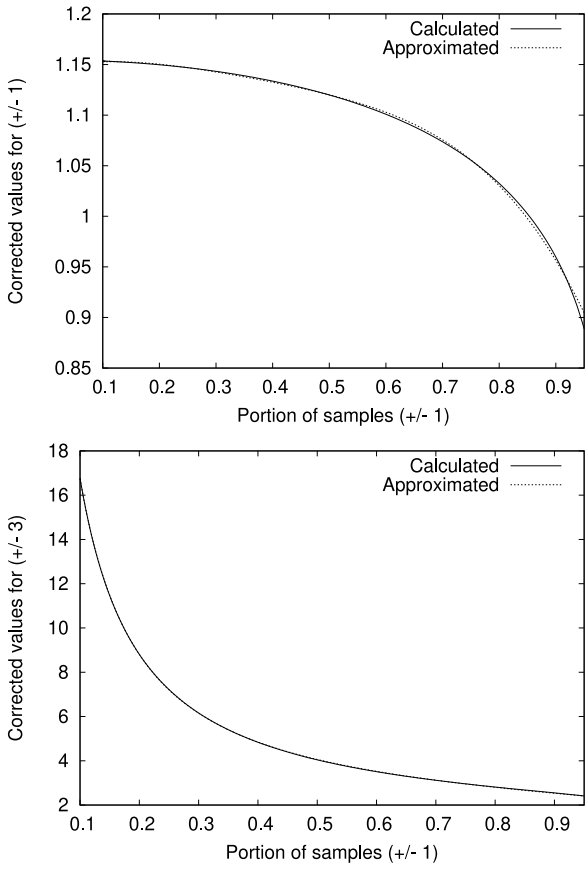

Fig. 1. The corrected values $Y1$ (upper panel) and $Y3$ (lower panel) versus observed fraction of samples equal ± 1.

where D is the dispersion constant, f_0 is the lowest observing frequency.

Consequently, the Fourier transform of the dedispersed signal may be computed by shifting the phase of each Fourier component of the observed signal by $-\phi(f)$.

In the observations of pulsars, this procedure is complicated by the fact that the function sought to be measured is not the Fourier spectrum of the signal in the strict mathematical sense, but the so-called dynamic spectrum. Computation of the dynamic spectrum consists of dividing the observing session into many non-intersecting time intervals, usually of equal duration and covering the whole session, and performing the Fourier analysis of the signal at each interval separately. Because of dispersion, the interval at which the signal is received depends not only on the moment of emission, but also on the frequency. For this reason, coherent dedispersion cannot be performed independently at every interval. An approach to overcome this difficulty was proposed by Hankins and Rickett (1975), who presented an elaborate description of the coherent dedispersion technique that we follow in our study.

3. Correction for two-bit sampling

Ground telescopes operating in VLBI mode record the signal using two-bit (four-level) digitizing. Four digitizing levels are prescribed: -3, -1, $+1$, and $+3$.

The transition threshold between 1 and 3 is supposed to be close to the original analog signal RMS (σ). The optimal threshold value is $t = 0.9674\sigma$ in two-bit sampling (Jenet and Anderson, 1998).

Besides, it is critical to switch off the telescope's automatic gain control (AGC) system in pulsar observations because the system's inertia would not operate properly at the ON-pulse and OFF-pulse stages of pulsar observation. Hence a pulsar signal is an example of a non-stationary noise process. Records of such signals must be corrected for two-bit digitizing. The problem was considered by Jenet and Anderson (1998). They have demonstrated that digitizing the signal before removing dispersive effects generates unwanted systematic artifacts in the data. Namely, the "negative" dips appear around the pulse in the average pulse profile.

We follow the method described in Jenet and Anderson (1998). Description of the applied technique is given in Appendix A.

Fig. 1 displays the corrected values $Y1$ (upper panel) and $Y3$ (lower panel) versus the observed fraction of samples equal ± 1. The method of these calculations explained in Appendix A. We recommend using the approximation functions to calculate $Y1$ and $Y3$ from the measured fraction of samples Φ with digitized values as ± 1. Let us define this independent variable determined from the recorded data as X.

$$Y1(X) = a + b \cdot X + c \cdot X^2 + d \cdot X^3 + e \cdot X^4, \tag{5}$$

$$Y3(X) = A \cdot \exp\left(-\frac{X^B}{D}\right) + C + E \cdot X. \tag{6}$$

For example, $Y1(0.67) = 1.085$ and $Y3(0.67) = 3.23$ for $\Phi = X = 0.67$ with $t = 0.9674\sigma_0$ as an optimal threshold.

The polynomial approximation is sufficient for $Y1$, while an approximation by a combination of an exponent and linear term is more suitable for $y3$. The coefficients used in approximating functions are the following:

$$a = 1.1438(5),\ b = 0.169(7),\ c = -0.96(3),\ d = 1.62(4),$$
$$e = -1.13(2);$$

$$A = 1920(200),\ B = 0.228(5),\ C = 3.24(4),\ D = 0.119(1),$$
$$E = -1.39(3).$$

RMS uncertainties of the coefficients are given in brackets. The difference between the calculated and approximated values cannot be distinguished due to the coarse scale in Fig. 1. In fact, RMS residuals are 0.0012 for $y1$ approximation and 0.011 for $y3$. In regular pulsar observations multilevel recording systems are used at radio telescopes to allow high dynamic ranges typical for pulsar radio emission. In VLBI recorders, two-bit digitizing is the normal mode of observation. For such recording modes, pulsar observers traditionally applied corrections for sampling inaccuracy, described in our paper. See for example Popov et al. (2002, 2008, 2009), Zhuravlev et al. (2011).

4. Implementation of coherent dedispersion

VLBI data processing uses correlators. The Astro Space Center developed its own software correlator, which was used to process almost all Radioastron mission data (Likhachev et al., 2017). This software correlator runs on a cluster and supports several data processing modes: continuum, spectral lines, and pulsars, including the ability to search and correlate giant pulsar pulses. It supports the widely recognized nomenclature of VLBI baseband data formats such as MarkIV/A, Mark5B, VDIF, RDF, LBA, etc.

Until now, the ASC correlator used incoherent dedispersion to process pulsars. Using the implemented method, the pulsar period was divided into bins. The correlator calculated spectra for each bin, then dispersion-related delay compensation was applied via the TEMPO2 software package (Hobbs et al., 2006). In addition, each bin is averaged over the observations. This will

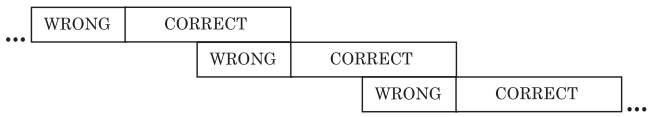

Fig. 2. The scheme of sequential sampling of recording sections with a duration T from a continuous record.

result in each bin containing a signal with no dispersion. As the number of bins increases, the dispersion effect decreases linearly. Therefore, it is necessary to sum only those data with pulse fringes and discard the rest.

In order to implement coherent dedispersion, a data converter was developed that coherently removes dispersion effects and compensates for two-bit sampling. The converter reads raw VLBI baseband data in known formats and converts it into an updated complex format supported by the updated ASC correlator. The converted data is then processed in standard mode without dedispersion. An approach such as this minimizes the number of changes to the correlator structure. This is how the converter operates step by step:

1. Read raw baseband data sample of M size. Supported data formats are the same as for ASC correlator: RDF (Radioastron Data Format), MarkIV/A, Mark5B, K5 A, LBA, and VDIF.
2. Split the sample M into P pieces, each with a N size ($P = M/N$).
3. Determine a parameter X by counting the portion of samples equal to ± 1 and calculate the values $Y1$ and $Y3$ for each data sample in the Pth piece using Eq. (5),(6);
4. Replace every sample s_i by the values $Y1$ or $Y3$, thus performing two-bit sampling correction;
5. Calculate the matrix ϕ_i of the phase shifts depending on the frequency f_i:

$$\phi_i = \frac{2\pi D f_i^2}{f_0^2(f_0 + f_i)} \tag{7}$$

6. Perform a complex M-point Fourier transform for each data chunk numbered $k = 1, \ldots, P$;
7. Multiply Fourier spectra by a correction complex function $R_i = e^{-\phi_i}$;
8. Add M more zeros to the end of the array to get an array of $2M$ values;
9. Perform inverse Fourier transform to obtain complex M samples;
10. Write the converted data to the output file in half-precision float (2 bytes) format.

In the converter, the Fourier transform is implemented as a Fast Fourier Transform (FFT) (Lovelace and Sutton, 1969) using the FFTW library (Frigo and Johnson, 2005). The converter can adjust the FFT size (M) and the number of data samples (N).

At first, one has to calculate the discrete Fourier transform of the recorded signal. Next, the Fourier transform result is multiplied by the complex frequency response function given by Eq. (4). An inverse Fourier transform will yield the recovered signal. It is equivalent to convolution of an impulse response function with the recorded voltage signal in the time domain. The discrete Fourier transform is represented by the periodic assumption. Thus, the first n points corresponding to the time interval of τ_d will be erroneously multiplied with the data wrapped around from the end of the segment and it must be discarded. So in this approach the first portion of the signal will be lost in the interval τ_d.

Fig. 2 illustrates the scheme of sequential sampling of signal sections of duration T from a continuous recording. Sequential values of starting time t_i are determined by the relation $t_i = t_0 + i(T - \tau_d)$. It is clear that T must be larger than τ_d and data reduction efficiency is $1 - \tau_d/T$.

Bit statistics (the number of ± 1 and ± 3) are estimated with a floating average method for P data piece size (N). The statistics are calculated for each data point in the interval [current position $-N/2$, current position $+N/2$]. Further, the two-bit sampling correction coefficient is considered as a function of estimated statistics ((5) and (6)).

The correlated data will be in a UVX format where cross and auto spectra are represented as a float (4 bytes number format). UVX files can be converted directly to IDI-FITS format for further analysis. Finally, bandpass correction and noise cleaning can be applied to the correlated data using the corresponding calibration tools of ASL software package (Likhachev et al., 2020).

4.1. Output data format

As it was mentioned, the developed converter uses a specific format as output (No Packet Data or NPD). The first 512 bytes are dedicated to the header, which is a string that contains information about the observation parameters separated by commas and includes: version, date (DD/MM/YYYY), time (HH:MM:SS.sss..), bandwidth ($1 \cdot 10^{-6} \times 1/dT$ — full bandwidth including all sub-bands), number of channels (NCH, polarization channels and sub-band channels), number of bits per sample.

The header is followed by half-precision float FP16×NCH data. The single-time data sample is (2×NCH) bytes. The order of channels is the same as for the initial input raw baseband data. At the same time, the lower sub-band is converted to the upper one, which requires a shift in the reference frequency when correlating the converted data. NPD format description is provided in Appendix B.

5. Results

We have performed several data processing experiments with the Radioastron observations of PSR J1239+2453 to test the effectiveness of the developed coherent dedispersion data converter. The observations were conducted at $f_0 = 324$ MHz with an IF bandwidth $B = 16$ MHz. Arecibo and Green Bank radio telescopes supported the session. The pulsar dispersion measure is 9.3 pc cm^{-3}. The estimated smearing time is 36.35 ms in the band of 316–332 MHz. For this test, we used the following sizes of FFT and data samples: $M = 10^6$ and $P = 10^3$.

Fig. 3 illustrates the reconstructed signals smoothed by 1 ms for one individual pulse (left panel) recorded at the Green Bank radio telescope. The right panel presents a comparison of pulse profiles averaged over 10 s (7 individual pulses). The middle panels show the signal after coherent dedispersion, but without two-bit sampling correction (raw data). The top panels present the signal with sampling correction applied (corrected data). The differences between corrected and raw profiles are shown in the bottom panels. Flux densities in the figure were calculated with the value of system equivalent flux density (SEFD) equal to 70 Jy for the Green Bank radio telescope. The difference is expressed in fractions of the maximum of the corresponding adjusted profile. One can see that the applied two-bit correction eliminates parasitic signal dips around the pulses and improves signal-to-noise ratio (SNR) by about 10%.

Two-bit sampling of the signal causes additional distortion in pulsar pulse radio spectra. These radio spectra contain significant information about the scattering and scintillation parameters of the interstellar plasma inhomogeneities, namely, the so-called diffraction pattern. From an astrophysical point of view, it is of interest to compare the diffraction pattern for different parts of

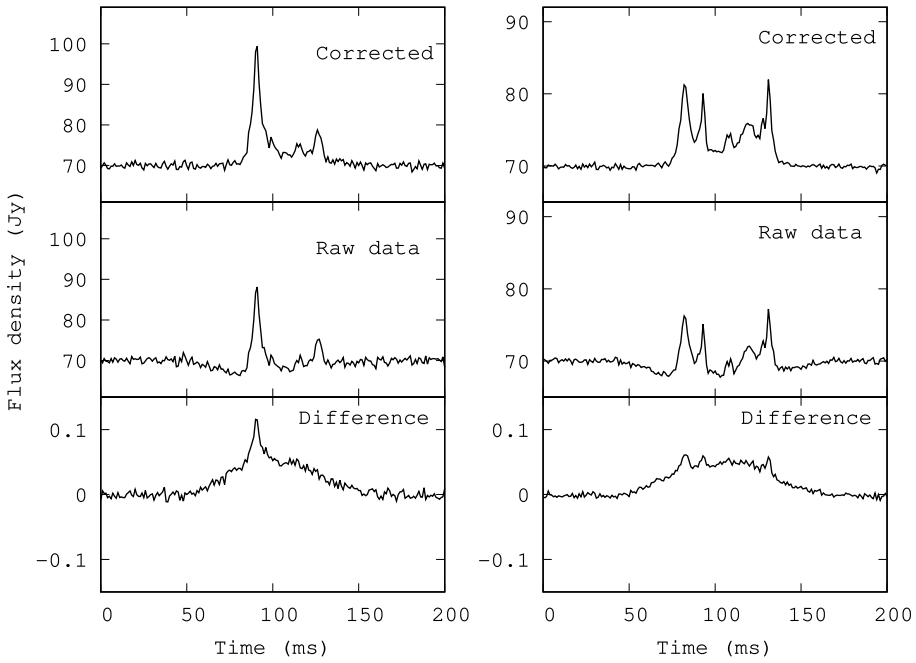

Fig. 3. Comparison of pulsar signals for PSR J1239+2453 restored without correction for two-bit sampling (middle panels), and with such correction applied (top panels). The differences between corrected and raw profiles are shown in the bottom panels. The left panel shows a single pulse, while the right panel presents a pulse averaged over 10 s (7 pulses). Time resolution is one millisecond.

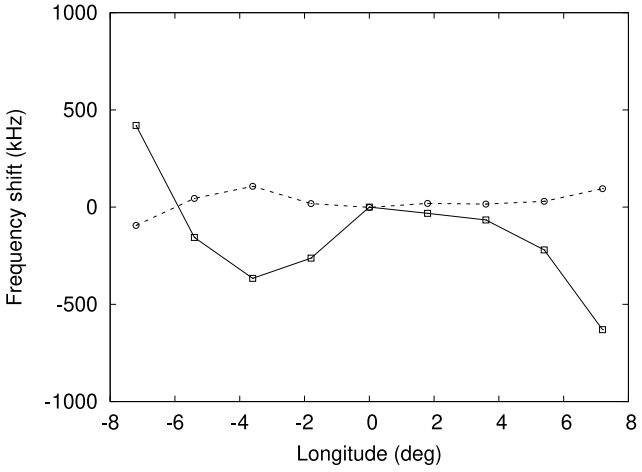

Fig. 4. The frequency shift between dynamic spectra obtained at different longitudes of the mean profile of PSR J1239+2453. The solid curve passing through the squares corresponds to the analysis of spectra obtained without two-bit digitization correction. The dotted line passing through the circles corresponds to the spectra calculated from the corrected signal.

the pulsar averaged profile. The detection of a shift in the diffraction pattern with longitude would indicate that radio emission at distinct longitudes occurs in spatially different parts of the pulsar's magnetosphere, i.e. it would be possible to measure such localization.

Fig. 4 shows the frequency shift between dynamic spectra obtained at different longitudes of the mean profile of PSR J1239+2453. The observations of 20 min were carried out with the Arecibo radio telescope. The solid curve with squares corresponds to the measurements obtained without two-bit digitization correction and the dotted line with circles corresponds to the measurements for the corrected signal. The notable effect of the diffraction pattern's longitude drift disappears after correction. This case was considered in a separate study (Popov et al., 2023).

The program supports multi-threading and we have estimated the approximate performance on the VDIF data of the Green Bank telescope for PSR J1239+2453 which had two polarization channels (LCP and RCP) and one sub-band with a bandwidth of 16 MHz. The measured time ratio was 1:14, thus it takes ≈14 s of real-time to convert one second of such data.

6. Conclusions

Besides coherent dedispersion, the developed converter also corrects digitized signals for two-bit sampling. This tool, together with the ASC software correlator and the ASL software package, forms the pipeline for processing the Radioastron mission pulsar data. The converter and the pipeline allows us to study the peculiarities of the pulsar radio emission that have high brightness temperature with a high time resolution of 62.5 ns provided by Radioastron mission.

Software was tested on the observational data of PSR J1239+2453 observed by the Radioastron space-ground interferometer. The corrected data then were used in the studies in Popov et al. (2023).

Processing of pulsar data with compensation for both dispersion and two-bit sampling is one of the most significant aspects. This allows the removal of parasitic effects in the data that interfere with their correct interpretation. It was shown that two-bit sampling distortions can introduce interference leading to false pulsar diffraction pattern drift in longitude disappear after bit statistic compensation.

Further, the converter and the pipeline will be used to reprocess the Radioastron pulsar raw baseband data with the maximum available temporal resolution. This data will be not smeared. In addition, it will be used to process pulsar VLBI data in modern known and supported formats.

The converter is available at GitHub repository.[1] This utility is a part of ASL software package.[2]

[1] https://github.com/Andrey-Andrianov/NPD_convertor

[2] https://millimetron.ru/en/for-scientists/astro-space-locator

Table B.1
General structure of NPD data format.

Record type	Size, (bytes)
Header	512
Binary data blocks	NCH * 2

CRediT authorship contribution statement

I.A. Girin: Software. **S.F. Likhachev:** Supervision, Conceptualization, Methodology. **A.S. Andrianov:** Software. **M.S. Burgin:** Writing – original draft, Writing – review & editing. **M.V. Popov:** Conceptualization, Methodology, Validation, Visualization. **A.G. Rudnitskiy:** Writing – original draft, Writing – review & editing, Visualization. **V.A. Soglasnov:** Conceptualization, Methodology, Validation. **V.A. Zuga:** Investigation, Validation, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Appendix A. Method of correction two-bit digitizing

First, it is necessary to estimate the undigitized power level σ_0 using a fraction of samples Φ with values of $x_i = \pm 1$ in the specified portion of the signal record. Let t be the threshold established at a given observation. Then the fraction of samples $-t < x_i < t$ is given by

$$\Phi(t) = \frac{1}{\sqrt{2\pi}\sigma}\int_0^t \exp\left(-\frac{x^2}{2\sigma^2}\right)dx = \mathrm{erf}\left(\frac{t}{\sqrt{2}\sigma}\right) \tag{A.1}$$

Calculate the values of the digitizer using the equations below

$$Y3 = \frac{s}{t}\sqrt{1 + \frac{\alpha t}{s}\frac{e^{-\frac{t^2}{2s^2}}}{1 - \mathrm{erf}(t/\beta s)}} \tag{A.2}$$

$$Y1 = \frac{s}{t}\sqrt{1 - \frac{\alpha t}{s}\frac{e^{-\frac{t^2}{2s^2}}}{\mathrm{erf}(t/\beta s)}} \tag{A.3}$$

We substituted $\alpha = \sqrt{\frac{2}{\pi}}$, and $\beta = \sqrt{2}$ to simplify the form of the above-given equations. With these substitutions the Eqs. (A.2), (A.3) are equivalent to the Eqs. (40) and (41) in Jenet and Anderson (1998).

Variable s corresponds to the current value of σ (RMS) during observations. The threshold level t is fixed when the AGC is turned off. But the RMS level s varies between the ON-pulse and OFF-pulse windows.

Appendix B. Description of the NPD format

Below is a detailed description of the NPD format. The data file begins with a header (see Table B.1), followed by data blocks. Time resolution and the duration of recorded observations determine the number of data blocks, where time resolution $dT = 1/(B \cdot 1\,000\,000)$.

The header can be interpreted as a 512-byte string. It contains the data parameters separated by commas (see Table B.2). The start time of the data is the same as in the original file and preserves an initial precision of 1 ns.

Table B.2
Structure of NPD header.

Variable	Description	Value/Format
VER =	Version number	1
DATE =	Start date of the data	DD/MM/YYYY
TIME =	Start time of the data	HH:MM:S.ssss...
BW =	Bandwidth, (MHz)	16
NCH =	Number of channels	4
BZERO =	Zero-baseline flag	0
BSCALE =	Baseline scale factor	1
SAMPLEBIT =	Size of single data record (bits)	16

Example of header: VER = 1, DATE = 22/12/2017, TIME = 10:00:00.15625, BW = 32, NCH = 4,BZERO = 0.0000000e+00, BSCALE = 1.0000000e+00, SAMPLEBIT = 16.

The header is followed by data blocks. Each data block contains signal information for each NCH channel in half-precision format (2 bytes). The number of channels NCH is determined by the number of sub-bands and polarization channels used in the original observations. As an example, RadioAstron recorded the data in two sub-bands of 16 MHz each and two circular polarization channels, RCP and LCP. In this case, NCH = 4 for Radioastron observations.

Time	Format	Value
TIME+dT	NCH*2 bytes	ch1 (2 bytes), …., NCH (2 bytes)
...	...	...
TIME+ND*dT	NCH*2 bytes	ch1 (2 bytes), …., NCH (2 bytes)

Therefore, the blocks containing data will be arranged as follows: where ND is the total duration of observations in seconds divided by time resolution dT, and TIME is the start time of the header data. The data blocks themselves contain no time stamps.

References

Baan, W.A., An, T., Henkel, C., et al., 2022. H_2O MegaMaser emission in NGC 4258 indicative of a periodic disc instability. Nat. Astron. 6, 976–983. doi:10.1038/s41550-022-01706-y.

Frigo, M., Johnson, S.G., 2005. The design and implementation of FFTW3. Proc. IEEE 93 (2), 216–231, Special issue on "Program Generation, Optimization, and Platform Adaptation".

Hankins, T.H., Eilek, J.A., 2007. Radio emission signatures in the crab pulsar. Agron. J. 670 (1), 693–701. doi:10.1086/522362, arXiv:0708.2505.

Hankins, T.H., Rickett, B.J., 1975. Pulsar signal processing. Methods Comput. Phys. 14, 55–129.

Hobbs, G.B., Edwards, R.T., Manchester, R.N., 2006. TEMPO2, a new pulsar-timing package - I. An overview. Mon. Not. R. Astron. Soc. 369 (2), 655–672. doi:10.1111/j.1365-2966.2006.10302.x, arXiv:astro-ph/0603381.

Jenet, F.A., Anderson, S.B., 1998. The effects of digitization on nonstationary stochastic signals with applications to pulsar signal baseband recording. Publ. Astron. Soc. Pac. 110 (754), 1467–1478. doi:10.1086/316273.

Kardashev, N.S., Alakoz, A.V., Andrianov, A.S., Artyukhov, M.I., et al., 2017. RadioAstron science program five years after launch: Main science results. Sol. Syst. Res. 51 (7), 535–554. doi:10.1134/S0038094617070085.

Kardashev, N.S., Khartov, V.V., Abramov, V.V., Avdeev, V.Y., et al., 2013. "RadioAstron"-A telescope with a size of 300 000 km: Main parameters and first observational results. Astron. Rep. 57 (3), 153–194. doi:10.1134/S1063772913030025, arXiv:1303.5013.

Lawson, J., Uhlenbeck, G., of Technology. Radiation Laboratory, M.I., 1965. Threshold Signals. In: Dover Books on Engeineering and Engineering Physics, Dover Publications.

Likhachev, S.F., Girin, I.A., Avdeev, V.Y., Andrianov, A.S., et al., 2020. Astro Space Locator — A software package for VLBI data processing and reduction. Astron. Comput. (ISSN: 2213-1337) 33, 100426. doi:10.1016/j.ascom.2020.100426, URL: https://www.sciencedirect.com/science/article/pii/S2213133720300809.

Likhachev, S.F., Kostenko, V.I., Girin, I.A., Andrianov, A.S., et al., 2017. Software correlator for radioastron mission. J. Astron. Instrum. 6, 1750004–1750131. doi:10.1142/S2251171717500040, arXiv:1706.06320.

Lin, R., van Kerkwijk, M.H., Main, R., Mahajan, N., et al., 2023. Resolving the emission regions of the crab pulsar's giant pulses. II. Evidence for relativistic motion. Astrophys. J. 945 (2), 115. doi:10.3847/1538-4357/acba95.

Lorimer, D.R., Kramer, M., 2012. Handbook of Pulsar Astronomy.

Lovelace, R.V.E., Sutton, J.M., 1969. Digital search methods for pulsars. Naute 222 (5190), 231–233. doi:10.1038/222231a0.

Main, R., Lin, R., van Kerkwijk, M.H., Pen, U.-L., Rudnitskii, A.G., et al., 2021. Resolving the emission regions of the crab pulsar's giant pulses. Astrophys. J. 915 (1), 65. doi:10.3847/1538-4357/ac01c6.

Popov, M.V., Bartel, N., Andrianov, A.S., Burgin, M.S., et al., 2023. Technical constraints on interstellar interferometry and spatially resolving the pulsar magnetosphere. Astrophys. J. 954 (2), 126. doi:10.3847/1538-4357/ace961.

Popov, M.V., Bartel, N., Cannon, W.H., Novikov, A.Y., et al., 2002. Pulsar microstructure and its quasi-periodicities with the S2 VLBI system at a resolution of 62.5 nanoseconds. Astron. Astrophys. 396, 171–187. doi:10.1051/0004-6361:20021402, arXiv:astro-ph/0107073.

Popov, M.V., Soglasnov, V.A., Kondrat'ev, V.I., Bilous, A.V., et al., 2008. Results of three-frequency monitoring of giant pulses from the Crab pulsar. Astron. Rep. 52 (11), 900–909. doi:10.1134/S1063772908110048.

Popov, M., Soglasnov, V., Kondratiev, V., Bilous, A., et al., 2009. Multifrequency study of giant radio pulses from the crab pulsar with the K5 VLBI recording terminal. Publ. Astron. Soc. Japan 61, 1197. doi:10.1093/pasj/61.6.1197, arXiv:0903.2652.

Shatskaya, M.V., Abramov, A.A., Fedorov, N.A., Kostenko, V.I., et al., 2020a. Data processing center of RadioAstron space VLBI project. Adv. Space Res. 65 (2), 813–820. doi:10.1016/j.asr.2019.05.043.

Shatskaya, M.V., Andrianov, A., Likhachev, S.F., Seliverstov, S., et al., 2020b. In: Pizzo, R., Deul, E.R., Mol, J.D., de Plaa, J., Verkouter, H. (Eds.), Data Processing Center for Space VLBI Missions. In: Astronomical Society of the Pacific Conference Series, vol. 527, p. 21.

Soglasnov, V., 2007. Amazing properties of giant pulses and the nature of pulsar's radio emission. In: Becker, W., Huang, H.H. (Eds.), WE-Heraeus Seminar on Neutron Stars and Pulsars 40 Years After the Discovery. p. 68. doi:10.48550/arXiv.astro-ph/0701190, arXiv:astro-ph/0701190.

Soglasnov, V.A., Popov, M.V., Bartel, N., Cannon, W., Novikov, A.Y., Kondratiev, V.I., Altunin, V.I., 2004. Giant pulses from PSR B1937+21 with widths <=15 nanoseconds and T_b>=5$\times 10^{39}$ K, the highest brightness temperature observed in the universe. Agron. J. 616 (1), 439–451. doi:10.1086/424908, arXiv:astro-ph/0408285.

Thompson, A.R., Moran, J.M., Swenson, J., 2017. Interferometry and Synthesis in Radio Astronomy, third ed. doi:10.1007/978-3-319-44431-4.

Van Vleck, J., Middleton, D., 1966. The spectrum of clipped noise. Proc. IEEE 54 (1), 2–19. doi:10.1109/PROC.1966.4567.

Zhuravlev, V.I., Popov, M.V., Kondrat'ev, V.I., Kovalev, Y.Y., et al., 2011. Parameters of giant pulses from the Crab pulsar measured with the Mark5A VLBI system. Astron. Rep. 55 (8), 724–732. doi:10.1134/S1063772911080105.