VOLKS2: A Transient Search and Localization Pipeline for VLBI Observations

Lei Liu1,2,3,4, Zhijun Xu1,2, Zhen Yan1,2, Weimin Zheng1,2,3,4, Yidan Huang1,2, and Zhong Chen1,2
1 Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, People’s Republic of China; liulei@shao.ac.cn
2 Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210008, People’s Republic of China
3 Shanghai Key Laboratory of Space Navigation and Positioning Techniques, Shanghai 200030, People’s Republic of China
4 National Basic Public Service Data Center, Beijing 100190, People’s Republic of China

Received 2021 June 18; revised 2021 August 10; accepted 2021 August 11; published 2021 September 23

Abstract

We present VOLKS2, the second release of “VLBI Observation for transient Localization Keen Searcher.” The pipeline aims at transient search in regular VLBI observations as well as detection of single pulses from known sources in dedicated VLBI observations. The underlying method takes the idea of geodetic VLBI data processing, including fringe fitting to maximize the signal power and geodetic VLBI solving for localization. By filtering the candidate signals with multiple windows within a baseline and by cross matching with multiple baselines, RFIs are eliminated effectively. Unlike the station autospectrum-based method, RFI flagging is not required in the VOLKS2 pipeline. EVN observation (EL060) is carried out, so as to verify the pipelines detection efficiency and localization accuracy in the whole FoV. The pipeline is parallelized with MPI and further accelerated with GPU, so as to exploit the hardware resources of modern GPU clusters. We can prove that, with proper optimization, VOLKS2 could achieve comparable performance as autospectrum-based pipelines. All the code and documents are publicly available, in the hope that our pipeline is useful for radio transient studies.

Unified Astronomy Thesaurus concepts: Very long baseline interferometry (1769); Publicly available software (1864); Transient detection (1957); Distributed computing (1971); GPU computing (1969); Pulsars (1306)

1. Introduction

The search of radio transients is a hot topic in modern astronomy. Fast radio burst (FRB), as one of the most mysterious transient events, was first discovered in 2007 (Lorimer et al. 2007). Since then, more than 300 events have been recorded in TNS.5 Recently, CHIME present a catalog that contains 535 FRBs between 400 and 800 MHz from 2018 July 25–2019 July 1 (The CHIME/FRB Collaboration et al. 2021). FRB 20200428, discovered by CHIME (The CHIME/FRB Collaboration et al. 2020) and STARE2 (Bochenek et al. 2020), is assumed as the first event detected inside the Milky Way. The simultaneous detection in high energy band (Mereghetti et al. 2020; Li et al. 2021) and the accurate localization by FAST (Zhang et al. 2020) confirm the burst is from the known Galactic magnetar SGR 1935+2154, which strongly supports the magnetar origin of FRB (Yuan et al. 2020).

Although the mainstream of transient search is aperture arrays, e.g., UTMOST (Caleb et al. 2016), CHIME (Amiri et al. 2018), LOFAR (ter Veen et al. 2019), and the popular search pipelines are station autospectrum-based (Ransom 2011), we think it is still worthwhile to develop a pipeline that is based on cross spectrum and dedicated to Very Long Baseline Interferometry (VLBI) observations. For one thing, transient search with unknown position could be conducted as commensal task in regular VLBI observations, e.g., V-FASTR (Wayth et al. 2011) and LOCATe (Paragi 2016). Moreover, for those repeating sources, e.g., repeating FRB, magnetas, pulsars, etc, dedicated VLBI observations are usually required for following monitoring and high accuracy localization (Chatterjee et al. 2017; Marcote et al. 2017). Therefore, a pipeline that is able to detect single pulse and carry out localization is very necessary for VLBI observations.

Liu et al. (2018a) propose a cross-spectrum-based single pulse-search method. It takes the idea of geodetic fringe fitting, which maximizes the single pulse signal by fully utilizing the cross spectrum fringe phase information. Once the single pulse is detected, its position can be derived immediately with geodetic VLBI solving (Liu et al. 2019). The cross-spectrum-based search method outperforms the autospectrum method by its ability to extract single pulse signal from highly RFI contaminated data (Liu et al. 2018b). Moreover, as the concept of “Satellite Internet Access” becomes popular, a large amount of satellites are sent to the Earth orbit, e.g., the “Starlink” project initiated by SpaceX. Those satellites not only transmit radio signals at a wide range of frequencies, but also reflect FM broadcastings from the ground. Recent studies suggest those artificial signals are huge threats for radio astronomy especially transient studies: it is almost impossible for station autospectrum-based methods to differentiate all of these signals. For the detected candidates, cross matchings with NORAD database are even required for further confirmation (Clery 2020; Prabu et al. 2020). In contrast, RFI contamination is not a big problem for the cross-spectrum-based method. Since the baselines are usually thousand kilometers long, the probability that RFIs are simultaneously detected by multiple stations and aligned in frequency domains is extremely small. Moreover, RFIs could be eliminated effectively by filtering the candidate signals with multiple windows inside a baseline and by cross matching with multiple baselines.

In Liu et al. (2019), we present the first release of VOLKS pipeline that implements the cross spectrum based method. However, the pipeline is still in its very preliminary stage, e.g., no parallel support, no hardware acceleration, dependence of external software (AIPS6), etc. What’s more, the pipeline is only tested with the Chinese VLBI Network (Zheng 2015).
Figure 1. Data flow of the VOLKS2 pipeline. Visibilities are first loaded into memory for all baselines in a given time range. Then initial calibration is carried out for every polarization and intermediate frequency (IF). Then visibilities from multiple polarizations are combined together. Dedispersion is carried out for every dispersion measure (DM). After fringe fitting, single baseline window matching and multiple baselines cross matching, a single pulse (SP) list is obtained for each DM. For every single pulse, residual delay of each baseline is derived via fringe fitting. Finally the offset to a priori position is derived with geodetic VLBI solving.

In this paper, we introduce the second release of the pipeline: VOLKS2 (Liu 2021). Compared with its first release, it implements many features that have been verified in the actual data processing. We can demonstrate that, this release is more suitable for productive deployment.

This paper is organized as follows: Section 2 introduces the VOLKS2 pipeline, including its algorithm and performance improvement. In Section 3, the pipeline is applied to EVN observation EL060. The single pulse search and localization result is presented. Section 4 presents the conclusion.

2. The VOLKS2 Pipeline

The detailed introduction to the VOLKS2 pipeline has been given in Liu et al. (2018a, 2019). Significant improvements have been made since the first release of the pipeline. Figure 1 present the data flow of current pipeline. In this section, we will focus on the updates of the pipeline and the underlying algorithms.

2.1. Calibration

The calibration is divided into two parts, namely initial calibration and fine calibration.

2.1.1. Initial Calibration

The purpose of initial calibration is to correct the initial delay \( \Delta x_n^0 \) and phase \( \Delta \phi_n^0 \) for every IF (n for IF index), baseline, and polarization. The calibration values (delay and phase) are derived by carrying out fringe fitting on the strong calibration source. We assume these quantities reflect the status of the VLBI system and are kept unchanged across the observation. Figure 2 demonstrates the fringe phase before and after initial calibration for the calibration source itself. The large IF delay is mainly due to station clock offsets. The discrepancies of initial delay and phase between IFs are small; however, they still exist. After calibration, the fringe phase becomes flat and aligned across IFs.

2.1.2. Fine Calibration

Fine calibration is carried out with reference sources close to the target. We assume the position of the reference source is accurate. Therefore the nonzero residual delay across all IFs (upper panel of Figure 3) is caused by instruments and atmosphere. For each baseline, a residual delay is fitted. The fitting scheme is identical to the target source as described in Section 2.2. The fitted value will be used to correct the baseline residual delay. Unlike phase reference calibration, which requires delay, delay rate, and phase information, fine calibration uses only delay information. This is demonstrated in Equation (2) of Section 2.3.

2.2. Fringe Fitting

Figure 4 presents the cross spectrum (fringe phase and amplitude) of a single pulse detected in the pulsar observation. Following the Mk4 format adopted in the geodetic VLBI postprocessing package HOPS7, the amplitude of each frequency point is normalized with the autospectrum dispersion of the two stations: \( |\tilde{S}_f| = |\tilde{S}_f| / (\sigma^2 f^2)^{1/4} \). We would like to demonstrate that the fringe fitting process is effective in extracting single pulses from background noise. In this example, the fringe phase is clearly detected with just 4 milliseconds integration. A detailed explanation of the procedure is presented in Liu et al. (2018a). For the target single pulse search scan, the length of APs8 are in the level of milliseconds, e.g., 1.024 ms for EL060. For each baseline, several APs are summed together along the time axis to construct time segments with given window sizes. Fringe fittings are carried out for those time segments in each window. Window sizes are set to be comparable with the pulsar width, e.g., for PSR J0332+5434, the pulsar width is around 7 ms. The corresponding window sizes are 2 ms, 4 ms, 8 ms, and 16 ms. A single pulse is regarded as detected if it appears in at least two windows.

One thing we have to emphasize is the slightly different fringe fitting scheme. In Liu et al. (2018a), standard geodetic
VLBI fringe fitting is used, which involves searching two quantities namely single-band and multiband delay. However this scheme is designed for legacy system, in which the bandwidth of each IF is much narrower than IF intervals. In the typical astrophysical observation and in the next generation geodetic VLBI system (Petrachenko et al. 2013), IFs are much wider and always continuous. After calibration, it is more suitable to treat them as one fringe. Therefore, in VOLKS2, we search only one delay $\Delta \tau$ across all IFs inside the given band:

$$ G(\Delta \tau) = \sum_{n=0}^{N-1} \sum_{j=1}^{J-1} S(n, j) e^{-i\Phi(n, j)} $$

with

$$ \Phi(n, j) = 2\pi(f_0^n + f_j - f_{\text{ref}})\Delta \tau + 2\pi f_j \Delta \tau_0^n + \Delta \phi_0^n, $$

where $G(\Delta \tau)$ represents the searching function of the corresponding time segment $S$, $f_j$ is the $j$th frequency point in the $n$th IF, $f_{\text{ref}}$ is the reference frequency for fringe fitting, and $\Delta \tau_0^n$ and $\Delta \phi_0^n$ are the initial delay and phase calibration values of the corresponding IF.

In the actual implementation, delay search is carried out with fast-Fourier transform (FFT). To reduce the effect of cyclic convolution, padding is required. We choose a padding factor of 4 in the current pipeline. Standard geodetic fitting scheme conducts 2D FFT, which requires a memory size of at least 16 times ($4 \times 4$) that of the actual data. VOLKS2 conducts 1D FFT, which reduces memory consumption by a factor of 4. This is especially important for graphics processing unit (GPU) computing: compared with CPU, the size of GPU memory is quite limited. Lower memory consumption means we can initiate more computing context within one card, and therefore achieve higher performance.

Figure 2. Demonstration of initial calibration. Upper and lower panels demonstrate the fringe phase before and after calibration. The initial phase and delay are derived for every individual polarization, baseline, and IF. Parameters: EL060 (EVN), J1800+3848, Scan 36, LL polarization, Ir-Mc baseline, 30 s integration.

Figure 3. Demonstration of fine calibration. Upper and lower panels demonstrate the fringe phase before and after calibration. The residual delay is derived by carrying out fringe fitting across all IFs. Parameters: EL060 (EVN), J0346+5434, Scan 37, Ir-Mc baseline, 60 s integration.
Moreover, the current 1D scheme provides much higher delay-search resolution, which helps enhancing the SNR of the detected signal.

### 2.3. Localization

Localization is an important step in single pulse search. For VLBI observation, the radio-imaging method is usually adopted, such as the famous “realfast” pipeline (Law et al. 2015), which plays an important role in the first successful localization of repeating burst FRB121102. However, such kind of method carries out single pulse search in every snapshot, which is time consuming. As pointed out by Liu et al. (2019), by assuming the single pulse is a point source, it is possible to derive the single pulse position given the relation between the residual delay and the offset to a priori position. To improve the localization result, the single pulse fringe of each baseline must be calibrated. In the first release of VOLKS, this is conducted in AIPS. The pipeline at that time provides tools for preparing AIPS inputs and output. The calibration is conducted manually in AIPS.

For the design of VOLKS2, we want to get rid of the dependence of external software, and make the calibration and localization process simple and automatic. Therefore, we propose to conduct fine calibration, in which the residual delay is calibrated with the fringe fitting result of nearby reference sources:

\[
\tau - \tau_{\text{cal}} = \frac{\partial \tau}{\partial \alpha} \Delta \alpha + \frac{\partial \tau}{\partial \delta} \Delta \delta,
\]

where \(\tau\) and \(\tau_{\text{cal}}\) are the single pulse residual delay and the fine calibration delay for this baseline, \(\frac{\partial \tau}{\partial \alpha}\) and \(\frac{\partial \tau}{\partial \delta}\) are partial derivatives of delay by Ra and Dec at a priori position, and \(\Delta \alpha\) and \(\Delta \delta\) are offsets to a priori position. The derived position is the a priori position corrected by offsets.

According to above equation, by assuming the single pulse as a point source and taking the idea of geodetic VLBI solving, we may derive the single pulse position with delay information only. This is different from phase reference calibration carried out by AIPS, which uses all information including delay, rate, and phase. Compared with Equation (2), phase reference calibration might be more accurate. However, it is more complicated and therefore not suitable for automatic implementation. Moreover, VOLKS2 pipeline is designed for regular VLBI observations; phase reference calibration requires the reference source close to the target source, which is not always available.

### 2.4. Performance Improvement

The purpose of the VOLKS2 pipeline is to carry out single pulse search in VLBI observations. The designed running platform is modern GPU cluster. As the most time consuming part, fringe fitting is our concern for speedup. This part is both data and computationally intensive. First of all, loading data into memory is already a huge challenge: in EL060, which is used for single pulse search testing, the data size of a target search scan is around 650 GB (1.024 ms AP, 190 s duration, 7 stations, and 21 baselines). In contrast, the memory size of one computing node is in the order of 100 GB. Even if using RAID to improve disk IO, the disk reading speed is still no higher than 1 GB s\(^{-1}\). Moreover, the fitting process itself is time consuming: in the above scan, for one DM, we have to carry out 4 million 32768 sized FFT, and a comparable number of matrix sum and finding maximum amplitude operations. To make the time consumption of fringe fitting acceptable, in the development stage, a lot of efforts have been made to improve the performance and to fully exploit the hardware resources, such as parallelization and code optimization.

#### 2.4.1. Parallelization

The fringe fitting part of the pipeline is fully parallelized with the message passing interface (MPI). When the program starts, one assignment process and multiple calculation processes are initiated. The number of latter one depends on the CPU core number. For each scan, when a calculation process is idle, it sends task request to assignment process and obtains the time range. Then it loads the corresponding data into memory, carries out fringe fitting for all baselines in this
time range, and saves the results in disk. This scheme is easy to understand and implement. No data communication is required between calculation processes, which keeps the architecture simple: the MPI part and the fringe fitting part are fully decoupled. When some processes are loading data, others carry out calculation. In this way the overhead of disk reading is reduced.

2.4.2. Code Optimization

The fringe fitting part could be divided into six steps:

1. Disk to memory. Load visibilities of given time range to memory.
2. Baseline load. Load the required baseline data to the buffer array.
3. Initial calibration. Carry out IF initial phase and delay calibration for buffer data.
4. Dedispersion. For buffer data with a given DM value.
5. FFT. For each AP, carry out FFT along the frequency axis.
6. Find max. For each window size, combining the FFTed APs along the time axis accordingly; find out the maximum amplitude along the delay axis (frequency axis after FFT) for each combined AP.

In the above steps, step 1 is conducted once for a given time range; steps 2–6 are repeated for every baseline; and steps 4–6 are repeated for every DM value and baseline. To investigate the time consumption of each step, we carry out benchmarks. Figure 5 demonstrates the result before GPU acceleration. For baseline load (step 2), both NumPy and C results are presented (dashed and solid lines). Other steps are conducted with NumPy.

![Figure 5. Benchmark results of fringe fitting part before GPU acceleration. For baseline load (step 2), both NumPy and C results are presented (dashed and solid lines). Other steps are conducted with NumPy.](image1)

Figure 5. Benchmark results of fringe fitting part before GPU acceleration. For baseline load (step 2), both NumPy and C results are presented (dashed and solid lines). Other steps are conducted with NumPy.

target record with \(\sim\log_2 N\) read requests (\(N\) is the total number of records). This routine greatly reduces the records search time. However, the disk reading speed is still determined by the hardware configuration. As a result, the time consumption of step 1 is large and cannot be easily reduced. Fortunately this step is conducted only once. When averaged to every individual baseline and DM value, the portion is small.

After loading the records to memory, for each record, one has to parse the record header, find out the necessary information (time, polarization, baseline number, and frequency), then copy the data to the right place in the buffer array. This is summarized as step 2, which involves a large number of conditional jump and memory read/write operations, and is usually not suitable for script languages such as Python. Therefore, we rewrite this part in C, compile it to dynamic loading library, and call it in Python with “ctypes.” With this optimization, the time consumption of this step is reduced by an order (dotted and solid gray line in the figure), which makes the following optimization worthwhile.

After rewriting step 2 in C, the actual fringe fitting calculation, steps 4–6 become the bottleneck. We use GPU to achieve further speedup. Two GPU computing frameworks, Torch and CuPy, are used to call GPU resources. The benchmark is carried out with the NVIDIA RTX 3080 in the server. The result is demonstrated in Figure 6. With the help of GPU, steps 5 and 6 (FFT and Find max) achieve more than 2 orders speedup. The dedispersion step (step 4) is also accelerated when \(t_{\text{seg}} \geq 1\) s. However, the time consumption of this step does not change with data length. The reason is, for each frequency point, dedispersion invokes the GPU kernel once and applies a specific time shift to the data. In EL060, which is used for the benchmark, the GPU kernel is invoked 8192 times. Since the time shift operation is simple, the kernel running time is negligible. However the accumulated kernel launch overhead, which is independent of data length, becomes significant. This part will be further optimized with the update of the pipeline.

One important parameter for optimization is the overall speedup. According to the analysis above, one may realize that the value depends on the specific application scenarios: for a typical configuration of EL060 (data length 4 s, 21 baselines), with the known DM, the speedup is 7.9 X (209.1 s versus 26.6 s). However, if the DM search is required, in the case of
20 DM values, the speedup is 23.7 X (1876.1 s versus 79.0 s). We would like to point out that some steps, e.g., initial calibration and dispersion, can be further optimized, so as to achieve even higher speedup.

3. Single Pulse Detection in EL060

We carry out single pulse detection with EL060 (EVN). The purpose of the observation is to verify the cross spectrum based single pulse detection scheme and the corresponding VOLKS2 pipeline. The main parameters of the observation are listed in Table 1. The observation is divided into pulsar part and the rotating radio transient (RRAT; McLaughlin et al. 2007) part. In general, single pulses are successfully detected and localized in the pulsar observation. However, no single pulse is detected in the RRAT observation. We will present the detailed result and analysis in this section.

3.1. Pulsar Observation

This part observes pulsar J0332+5434 (B0329+54). By placing the pulsar in the FoV with different offsets to the FoV center, we could investigate the detection efficiency and the localization accuracy in the whole FoV. The size (diameter) of the FoV is estimated as \(\lambda/D\), where \(\lambda\) is set to 18 cm (L band) according to the observation. \(D\) is the antenna diameter. A value of 32 m is selected, which is common for EVN stations (Mc, Tr, Ir, Sv, Zc, Bd). Above the yield a FoV radius of \(r = 9.7\). When preparing for the observation schedule, the pulsar is placed at four different places in the FoV by adjusting the offset between the pulsar and the antenna pointing center. The details of point centers together with the detection results at that place are listed in Table 2. Among those four pointing centers, P0 points to pulsar’s a priori position (R.A. 03\(^{\text{h}}\)32\(^{\text{m}}\)59\(\rightleftharpoons\)368, decl. 54\(^{\circ}\)34\(\rightleftharpoons\)43\(\rightleftharpoons\)57). Of1, Of2, and Of3 point to three different offsets by adjusting decl. only (R.A. is kept unchanged).

3.1.1. Detection Rate

To evaluate the detection rate of each source under the same configuration, we exclude the Onsala (“O”) station, since it is not available in the observation of OF2 and OF3. Moreover, for each scan, the on-source time of each telescope is different. Therefore, we only include single pulses in the time range that all the considered stations are available for observation.

In total, we have seven stations (Bd, Ir, Mc, Sr, Sv, Tr, Zc) that consist of 21 baselines. For each baseline, several APs are summed together along the time axis to construct time segments of different window sizes. We choose window sizes of 2, 4, 8, and 16 APs. A single pulse is assumed to be detected when it fulfills the follow criterion:

1. For the time segments series of given window size, a single pulse is selected if its normalized power exceeds the threshold, which is set to 3 in this observation.
2. For a give baseline, match single pulses from four windows. A single pulse is selected if it appears in at least two windows.
3. Cross-match single pulses from multiple baselines. A single pulse is identified if it is detected with at least five baselines.

Note that above criterion, together with the selection of windows sizes, could be adjusted according to observation.

According to Table 2, for each target, we calculate the total observation time and the “expected” single pulse number. Once a single pulse is detected, we could differentiate whether it is from the pulsar based on its pulsar phase. Real signals from pulsar and RFIs are labeled as “detected” and “invalid,” respectively. We get a lower detection rate and an invalid rate when the pulsar is close to the FoV center (P0 and Of1). However this does not agree with our expectation: the central region of the FoV means a higher antenna response, which should yield a higher detection rate. By investigating the signal-to-noise ratio (S/N) as a function of time, the decrease of the S/N in one of the P0 scan is observed in almost all baselines, which suggest that this is due to the flux variation of the pulsar. The most possible explanation of this phenomenon is refractive interstellar scintillation, which is caused by the small scale density fluctuations and appears as flux density variations in both time and frequency domains. According to Wang et al. (2008), the scintillation timescale of the PSR J0332+5434 is 17.1 minutes, which is consistent with the dwelling time of each pointing center. Therefore, our postulation is the detection efficiency is roughly constant in the whole FoV. The variation of detection efficiency from 70%~90% is mainly caused by interstellar scintillation.

3.1.2. Localization Accuracy

For single pulses detected with the above criterion, we carry out localization using the method described in Liu et al. (2019). As an improvement of the pipeline, fine calibration is proposed and introduced in Section 2.3. When processing the data, we find that some further refinements are required to give reasonable localization result. First of all, Bd-related baselines are excluded. In this work, the delay search range and the related FFT size is configured for projected baselines shorter than 3000 km. Bd-related baselines exceed this limit. We can still detect single pulses with these baselines. However, due to the ambiguity of the delay search range, the fitted residual delay cannot be used for localization. Second, baselines with lower S/Ns are excluded. For localization, visibilities that contain the single pulse signal are extracted and fringe fitted again by following the normalization and fitting procedure used in HOPS. Besides a more accurate fitting result, the S/N is
derived accordingly. We take a threshold of S/N > 6, which is similar as that adopted in geodetic VLBI solving. For each single pulse, we exclude baselines that do not fulfill the above two requirements. Those with at least three baselines are used for localization.

We have to point out that one additional data point (OF3, Scan 40, 31.996 s, Mc-Sr baseline), is excluded from position solving. Although its S/N is high enough, the fitting result is wrong (probably due to RFI). The position of the corresponding single pulse solved by including this baseline deviates significantly from that of other single pulses. Actually this feature can be used to identify the incorrect baseline fitting result:

1. For the given single pulse, carry out solving several times, and each time exclude one baseline.
2. Check the solving result; the solution that includes the incorrect baseline fitting result deviates significantly from other solutions.

Figure 7 demonstrates the localization result. By placing the pulsar in the FoV with different offsets to the pointing center, we may investigate the corresponding localization accuracy. Clearly larger offset yields larger discrepancy and scatter. This can be explained with the underlying algorithm of geodetic VLBI solving: the linear relation between the position offset and the delay is valid only if the target is close enough to a priori position (small offset). As the offset becomes large, the linear approximation is not enough to describe its relation with delay, which leads to the large discrepancy. Meanwhile, scatters are amplified correspondingly. Also note that in EL060, the offset is always along the decl. direction, thus leading to the systemic trends of discrepancy. In Figure 7, we present the localization result with and without fine calibration. A systematic shift is clearly observed between the two data sets. With calibration, the localization result derived with P0 as the pointing center gets closer to the reference position, which is consistent with our analysis: small offset yields small discrepancy. This also proves that although simple and therefore easy to implement, our fine calibration scheme is effective in improving the localization accuracy. In summary, the localization discrepancy is around 1 arcsec when the target appears at the edge of the FoV, and quickly goes to well within 100 mas at the FoV center.

3.2. RRAT Observation

In this part, we observe RRAT source J1819–1458 and J1854+0306, so as to verify the DM search capability of the pipeline. However, in the 80 minutes observation of 2 RRAT sources, no single pulse is detected. According to the RRAT catalog,9 the flux and width of the RRAT sources, in particular J1854+0306, are quite similar with that of pulsar J0332+5434. It is unreasonable that all single pulses are missed. Our first suspicion is the newly developed pipeline. Therefore, we extract the autospectrum of each station from the visibility output of the correlator, convert them to the filterbank format, and carry out single pulse search with STEP. However, no single pulse is detected. The STEP transient search pipeline (Z. Xu et al. 2021, in preparation) is developed in SHAO and tested using the ASKAP FRB data (Shannon & Bannister 2018). In particular, since Ef and Sr are larger than other telescopes, higher sensitivities are expected. Therefore, we further inspect the data by eyes and still do not detect a single pulse.

By looking up references, one may find that the detected burst rate of RRAT source strongly depends on telescope sensitivity, e.g., the L-band burst rates of J1854+0306 are 8.9, 84, and 102 ± 10 per hour for Parkes (Keane et al. 2011), Arecibo (Deneva et al. 2009), and FAST (Lu et al. 2019), respectively. Moreover, the intensity of single pulse varies significantly with time. In some epochs, no single pulse is ever detected (Keane et al. 2011). Note that the size of Ef is larger than that of Parkes, we may expect a similar or higher detection rate. Paragi (2016) do report the detection of single pulse in the e-EVN data of J1819–1458, which is the first target of EL060. However, in our observation the Ef only observed the second target J1854+0306 for 25 minutes. No single pulse is detected in that short session. One thing we have to admit is, the possibility that the failed detection is caused by the VLBI mode observation still cannot be excluded: in our current treatment,

Table 2

| Pointing center | Offset | Dec    | Data length | Expected | Detected | Invalid |
|-----------------|--------|--------|-------------|----------|----------|---------|
| P0              | 0.0    | 54°34′43″57 | 340 s      | 476      | 345 (72.5%) | 1 (0.3%) |
| OF1             | 0.2    | +1°56′02″   | 340 s      | 476      | 424 (89.1%) | 16 (3.6%) |
| OF2             | 0.5    | +4°50′06″   | 360 s      | 505      | 406 (80.7%) | 0 (0.0%) |
| OF3             | 0.9    | +8°42′11″   | 306 s      | 429      | 388 (90.4%) | 0 (0.0%) |

Note. “Offset” is in the fraction of FoV radius. The R.A. of the 4 point centers is set to 03h32m59.5368 and is kept unchanged.
raw data is recorded with VLBI backend. The single pulse detection is based on the correlated autospectrum. Direct recording in pulsar mode might yield a better result.

4. Conclusion

In this work, we introduce the VOLK2 pipeline, which takes the idea of geodetic VLBI fringe fitting and carries out single pulse search in the VLBI cross spectrum. In the VOLKS2 pipeline, single pulses are identified by fully utilizing the cross-spectrum fringe phase information. By filtering candidate signals with multiple window sizes inside one baseline and by cross matching with multiple baselines, RFIs are eliminated effectively. Compared with station based single pulse search cross matching with multiple baselines, the VOLKS2 pipeline does not require extra RFI flagging.

Once the single pulse is detected, its position could be derived by geodetic VLBI solving with reasonable accuracy. The pipeline is designed for the transient search and localization in regular VLBI observations, as well as single-pulse detections with a known target (repeating FRBs, pulsars, etc.) in dedicated VLBI observations.

To verify the pipeline and the underlying cross spectrum based single pulse search method, we carry out EVN observation EL060. By placing the target pulsar in the FoV with different offsets to the FoV center, the detection efficiency and localization accuracy are investigated. The variation could be explained with the flux change caused by interstellar scintillation. Moreover, a higher localization accuracy is achieved when the single pulse appears in the FoV center.

By utilizing fine calibration proposed in this work, a localization accuracy of better than 100 mas is achieved. Also in this observation, no single pulse from RRAT sources is detected. We postulate this is because the telescopes are not sensitive enough to typical RRAT pulses. Therefore some further investigation is still required.

To fully utilize the hardware resources of modern GPU clusters, the pipeline is parallelized with MPI. Moreover, some steps in the fringe fitting part are accelerated with GPU. Depending on the specific search requirement and hardware configuration, the pipeline achieves a speedup from 7 to 23.7 X. One thing we want to point out is, some unique steps, “FFT” and “Find max,” of our cross spectrum based pipeline achieves remarkable speedup (>100 X) with the use of GPU, which makes their time consumption negligible compared with that of disk read and dedispersion. This suggest that with proper optimization, the cross spectrum based pipeline could run as fast as the popular autospectrum-based pipeline. All the code and document are available in GitHub repository. We hope the VOLKS2 pipeline is useful for radio transient studies.

L.L. would like to thank Dr. Wu Jiang for the kind support in the preparing and processing of the observation data. This work is supported by the National Natural Science Foundation of China (grant Nos. 11903067, 12041301, U2031119, 11973011), CAS Key Technology Talent Program, Shanghai Leading Talents program, National Basic Public Science Data Center “Radio astronomy and deep-space exploration data-base” (NBSDC-DB-11), and the Natural Science Foundation of Shanghai (grant No. 20ZR1467600). The European VLBI Network (www.evlbi.org) is a joint facility of independent European, African, Asian, and North American radio astronomy institutes. Scientific results from data presented in this publication are derived from the following EVN project code: EL060.

ORCID iDs

Lei Liu https://orcid.org/0000-0002-2920-1880
Zhijun Xu https://orcid.org/0000-0003-4853-7619
Zhen Yan https://orcid.org/0000-0002-9322-9319

References

Bochenek, C. D., Ravi, V., Belov, K. V., et al. 2020, Natur, 587, 59
Caleb, M., Flynn, C., Bailes, M., et al. 2016, MNARS, 458, 718
Chatterjee, S., Law, C. J., Wharton, R. S., et al. 2017, Natur, 541, 58
CHIME/FRB Collaboration, Andersson, B. C., Bandura, K. M., et al. 2020, Natur, 587, 54
Clery, D. 2020, Sci, 370, 274
Deller, A. T., Brinken, W. F., Phillips, C. J., et al. 2011, PASP, 123, 275
Deneva, J. S., Cordes, J. M., McLaughlin, M. A., et al. 2009, ApJ, 703, 2259
Keane, E. F., Kramer, M., Lyne, A. G., et al. 2011, MNARS, 415, 3065
Law, C. J., Bower, G. C., Burke-Spolaor, S., et al. 2015, ApJ, 807, 16
Li, C. K., Liu, L., Xiong, S. L., et al. 2021, NatAs, 5, 378
Liu, L. 2021, VOLKS2: VLBI Observation for single pulse localization Keen Searcher, 2nd release, v0.7, Zenodo doi:10.5281/zenodo.5168951
Liu, L., Jiang, W., Zheng, W., Yan, Z., Zhang, J., Ma, M., & Luo, W. 2019, AJ, 157, 136
Liu, L., Tong, F., Zheng, W., Zhang, J., & Tong, L. 2018a, AJ, 155, 98
Liu, L., Zheng, W., Yan, Z., & Zhang, J. 2018b, RAA, 18, 069
Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, Sci, 318, 777
Lu, J., Peng, B., Liu, K., et al. 2019, SCPMA, 62, 95903
Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, AJ, 129, 1993
Marcote, B., Paragi, Z., Hessels, J. W. T., et al. 2017, ApJL, 834, L8
McLaughlin, M. A., Rea, N., Gaensler, B. M., et al. 2007, ApJ, 670, 1307
Mereghetti, S., Savchenko, V., Ferrigno, C., et al. 2020, ApJL, 898, L29
Paragi, Z. 2016, arXiv:1612.00508
Petrochuk, W., Behrend, D., Hase, H., et al. 2013, EGU General Assembly Conf. Abstracts, 15, EGU2013-12867
Prabu, S., Hancock, P., Zhang, X., et al. 2020, PASA, 37, e052
Ransom, S. 2011, Astrophysics Source Code Library, ascl:1107.017
Shannon, R., & Bannister, K. 2018, Data from the ASKAP latitude 50 Fast Radio Burst (FRB) sample. v3. CSIRO. Data Collection
The CHIME/FRB Collaboration, Amiri, M., Andersen, B. C., et al. 2021, arXiv:2106.04352
The CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2018, ApJ, 865, 48

L.L. would like to thank Dr. Wu Jiang for the kind support in the preparing and processing of the observation data. This work is supported by the National Natural Science Foundation of China (grant Nos. 11903067, 12041301, U2031119, 11973011), CAS Key Technology Talent Program, Shanghai Leading Talents program, National Basic Public Science Data Center “Radio astronomy and deep-space exploration data-base” (NBSDC-DB-11), and the Natural Science Foundation of Shanghai (grant No. 20ZR1467600). The European VLBI Network (www.evlbi.org) is a joint facility of independent European, African, Asian, and North American radio astronomy institutes. Scientific results from data presented in this publication are derived from the following EVN project code: EL060.

https://github.com/liulei/volks2