

Fast radio burst search: cross spectrum vs. auto spectrum method

Lei Liu^{1,2,3}, Weimin Zheng^{1,2,3}, Zhen Yan^{1,2} and Juan Zhang^{1,3,2}

¹ Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China; liulei@shao.ac.cn

² Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210008, China

³ Shanghai Key Laboratory of Space Navigation and Positioning Techniques, Shanghai 200030, China

Received 2018 January 23; accepted 2018 March 15

Abstract The search for fast radio bursts (FRBs) is a hot topic in current radio astronomy studies. In this work, we carry out a single pulse search with a very long baseline interferometry (VLBI) pulsar observation data set using both auto spectrum and cross spectrum search methods. The cross spectrum method, first proposed in Liu et al., maximizes the signal power by fully utilizing the fringe phase information of the baseline cross spectrum. The auto spectrum search method is based on the popular pulsar software package PRESTO, which extracts single pulses from the auto spectrum of each station. According to our comparison, the cross spectrum method is able to enhance the signal power and therefore extract single pulses from data contaminated by high levels of radio frequency interference (RFI), which makes it possible to carry out a search for FRBs in regular VLBI observations when RFI is present.

Key words: techniques: interferometric — radio continuum: general — methods: data analysis — pulsars: general

1 INTRODUCTION

A fast radio burst is a kind of high flux radio burst that is characterized by its high dispersion measure (DM) and millisecond duration. It was first reported by Lorimer et al. (2007). Until now, about 20 such events have been discovered with large single dish telescopes (Thornton et al. 2013; Spitler et al. 2016) and specially designed interferometers (Caleb et al. 2016). Current studies can almost confirm their extragalactic origin. However, their burst mechanism is still not clear. According to Katz (2016), non-repeating and repeating bursts might have different origins.

One big challenge in FRB studies is their precise localization, which is extremely important for discovering their possible afterglows and background counterparts in multiple wavelengths. It is expected that various kinds of high angular resolution interferometers, e.g., UTMOST (Caleb et al. 2016) and CHIME (Ng et al. 2017), will be the main facilities for FRB search in the near future. In addition, very long baseline interferometry (VLBI, Thompson et al. 2001), as the astronomical technique that achieves the highest angular resolution, has been used in the direct localization of FRB events. The joint

observation with VLA, Arecibo, EVN and instruments in other wavelengths has revealed the precise localization of the repeating burst FRB 121102 and detected its possible counterpart in radio and optical bands (Chatterjee et al. 2017; Marcote et al. 2017; Tendulkar et al. 2017). Astronomers also try to carry out FRB searches with legacy VLBI raw data and ongoing VLBI observations, e.g. the V-FASTR project with VLBA (Wayth et al. 2011; Thompson et al. 2011) and the LOCAtE project with EVN (Paragi 2016).

In general, there are three kinds of VLBI observation data: astrophysical, geodetic and data used for deep space exploration. Most of them, if not all, can be applied for FRB search. Because of expensive storage, most of these raw data will be deleted immediately after correlation. For us, these data are precious and deserve further investigation. Our plan is to develop a pipeline to carry out FRB search before data deletion. Initially, we chose the popular auto spectrum based single pulse search algorithm provided by PRESTO (Ransom 2001). However, soon we realized that the auto spectrum method did not work when radio frequency interference (RFI) was present. To fully exploit such kind of data, we have developed a new method. In Liu et al. (2018), we present a

cross spectrum based single pulse search method. It utilizes the fringe phase information of the baseline cross spectrum, so as to maximize the power of single pulse signals. We will introduce this method in Section 2.1.

To evaluate the performance of both auto spectrum and cross spectrum based single pulse detection methods, we have carried out a single pulse search on a VLBI pulsar observation data set using both methods. The advantage of using pulsar data is the arrival time (pulse phase) of a pulsar signal is predictable, which makes it possible to differentiate if a single pulse is a pulsar signal or not.

This paper is organized as follows: In Section 2, we introduce the auto and cross spectrum based single pulse search methods. In Section 3, we present the single pulse detection result using both methods. In Section 4, we summarize the whole work.

2 THE CROSS SPECTRUM AND AUTO SPECTRUM BASED METHODS

2.1 Cross Spectrum Method

The cross spectrum based single pulse search method was first proposed in Liu et al. (2018). It takes the idea of fringe fitting in geodetic VLBI data postprocessing, which fully utilizes the fringe phase information to maximize the signal power (Takahashi 2000; Cappallo 2014). We make special optimizations for the original fringe fitting scheme, so as to achieve higher performance and signal power with a cross spectrum of millisecond duration. The method itself is fully described in Liu et al. (2018). Below we give a brief summary:

- (a) VLBI correlation of raw data. It is recommended that the station clocks are well adjusted, so that the residual delay is limited to one sample period and the fringe rate is within 10^{-2} Hz. The accumulation period (AP) of output cross spectrum should be sufficiently small, e.g., 1 millisecond, so as to resolve a typical FRB.
- (b) Dedispersion and construction of time segments. In the cross spectrum method, we carry out incoherent dedispersion on the cross spectrum with millisecond duration. Then several such kinds of dedispersed cross spectra are combined to construct a time segment with different window sizes (APs). After this step, several lists of time segments with different window sizes are constructed.
- (c) Fringe fitting. For each time segment, we find out the specific multi band delay and single band delay that maximize the delay resolution function. In the actual implementation, we use a 2D fast Fourier transform (FFT) to speed up the search process.

- (d) Single pulse extraction on one baseline. For each time segment list with different window sizes on one baseline, after fringe fitting, signal powers are normalized according to power fluctuation; then single pulses are extracted according to a given threshold. After that these single pulses are filtered in multiple windows to further exclude RFIs. In the current scheme, single pulses that are detected on at least three windows are selected as candidate signals.
- (e) Cross matching candidate signals from multiple baselines.

2.2 Auto Spectrum Method

The famous pulsar search software package PRESTO provides support for auto spectrum based single pulse search. The whole process can be divided into several steps:

- (a) For each station, carry out incoherent dedispersion on the input auto spectrum.
- (b) Subdivide the auto spectrum into small pieces with given time duration. For each piece, remove the trend and normalize the spectrum with standard deviation; smooth the sample points with multiple down factors.
- (c) For sample points in each down factor, pick up candidate signals according to the given threshold.
- (d) Walk through the candidate lists of different down factors and remove candidates that are close to other candidates but are less significant.
- (e) Cross match candidate signals detected from multiple stations. Two candidate signals are assumed to match if their time range overlaps with each other.

The algorithm of auto spectrum based search method is simple and easy to implement. Therefore it is widely used in various kinds of FRB search projects. By cross matching candidate signals from multiple stations, a significant amount of RFI can be excluded. However, one big disadvantage of this method is no valid single pulse signals can be extracted from the corresponding station when the radio interference is strong or the sensitivity of the station is low. This is clearly demonstrated in Section 3.

One thing we do not mention is the DM search scheme. For both methods, we have to divide the target DM search range into several DM bins, and carry out single pulse search in each of these DM bins. For the auto spectrum method, there is an optimized DM search scheme provided by PRESTO. For the cross spectrum method, the bin width is determined by both the window size and frequency range as proposed in Liu et al.

(2018). This work does not involve the DM search. The main reason is the DM of the pulsar data set (described in the next section) is just $26.833 \text{ pc cm}^{-3}$, which is too low to carry out effective DM search.

3 COMPARISON OF DETECTION RESULTS

3.1 Pulsar Data Set

The VLBI pulsar data set used in this work is taken from the Chinese VLBI Network (CVN, Zheng 2015) observation of pulsar PSR J0332+5434 (Chen et al. 2015). The three CVN telescopes, Shanghai (Sh), Kunming (Km) and Urumqi (Ur), took part in the observation. The system equivalent flux density (SEFD) values of the three telescopes are 800 Jy, 350 Jy and 560 Jy, respectively. The target source, PSR J0332+5434, is one of the brightest pulsars ever identified. The average flux is around 0.1 Jy at S band (Kramer et al. 2003). According to the ATNF Pulsar Catalogue (Manchester et al. 2005), the DM value is $26.833 \text{ pc cm}^{-3}$ and the period is 0.714 s. The 96 MHz observation bandwidth in S band (2192 MHz - 2288 MHz) is equally divided into six 16 MHz frequency channels. For correlation, we use 64 points in the FFT, which corresponds to 32 frequency points in each frequency channel. The observation was carried out on 2015 Feb 15 and lasted for 12 hours. In this work, we use pulsar observation scans 69, 71 and 73 for single pulse search. Since the starting and ending times of the raw data are different for each station and scan, to keep consistency, for each scan, we use the data between 10 s and 170 s.

The three panels in Figure 1 display the folding profiles of PSR J0332+5434 acquired from the three stations. To obtain a profile, we first carry out time shifts on the raw data, so that data from the three stations are in the same geocentric reference frame. Then those data are Fourier transformed to the frequency domain. We calculate the pulse phase for each frequency point and assign it to the corresponding pulse phase bin. Usually, the profile appears after enough time has elapsed for accumulation. The Km and Ur panels show a clear pulse profile. In contrast, the strong 99.9475 Hz RFI makes it impossible to extract any valid pulsar signal from Sh station. The peak in Km station is higher, which corresponds to its higher sensitivity (low SEFD). The pulsar phase ranges for the two stations are almost overlapped with each other. According to Liu et al. (2018), we set it to 0.973–0.983. A single pulse is assumed to be a “high probability pulsar signal” if its time range is overlapped with the pulsar phase range. We have to point out that the pulse phase information itself cannot exclude the possi-

bility of false detection. However, it is still a good criterion to distinguish pulsar signals since single pulses outside this phase range are definitely RFI.

3.2 Detection Results

In this section, we present the single pulse detection results using both the cross spectrum and auto spectrum methods.

For the cross spectrum method, we use the CVN software correlator (Zheng et al. 2010) for VLBI correlation. The output AP is set to 1.024 ms. For fringe fitting, we choose 3C 273 in scan 293 as the calibration source. For multiple window filtering as described in step d in Section 2.1, we choose the window lengths of 4, 8, 16, 24 and 32 APs.

The single pulse detection result is presented in Table 1. We define the detection accuracy as the fraction of high probability pulsar signals among all the detected signals. From the table, the Km-Ur baseline yields the highest detection accuracies and the largest number of high probability pulsar signals, which is consistent with its high sensitivity. In contrast, the detection accuracy of baselines related to Sh station is much lower, which is due to the strong surrounding RFI. In Liu et al. (2018), we also present the multiple baseline cross matching result. Single pulses detected simultaneously on two or three baselines can almost exclude the possibility of false detection.

Table 1 Cross spectrum search results. The numbers in parentheses correspond to single pulses for which the pulse time range is overlapped with the pulsar phase range (high probability pulsar signal).

Scan No.	Cross spectrum		
	Sh-Km	Sh-Ur	Km-Ur
69	37 (12)	33 (2)	49 (40)
71	26 (8)	35 (3)	57 (41)
73	29 (7)	34 (4)	51 (36)

For the auto spectrum method, we first convert the Mark5b (Whitney 2003) format raw VLBI observation data to the filterbank format which is readable by PRESTO. Raw data are time shifted according to delay models, so that the filterbank data and the VLBI cross spectrum output are in the same geocentric reference frame. Filterbank files are generated for scans 69, 71 and 73 from Sh, Km and Ur stations respectively. Parameters for these filterbank files are listed in Table 2.

For single pulse detection with PRESTO, we set a detection threshold of 3 (defaults to 5 in the original pro-

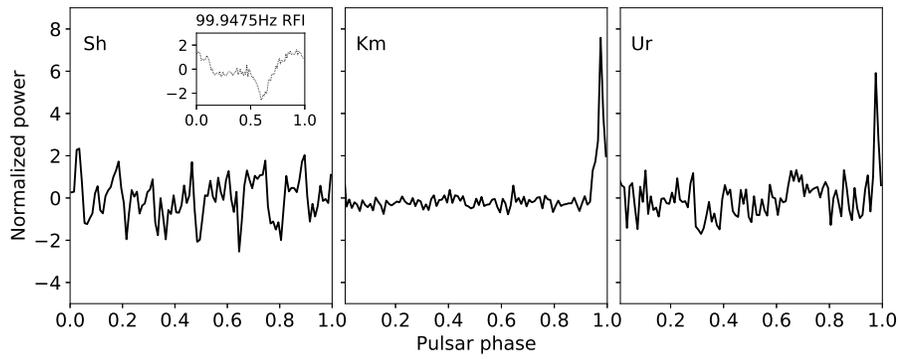


Fig. 1 Pulse profile of PSR J0332+5434. The profiles are derived by folding the data between 10 s and 170 s in scan 73 of CVN observation psrf02. The *thin dotted line* in the Sh panel corresponds to 99.9475 Hz RFI at Sh station. According to Liu et al. (2018), the peaks in the Km and Ur profiles correspond to a pulsar phase range from 0.973 to 0.983.

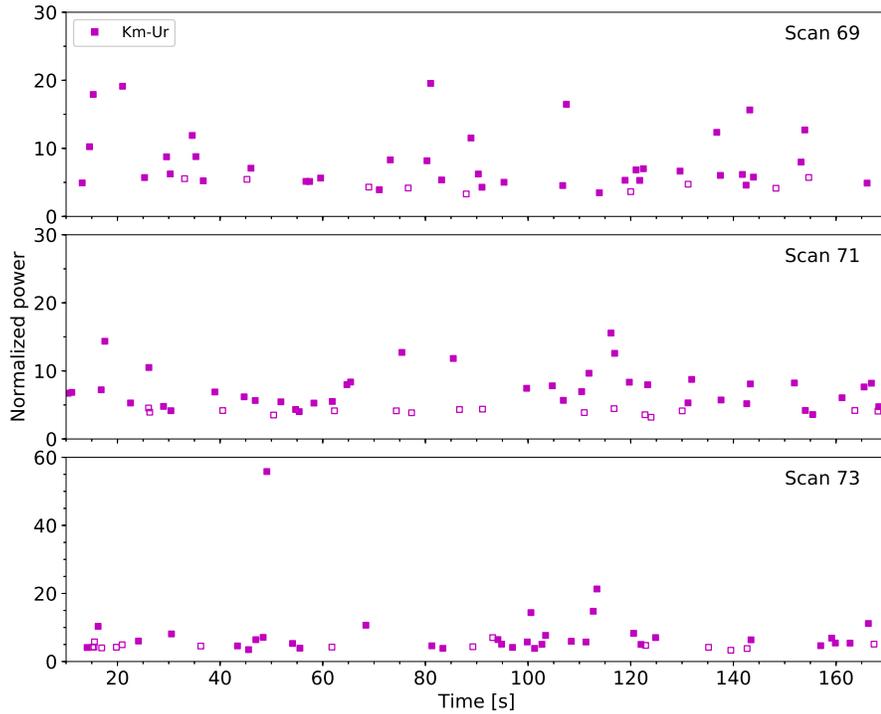


Fig. 2 Cross spectrum detection results of the Km-Ur baseline. *Filled and empty squares* correspond to high probability pulsar signals and false detections, respectively. The “normalized power” is defined as the signal power subtracted by the average level and then normalized with the standard deviation (Liu et al. 2018). The total numbers of detected signals in the three scans are 49, 57 and 51 from top to bottom respectively. The corresponding numbers of high probability pulsar signals are 40, 41 and 36.

gram) and a maximum downfactor of 490¹. Because of the strong RFI at Sh station, it is impossible to detect any real signal. Therefore, we only present the single pulse detection results for Km and Ur stations.

The Km and Ur single pulse search and two station cross matching result are presented in Table 3. For

¹ In the program, the maximum supported value is 300. We modify it to 490 to yield a maximum width of 31.36 ms, so that it is comparable with the maximum window length of 32.768 ms for cross spectrum search in Section 2.1.

each station, a large number of single pulses is detected. However, the detection accuracy is just slightly higher than 1%, which means most of the detected signals are RFIs. By cross matching detection results from the two stations, detection accuracy becomes higher. As a comparison, the cross spectrum method identifies more high probability pulsar signals with much higher detection accuracy, which demonstrates that the cross spectrum based method is better at extracting single pulses from RFI contaminated data.

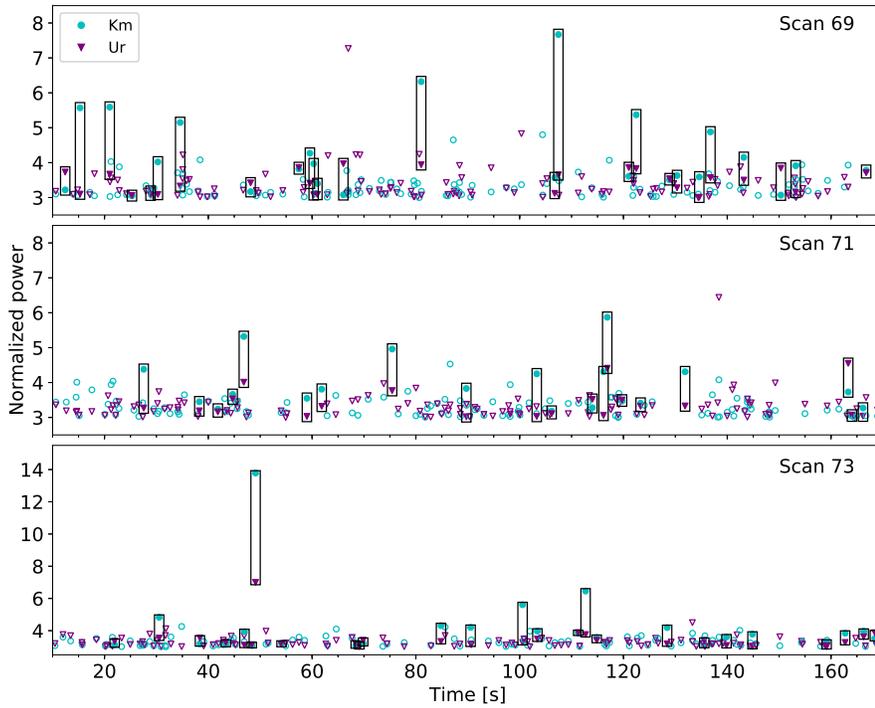


Fig. 3 Cross matching result of Km and Ur stations with the auto spectrum method. Single pulses presented in the figure are detected simultaneously (time ranges are overlapped with each other) by two stations. *Filled* and *empty* symbols correspond to high probability pulsar signals and false detections, respectively. For clarity, high probability pulsar signals are enclosed with *black rectangular boxes*. The “normalized power” is defined as the detrended signal power normalized with the standard deviation, as proposed in the PRESTO package. The total numbers of cross matched signals in the three scans are 140, 136 and 129 from top to bottom respectively. The corresponding numbers of high probability pulsar signals are 26, 20 and 26.

Table 2 Parameter settings for filterbank files. Low and high channel frequencies correspond to the frequency in the middle of the respective channel. In the filterbank format, file time must be divisible by subint time, therefore it is slightly shorter than 160 s.

Parameter	Setting
Sample time	64 μ s
Low channel	2192.5 MHz
High channel	2287.5 MHz
Channel width	1 MHz
Channel number	96
Spectra per subint	2400
Spectra per file	2498400
Time per subint	0.1536 s
Time per file	159.8976 s
Sample bits	8

Figures 2 and 3 present the Km-Ur baseline detection result with the cross spectrum method and the cross matching result of Km and Ur stations with the auto spectrum method. For the auto spectrum result, the signal powers of Km station are usually higher than those of Ur station, which is consistent with their sensitivity. By comparing the two figures, we may find that the nor-

Table 3 Auto spectrum search results. For comparison, the single pulse detection result of Km-Ur baseline (cross spectrum) is also presented. The numbers in parentheses correspond to single pulses for which the pulse time range is overlapped with the pulsar phase range (high probability pulsar signal).

Scan No.	Auto spectrum			Cross spectrum
	Km	Ur	Cross matching	Km-Ur
69	6334 (151)	6549 (87)	140 (26)	49 (40)
71	6296 (129)	6390 (91)	136 (20)	57 (41)
73	6355 (141)	6483 (94)	129 (26)	51 (36)

malized powers of the cross spectrum result are usually higher than those of the auto spectrum result. This is because the cross spectrum method fully utilizes the cross spectrum fringe phase information, which enhances the signal power. By utilizing this feature, the cross spectrum method is able to extract more single pulses with higher accuracy.

4 SUMMARY

In this work, we present the single pulse detection result on a VLBI pulsar observation data set using both cross spectrum and auto spectrum methods.

Compared with the auto spectrum method, the cross spectrum method is able to extract more signal pulses with higher detection accuracy. The signal power of the cross spectrum method is higher than that of the auto spectrum method, which leads to a higher confidence level. The cross spectrum method is able to extract single pulses from highly RFI contaminated data. According to the comparison, we may find that the cross spectrum method makes it possible to carry out FRB search in VLBI observation with low sensitivity telescopes, even when RFI is present.

Due to the limitation of currently available data, our comparisons are only limited to a low DM environment and do not involve DM search. It has been demonstrated that the auto spectrum method is very effective at excluding RFIs by a large number of DM trials. We still have to verify the performance of the cross spectrum method in a high DM environment. To obtain a high DM data set, a VLBI observation of a Rotating Radio Transit (RRAT, McLaughlin et al. 2006) source is already in our plan. One possible choice is J1819–1458, the DM value of which is 196 pc cm^{-3} with flux 3.6 Jy at 1.4 GHz (Keane et al. 2011). We will present the cross spectrum method single pulse search result with this source in our future work.

Acknowledgements This work was funded by the National Natural Science Foundation of China (Grant Nos. 11373061, 11573057 and U1631122), the Key Laboratory of Radio Astronomy of Chinese Academy of Sciences (CAS), Shanghai Key Laboratory of Space Navigation and Positioning Techniques (ZZXT-201702), the CAS Key Technology Talent Program and the National R&D Infrastructure and Facility Development Program of China, “Fundamental Science Data Sharing Platform” (DKA2017-12-02-09). The authors appreciate the support of the CVN data processing center.

References

- Caleb, M., Flynn, C., Bailes, M., et al. 2016, *MNRAS*, 458, 718
- Cappallo, R. 2014, in *International VLBI Service for Geodesy and Astrometry 2014 General Meeting Proceedings: “VGOS: The New VLBI Network”*, eds. Dirk Behrend, Karen D. Baver, Kyla L. Armstrong, 91 (Beijing: Science Press)
- Chatterjee, S., Law, C. J., Wharton, R. S., et al. 2017, *Nature*, 541, 58
- Chen, W., Jiang, W., Li, Z. X., Xu, Y. H., & Wang, M. 2015, *Acta Astronomica Sinica*, 56, 658
- Katz, J. I. 2016, *Modern Physics Letters A*, 31, 1630013
- Keane, E. F., Kramer, M., Lyne, A. G., Stappers, B. W., & McLaughlin, M. A. 2011, *MNRAS*, 415, 3065
- Kramer, M., Karastergiou, A., Gupta, Y., et al. 2003, *A&A*, 407, 655
- Liu, L., Tong, F., Zheng, W., Zhang, J., & Tong, L. 2018, *AJ*, 155, 98
- Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, *Science*, 318, 777
- 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, *ApJ*, 834, L8
- McLaughlin, M. A., Lyne, A. G., Lorimer, D. R., et al. 2006, *Nature*, 439, 817
- Ng, C., Vanderlinde, K., Paradise, A., et al. 2017, arXiv:1702.04728
- Paragi, Z. 2016, arXiv:1612.00508
- Ransom, S. M. 2001, *New Search Techniques for Binary Pulsars*, PhD thesis, Harvard University
- Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, *Nature*, 531, 202
- Takahashi, F., ed. 2000, *Very Long Baseline Interferometer*, in: *International VLBI Service for Geodesy and Astrometry 2000 General Meeting Proceeding*
- Tendulkar, S. P., Bassa, C. G., Cordes, J. M., et al. 2017, *ApJ*, 834, L7
- Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2001, *Interferometry and Synthesis in Radio Astronomy*, (2nd edn.; New York: Wiley)
- Thompson, D. R., Wagstaff, K. L., Brisken, W. F., et al. 2011, *ApJ*, 735, 98
- Thornton, D., Stappers, B., Bailes, M., et al. 2013, *Science*, 341, 53
- Wayth, R. B., Brisken, W. F., Deller, A. T., et al. 2011, *ApJ*, 735, 97
- Whitney, A. R. 2003, in *Astronomical Society of the Pacific Conference Series*, 306, *New Technologies in VLBI*, ed. Y. C. Minh, 123
- Zheng, W. 2015, *IAU General Assembly*, 22, 2255896
- Zheng, W., Quan, Y., Shu, F., et al. 2010, in *Sixth International VLBI Service for Geodesy and Astronomy, Proceedings from the 2010 General Meeting, “VLBI2010: From Vision to Reality”*, eds. D. Behrend and K. D. Baver, *NASA/CSP 2010–215864*, 157