Mitigation Methods for Radio Frequency Interference in Fast Radio Burst Observational Data: A Postprint

Authors:

Date: 2024-06-03T00:00:00+00:00

Abstract

Rapidly searching for rare Fast Radio Burst (FRB) events from massive astronomical observation datasets, interference mitigation represents a critical and challenging task. Radio Frequency Interference (RFI) can overwhelm genuine astronomical events and also lead search algorithms to produce numerous false-positive candidates. Due to the complexity of interference sources and varieties, there is currently no universal solution to this problem. To mitigate the impact of interference on FRB searches, we analyzed and investigated the interference characteristics in L-band observations from the Nanshan 26;m radio telescope, establishing a three-tiered interference mitigation processing pipeline targeting both narrowband and broadband interference, which effectively reduced interference contamination in the observational data. Integrating this pipeline into the FRB Dispersed Dynamic Spectra Search (DDSS) pipeline, experimental results demonstrated that both the detection rate and accuracy of the search algorithm were significantly enhanced. This approach offers valuable insights for interference mitigation in FRB observational data.

Full Text

Radio Frequency Interference Mitigation Methods for Fast Radio Burst Observational Data

LIU Yan-ling[1;2;3], **CHEN Mao-zheng**[1;2;3], **YUAN Jian-ping**[1;2], **LI Jian**[1;3], **YAN Hao**[1;2;3], **ZHAI Nan-nan**[1;2]

(1 Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011)

(2 University of Chinese Academy of Sciences, Beijing 100049)

(3 Xinjiang Key Laboratory of Microwave Technology, Urumqi 830011)

Abstract

Searching for rare fast radio burst (FRB) events from massive astronomical observational datasets requires effective radio frequency interference (RFI) mitigation, which remains a crucial yet challenging task. RFI can drown out genuine astronomical signals and generate numerous false-positive candidates in search pipelines. Due to the complexity of interference sources and types, no universal solution currently exists. To reduce RFI impact on FRB searches, we analyzed interference characteristics in L-band observations from the Nanshan 26 m radio telescope and developed a three-level RFI mitigation processing workflow targeting dominant narrowband and broadband interference. This approach effectively alleviates data contamination. When embedded into the FRB Dispersed Dynamic Spectra Search (DDSS) pipeline, experimental results demonstrate improved detection rates and accuracy. This method provides a valuable reference for RFI mitigation in FRB observational data.

Key words: radio continuum: transients, radio frequency interference (RFI), methods: data analysis

Radio astronomy relies on radio telescopes to passively receive extremely faint cosmic radio signals. The received electromagnetic signals are extraordinarily weak, approximately 150–220 $\mathrm{dB \cdot W \cdot m^{-2}}$. For example, the strongest radio sources exhibit a power spectral density of about 10^5 Jy at 300 MHz (1 Jy = 10^{-26} $\mathrm{W \cdot m^{-2} \cdot Hz^{-1}}$) [1]. Consequently, radio astronomical observations are highly vulnerable to anthropogenic radio interference from mobile phones, satellites, aircraft, broadcast television, and electronic equipment within observatory premises [2]. Furthermore, advances in radio astronomy technology have continuously improved receiver sensitivity and bandwidth, often exceeding protected radio astronomy spectrum allocations. Simultaneously, the proliferation of digital receivers, backends, commercial equipment, electrical devices, various cables, and aerospace satellite communications has complicated the electromagnetic environment at radio observatories. RFI mitigation will remain one of the primary challenges in present and future radio astronomy. Single-dish observations are particularly susceptible to interference because they offer little discrimination against it, allowing both astronomical and interfering signals to enter the system [3].

Fast radio bursts (FRBs) are bright, millisecond-duration transient radio phenomena [4–6]. Unlike pulsars with their short, stable periodic pulses, capturing rare FRBs presents considerably greater difficulty. In radio transient and pulsar searches, ideal data consist of weak astrophysical signals superimposed on uncorrelated Gaussian noise. Interference complicates this scenario by submerging FRB signals, reducing signal-to-noise ratios, and causing detection algorithms to miss astronomical events. Additionally, interference triggers search algorithms to generate false candidate samples, increasing the difficulty of subsequent screening. Therefore, RFI mitigation is essential for FRB search observations.

RFI mitigation has long accompanied radio astronomical research. Although observatories are generally built far from urban centers, urbanization and expanding human activities continuously degrade site electromagnetic environments. Figure 1 [Figure 1: see original paper] illustrates a schematic framework of electromagnetic interference sources and mitigation strategies. With the development of space radio technology and increasing on-site electronic equipment, management efforts such as establishing radio-quiet zones and controlling observatory electronic devices alone cannot eliminate RFI effects on astronomical data. Consequently, mitigation techniques must be implemented in receiver chains and data post-processing stages. In analog RF front-ends, appropriate filters remove strong interfering signals that could saturate amplifiers [7-8]. In digital domains, both real-time online processing and offline post-processing reduce weaker interference through time-domain excision [9–10], frequency-domain excision [11], spatial-domain excision [12–13], adaptive filtering [14–16], and machine learning [17]. Additionally, manual visual inspection for RFI marking remains common in fine data analysis. However, this approach becomes impractical with increasing data volumes, particularly for astrophysical transient search observations. Overall, no universal method can eliminate all interference; RFI mitigation in radio astronomy represents a complex engineering task.

Recent intensive observational studies of transient signals like FRBs have spurred increased research into RFI mitigation methods for single-pulse data. However, real-time processing requirements render many algorithms based on kurtosis, surface fitting, or Fourier transforms impractical due to computational complexity. Therefore, transient signal processing prioritizes simplicity and speed. Men et al. [18] proposed the Zero-DM matched filter (ZDMF) method for mitigating instantaneous broadband interference, applying it alongside fixed narrowband interference removal in their BEAR (Burst Emission Automatic Roger) search pipeline. Boyle et al. [19] presented an edge-threshold segmentation algorithm combined with time-domain threshold excision to reduce false-positive FRB detections. Sclocco et al. [20] developed the open-source RFIm library for time- and frequency-domain threshold excision to eliminate broadband and narrowband interference in transient observations. Morello et al. [21] developed the Inter-Quartile Range Mitigation (IQRM) algorithm for narrowband interference in astrophysical transient and pulsar search data. In summary, FRB observation data RFI mitigation requires comprehensive analysis of interference characteristics and the application of multiple methods. This also represents a primary means of distinguishing terrestrial anthropogenic interference from distant cosmic radio signals.

This paper analyzes interference characteristics in observational data from the FRB observation terminal at the Nanshan 26 m radio telescope of the Xinjiang Astronomical Observatory, Chinese Academy of Sciences. By designing a multi-level RFI mitigation strategy for data preprocessing, we provide high-quality observational data to FRB search pipelines, thereby improving detection rates and accuracy. Subsequent sections detail interference classification and characteristics, multi-level mitigation measures, experimental results, and applications.

2.1 FRB Signal Characteristics

Similar to single pulses from pulsars, when FRBs from distant space propagate through the intergalactic medium (IGM) and are received by ground-based telescopes, the radio pulses exhibit dispersion, scattering, and scintillation effects, causing pulse broadening, distortion, intensity variations, and frequency drift. Dispersion has the most significant impact, causing energy signals at different frequencies to arrive at the telescope at different times—specifically, higher-frequency signals arrive before lower-frequency ones. The arrival time delay due to dispersion is expressed as:

$$\Delta t = 4.15 \times 10^6 \times (\nu_1^{-2} - \nu_2^{-2}) \times \mathrm{DM}$$

where ν_1 and ν_2 are the central frequencies of two frequency channels in MHz, and DM (Dispersion Measure) represents the integrated electron density along the line of sight, with units of $\mathrm{cm}^{-3} \cdot \mathrm{pc}$, expressed as:

$$\mathrm{DM} = \int n_e \, dl$$

where n_e is electron number density, d is the FRB distance, and l is the line-of-sight path length.

The dispersion measure is an important astrophysical quantity obtained by measuring the frequency-dependent arrival time of celestial signals. Larger DM values indicate more distant sources, enabling distance estimation. In pulsar and FRB observations, dedispersion is the most important and characteristic technique.

2.2 RFI Signal Characteristics and Impact on FRB Searches

Interference can be classified in multiple ways: impulsive or continuous, narrowband or broadband, periodic or aperiodic, external or internal to the observatory. In actual observational data, the superposition of narrowband persistent interference and impulsive broadband interference is most common. FRBs are typically one-off, unpredictable, rarely repeating bursts. Among over 700 observed FRBs, only about 60 are repeaters, with most detected only twice. Although FRBs radiate enormous energy, they remain extremely weak compared to anthropogenic interference such as aircraft radar signals, mobile communications, satellite signals, and observatory electronic equipment emissions, making detection exceptionally difficult. Current challenges in FRB observation RFI mitigation include: (1) complex interference sources with diverse manifestations in data; (2) dynamic interference behavior showing significant variation across time and frequency; and (3) progressively worsening interference conditions over the long term. RFI mitigation research has continuously accompanied radio astronomy. For FRB observations, while minimizing RFI contamination is desirable, excessive excision risks data loss and potential removal of astronomical events. Additionally, overly complex mitigation algorithms in real-time searches increase pipeline processing time. Therefore, balancing accuracy and speed is essential when implementing RFI mitigation measures.

2.3 Characteristics of Perytons

Perytons, named after a mythical half-deer, half-bird creature with a human shadow, refer to ground-based radio interference signals that mimic FRBs in radio astronomy. The Parkes 64 m radio telescope observed such signals, later identified as interference from microwave ovens [22]. These signals exhibit dispersion characteristics that can mislead FRB detection algorithms. However, due to their low occurrence probability, current search algorithms do not specifically distinguish them from genuine FRB events, and this study does not address their differentiation further.

3. Main Sources and Classification of L-Band Interference at Nanshan

The Nanshan station of Xinjiang Astronomical Observatory, established in 1991, currently operates a 26 m radio telescope system, 40 cm precision photoelectric telescope, 1 m optical telescope, multiple GPS receivers, and a solar chromospheric telescope, with several important instruments under construction or planned. Expanding anthropogenic radio emissions will inevitably complicate the site’ s electromagnetic environment. As shown in Figure 1, interference sources at radio observatories divide into internal and external categories. Internal interference originates from digital clocks, power switches, relays, digital processing equipment, and residential electrical devices (wireless mice, TV set-top boxes, laptops, microwave ovens), characterized by irregular frequency and timing but relatively fixed locations. External interference includes mobile communications, TV broadcasting, aeronautical navigation, aeronautical mobile services, radar, satellites, and amateur radio, featuring relatively fixed transmission bands and higher power. Observational data often contain information from multiple interference sources, making visual identification of interference origins difficult based solely on signal morphology changes. The Nanshan L-band receiver primarily conducts pulsar and single-pulse searches, so this paper focuses on RFI mitigation strategies for L-band observational data. The Nanshan L-band receiver covers 1400–1720 MHz, with known external interference sources in this band listed in Table 1 .

Table 1 The sources of external RFI in the frequency range of 1400–1720 MHz [23]

RFI Source	Frequency/MHz	Characteristic
GPS	1227.6, 1575.42	Persistent
Inmarsat	1526-1559	Persistent
GLONASS	1598-1606	Persistent
Iridium	1618.85-1626.5	Persistent

From the perspective of human activity patterns, interference types, distribution, quantity, and intensity vary across observation periods. Local working hours in Urumqi run from 10:00 to 20:00 Beijing time, during which interference activity is high. From 20:00 to 10:00 the next day, interference is relatively reduced. Consequently, nighttime observational data contain less interference.

4.1 Observational Data Analysis

RFI mitigation for FRB search data aims to improve signal-to-noise ratios for easier FRB detection and reduce false-alarm probabilities by minimizing interference. Figure 2 [Figure 2: see original paper] shows a frequency-time intensity plot of L-band observations from the Nanshan 26 m radio telescope. Since the digital backend bandwidth exceeds the L-band receiver' s operating bandwidth, the saved data include signals outside the effective observation band, which can be removed during processing if necessary. The figure primarily contains narrowband persistent interference, multiple broadband impulsive interference events, and short-duration narrowband interference. While short-duration narrowband interference is not easily visible in Figure 2, it becomes apparent in the downsampled data after standardized processing described later. This paper implements hierarchical mitigation measures targeting these three interference types.

4.2.1 Narrowband Persistent Interference Elimination

We utilized frequency-time two-dimensional spectral data of PSR J1935+1616 observed by the Nanshan 26 m radio telescope on May 3, 2020. The L-band receiver operating bandwidth is 320 MHz, with a digital backend bandwidth of 512 MHz, time resolution of 64 s, and frequency resolution of 1 MHz. During processing, only data within the receiver' s operating bandwidth were retained. The two-dimensional spectral data were segmented along the time axis into samples of 2048 time bins, yielding 320 × 2048 pixel spectral images. The L-band receiver operates from 1400-1720 MHz, with persistent narrowband interference appearing at 1450-1465 MHz and 1535-1545 MHz. Conventional processing of contaminated frequency channels involves direct removal. Various replacement schemes exist: (i) substituting with a constant value; (ii) replacing with the mean of median values from remaining channels; or (iii) replacing with artificial Gaussian noise.

As interference mitigation serves as a data preprocessing step in the FRB search pipeline, and subsequent processing requires standardization or normalization, this study adopts artificial Gaussian noise replacement. Additionally, to improve signal-to-noise ratio and visibility, we applied 16-fold downsampling, yielding final 320 × 128 pixel spectral images for processing.

4.2.2 Narrowband Short-Duration Interference Elimination

Unlike persistent narrowband interference discussed in Section 4.2.1, this refers to short-duration narrowband interference appearing only in portions of the data. We applied the IQRM algorithm proposed by Morello et al. [21] to address this interference type. IQRM identifies all frequency channels affected by narrowband interference within a given dynamic spectrum data block. The algorithm first performs spectral statistics for each frequency channel using channel standard deviation, then determines potentially contaminated channels through comparative analysis. The algorithm is non-parametric and robust against trends in sequential data.

The frequency channel standard deviation sequence is represented as $(x_i)_{i=0}^{n-1}$, where x_i denotes the standard deviation of the i-th frequency channel and n is the number of channels. IQRM has two adjustable parameters: radius r and threshold t. Radius r defines the comparison range when evaluating the anomalous status of a given data point. Threshold t represents the significance level as a multiple of Gaussian σ (the standard deviation of the comparison sequence), where $t\sigma$ serves as the threshold for identifying outliers. The procedure comprises:

First, compute the difference between each frequency channel' s standard deviation and those of neighboring channels within the specified radius:

$$\Delta_i = x_i - x_{i-k}$$

where k ranges from $[-r, +r]$ and $k \neq 0$. For the resulting Δ_i values, denoted as Q_k, we identify outliers using a Tukey-like rule [24]. The outlier condition is:

$$|\Delta_i - \tilde{\mu}_k| > t \times \tilde{\sigma}_k$$

where $\tilde{\mu}_k$ is the median of the sequence ($\tilde{\mu}_k = Q_k(0.5)$) and $\tilde{\sigma}_k$ is the standard deviation calculated using the normalized interquartile range:

$$\tilde{\sigma}_k = \frac{1}{1.349} \times \mathrm{IQR} = \frac{Q_3 - Q_1}{1.349}$$

where IQR is the interquartile range, Q_3 is the upper quartile, and Q_1 is the lower quartile.

If condition (4) holds, x_i exhibits greater anomaly than x_{i-k}. However, this does not automatically classify x_i as a high outlier and x_{i-k} as a negligible low outlier; it represents a relative comparison. Therefore, we must synthesize results from multiple k value calculations, eliminate low outliers, and retain only high outliers for mitigation. A channel x_i is classified as a high outlier if, across all rounds of computation, condition (4) holds for x_i more frequently than for x_{i-k}.

For identified frequency channels, we apply the same Gaussian noise replacement scheme described in Section 4.2.1. Figure 3 [Figure 3: see original paper] presents two processing examples, with raw observational data shown in the upper panels. Narrowband interference mitigation significantly improves single-pulse image quality.

If short-duration narrowband interference exists in a frequency channel within a data block, IQRM identifies and mitigates it. While persistent narrowband interference is also accurately identified and processed, we conducted an experiment demonstrating that applying Gaussian noise replacement for interference in the 1450–1465 MHz range before IQRM processing further improves single-pulse visibility for some images. Figure 4 [Figure 4: see original paper] shows the original unmitigated image (top), IQRM-only mitigation (middle), and the result of first removing persistent narrowband interference then applying IQRM (bottom). The combined approach yields significantly enhanced signal quality.

4.2.3 Broadband Instantaneous Interference Mitigation

Broadband instantaneous interference is the most common type and most likely to cause false detections in FRB search software. Its key distinction from FRB signals is the lack of dispersion. In 2009, Eatough et al. [25] proposed the Zero-DM Filter (ZDF) method for eliminating such interference. ZDF removes interference by subtracting the mean of all frequency channel data at each sampling point from every data point. This method must be applied before dedispersion processing:

$$S'(i,j) = S(i,j) - \frac{1}{n_{\text{chans}}} \sum_{i=1}^{n_{\text{chans}}} S(i,j)$$

where $S(i,j)$ is the data at frequency channel i and time sample j, n_{chans} is the number of frequency channels, and $S'(i,j)$ is the mitigated data.

Theoretically, this method perfectly eliminates instantaneous broadband interference. However, practical applications reveal that broadband interference intensity is not uniform across the entire bandwidth, with extreme cases showing interference in only partial bands. This can lead to excessive signal excision. Men et al. [18] proposed ZDMF as an upgraded ZDF version that calculates proportional contributions from each frequency channel to remove only the interference component. Figure 5 [Figure 5: see original paper] compares both methods on identical data: the original unprocessed image (top), ZDF mitigation (middle), and ZDMF mitigation (bottom). ZDMF demonstrates superior performance.

The ZDMF approach minimizes residuals by fitting Zero-DM waveform sequences to residuals of given frequency channel i sequences to obtain baseline β_i and scaling factor α_i. The residual is defined as:

$$\chi^2 = \sum (S_i - \alpha_i S_{\mathrm{dm=0}} - \beta_i)^2$$

where S_i is the time series of frequency channel i and $S_{\mathrm{dm=0}}$ is the mean of all frequency channel time series (the Zero-DM waveform sequence). The scaling factor α_i is calculated as:

$$\alpha_i = \frac{S_{\mathrm{dm=0}} \cdot S_i - \frac{1}{N} \sum S_{\mathrm{dm=0}} \sum S_i}{S_{\mathrm{dm=0}} \cdot S_{\mathrm{dm=0}} - \frac{1}{N} (\sum S_{\mathrm{dm=0}})^2}$$

where "·" denotes the inner product of time series and N is the number of data points. The interference contribution is removed using:

$$S_i' = S_i - \alpha_i S_{\mathrm{dm=0}}$$

where S_i' is the mitigated data for frequency channel i. We do not remove the offset β_i as it does not affect FRB dynamic spectrum sample detection. Figure 6 [Figure 6: see original paper] shows the application results for our observational data.

4.3 Processing Speed and Detection Experiments

To validate the necessity of our three-level mitigation strategy, we conducted comparative experiments using IQRM and ZDMF individually versus the combined IQRM+ZDMF workflow. Results are shown in Figures 7 [Figure 7: see original paper] and 8 [Figure 8: see original paper], with raw frequency-time dynamic spectra (top), single-method mitigation (middle), and combined mitigation (bottom). The IQRM+ZDMF combination consistently outperforms individual methods.

We integrated this mitigation workflow into our FRB DDSS processing pipeline [26]. Processing 268.44 seconds of PSR J1935+1616 data for single-pulse searching yielded the results in Table 2 . The pipeline with ZDMF requires over ten times longer than without mitigation. Note that execution speeds vary significantly across hardware platforms, programming languages, and algorithm designs. Our reported test results reflect a single-threaded pipeline; the comparative data between the two processing workflows are more meaningful. When applying only narrowband interference mitigation, the FRB search pipeline processes data 6.8 times faster than the observation generation rate. However, achieving real-time search with the full three-level mitigation workflow requires targeted design optimizations, such as multi-threaded parallel processing, to meet speed requirements. The ZDMF-enhanced pipeline detects more single pulses, demonstrating that improved data quality through three-level mitigation enhances detection rates. While we could calculate the expected number of single pulses in the data packet based on PSR J1935+1616' s period, we cannot use this to further calculate and evaluate the pipeline' s detection rate because individual pulses vary in intensity, with many buried in observational background noise.

Table 2 The execution speed of RFI mitigation algorithm and pulse detection number produced by the FRB DDSS pipeline

Methods	Processing time/s	Pulse detections
IQRM+ZDMF		

5 Summary and Outlook

This paper discusses challenges posed by the interference environment in radio astronomy and major mitigation measures, analyzes FRB signal characteristics and RFI impacts on FRB searches, and reviews the electromagnetic interference environment, primary sources, and distribution at the Nanshan 26 m radio telescope site. To improve FRB detection algorithm accuracy, we classified dominant interference types in Nanshan L-band observations and established a multi-level mitigation processing workflow. Experimental results demonstrate significant mitigation effectiveness. However, this approach increases data processing time, requiring careful consideration of real-time processing requirements when directly applying the workflow to FRB search pipelines. This method is generalizable and provides valuable reference for other radio telescope data processing.

Acknowledgments

We thank the reviewers and editors for their valuable suggestions that improved the rigor and completeness of this paper. We also thank the Nanshan station 26 m radio telescope for providing observational data for this research.

References

[1] Ford J M, Buch K D. Geoscience & Remote Sensing Symposium. Quebec City: IEEE, 2014: 231
[2] Hogden J, Wiel S V, Bower G C, et al. AJ, 2012, 747:
[3] Baan W A. General Assembly and Scientific Symposium. Istanbul, Turkey: IEEE, 2011: 1
[4] Lorimer D R, Bailes M, Mclaughlin M A, et al. Science, 2007, 318: 777
[5] Thornton D, Stappers B, Bailes M, et al. Science, 2013, 341: 53
[6] Petroff E, Hessels J W T, Lorimer D R. A&ARv, 2019, 27: 4

[7] Pal S, Lancaster M J, Norrod R D. IEEE Microwave & Wireless Components Letters, 2012, 22: 236
[8] Tuccari G, Caddemi A, Barbarino S, et al. arXiv preprint astro-ph/0501019, 2005
[9] Fridman P A. AJ, 2008, 135: 1810
[10] Lane W M, Cohen A S, Kassim N E, et al. RaSc, 2005, 40: 2189
[11] Nita G M, Gary D E, Liu Z, et al. PASP, 2007, 119: 805
[12] Nagel J R, Bradley R, Warnick K F, et al. RaSc, 2007, 42: 553
[13] Hellbourg G, Trainini T, Weber R, et al. 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO). Bucharest, Romania: IEEE, 2012: 200
[14] Ellingson S W, Bunton J D, Bell J F. ApJS, 2001, 135:
[15] Poulsen A J, Jeffs B D, Warnick K F, et al. AJ, 2005, 130:
[16] Finger R, Curotto F, Fuentes R, et al. PASP, 2018, 130:
[17] Vos E E, Francois Luus P S, Finlay C J, et al. 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP). Pittsburgh, PA, USA. IEEE, 2019: 1
[18] Men Y P, Luo R, Chen M Z, et al. MNRAS, 2019, 488:
[19] Boyle J, Sclocco A. 2019 RFI Workshop–Coexisting with Radio Frequency Interference (RFI). Toulouse, France. IEEE, 2019: 24
[20] Sclocco A, Vohl D, van Nieuwpoort R V. 2019 RFI Workshop–Coexisting with Radio Frequency Interference (RFI). Toulouse, France. IEEE, 2019: 1
[21] Morello V, Rajwade K M, Stappers B W. MNRAS, 2021, 510: 1393
[22] Petroff E, Keane E F, Barr E D, et al. MNRAS, 2015, 451: 3933
[23] An T, Chen X, Mohan P, et al. Acta Astronomica Sinica, 2017, 58: 20
[24] Tukey J W. Exploratory Data Analysis. Boston, US: Addison-Wesley, 1977: 39-56
[25] Eatough R P, Keane E F, Lyne A G. MNRAS, 2009, 395:
[26] Liu Y L, Li J, Liu Z Y, et al. RAA, 2022, 22: 15289

Note: Figure translations are in progress. See original paper for figures.