

WAVELET DENOISING OF RADIO OBSERVATIONS OF ROTATING RADIO TRANSIENTS (RRATS): IMPROVED TIMING PARAMETERS FOR EIGHT RRATS

M. JIANG

Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, WV 26506

B.-Y. CUI¹

Department of Physics and Astronomy
West Virginia University
Morgantown, WV 26506

N.A. SCHMID¹

Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, WV 26506

M.A. MCLAUGHLIN¹

Department of Physics and Astronomy
West Virginia University
Morgantown, WV 26506

Z.-C. CAO

Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, WV 26506

¹Center for Gravitational Waves and Cosmology, West Virginia University, Morgantown, WV 26505

ABSTRACT

Rotating radio transients (RRATs) are sporadically emitting pulsars detectable only through searches for single pulses. While over 100 RRATs have been detected, only a small fraction (roughly 20%) have phase-connected timing solutions, which are critical for determining how they relate to other neutron star populations. Detecting more pulses in order to achieve solutions is a key to understanding their physical nature. Astronomical signals collected by radio telescopes contain noise from many sources, making the detection of weak pulses difficult. Applying a denoising method to raw time series prior to performing a single-pulse search typically leads to a more accurate estimation of their times of arrival (TOAs). Taking into account some features of RRAT pulses and noise, we present a denoising method based on wavelet data analysis, an image-processing technique. Assuming that the spin period of an RRAT is known, we estimate the frequency spectrum components contributing to the composition of RRAT pulses. This allows us to suppress the noise, which contributes to other frequencies. We apply the wavelet denoising method including selective wavelet reconstruction and wavelet shrinkage to the de-dispersed time series of eight RRATs with existing timing solutions. The signal-to-noise ratio (S/N) of most pulses are improved after wavelet denoising. Compared to the conventional approach, we measure 12% to 69% more TOAs for the eight RRATs. The new timing solutions for the eight RRATs show 16% to 90% smaller estimation error of most parameters. Thus, we conclude that wavelet analysis is an effective tool for denoising RRATs signal.

Keywords: pulsars: general – stars: neutron – methods: data analysis – techniques: image processing

1. INTRODUCTION

Rotating Radio Transients (RRATs) are sporadic pulsars that were detected through a search for single pulses and are not detectable through their time-averaged emission. Their profiles are typically a few milliseconds in duration, with intervals between detected pulses ranging from minutes to hours (McLaughlin et al. 2006). They were first discovered through single-pulse search reprocessing of the Parkes Multibeam Pulsar Survey data (Manchester 2001; Keith et al. 2009; Keane & McLaughlin 2011) and currently over 100 such objects are known¹. It is clear that RRATs represent one extreme end of the pulsar intermittency spectrum, with some objects detected as RRATs appearing as normal pulsars when observed with a more sensitive telescope or at a different frequency. However, it is not yet clear what physical or environmental properties determine the emission properties of neutron stars. The discovery of more RRATs, followed by long-term observations resulting in phase-connected solutions, is essential to answer this question.

For most pulsars, in order to calculate times of arrival (TOAs) the time series data are folded (Larsson 1996) at the identified period to amplify the signal-to-noise ratio (S/N) of the pulse and to calculate a composite profile. However, the sporadic emission from RRATs means that they are usually not detectable through folding. Therefore, it is necessary to calculate the TOA of each individual pulse to determine timing solutions. However, astronomical signals collected by radio telescopes contain many types of noise including radiometer noise, sky noise, and radio frequency interference (RFI). Even in remote areas, where most radio telescopes are located, terrestrial sources of RFI can have a significant impact on our ability to detect single pulses, and in the most extreme cases, terrestrial RFI is so strong that it makes the signals completely undetectable. In order to conduct a successful search for single pulses in the time domain, it is desirable to decrease this background noise in the data before applying a single-pulse search procedure. Applying an efficient denoising method to data containing RRAT signals may lead to an increased number of detected single pulses and improved accuracy of TOAs and timing solutions.

A conventional denoising method was proposed by Cordes & McLaughlin (2003), which presented a framework for processing radio data to detect single pulses. A boxcar filter (matched filter) was utilized to detect pulse candidates in de-dispersed time series, which were then plotted for manual inspection. This conventional method was employed among others by Deneva et al. (2009), Keane et al. (2010), Bagchi et al. (2012) and Cui et al. (2017) for single-pulse searches.

In this work, we describe a wavelet-based denoising algorithm that may be applied to de-dispersed time series data containing RRAT pulses. We compare the pulsar fitting results based on the proposed method with the results in Cui et al. (2017) which employed a conventional denoising method. The method builds on the assumption that for pulses collected by a radio telescope, the noise contributes to the entire frequency spectrum, while the RRAT pulses mostly contribute to lower frequency bands. Here, the frequency is not a radio frequency or a spin frequency. The frequency mentioned throughout this paper, without any other specification, is the frequency of the Fourier spectrum of the time series. Assuming that the spin period of an RRAT is known, we estimate the range of frequencies contributing to a description of an RRAT signal and use it to develop an effective denoising approach, which leads to an improved timing solution compared to the conventional method.

Wavelets are functions localized in time and space (Daubechies et al. 1992). Similar to Fourier Transforms (FTs), wavelet transforms are linear transforms used to approximate (decompose) any unknown function (signal). Wavelet transforms (WTs) can be used for data compression, signal analysis, data transformation and other applications. In the last decade wavelet analysis had a great impact on mathematics in the fields of approximation theory and harmonic analysis (Christensen & Christensen 2004), as well as on engineering in the fields of signal processing, image processing, and computer graphics (Mallat 1999; Chan & Shen 2005). Different from Fourier analysis, WTs involve functions described by two parameters, delay and scale, leading to comprehensive joint time- and frequency-domain analyses. This feature of wavelets provides us with an efficient way to extract RRAT pulses from astronomical data. In spite of our inability to entirely remove noise from the data, we can decrease its magnitude and thus increase the S/N of the RRAT pulses.

The remainder of the paper is organized as follows. Section 2 describes the eight RRATs analyzed in this paper. In Section 3, we introduce WTs. Based on the a priori knowledge of the spin period of a RRAT signal, we develop a frequency description of the RRAT and then apply wavelet denoising using selective wavelet reconstruction and wavelet shrinkage to the data containing the RRAT signal. The denoised time series are then subjected to the conventional single-pulse searching. In Section 4, we discuss metrics that we apply to evaluate the performance of the wavelet-based denoising approach. The results of denoising, timing analyses, and pulse profiles are summarized in Section 5. Section 6 analyzes the S/N distribution of RRAT data before and after the application of the wavelet denoising approach to data and the factors for timing solution improvement. Finally, planned future work and conclusions are presented in

¹ See <http://astro.phys.wvu.edu/rratalog>

PSR Name	Telescope	Data Machine	Frequency (MHz)	Bandwidth (MHz)	Sample Time (μ s)	Time Span (Years)	Number of Observations	Burst Rate (hr^{-1})	Mean Flux Density for a Single Pulse (mJy)
J0735–6302	Parkes	BPSR	1400	256	64.0	2.06	22	62.5	0.4
J1048–5838	Parkes	SCAMP	1400	256	100.0	15.2	52	6.0	2.4
J1226–3223	Parkes	BPSR	1400	256	64.0	1.94	19	49.3	2.8
J1623–0841	GBT	GUPPI	350/820	100/200	245.76	2.43	24	45.0	11 (820 MHz); 84 (350 MHz)
J1739–2521	GBT	GUPPI	820	200	491.52	1.60	25	38.3	49
J1754–3014	GBT	GUPPI	350/820	100/200	245.76	1.56	25	120.3	21 (820 MHz); 410 (350 MHz)
J1839–0141	GBT	GUPPI	820	200	491.52	2.50	38	19.5	18
J1848–1243	GBT	GUPPI	820	200	245.76	1.80	29	34.0	13

Table 1. Observing parameters for the eight RRATs described in this paper.

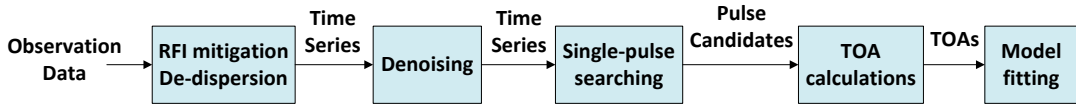


Figure 1. Block diagram of the conventional single-pulse search processing applied to a time series containing an RRAT’s signal. In this paper, we focus on the last four blocks.

Section 7.

2. OBSERVATIONS

Five (J1048–5838, J1739–2521, J1754–3014, J1839–0141, and J1848–1243) RRATs discussed in this paper were discovered in the process of re-analyzing the 1.4-GHz Parkes Multibeam Pulsar Survey data (Manchester 2001; Keith et al. 2009), PSR J1623–0841 was discovered through the 2007 GBT 350-MHz drift-scan survey (Boyles et al. 2013) and the remaining two (J0735–6302 and J1226–3223) were discovered in re-analyses of the 2009 1.7 GHz southern-sky high Galactic latitude survey data (Jacoby et al. 2009; Burke-Spolaor & Bailes 2009). The follow-up timing and monitoring observation used the same two telescopes and are reported in Cui et al. (2017). The eight RRATs have spin periods ranging from 0.4 to 6.2 s. The observation frequencies of the follow-up observations cover the range from 350 MHz to 1.4 GHz and are summarized in Table 1.

3. WAVELET DENOISING

Figure 1 shows a block diagram of the conventional single-pulse processing applied to a time series containing an RRAT signal. In this paper, we focus on the last four blocks in Figure 1, assuming that the data have been already de-dispersed. We will later demonstrate in Section 5 that the inclusion of wavelet denoising leads to improved detection of single pulses and allows for more precise timing results compared to those obtained with the conventional approach.

3.1. Wavelet Transform

The wavelet denoising approach proposed in this paper includes two steps. Selective wavelet reconstruction, or selecting the proper wavelet decomposition sub-bands for reconstructing the signal, is the first step. The second step is applying wavelet shrinkage (thresholding) to the wavelet coefficients of the sub-bands selected in the first step. A detailed description of both steps is provided in Sections 3.2 and 3.3, respectively.

Wavelets are a set of functions localized in time and space. Similar to sine and cosine functions, wavelets are used as basis functions for a linear decomposition of non-stationary or time varying signals. A wavelet is defined as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \quad (1)$$

where $\psi(\cdot)$ is a known function, a is a time scale parameter, and b is a delay parameter. The parameter a is a positive real number, selected from a set A ; b is a real number, selected from a set B . Note that $\psi_{1,0}(t)$ is known as the mother wavelet, selected to satisfy (see Daubechies (1988) for details):

$$\int_{-\infty}^{\infty} \psi(t)dt = 0. \quad (2)$$

A great multiplicity of mother wavelet functions has been developed in the past two decades. The choice of a mother wavelet function depends on a specific application.

Daubechies 5



Figure 2. The profile of the *Daubechies 5* wavelet. Based on our numerical analysis, a mother wavelet with vanishing moment 5 was found to provide a good representation for a RRAT’s pulse.

Similar to a Fourier decomposition, which approximates a function by a linear combination of trigonometric functions with different frequencies, wavelets employ the scale parameter a to analyze the frequency content of a signal. Given a mother wavelet and a set of scaled and delayed wavelet functions, the continuous WT (Daubechies 1990) of a function $f(t) \in L^2$ at scale a and delay b is given as:

$$W_\psi f(b, a) = \int_{-\infty}^{\infty} f(t) \cdot \psi_{a,b}^*(t) dt, \quad (3)$$

where $W_\psi f(b, a)$ are wavelet coefficients, and “*” denotes the operation of complex conjugate. The scale and delay parameters are often set to powers of 2 such that $a = 2^{-s}$ and $b = k2^{-s}$, where k and s are integers.

Because the WT has two parameters—scale and delay—a signal processed by the WT yields an output signal with two dimensions: frequency and time, respectively. This explains how wavelet analysis combines both the frequency- and time-domain analyses (Strang & Nguyen 1996). Equation (3) shows how WT coefficients are obtained. Here, the collection of all possible pairs (a, b) is denoted by the set $A \times B$, and consequently, the inverse WT (Goswami & Chan 1999) can be expressed as

$$\hat{f}(t) = \frac{1}{C_\psi} \iint_{(A \times B)} \frac{1}{a^2} [W_\psi f(a, b)] \cdot \psi_{a,b}(t) da db, \quad (4)$$

where $\hat{f}(t)$ is the selective wavelet reconstruction for $f(t)$. The factor C_ψ is a constant depending on the choice of the wavelet and is given by:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad (5)$$

where the wavelet window $\hat{\psi}(0) = 0$, “ $\hat{\cdot}$ ” denotes the operation of FT.

When developing a wavelet denoising approach for a particular application, one is expected to (a) select a mother wavelet, (b) summarize selective wavelet reconstruction and (c) specify a wavelet shrinkage (thresholding) approach. Among the most popular mother wavelets are Haar, Daubechies, biorthogonal, Mexican hat, etc. (see Haar (1910); Grossmann & Morlet (1984); Daubechies et al. (1992); Mallat (1999)). They have been utilized for various applications due to their different mathematical properties. Daubechies et al. (1992) proposed a series of mother wavelets called *Daubechies n*, where n is from 1 to 20, where the integer n stands for numbers of vanishing moments. Vanishing moments determine the profile of wavelet. Based on our numerical analysis, a mother wavelet with vanishing moment 5 was found to provide a good representation for an RRATs pulse, and thus we selected *Daubechies 5* as a mother wavelet. Figure 2 shows the profile of the *Daubechies 5* wavelet. Other wavelet families may be also suitable for representing some RRATs’ pulses due to the various single-pulse shapes. The comparison of denoised results of RRATs’ data processed with different wavelet families is beyond the scope of this paper.

3.2. Selective Wavelet Reconstruction

After applying the WT with varying values of the scale parameter a to noisy data, we represent the data in terms of “frequency bands,” with each band parameterized by a value of a . Therefore, if we know the frequency range of an

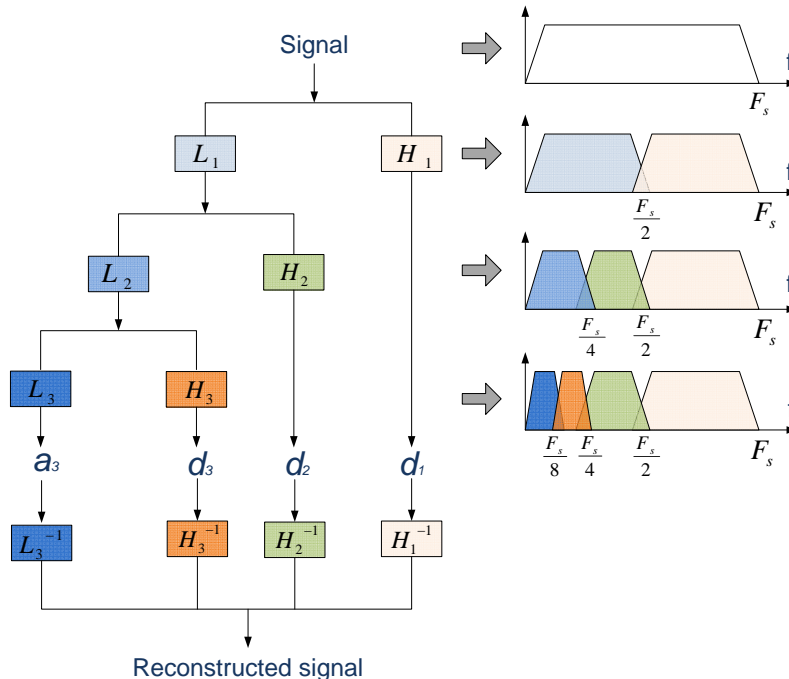


Figure 3. The left side of the figure is a discrete wavelet decomposition and reconstruction tree. Here, we chose a three-level tree for our example. The input signal is a time series. L_n and H_n ($n = 1, 2, 3$) are “high pass” and “low pass” filters with different cut-off frequencies used to decompose the signal at different scales. Here, d_n is known as the detail wavelet coefficient and a_n is an approximation wavelet coefficient. They are the outcomes of discrete wavelet transformation. L_n^{-1} and H_n^{-1} are filters used for signal reconstruction (inverse DWT). The right side of the figure shows how the signal is being decomposed by each filter.

astrophysical pulse of interest, we can extract it from the noisy data by keeping only the appropriate sub-bands and filtering out everything else.

In practice, we usually encounter discrete signals. Therefore, we consider an effective implementation applicable to discrete signals called the Discrete Wavelet Transform (DWT). The DWT uses multiresolution analysis (MRA) and synthesis for signal decomposition and reconstruction (Daubechies 1990). MRA allows an efficient implementation of the WT (similar to the FFT for FT). In MRA, a function is viewed at various levels of resolutions, which can be simply understood as the function being represented by wavelets with different scale parameters and delay parameters (Goswami & Chan 1999; Mallat 1989).

Figure 3 shows a discrete wavelet decomposition and reconstruction tree. Here, we selected a three-level tree as an example. The input signal is a time series. L_n and H_n ($n = 1, 2, 3$) are “high pass” and “low pass” filters with different cut-off frequencies used to decompose the signal at different scales. In equation (3), $\psi_{a,b}^*(n)$ is a function used for transforming (decomposing) $f(n)$ at various scales a and different delays b . Here, in discrete wavelet decomposition, L_n and H_n are the filters (functions) used for decomposing the signal at various levels of resolution and different delays. They act as $\psi_{a,b}^*(t)$ in equation (3). Also, d_n is known as the detail wavelet coefficient and a_n is an approximation wavelet coefficient. They are the outcomes of discrete wavelet transformation. L_n^{-1} and H_n^{-1} are filters used for signal reconstruction (inverse DWT). Different from continuous-time wavelet decomposition, discrete-time wavelet decomposition can be represented by a logarithmic filter tree.

A single pulse from an RRAT has a shape and duration. Astronomers model the pulse as an intrinsic Gaussian function convolved with an exponential, as the pulse will be scattered due to propagation through the interstellar medium. Due to its shape and duration, a pulse can be described by a narrow range of frequencies. Therefore, after the application of wavelet decomposition, the RRAT signal will be represented by only a few sub-bands in the wavelet decomposition. In order to decide which wavelet sub-bands contribute to an RRAT pulse and thus should be used for reconstruction, we need to first estimate the range of frequencies that contains most of the RRAT signal.

Consider a time series of length T with samples spaced t_s apart (in our case t_s and T are typically around $100 \mu\text{s}$ and 30 minutes). If the series contains a few pulses, we can estimate its period P . In FT space, the interval between two samples of Fourier spectrum is $1/T$. The fundamental frequency of the RRAT signal is $f_{min} = 1/P$ and the total

number of harmonics of the RRAT signal in the frequency range from 0 to $1/t_s$ is P/t_s . Because the Fourier spectrum (harmonics) of the RRAT pulses rapidly decrease with increasing frequency, the last (or highest frequency) harmonics are completely dominated by noise. Thus, removing this noise will denoise the RRAT signal. Here, we recommend keeping the 200 largest frequency components (harmonics) composing of the RRAT signal. This number is selected empirically by using a maximum S/N criterion based on tests of RRATs data. Thus, the main frequency components contributing to the description of RRAT pulses are between $f_{min} = 1/P$ and $f_{max} = 200/P$. Now we can estimate the decomposition levels in the logarithmic filter tree (see Figure 3) containing the RRAT signal following the rules:

$$\begin{aligned} N_{max} &= \lfloor \log_2(F_s/f_{min}) \rfloor + 1 = \lfloor \log_2(P/t_s) \rfloor + 1, \\ N_{min} &= \lfloor \log_2(F_s/f_{max}) \rfloor + 1 = \lfloor \log_2(P/(200 \times t_s)) \rfloor + 1, \end{aligned} \quad (6)$$

where N_{min} is the lowest wavelet decomposition sub-band (level) retained for reconstruction, N_{max} is the highest sub-band (level) retained for reconstruction, and F_s is the sample frequency of the observation data.

As an example, the sampling time for the observation data of J1048–5838 is $100 \mu s$ and its rotational period is 1.23 s. Therefore, $f_{min} = 0.82$ Hz and $f_{max} = 164$ Hz, and according to equation (6), $N_{max} = 14$ and $N_{min} = 7$. These levels between 7 and 14 are expected to contribute to the description of the RRAT's pulses. Figure 4 shows a short slice of J1048–5838 de-dispersed time series in which two pulses are detected and signals reconstructed from the detail coefficients of levels 1 through 14. There are obvious pulses observed in the panels labeled as S_{D_7} , S_{D_8} and $S_{D_{10}}$. They are the three main components that contributed to the reconstructed pulse signal. The negative intensities in the pulse signal are due to the selected profile of the mother wavelet as shown in equation (2); we will show how to remove the negative components in Section 5.3.

3.3. Wavelet Shrinkage

Once the levels contributing to signal reconstruction are identified, the next step is to estimate the weights of each contributing level. We apply wavelet shrinkage, developed by Donoho & Johnstone (1994), which selects wavelet coefficients based on thresholding and uses them for reconstruction. There are two types of thresholding used in practice: hard thresholding and soft thresholding (Bruce & Gao 1996; Fodor & Kamath 2003). The output DWT coefficients $W'_{hard}(k)$ and $W'_{soft}(k)$ based on these two types of thresholding with threshold η are:

$$W'_{hard}(k) = \begin{cases} W(k), & |W(k)| > \eta \\ 0, & |W(k)| \leq \eta \end{cases} \quad (7)$$

$$W'_{soft}(k) = \begin{cases} \text{sign}(W(k))(W(k) - \eta), & |W(k)| > \eta \\ 0, & |W(k)| \leq \eta \end{cases} \quad (8)$$

where $W(k)$ is the DWT coefficient before the application of thresholding. Function $\text{sign}(\cdot)$ is defined as:

$$\text{sign}(x) = \begin{cases} -1, & \text{if } x < 0 \\ 0, & \text{if } x = 0 \\ 1, & \text{if } x > 0 \end{cases} \quad (9)$$

The observed noisy time series $Y(t)$ at time t containing an RRAT signal can be modeled as an RRAT signal $X(t)$ contaminated by an independent additive noise $N(t)$:

$$Y(t) = X(t) + N(t), \quad (10)$$

where for a fixed value of t , $N(t)$ is a Gaussian random variable with mean zero and variance σ^2 . Noise samples are independent and identically distributed. In this paper, rule (7) is applied to wavelet coefficients. The threshold η is estimated based on the noise variance in each sub-band. In Figure 4, the signal is decomposed into 14 levels (sub-bands), where levels 7 through 14 are expected to contribute to the RRAT signal description. Figure 5 shows the outcome of applying wavelet shrinkage to sub-bands 7 - 14. As levels 1 to level 6 have no contribution to the RRAT signal, they are not used for wavelet reconstruction and are shown as a zero signal. The top left panel is the reconstructed signal after wavelet denoising. Comparing with the original signal in Figure 4, we can see the noise is suppressed.

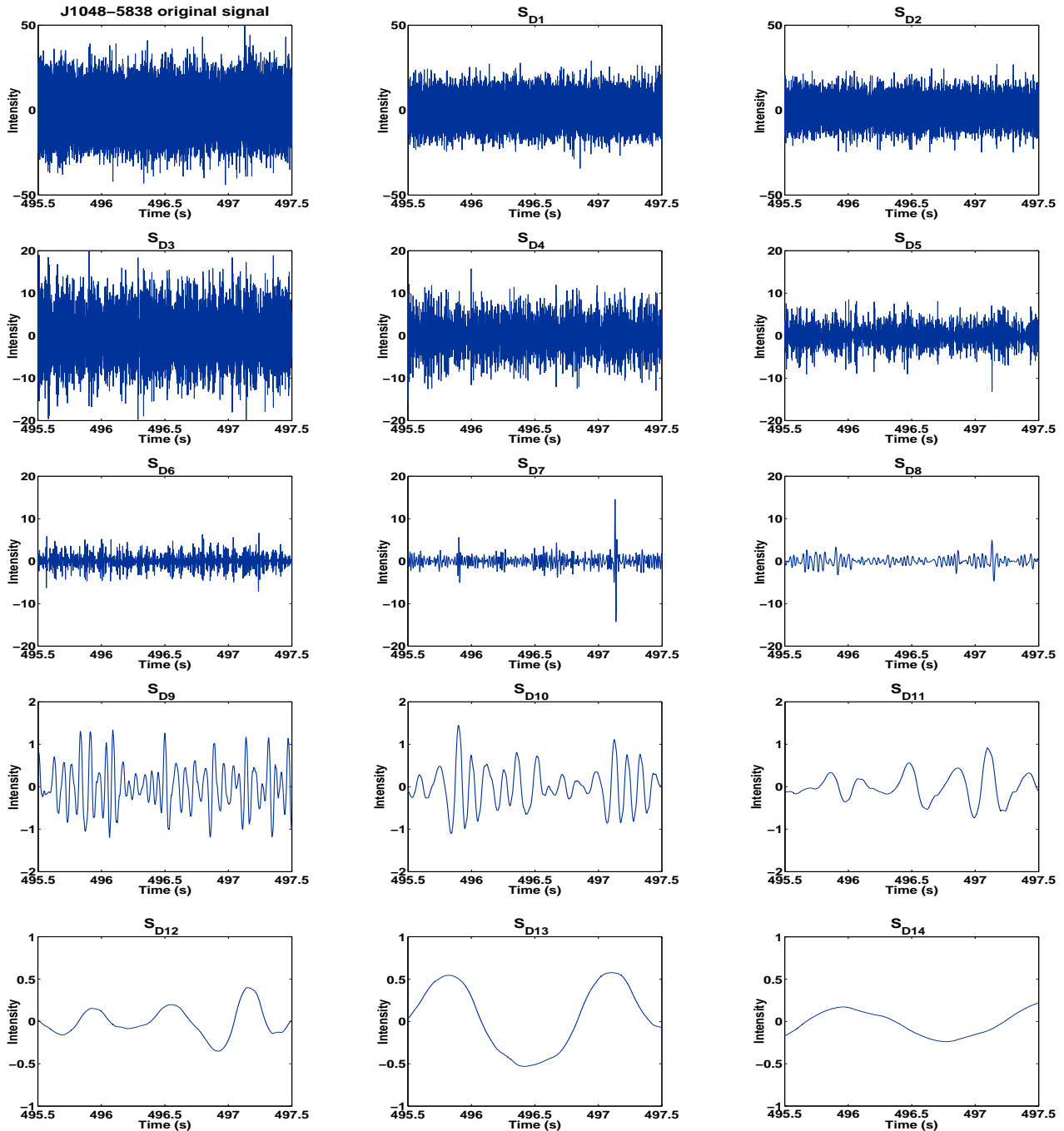


Figure 4. A short slice of J1048–5838 de-dispersed time series in which two pulses are detected and signals reconstructed from the detail coefficients of levels 1 through 14. From equation (6), $N_{max} = 14$ and $N_{min} = 7$, which means levels 7 through 14 are expected to contribute to the signal description. Note that there are obvious pulses observed in the panels S_{D7} , S_{D8} and S_{D10} . They are the three main components that contributed to the reconstructed pulse signal. The negative intensities in the pulse signal are due to the profile of the selected mother wavelet shown in Figure 2. This artifact will be removed due to a post-processing described in Section 5.3.

Figure 6 shows the dependence of the S/N value on the threshold η in the process of selecting the weight of the detail coefficients at level 7 for J1048–5838’s time series. Note that S/N peaks at $\eta = 25$, which is selected as the optimal threshold at level 7. RRATs have various optimal η for different sub-bands.

In this paper, we first apply wavelet analysis (decomposition) to time series containing an RRAT signal, then select the decomposition levels over the range of frequencies containing the RRAT signal and apply optimal thresholds to

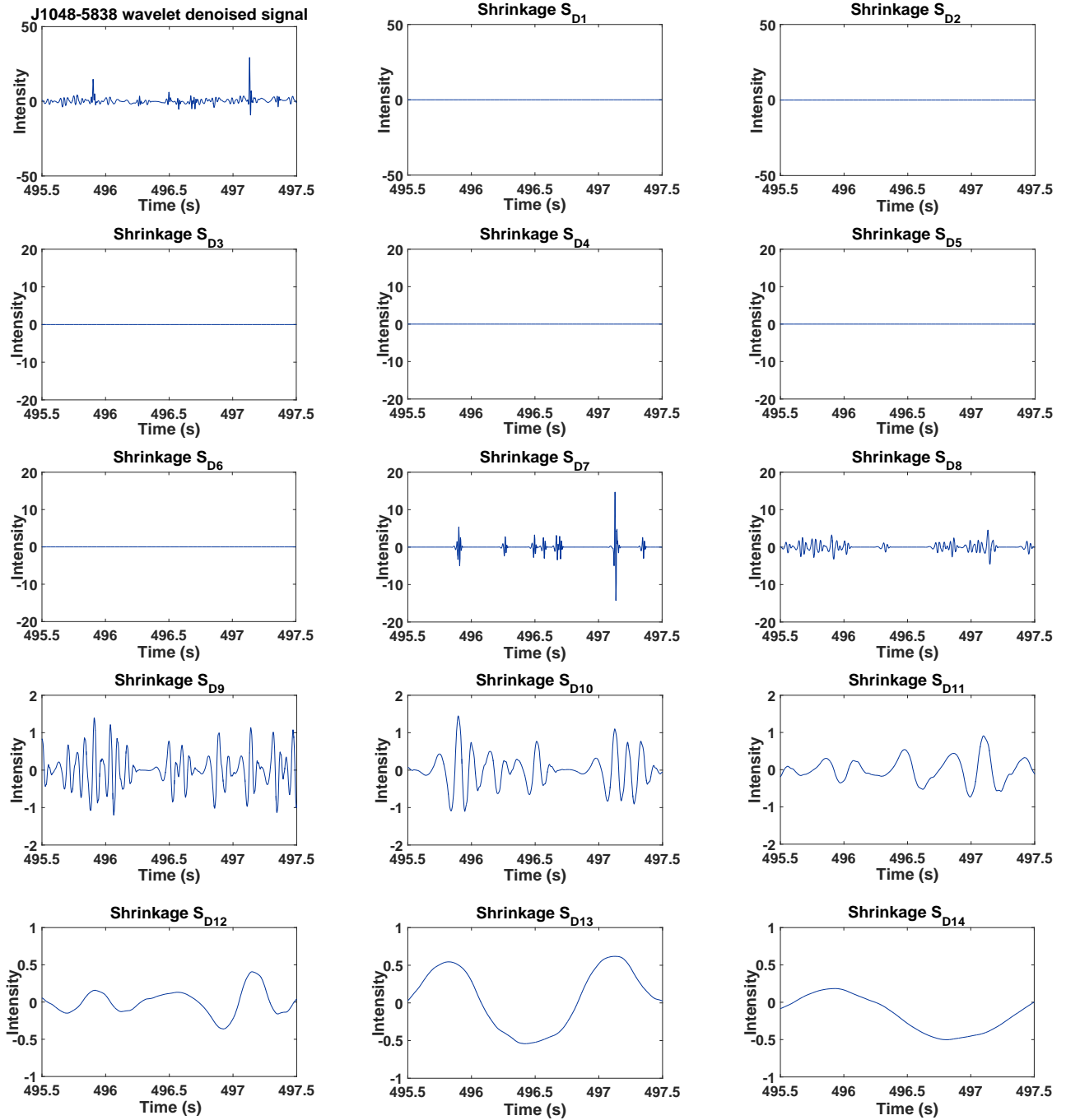


Figure 5. The outcome of applying wavelet shrinkage to sub-bands 7 - 14 (corresponding to the wavelet decomposed levels shown in Figure 4.). As levels 1 to level 6 have no contribution to the RRAT signal, they are not used for wavelet reconstruction and are shown as a zero signal. The top left panel is the reconstructed signal after wavelet denoising.

the detail coefficients at the selected levels. The last step is the reconstruction of the signal from the coefficients of the selected levels. We use the Wavelet Analysis toolbox from Matlab (Matlab 2015) for wavelet decomposition and thresholding. We use the *wavedec()* function for multilevel wavelet decomposition, the function *wthcoef()* for the wavelet coefficients thresholding, and the *wrccoef()* function for reconstructing single pulses from the wavelet coefficients.

4. PERFORMANCE MEASURES

We compare the conventional denoising method currently used in single-pulse searching—boxcar smoothing (Cordes & McLaughlin 2003) with the wavelet denoising approach. After applying wavelet denoising to a time series, we detect

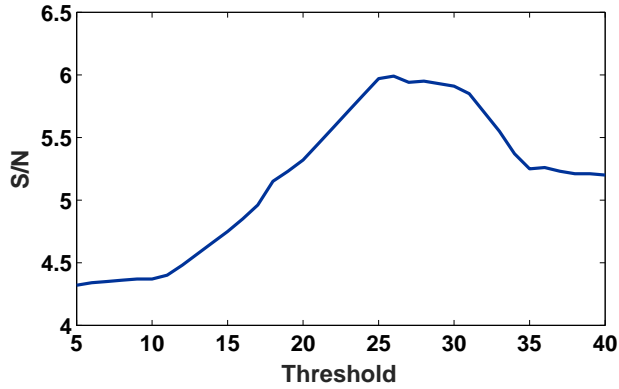


Figure 6. Shown is the dependence of the S/N value on the threshold η in the process of selecting the weight of the detail coefficients at level 7 for J1048–5838’s time series. The x-axis is the threshold η applied to the wavelet coefficients, the y-axis is an S/N of the pulse after thresholding. Here, η is estimated based on the noise variance in each sub-band.

candidate pulses with peak S/N exceeding a preset threshold. The peak S/N is calculated from the single-pulse profile and is defined as the peak divided by the root mean square of the noise (the profile amplitude in the off-pulse region). The S/N threshold is typically around 5σ or 6σ , where σ is the root mean square of the noise in each period. In this paper, the S/N threshold is set to either 4.5σ or to 5σ , depending on the data. This indicates that we are able to detect pulses from very faint sources. After identifying candidate pulses, we proceed with the periodicity check. We use temporal locations of peaks as the times of arrival. If the times of arrival t_1 and t_2 of an arbitrary pair of pulses satisfy the following inequality

$$\left| \frac{t_1 - t_2}{P} - \text{round} \left(\frac{t_1 - t_2}{P} \right) \right| < \varepsilon, \quad (11)$$

then the pulses are retained for further analysis. Here, P is the spin period of an RRAT and ε is a tolerance for the periodicity check.

We use False Alarm Rate (FAR) as a measure of detection performance. Here FAR is defined as the fraction of pulses assumed to be real, which may actually occur due to RFI accidentally falling into the periodicity window (2ε). It is defined as:

$$\text{FAR} = \frac{(N_{\text{candidates}} - N) \cdot 2\varepsilon}{N}, \quad (12)$$

where $N_{\text{candidates}}$ is the number of pulses after denoising and N is the number of pulses remaining after the periodicity check.

At last, to determine which pulses are due to an RRAT and which are spurious, we perform a timing analysis, as described in Cui et al. (2017). We delete data points with large residuals outside of a reasonable range of jitter (generally the pulses with residuals that are larger than the average TOA error), which produces errors due to the intrinsic pulsar emitting mechanism, or intrinsic inconsistencies in pulse phase. The number of real pulses determined by this method is listed in the sixth column in Table 2 .

5. NUMERICAL RESULTS

In this section, we first present the result of wavelet denoising applied to the time series of eight RRATs. Then we compare two timing solutions - one using TOAs calculated from data that have been denoised using the wavelet approach and the other using the conventional method (boxcar) in Cui et al. (2017). Along with the comparison of the timing solution, we compare the shape of denoised RRAT pulses.

5.1. Denoising

Smoothing data with a boxcar filter is a conventional denoising approach used for pulsar radio observation data. It lowers the effective sampling rate of the time series by binning the data (resampling the data with a lower sample rate), typically using 128, 256, or 512 bins per period. The boxcar filter is usually applied in two stages. In the single-pulse search stage (without knowing the pulsar spin period), an N -sample boxcar (the width of boxcar is N times the sampling interval, here N is optimized by maximizing the S/N) is applied to the original time series to identify the

PSR Name	Number of Pulses After Denoising ($N_{candidates}$)	Number of Pulses After Periodicity Check (N)	Tolerance (ϵ)	FAR (%)	Number of Real Pulses	Number of Real Pulses with Boxcar Denoising	Increase in Detected Pulses	Number of Epochs for Fitting	Additional Epochs
J0735–6302	1613	497	0.01	2.25	479	304	58%	6	1
J1048–5838	799	374	0.01	2.41	310	207	51%	24	5
J1226–3223	1419	469	0.01	4.05	430	360	19%	15	0
J1623–0841	7704	1693	0.01	7.10	1513	1202	26%	16	0
J1739–2521	1955	580	0.015	7.20	543	321	69%	16	2
J1754–3014	2460	862	0.03	11.12	829	550	51%	19	2
J1839–0141	1651	500	0.01	4.80	481	386	25%	14	1
J1848–1243	2133	522	0.015	9.26	440	393	12%	20	1

Table 2. Wavelet denoising results for time series of eight RRATs based on the observations in Table 1. Note. The second column of Table 2 lists the number of pulses detected in the time series after the data were processed with wavelets. The third column in the table is the number of pulses remaining after the periodicity check. The fourth column shows the tolerance for the periodicity check. The fifth column lists FARs for the eight RRATs. Here FAR is the fraction of pulses assumed to be real which may actually be due to RFI falling into the periodicity window (2ϵ) randomly. The sixth column summarizes the number of remaining pulses that are determined as real and used for pulsar model timing. The seventh column lists the number of detected pulses when the boxcar denoising method has been effectively applied. The eighth column presents the difference of the results in the sixth and seventh columns, which corresponds to the difference in the detected pulses in the case of the wavelet based and boxcar denoising. Note that the number of correctly detected pulses for the case of wavelet denoising is higher compared to the case of denoising with the conventional boxcar filter. The ninth column lists the number of epochs used for pulsar parameter fitting. The last column is the number of additional epochs which had zero detected pulses based on boxcar denoising, but have a few detected pulses with wavelet denoising.

pulse candidates. Then another boxcar (the width of boxcar is different from the one utilized at the previous stage, generally 128, 256, or 512 bins per period) is applied to time series for calculating the profile of the pulse.

On the contrary, wavelet denoising will not decrease the sampling rate of the time series. The comparison between the results of denoising with wavelets and boxcar applied to time series of eight RRATs is given in Table 2. Figure 7 shows a comparison between boxcar denoised data and wavelet denoised data on a short slice of a J1048–5838 time series, which contains a strong pulse. The first panel shows how we obtain the optimal width of the boxcar for these data. The x-axis is the width of the boxcar used for denoising, and the y-axis is the S/N of the pulse for the corresponding width of boxcar. The red curve is the polynomial fitting curve. Here, we use 30 time samples as the optimal width of the boxcar. However, in the pulse detection stage, it is hard to guarantee that the optimal boxcar is applied to every pulse, as they are applied in discrete steps. The second panel shows boxcar denoised data and the last panel shows the wavelet denoised data. The S/N of the pulse is improved from 10.54 to 13.34. Such improvement in S/N is particularly meaningful for weak pulses, as they may be undetectable in the case of boxcar denoising and become detectable after denoising with wavelets. This is why most of the newly found pulses in the eight time series have S/N around five. We will further discuss the distribution of S/N in Section 6. The width of the pulse ($\overline{W50}$) is calculated at 50% of the peak intensity. The width of the pulse is labeled in the figure, with pulses in data denoised with wavelets appearing narrower compared to the same pulses in the data denoised with the boxcar.

5.2. Timing Solution

We calculate TOAs for the RRAT pulses by calculating the time of the peak of the pulse. We do this instead of cross-correlating with a template, due to the variable single-pulse shapes. We then use the pulsar timing TEMPO2 (Hobbs et al. 2006) software to fit a timing model by minimizing the chi-square criterion of the timing residuals. The results are listed in Table 3. Comparing with the previous results based on the conventional boxcar approach given in Cui et al. (see 2017, Table 2), which is shown in Table 4, we see that most of the solutions exhibit smaller errors of right ascension (R.A.), declination (decl.), and period derivative (\dot{P}), but larger errors in period. We will discuss this in more detail in Section 6.2.

5.3. Pulse Profiles

We generate the composite pulse profiles by summing all detected individual single pulses, assuming the timing model was derived as shown in Section 5.2. After being processed by wavelet denoising, there are negative lobes on both sides of the pulse profile. These negative lobes are not an intrinsic feature of the pulse profile. They are due to the choice of the mother wavelet shown in Figure 2. Here, we use a high-pass filter to remove these lobes. For instance, we use a high-pass filter with a cut-off frequency of 34 Hz to remove the negative fluctuations in the pulse profile of J1839–0141. Figure 8 shows the composite pulse profiles of eight RRATs based on the wavelet denoising results in

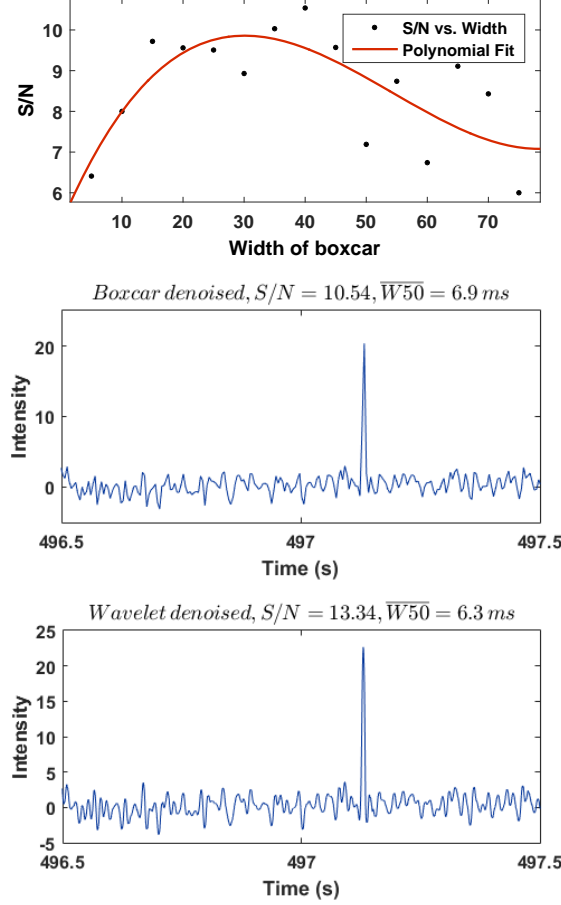


Figure 7. The comparison between a boxcar denoised profile and the same pulse denoised with wavelets on a short slice of a J1048–5838 time series that contain a strong pulse. The top panel shows how we select the optimal width of the boxcar for these data. The x-axis is the width of the boxcar used for denoising, and the y-axis is the corresponding S/N of the pulse. Here, we use 30 time samples as the optimal width of the boxcar. The middle panel shows the boxcar denoised data and the panel at the bottom shows the wavelet denoised data. Here, the S/N of the pulse is improved from 10.54 to 13.34.

PSR Name	R.A. (J2000)	Decl. (J2000)	P (s)	\dot{P} (10^{-15} s s $^{-1}$)	RMS (ms)	DM (pc cm $^{-3}$)	Epoch (MJD)	Width (ms)	B (10^{12} G)	\dot{E} (10^{31} ergs s $^{-1}$)	τ (Myr)
J0735–6302	07:36:20.01(27)	–63:04:16(2)	4.8628739612(7)	151.89(25)	6.32	19.4	56212	30	27.5	5.2	0.5
J1048–5838	10:48:12.561(14)	–58:38:19.02(10)	1.231304776630(4)	12.19375(7)	2.47	70.7(9)	53510	7	3.9	26	1.6
J1226–3223	12:26:46.628(40)	–32:23:01(1)	6.1930040852(5)	7.05(1)	10.08	36.7	56114	34	6.7	0.12	13
J1623–0841	16:23:42.6827(97)	–08:41:36.6(5)	0.50301500560(1)	1.9556(7)	3.74	59.79(2)	55079	11	1.0	61	4.1
J1739–2521	17:39:32.629(54)	–25:21:56(15)	1.81846119286(24)	0.24(2)	19.17	186.4	55631	43	0.7	0.7	119
J1754–3014	17:54:30.181(40)	–30:15:03(5)	1.3204904144(3)	4.428(19)	17.19	89.70(7)	55292	36	2.4	7.6	4.7
J1839–0141	18:39:06.9848(85)	–01:41:56.0(2)	0.93326558076(2)	5.944(1)	2.53	293.4	55467	13	2.4	29	2.5
J1848–1243	18:48:18.026(14)	–12:43:30(1)	0.41438335440(2)	0.4405(8)	3.16	91.96(7)	55595	9	0.4	24	15

Table 3. Timing solutions and derived parameters after using the wavelet denoising method for times series of eight RRATs: R.A., decl., spin period, period derivative, root-mean-square of residuals, DM, epoch of period, width of composite profile, the inferred magnetic field, and spin-down energy loss rate and characteristic age are listed. Note. The width is calculated at 50% of the peak intensity ($\overline{W50}$). The magnetic field is at the pulsar surface and assumes alignment between spin and magnetic axis ($B = 3.2 \times 10^{19} \sqrt{P\dot{P}}$). The pulsar spin-down luminosity is calculated by $\dot{E} = -4 \times 10^{46} \dot{P}/P^3$.

Table 2 and the previous boxcar denoised data from Cui et al. (2017). The width and S/N values for these RRATs vary due to different intrinsic properties. Most of the pulse profiles become narrower after wavelet denoising (refer to the columns of Table 3). Initial tests show that the narrowness of the wavelet reconstructed profiles is not an intrinsic

PSR Name	R.A. (J2000)	Decl. (J2000)	P (s)	\dot{P} ($10^{-15} \text{ s s}^{-1}$)	RMS (ms)	DM (pc cm^{-3})	Epoch (MJD)	Width (ms)	B (10^{12} G)	\dot{E} ($10^{31} \text{ ergs s}^{-1}$)	τ (Myr)
J0735–6302	07:35:5(2)	–63:02:0(4)	4.862873966(3)	159(5)	6.38	19.4	56212	24	28	5.5	0.5
J1048–5838	10:48:12.57(2)	–58:38:18.58(15)	1.231304776631(3)	12.19369(7)	2.22	70.7(9)	53510	7	3.9	26	1.6
J1226–3223	12:26:45.9(4)	–32:23:14(5)	6.1930029826(6)	7.68(11)	7.55	36.7	56114	57	6.0	0.08	20
J1623–0841	16:23:42.711(10)	–08:41:36.4(5)	0.503014992514(6)	1.9582(6)	0.76	60.433(16)	55079	13	1.0	61	4.1
J1739–2521	17:39:32.83(10)	–25:21:2(2)	1.8184611641(2)	0.29(3)	4.85	186.4	55631	68	0.7	0.7	99
J1754–3014	17:54:30.08(5)	–30:14:42(6)	1.3204902915(3)	4.424(12)	3.61	99.38(10)	55292	62	2.4	7.6	4.7
J1839–0141	18:39:07.03(3)	–01:41:56.0(9)	0.93326564072(6)	5.943(3)	1.80	293.2(6)	55467	17	2.4	29	2.5
J1848–1243	18:48:17.980(8)	–12:43:26.6(5)	0.41438334869(2)	0.440(2)	1.36	88.0	55595	9	0.4	24	15

Table 4. Timing solutions and derived parameters for eight RRATs based on the conventional boxcar approach. Note. See Table 3 for a description of columns.

property of these RRATs but due to the removal of some parts of the full spectrum.

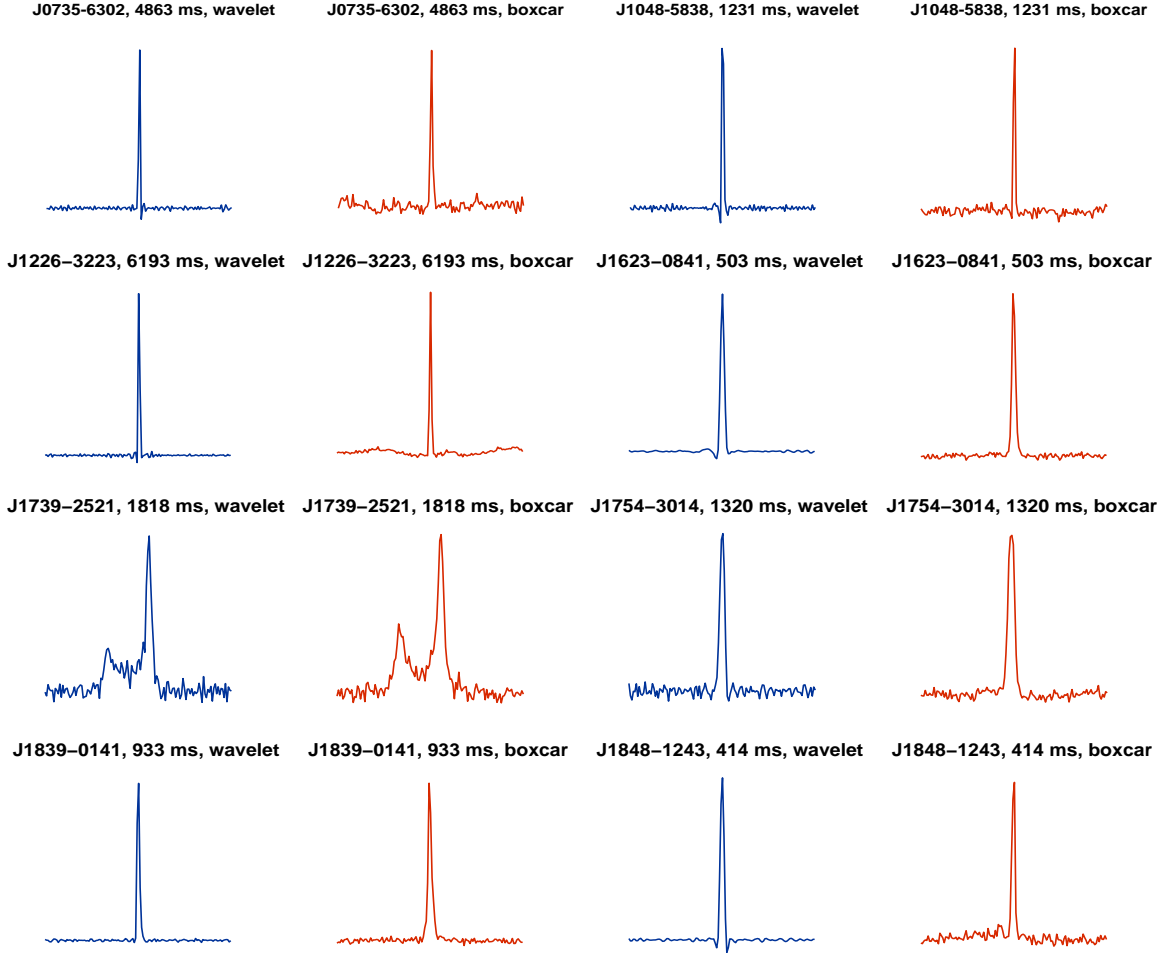


Figure 8. Comparison of composite pulse profiles of eight RRATs generated from wavelet denoised data and boxcar denoised data. The wavelet denoised pulse profiles of eight RRATs are based on the denoising results in Table 2. Each profile is a sum of all detected individual single pulses. The spin period of each RRAT is listed above each profile. The pulse width of the profile is specified in Table 3. Note that after being wavelet denoised, a profile may have negative lobes on both sides. These are due to the selected shape of the mother wavelet. The composite pulse profiles which are generated from wavelet denoised data have been processed by a high-pass filter to remove the negative lobes.

6. DISCUSSION

6.1. *S/N Distribution*

Because wavelet denoising can improve the S/N of pulses by decreasing the weights of the noise against the signal, we can detect more pulses with relatively small S/N in wavelet denoised data. Figure 9 shows a comparison of single-pulse S/N for J1048–5838 calculated based on the two denoising methods (boxcar and wavelet). Here, the S/N is calculated from rebinned wavelet single-pulse profiles. This is generally larger than that calculated from boxcar single-pulse profiles. However, as shown in this figure, some weak pulses could have higher S/N with boxcar denoising due to the selection effect of the threshold. The S/N threshold of the wavelet is lower than that of the boxcar due to the different sample rate and number of bins per period. Thus, some pulses detected from wavelet denoised profile (data) that are undetectable in a boxcar denoised profile are not shown in this figure. The red solid line indicates a linear fit to the main distribution, the green dashed line shows where the two algorithms perform the same, and the brown dotted-dashed line shows a search threshold of 5σ . From the distribution and the linear fitting result, we can see that the wavelet method successfully increased most of the single-pulse S/N values, especially for those originally larger than seven and less than 20.

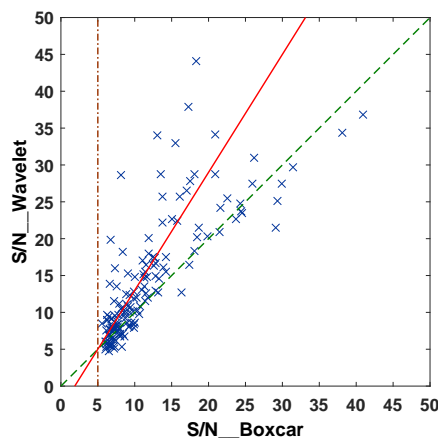


Figure 9. A comparison of single-pulse S/N calculated from J1048–5838’s wavelet denoised signal and boxcar denoised signal. The S/N here is calculated from rebinned wavelet single-pulse profiles, which is proved to be generally larger than that calculated from boxcar single-pulse profile. The red solid line indicates a linear fit to the main distribution, the green dashed line shows where the two algorithms perform the same, and the brown dotted-dashed line shows a search threshold of 5σ .

6.2. *Timing Solution Improvement*

In this work, there are three factors that may impact wavelet denoising’s ability to improve RRAT timing solutions: the first is more pulses detected in epochs that already had some detections; the second is pulses detected in epochs that had zero detections before; and the third is sharper pulse profiles generated from wavelet denoised data. We introduce a cross-test to identify which of these three factors is responsible for the decrease in pulsar parameter errors using data on J1048–5838. We first fit the pulses from the wavelet denoised data that have counterparts in the boxcar results (labeled as group A, which contains the fewest number of TOAs). Then we compare this timing solution with the timing solution obtained from the boxcar denoised data (labeled as group B) and wavelet denoised data (labeled as group C, which contains the largest number of TOAs). According to Table 5, the average error on parameters of group A is the largest of the three, while the average error on parameters of group C is the smallest. Such differences in the timing solution errors are caused by the different number of TOAs. Here, the improvement of timing solution is largely due to the increased number of TOAs, as the number of pulses detected in epochs which had zero detections before for J1048–5838 data is a small fraction. The improvement of timing solution for J1048–5838 is largely due to the first factor.

One interesting phenomenon is that, for J1048–5838, the errors on R.A., decl., and \dot{P} in the wavelet timing solution are all reduced compared with the timing solution of boxcar results, while the period error becomes larger. According to the pulse amplitude distribution of J1048–5838 (see Figure 5 in Cui et al. (2017) and Figure 8 in this paper), most

	R.A. Error (J2000)	Decl. Error (J2000)	P Error (10^{-12} s)	\dot{P} Error (10^{-20} s s $^{-1}$)	Number of TOAs	RMS (ms)
Cross test (A)	0.0194	0.156	5.09	7.06	183	2.54
Boxcar (B)	0.0174	0.157	3.37	7.21	207	2.33
Wavelet (C)	0.0137	0.102	4.01	6.67	316	2.47

Table 5. Comparison of errors in timing parameters of J1048–5838 calculated from the boxcar denoised data, cross-test data and wavelet denoised data. Note. Here we take a cross-test using pulses detected in both boxcar and wavelet methods. The result shows that the improvement to timing solutions using the wavelet denoising method is mainly due to the increased number of pulses detected.

newly detected pulses have relatively low S/Ns, so that the average S/N is generally lower in wavelet denoised data than that in boxcar denoised data. Because $\sigma_{TOA} \propto 1/(S/N)$ (Lorimer & Kramer 2005), lower S/N increases TOA errors. This results in a larger error on the fitted period, but lower errors on the other fitted parameters. We perform a simulation to show that this timing analysis is reasonable. The details of the simulation are given in the following paragraph.

In the simulation, we first create two sets of fake RRAT data with the same parameters (sample time, period, S/N, and number of TOAs). In set I, we randomly delete several TOAs (17.24%). For set II, we include all TOAs and set the S/N of the TOAs detected in set I to a lower value. Then we fit these two sets of data to a pulsar timing model and find that the error of the period in set II is larger. This simulation shows that even with a larger number of TOAs, the decrease in the S/N of the TOAs may still result in a larger error on the fitted period.

The second factor (pulses detected in epochs that had zero detections before) obviously improves the timing solution for J0735–6302. Five epochs were detected and used for fitting in the conventional approach, while six are detected and used for the wavelet timing solution (refer to the last two columns of Table 2). Because five epochs were inadequate for precise fitting, the previous fitting errors are larger. In this paper, the errors on the timing parameters are dramatically decreased (see Table 3).

7. CONCLUSIONS AND FUTURE WORK

In this paper, we applied a wavelet denoising using selective wavelet reconstruction and wavelet shrinkage to de-dispersed radio observation data for eight RRATs. This approach suppresses the noise in the data successfully. More TOAs are able to be measured using single-pulse searching of wavelet denoised data compared with the traditional denoising approach. More precise timing solutions are obtained based on wavelet denoised RRATs data, largely due to a larger number of TOAs and epochs available for use in fitting a timing solution. On the other hand, comparing with the boxcar approach, the wavelet approach can not perform better on accurately reconstructing the pulse profiles. For further research, we will compare the speed of the wavelet approach with the boxcar approach in the same runtime environment. Then we can extrapolate using the wavelet approach for single-pulse searching and fast radio burst (FRB) searching and apply wavelet denoising to other astronomical data.

ACKNOWLEDGMENTS

The work is supported by NSF Award 1458952. We would like to thank West Virginia University for its financial support for the Green Bank Telescope (GBT), which provided some of observation data in this paper. The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Parkes radio telescope is part of the Australia Telescope National Facility funded by the Australian Government for operation as a National Facility managed by CSIRO. We would also like to thank the editors and the reviewer for their comments and suggestions that improved the paper.

Facility: Green Bank Telescope, Parkes Telescope.

Software: Matlab (Matlab 2015), TEMPO2 (Hobbs et al. 2006)

REFERENCES

- Bagchi, M., Nieves, A. C., & McLaughlin, M. A. 2012, Monthly Notices of the Royal Astronomical Society, 425, 2501
- Boyles, J., Lynch, R. S., Ransom, S. M., et al. 2013, ApJ, 763, 80
- Bruce, A. G., & Gao, H. Y. 1996, Biometrika, 83, 727
- Burke-Spolaor, S., & Bailes, M. 2009, MNRAS, 1931
- Chan, T. F., & Shen, J. J. 2005, Image processing and analysis: variational, PDE, wavelet, and stochastic methods (Philadelphia: SIAM), doi:10.1137/1.9780898717877

- Christensen, O., & Christensen, K. L. 2004, Approximation theory: from Taylor polynomials to wavelets (New York: Springer Science & Business Media), doi:10.1007/978-0-8176-4448-2
- Cordes, J. M., & McLaughlin, M. A. 2003, 596, 1142
- Cui, B. Y., Boyles, J., McLaughlin, M. A., & Palliyaguru, N. 2017, *The Astronomical Journal*, 840, doi:10.3847/1538-4357/aa6aa9
- Daubechies, I. 1988, *Communications on Pure and Applied Mathematics*, 41, 909
- Daubechies, I., et al. 1992, *Ten lectures on wavelets* (Philadelphia: SIAM), doi:10.1137/1.9781611970104
- Daubechies, I. 1990, *Information Theory, IEEE Transactions on*, 36, 961
- Deneva, J. S., Cordes, J. M., McLaughlin, M. A., et al. 2009, *ApJ*, 703, 2259
- Donoho, D., & Johnstone, I. M. 1994, *Biometrika*, 81, 425
- Fodor, I. K., & Kamath, C. 2003, *Journal of Electronic Imaging*, 12, 151
- Goswami, J. C., & Chan, A. K. 1999, *Fundamentals of Wavelets: Theory, Algorithms, and Applications* (New York: Wiley-Interscience), doi:10.1002/9780470926994
- Grossmann, A., & Morlet, J. 1984, *SIAM journal on mathematical analysis*, 15, 723
- Haar, Z. 1910, *Mathematische Annalen*, 69, 331
- Hobbs, G. B., Edwards, R. T., & Manchester, R. N. 2006, *MNRAS*, 369, 655
- Jacoby, B. A., Bailes, M., Ord, S., et al. 2009, *ApJ*, 699, 2009
- Keane, E. F., Ludovici, D. A., Eatough, R. P., et al. 2010, *MNRAS*, 401, 1057
- Keane, E. F., & McLaughlin, M. A. 2011, *Bulletin of the Astronomical Society of India*, 39, 333
- Keith, M. J., Eatough, R. P., Lyne, A. G., et al. 2009, *MNRAS*, 395, 837
- Larsson, S. 1996, *Astronomy and Astrophysics Supplement Series*, 117, 197
- Lorimer, D., & Kramer, M. 2005, *Handbook of Pulsar Astronomy* (Cambridge: Cambridge University)
- Mallat, S. 1999, *A wavelet tour of signal processing* (Burlington: Academic press)
- Mallat, S. G. 1989, *Pattern Analysis and Machine Intelligence*, 11, 674
- Manchester, R. N. 2001, *PASA*, 18, 1
- Matlab. 2015, *MATLAB Wavelet Analysis Toolbox, v.8.5*, Matlab
- McLaughlin, M. A., Lyne, A. G., Lorimer, D. R., et al. 2006, *Nature*, 439, 817
- Strang, G., & Nguyen, T. 1996, *Wavelets and Filter Banks* (Wellesley: Wellesley-Cambridge Press)