The Astrophysical Journal, 962:131 (10pp), 2024 February 20 https://doi.org/10.3847/1538-4357/ad1ce7

Redeveloping a CLEAN Deconvolution Algorithm for Scatter-broadened Radio Pulsar Signals

Olivia Young[1,2,3] and Michael T. Lam[1,2,4]
[1] School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA; ory3002@rit.edu
[2] Laboratory for Multiwavelength Astrophysics, Rochester Institute of Technology, Rochester, NY 14623, USA
[3] National Radio Astronomy Observatory, 520 Edgemont Rd., Charlottesville, VA 22903, USA
[4] SETI Institute, 339 N Bernardo Ave. Suite 200, Mountain View, CA 94043, USA

Received 2023 June 9; revised 2024 January 8; accepted 2024 January 8; published 2024 February 14

Abstract

Broadband radio waves emitted from pulsars are distorted and delayed as they propagate toward the Earth due to interactions with the free electrons that compose the interstellar medium (ISM), with lower radio frequencies being more impacted than higher frequencies. Multipath propagation in the ISM results in both later times of arrival for the lower frequencies and causes the observed pulse to arrive with a broadened tail described via the pulse broadening function. We employ the CLEAN deconvolution technique to recover the pulse broadening timescale and by proxy the intrinsic pulse shape. This work expands upon previous descriptions of CLEAN deconvolution used in pulse broadening analyses by parameterizing the efficacy on simulated data and developing a suite of tests to establish which of a set of figures of merit leads to an automatic and consistent determination of the scattering timescale and its uncertainty. We compare our algorithm to the cyclic spectroscopy method of estimating the scattering timescale, specifically to the simulations performed in Dolch et al. (2021). We test our improved algorithm on the highly scattered millisecond pulsar J1903+0327, showing the scattering timescale to change over years, consistent with estimates of the refractive timescale of the pulsar.

Unified Astronomy Thesaurus concepts: Radio pulsars (1353); Interstellar medium (847); Deconvolution (1910); Interstellar scattering (854)

1. Introduction

Radio pulsars provide unique probes of the ionized interstellar medium (ISM) and allow us to gain insight into its structure and variability by modeling the effects of the delays and distortions on the emitted radio pulses as observed at the Earth (Lorimer & Kramer 2004). While delays due to dispersion are routinely modeled in pulsar timing experiments (e.g., Verbiest et al. 2016), distortions due to multipath propagation are not, and it can be difficult to do so (Shannon & Cordes 2017). Determining the distortion level is difficult due to both the intrinsic pulse shape and the underlying geometry and spectrum of the turbulent medium being unknown (Cordes et al. 1986; Cordes & Rickett 1998), and the time- and path-dependent variations in the observed pulse broadening function (PBF; Williamson 1972). Not only can separating these effects yield important insights into the nature of the ionized ISM but it can also provide proper mitigation of pulse-profile-based impacts on pulsar arrival times used in precision timing experiments such as low-frequency gravitational-wave detectors (Stinebring 2013).

CLEAN deconvolution, originally developed for radio interferometric imaging (Högbom 1974), was applied to radio pulses in Bhat et al. (2003) to recover both the pulse broadening (scattering) timescale τ_d and the intrinsic shape simultaneously via the use of an assumed PBF. Unlike in synthesis imaging where the positions of the array elements are known while the sky brightness distribution is not, neither the analogous PBF nor intrinsic pulse shape, respectively, are known. Bhat et al. (2003) introduced figures of merit (FOMs) to iteratively test trial values of τ_d under an assumed PBF, demonstrating variation in the rebuilt intrinsic pulses for PSR J1852+0031 for different PBFs and application to several other pulsars.

We expand upon the CLEAN deconvolution algorithm presented in Bhat et al. (2003) to prepare for *automated* deployment on data sets of significantly more pulsars. In this work, we primarily focus on the broadening effects of the ISM and recovering τ_d with the intention of applying the algorithm to the multifrequency profiles of pulsars distributed throughout the galaxy to understand both the bulk properties of the turbulence in the ISM and specific, unique lines of sight. Understanding these properties informs priors on pulsar timing arrays and other high-precision pulsar timing experiments in which scattering biases estimates of the arrival times (Lentati et al. 2017). This work is the first of several papers on robust method development and deployment on real data from a larger selection of pulsar observations.

In Section 2, we describe and expand upon the CLEAN deconvolution method introduced by Bhat et al. (2003). In Section 3, we perform systematic tests on simulated data, demonstrating the level of recall in the input τ_d values and quantifying our uncertainties in the estimates. We also compare our results with the cyclic spectroscopy (CS) deconvolution technique Demorest (2011) and discuss the trade-off of limitations in our method with the extensive computational complexity of the CS method. Finally, we apply our method to PSR J1903+0327 in Section 4 and discuss our future directions in Section 5.

2. The CLEAN Deconvolution Algorithm

CLEAN deconvolution for radio pulsars exploits the one-dimensional nature of pulsar profiles and differs from traditional CLEAN approaches where the instrumental response function is known. The analogous function in this work, the PBF, must be assumed from a priori models. Bhat et al. (2003) developed a method that can both determine the pulse broadening timescale $\tau_{\rm d}$ and recover the intrinsic pulse from observational pulsar profile data via the employment of the CLEAN deconvolution algorithm and FOMs. CLEAN can be applied using different models of the PBF of the ISM, making it a broadly encompassing method. In this work, we assumed the PBF for the commonly used thin-screen approximation for the ISM's geometry.

Bhat et al. (2003) describe the CLEAN algorithm for use in the deconvolution of radio pulsar pulses, along with the development of five FOMs used to determine the correct broadening timescale from a set of test values. In this section, we discuss the algorithm both as originally described and how the algorithm has been redeveloped for this work.

2.1. Modeling the Observed Pulse Profile

We assumed the observed pulse $y(t)$ to result from the convolution of the intrinsic pulse $x(t)$, the PBF $g(t)$, and the instrumental response function $r(t)$, given by

$$y(t) = x(t) \otimes g(t) \otimes r(t). \tag{1}$$

We simulated our intrinsic pulse $x(t)$ as a normalized, single-peaked Gaussian shape, which minimizes the asymmetry of the rebuilt pulse. We will further discuss the limitations of this choice in Section 2.3.1. This intrinsic pulse is by nature noiseless, as it simulates the pulse as seen at the pulsar.

The PBF for the ISM is commonly modeled as a thin screen (Cordes & Rickett 1998) for simplicity. The thin-screen approximation simplifies calculations, separating the physical turbulent processes from the geometry of the intervening gas, and, in the case of the PBF, simplifies the form as well; the thin-screen model works reasonably well for lines of sight with a single overdense region. We used this model in our work, given by

$$g(t|\tau_{\rm d}) = \frac{1}{\tau_{\rm d}} \exp\left(-\frac{t}{\tau_{\rm d}}\right) U(t), \tag{2}$$

where $U(t)$ is the Heaviside step function.

Lastly, the instrumental response function, denoted as $r(t)$, determines the resolution of the observed data. We assumed a delta function as an approximation for the instrumental response function with a width of one phase bin.[5]

2.2. CLEAN Deconvolution

CLEAN iteratively subtracts replicated components from an observed pulse until the residual structure falls below the rms of the off-pulse noise. As we do not know the value of $\tau_{\rm d}$ a priori, this iterative subtraction process is repeated for a range of test $\tau_{\rm d}$ values, with the assumed correct $\tau_{\rm d}$ chosen using FOMs. For the purposes of the algorithm, we treat $\tau_{\rm d}$ to be measured in time-bin resolution units as measured across the folded pulse's phase with N_ϕ total bins. We step through our CLEAN deconvolution process below.

1. *CLEAN component creation.* We first identify the location of the maximum of the deconstructed pulse after the ith iteration, $t_i \equiv \mathrm{argmax}\,[y_i(t)]$; our first iteration begins with the originally observed pulse $y_0(t)$. Each CLEAN component (CC) $y_c(t|t_i)$ starts with a delta function, $\delta(t - t_i)$, at the location of the maximum of the observed pulse, $\max\,[y_i(t)]$, multiplied by the loop gain value, γ, i.e.,

   $$y_c(t|t_i) = \gamma\{\max\,[y_i(t)]\}\,\delta(t - t_i) \equiv C_i\,\delta(t - t_i). \tag{3}$$

   Smaller loop gains result in a greater number of iterations before the stopping criterion is met but allow for finer intrinsic features to be resolved (Högbom 1974); in this work, we used $\gamma = 0.05$ (as in Bhat et al. 2003).
2. *Iterative subtraction off the main pulse.* After we construct $y_c(t|t_i)$, we convolve the CC with the instrumental response function $r(t)$ and the PBF with a given test $\tau_{\rm d}$, and then subtract this shape from the ith iteration pulse. The change in the profile at each iteration is described as

   $$\Delta y_i(t) = y_i(t) - \{y_c(t|t_i) \otimes [g(t|\tau_{\rm d}) \otimes r(t)]\}, \tag{4}$$

   with $y(t_i)$ as the input pulse profile to the ith iteration. The resulting subtracted profile then becomes the pulse profile for the next CLEAN iteration so that

   $$y_{i+1}(t) = \Delta y_i(t). \tag{5}$$
3. *Termination of the CLEAN algorithm.* The CLEAN algorithm is terminated when the maximum of the input pulse profile falls below the rms of the off-pulse noise, i.e., $\max\,[y_i(t)] \leqslant \sigma_{\rm off}$. Here, $\sigma_{\rm off}$ is calculated using the Python package PyPulse (Lam 2017), which calculates the off-pulse region by finding the region of the pulse profile that best minimizes the area under the curve.

The CLEAN algorithm above will provide the list of CCs along with the residual noise. The CCs can be used to reconstruct the intrinsic pulse shape, but for the purposes of this work, our final goal was to determine $\tau_{\rm d}$. The algorithm can run with any input value of $\tau_{\rm d}$; therefore our iterative method is repeated with different trial $\tau_{\rm d}$, from which we derived FOMs based on the reconstructed intrinsic pulse shape and the residual noise that resulted from each trial $\tau_{\rm d}$.

2.3. Figures of Merit

We employed six FOMs as follows: a measure of the positivity of the residual noise ($f_{\rm r}$), a measure of the skewness of the recovered intrinsic pulse (Γ), a count of the on-pulse-region residual points below the off-pulse noise level ($N_{\rm f}/N_\phi$), a measure of the ratio of the rms of the residual noise to the off-pulse noise rms ($\sigma_{\rm offc}/\sigma_{\rm off}$), a measure of the combined positivity and skewness measure ($f_{\rm c}$), and a count of the number of CCs each test $\tau_{\rm d}$ uses before the peak of the profile falls below the noise level ($\mathrm{N_{iter}}$). All except the last were used directly in Bhat et al. (2003), although this FOM was suggested by the authors. These six FOMs fall into three broad categories: figures based on the rebuilt intrinsic pulse, figures based on the residual noise after the CLEAN algorithm terminates, and a

---

[5] For clarity, we use the digital signal processing definition of the unit-height sample function being $\delta(t) = 1$ if $t = 0$, otherwise 0, which allows us to multiply by a constant as in Equation (3).

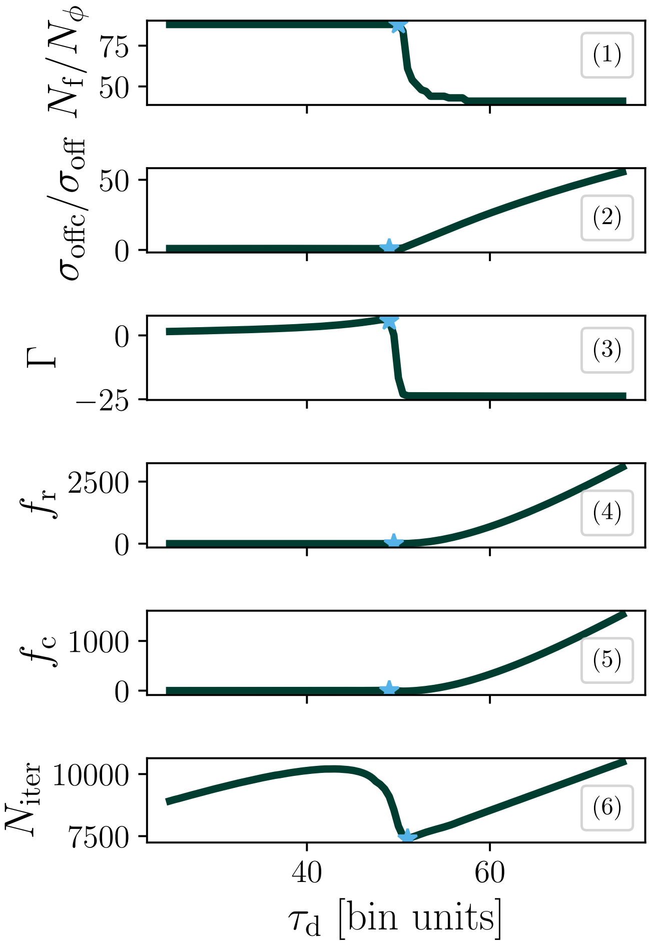

Figure 1. Summary of FOMs used in this work, for a pulse with simulated FWHM of 100 phase bin units, input $\tau_d = 50$ phase bins, and S/N = 100. We tested τ_d values ranging from 25 to 75 bin units with a step size of one bin. The blue star in each panel indicates the location of $\tau_d = 50$. Panel (1) shows the number of data points within a 3σ level of the noise FOM, panel (2) shows the rms FOM, panel (3) shows the skewness FOM, panel (4) shows the positivity FOM, panel (5) shows the combined skewness and positivity FOM, and panel (6) shows the number of iterations FOM.

figure based on the number of CCs generated before the algorithm terminates. We describe the FOMs grouped into these three categories in the sections below.

In Figure 1, we see the ideal result of the use of six FOMs and the methods for determining the "correct" τ_d.

2.3.1. FOM Measuring the Shape of the Rebuilt Intrinsic Pulse

We examine the CC amplitudes, C_i, and locations, t_i, found during the CLEAN process (e.g., see Equation (3)) to compute the Γ FOM. In our simulations, we created intrinsic pulses that are symmetric Gaussians, and therefore the correct rebuilt pulse should always be a perfectly symmetric Gaussian if the correct τ_d is used. In reality, intrinsic pulses may not be perfectly symmetric, and we discuss these implications in Section 5.

The Γ of the rebuilt pulses is calculated for each test τ_d by computing the third standardized moment, where n_c is the number of CCs:

$$\Gamma = \frac{\langle t^3 \rangle}{\langle t^2 \rangle^{3/2}}, \tag{6}$$

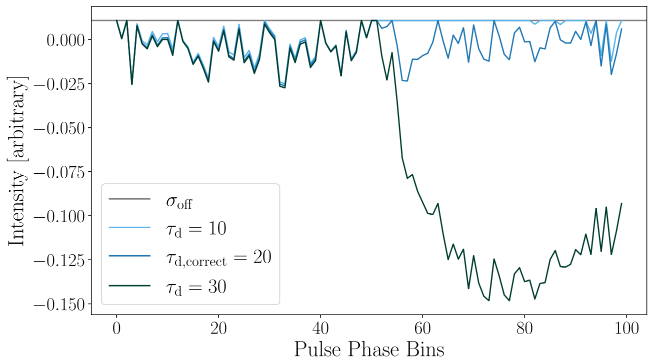

Figure 2. The residual noise left over after the CLEAN algorithm terminates for three test τ_d values of 10, 20, and 30 bins, where 20 is the simulated value. These time series are representative of the residuals used to calculate multiple FOMs. We can see the under- and oversubtraction for test τ_d values that are smaller than or larger than the true τ_d, respectively.

where $\langle t^n \rangle$ is

$$\langle t^n \rangle = \frac{\sum_{i=1}^{n_c} (t_i - \bar{t})^n C_i}{\sum_{i=1}^{n_c} C_i} \tag{7}$$

and $\bar{t}$ is

$$\bar{t} = \frac{\sum_{i=1}^{n_c} t_i C_i}{\sum_{i=1}^{n_c} C_i}. \tag{8}$$

The resulting Γ is ideally represented by the example in panel (3) of Figure 1, where the sharp fall-off point represents the general location of the correct τ_d.

2.3.2. FOMs Based on the Residual Noise

Three of our FOMs are built from measures of the residual noise after the completion of the CLEAN algorithm. We will also discuss a FOM that combines one of these FOMs (positivity) with the Γ FOM discussed previously—this is an important FOM as described in Bhat et al. (2003).

The residual noise is one of the end products of the CLEAN deconvolution process. A test τ_d that is larger than the correct value of τ_d results in a progressively larger oversubtraction, as shown in Figure 2. If the test τ_d is smaller than the correct value, it results in an unremoved noise floor in the baseline.

We can first calculate the rms of the residual noise,

$$\sigma_{offc} = \sqrt{\frac{1}{N_\phi} \sum_{j=1}^{N_\phi} [\Delta y_i(t_j|\tau_d)]^2}, \tag{9}$$

in comparison to the rms of the off-pulse region, σ_{off}, where this ratio, $\sigma_{offc}/\sigma_{off}$, will grow whenever over- or undersubtraction is performed and should otherwise approach a value of 1 for the appropriate subtraction. We find this is roughly equivalent to the single metric used to automatically determine τ_d used in Tsai et al. (2017) for multifrequency data from 347 pulsars.

Beyond the rms, we can count the total number of residual noise points, N_f, within a certain threshold level (we chose $3\sigma_{off}$) of the noise that satisfies the condition

$$|y_i - y_{off}| \leqslant 3\sigma_{off}. \tag{10}$$

As seen in Figure 2, for undersubtraction we expect all of the points to satisfy the condition and so the ratio $N_f/N_\phi = 1$, but the ratio will drop as oversubtraction occurs.

Besides these two metrics which measure deviations away from the rms noise, we also wish to enforce nonnegativity of the residual profile since we know pulsar signals must be above the baseline noise. A $f_{\rm r}$ FOM was defined by Bhat et al. (2003) in terms of a sum over the N_ϕ bins of the residual noise:[6]

$$f_{\rm r} = \frac{1}{N_\phi \sigma_{\rm off}^2} \sum_{j=1}^{N_\phi} [\Delta y_i(t_j|\tau_{\rm d})]^2. \tag{11}$$

If $\Delta y_i(t)$ is Gaussian white noise with an rms equal to $\sigma_{\rm off}$, then, as with the previous FOM, we would expect $f_{\rm r} \approx 1$, while oversubtraction would force the sum to increase well beyond 1.

Bhat et al. (2003) defined the $f_{\rm c}$ FOM, equally weighting the rebuilt intrinsic pulse shape and the residual noise by

$$f_{\rm c} = \frac{\Gamma + f_{\rm r}}{2}, \tag{12}$$

thus providing higher confidence in test $\tau_{\rm d}$ values with favorable values of both skewness and positivity. The typical shape of this FOM is shown in panel (5) of Figure 1.

2.3.3. FOM Measuring the Number of Iterations Performed

We developed this FOM to more directly measure the fit of the reconvolved CCs' broadening tails to the broadening of the observed pulse. As the amplitude for reconvolved CCs with larger broadening tails is smaller than those with smaller broadening tails (due to the normalization in Equation (2)), we expect a general increase in the number of iterations needed to deconvolve the observed pulse. Similarly, when reconvolved CCs with smaller broadening tails are subtracted from a pulse with a larger true broadening tail, more iterations will be required. However, when the CCs are convolved with the correct value of $\tau_{\rm d}$, neither under- nor oversubtraction occurs, resulting in fewer iterations being needed. Therefore, we expect a dip in our FOM around the correct value of $\tau_{\rm d}$.

2.4. Automating the Choice of the Correct τ_d Value

These FOMs were originally constructed to pinpoint the correct value of $\tau_{\rm d}$ by eye, with all previous implementations of CLEAN using by-eye methods of determining $\tau_{\rm d}$. This approach is impractical for large data sets, so we automated this process. We found that the simple approach of computing the numerical third derivative of each FOM with respect to $\tau_{\rm d}$ and finding the maximum has yielded good results, though the exact recall depends on both the value of $\tau_{\rm d}$ and the pulse signal-to-noise ratio (S/N). More complicated algorithms will be employed in future works, but the systematic error introduced by this choice is small in comparison to other noise sources, as shown next, so we opted to use it.

3. Automated Algorithm Performance

In this section, we will discuss the performance of our automated CLEAN method. An in-depth description of our redeveloped CLEAN algorithm in Python, as well as notes on how to use the open source versions available at doi:10.5281/zenodo.10064009, can be found in Young (2022).

We wished to robustly quantify the "correctness" of our $\tau_{\rm d}$ estimates in simulated data so that we could automatically assign uncertainties to our estimates on real data. To that end, we simulated multiple data sets with different input parameters to determine how these will affect the recall. Ideally, as in Dolch et al. (2021) for the CS algorithm (Demorest 2011), only the S/N and $\tau_{\rm d}$ of a profile should affect the recall accuracy of our CLEAN deconvolution, though we tested several other parameters as well.

To quantify the algorithm's performance, we computed a measure of the relative bias of the estimator.

Within this work, the values returned for each FOM have been weighted based on our parameterization work on simulated data.[7] When calculating our relative biases, defined as the fraction of the returned $\tau_{\rm d}$ to the correct injected $\tau_{\rm d}$, our relative biases are defined as

$$\epsilon_{\rm ave} = 1 - \frac{1}{N_{\rm runs}} \sum_{i=1}^{N_{\rm runs}} \left(\frac{\epsilon_i}{\tau_{\rm d}} \right), \tag{13}$$

where $N_{\rm runs}$ is the number of simulations for a given data set and ϵ_i, for readability, is defined as a weighted average of the $\tau_{\rm d}$ values returned by our FOMs (all six indexed with k) for each run ($i = 1 \ldots N_{\rm runs}$),

$$\epsilon_i = \frac{\sum_k w_k \tau_{k,i}}{\sum_k w_k}, \tag{14}$$

where our w_k values are normalized weights, defined as follows. We first calculated the weight by the standard $w_k = 1/\sigma_k^2$, where σ_k is the standard deviation of the 60 returned values for each FOM from a given run on simulated data—a greater spread in returned values for a given FOM downweights its importance in the end bias measurement. In practice, after our weights were calculated, we divided by the sum of the weights to normalize them, though this does not change the result in Equation (14).

3.1. Testing the Impact of S/N and τ_d on Recall

We first tested how CLEAN performs based on different injected pulse S/N (20, 70, 650, 2600) and $\tau_{\rm d}$ (1, 2, 4, 8, 16, 32, 64, 128, 256 μs) combinations. We simulated data sets using several of the characteristics of PSR B1937+21 (the first-known millisecond pulsar and a known scattered source) as follows; these parameters are also shown in Table 1. PSR B1937+21 has a spin period of 1.557 ms and an FWHM of 38.2 μs (Manchester et al. 2013). To reduce computing time, we used different numbers of phase bins depending on the injected $\tau_{\rm d}$ value, as shown in Table 2; we show in the next subsection that there is minimal impact in the recovery of $\tau_{\rm d}$ depending on the phase resolution of the pulses so long as the scattering tails are resolved. For each of our runs, we tested across 100 equally spaced $\tau_{\rm d}$ steps between 0.5 and 1.5 times the injected $\tau_{\rm d,correct}$. For each S/N–$\tau_{\rm d}$ pair, we simulated and ran CLEAN on 60 simulated pulse profiles. We computed the weighted average

---

[6] Bhat et al. (2003) introduced a multiplicative weight of order unity but did not specify the value. Here, we take that weight to be 1 and ignore introducing it in the main text. They also include a Heaviside step function, $U_{\Delta y}$. As this only changes the overall normalization of our FOM in our simulation runs, we ignore this in our work. It should be noted that this decision in practice makes $f_{\rm r}$ equivalent to $\sigma_{\rm offc}^2/\sigma_{\rm off}^2$.

[7] While each individual FOM described within this work is given its own weight based on our parameterization results, due to our $f_{\rm c}$ FOM being a combined measure of Γ and $f_{\rm r}$, the performance of these two FOMs are in practice weighted more heavily.

Table 1
Automated Algorithm Simulation Parameters

Parameter	Value
Spin period	1.557 ms
Pulse FWHM	38.2 μs
N_ϕ for $\tau_{\rm d} = 1, 2, 4\ \mu$s	2048
N_ϕ for $\tau_{\rm d} = 8, 16, 32\ \mu$s	1024
N_ϕ for $\tau_{\rm d} = 64, 128, 256\ \mu$s	512
Test $\tau_{\rm d}$ range	$(0.5\tau_{\rm d,correct}$–$1.5\tau_{\rm d,correct})$
Number of steps in test $\tau_{\rm d}$ array	100
S/N	20, 70, 650, 2600

bias as before but for all of the tests. We visualize the performance of each individual FOM in Figure 4.

We choose our S/N–$\tau_{\rm d}$ values in the same way as Dolch et al. (2021) to more directly compare our CLEAN deconvolution method with the CS algorithm. While the pulse profiles used for CLEAN deconvolution algorithms are created by folding filterbank data, CS uses raw voltage data to compute a folded cyclic spectrum, and is later deconvolved to recover an impulse response function using amplitude and phase information, making the latter more computationally intensive (though assuming no specific PBF). These methods are therefore difficult to directly compare, but we can expect to see improved performance for both methods as either S/N or $\tau_{\rm d}$ increase.

Indeed, after running our simulations, we see this expected behavior in Figure 3, where darker colors indicate better recovery of $\tau_{\rm d}$, which matches what is seen in Dolch et al. (2021) for CS. The numerical values shown in Figure 3 are in terms of the percentage of the correct $\tau_{\rm d}$, given by $\epsilon_{\rm ave}$.

We can further compare the performance of CS and our CLEAN deconvolution more quantitatively by finding how our algorithm's performance changes as a function of S/N and $\tau_{\rm d}$ as compared to CS. Dolch et al. (2021) performed an unweighted linear regression on their performance data with respect to S/N and $\tau_{\rm d}$, finding their cyclic merit quantity improves in performance equally when S/N or $\tau_{\rm d}$ is increased. Following suit, we find the slope of our algorithm's performance with respect to S/N and $\tau_{\rm d}$ via linear regression in the form of Equation (16) in Dolch et al. (2021):

$$\log_{10}(\epsilon_{\rm ave}) = \beta \log_{10}(\tau_{\rm d}) + \gamma \log_{10}({\rm S/N}) + {\rm C}. \tag{15}$$

After fitting our $\epsilon_{\rm ave}$ values, as in Figure 3, we found a slope ratio (γ/β) of 0.7, compared to ∼1 for CS. Therefore, unlike CS, CLEAN's performance improves more strongly in the $\tau_{\rm d}$ dimension than in the S/N dimension. While this may seem to imply that CLEAN will overtake CS in its accuracy for sufficiently large $\tau_{\rm d}$ and fixed S/N, both methods are bounded by scattering $\tau_{\rm d} \lesssim P$, the pulse period, and so in practice the observable $\tau_{\rm d}$ is limited.

In Figure 4, we show how well each FOM performs, with each panel showing the recovery over the full range of S/N–$\tau_{\rm d}$ pairs. Smaller dots indicate smaller relative biases, and thus a more accurate performance. Poor performance from one FOM will impact our averaged recall if an unweighted average is employed, therefore we used these results to implement weights for our FOMs. We see that, in general, the performance of each FOM improved with higher S/N or $\tau_{\rm d}$ like the average, though not all behave equally. For example, the $N_{\rm f}/N_\phi$ and $\sigma_{\rm offc}/\sigma_{\rm off}$ appear to perform better at somewhat lower $\tau_{\rm d}$ than the other FOMs. While the skewness Γ does not perform as well at the high S/N–$\tau_{\rm d}$ end, it does perform marginally better than the previously mentioned two FOMs. Our method of weighting these FOMs using the variance of the returned relative biases results in a moderate performance increase for lower values of S/N–$\tau_{\rm d}$, and up to 20% at large values of S/N–$\tau_{\rm d}$ when compared to an unweighted approach. This method in practice allows us to weight our $N_{\rm f}/N_\phi$ and $\sigma_{\rm offc}/\sigma_{\rm off}$ FOMs more heavily at values of S/N–$\tau_{\rm d}$ where we see better performance in Figure 4.

3.2. Testing Secondary Parameter Contributions to Recall Error

While we assumed the main contributors to the effectiveness of our algorithm to be our primary parameters, S/N and $\tau_{\rm d}$, we wanted to ensure that secondary parameters were not significant contributors to our recall error. We created a small-scale parameterization set via simulation of a base pulse profile with $\tau_{\rm d} = 256$ bins and S/N = 2600. We chose very large values for both $\tau_{\rm d}$ and S/N as the method was able to reliably recall the correct $\tau_{\rm d}$ for large $\tau_{\rm d}$ and S/N values (see Section 3.2). This set was used to determine how the number of bins in our observation, the FWHM of the intrinsic pulse, and the user-defined step size and range of the test $\tau_{\rm d}$ array affected the algorithm's performance. Additionally, our previous data set assumed an intrinsic pulse with similar parameters to B1937+21 only. Therefore, an additional motivation for probing these secondary parameters was to determine if we could extrapolate our results to observations of other pulsars with varying FWHMs and numbers of phase bins.

For these parameterization runs we used the base values shown in the second column of Table 2 and iterated over the values shown in the third column. We ran 20 simulations for each variation, which gave insight into these parameters' contribution to our recall and allowed for exploration into the expected larger contributions of $\tau_{\rm d}$ and S/N to the recall error.

As many pulse profiles are recorded with a different number of phase bins (see, e.g., Lorimer et al. 1997), we tested to see how the phase resolution of the observation affected our recall. We simulated data sets with N_ϕ ranging from 128 to 2048. In Figure 5, we see good agreement between the ranges of individual recall values returned for each phase bin number tested (green circles) and the averages (green squares) varying within only 1% across our tested values when employing our FOM weights found for S/N = 2600 and $\tau_{\rm d} = 256$. Thus, the minor variations in the average recalls could be explained by our limited number of runs resulting in incomplete coverage of the algorithm's performance. Therefore, within this limited number of performed simulations and phase bins tested, we assume that the number of phase bins in the observed pulse profile is not a large contributor to our total recall.

Results of testing how the FWHM of the intrinsic pulse affected our recall are shown in Figure 6, and reveal a large, though not unexpected, range in $\epsilon_{\rm ave}$ between the FWHMs tested. As the FWHM increases, the pulse takes up an increasing fraction of the number of bins in the pulse profile, thus making the CLEAN cutoff criterion of falling below the off-pulse noise level less effective. This result agrees with the findings of Jones et al. (2013), where they found the CS method to be less effective on wider pulses. We have tested the efficacy of our CLEAN algorithm over a large range of FWHMs, including values much larger than the average FWHM of known pulsars. Based on current values available

Table 2
Secondary Parameters Tested

Parameter	Base Value	Values
Number of phase bins	256	{128, 256, 512, 1024, 2048}
FWHM	$\frac{1}{8}N_\phi$	$[\frac{1}{64}, \frac{1}{32}, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}]N_\phi$
Range of τ_d	$(0.5\ \tau_{d,correct}–1.5\ \tau_{d,correct})$	$(0.1\ \tau_{d,correct}–\tau_{d,correct})$, $(0.1\ \tau_{d,correct}–2.0\ \tau_{d,correct})$, $(0.4\ \tau_{d,correct}–1.6\tau_{d,correct})$, $(0.5\ \tau_{d,correct}–1.5\ \tau_{d,correct})$
Number of steps in τ_d array	100	[10, 20, 50, 100, 200]

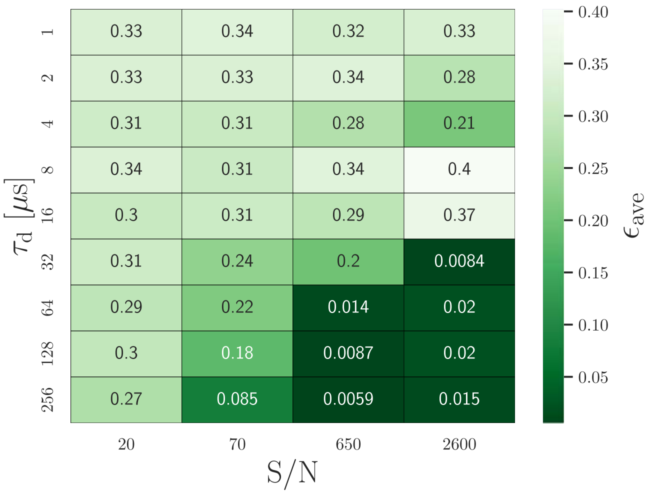

Figure 3. Average recall relative biases for CLEAN deconvolution in the same style as Figure 5 from Dolch et al. (2021). This plot gives an overview of the performance of the CLEAN algorithm by returning the average size of the biases with each S/N and τ_d pair. As smaller biases indicate better performance, CLEAN performs better on simulated data with larger values of both S/N and τ_d.

on the ATNF Pulsar Catalogue (Manchester et al. 2005) for around 2250 pulsars, 82% have FWHM values of one-sixteenth of the number of phase bins or less, 93% have FWHM values of one-eighth of the number of phase bins or less, and only one pulsar, J1806−1920, has an FWHM greater than half of the number of phase bins. While the values available are not all encompassing because FWHM is a radio-frequency-dependent parameter, the ranges that we choose to parameterize our algorithm over encompass the vast majority of known pulsars with FWHM measurements available in the ATNF Catalogue.

In Figure 7, we see the contribution of the number of steps or, interchangeably, the step size of the test τ_d array to our recall error. We included in this analysis a correction factor of $\Delta\tau_d/2$, the largest base error induced due to large step sizes resulting in the correct τ_d not being directly tested. For example, if the correct τ_d is 10.5 μs, and our test τ_d array only samples every $\Delta\tau_d = 1\ \mu$s, an error of 0.5 μs will be introduced; thus, we subtracted this factor of $\Delta\tau_d/2$ from our ϵ_{ave} to more conservatively estimate our uncertainties. The relative biases returned vary within 10%, therefore we concluded that the number of steps in the test τ_d array was not a large contributor to the overall recall error. Note that there is less than a 2% change to the recall between 50 and 200 steps, meaning that we can further decrease computation time without large impacts to our recall by decreasing the number of test τ_d values.

Finally, we parameterized the contribution of the range of test τ_d iterated over to our recall. In Figure 8, we can see that ranges which barely included the correct τ_d (first bar) result in poor performance, as expected. This results from the shapes of the FOMs not being fully covered over the correct injected τ_d. Other ranges that include the correct τ_d have recalls within less than 5% of each other, even when the ranges iterated over are much larger. Therefore, while more computationally intensive, we recommend running CLEAN over a large range of τ_d values to ensure the best estimate is chosen.

With the results of these runs, we see that these secondary effects have some small impact at large S/N and τ_d, but otherwise the most prominent influences on the recall of CLEAN deconvolution are the S/N and τ_d of the data. While there are some variations in the average recall for each of the parameters we tested, the average recalls varied within 10% or less for most tests, with the notable exceptions of large FWHMs and test τ_d ranges that barely include the correct value of τ_d—both as expected. We also conclude that for many typical pulse profile shapes and sizes there will not be a large effect on the recall with secondary parameters within the ranges tested in this work.

4. Applying CLEAN to PSR J1903+0327

To demonstrate the efficacy of our algorithm, we tested CLEAN on real data from the pulsar J1903+0327. PSR J1903+0327 is a millisecond pulsar with a spin period of 2.15 μs and a dispersion measure of 297 pc cm^{-3}. The pulsar also has a solar-mass companion and might be part of a hierarchical triple system (Champion et al. 2008). PSR J1903+0327 has been monitored by pulsar timing array collaborations such as the North American Nanohertz Observatory for Gravitational Waves (NANOGrav; Arzoumanian et al. 2021) in the effort to detect low-frequency gravitational waves. While these collaborations self-select for pulsars with low amounts of pulse broadening (narrower pulses have higher timing precision), PSR J1903+0327 has some of the most prominent scattering in these data sets, with the broadening tail visible by eye. With over a decade of timing data on this pulsar, we analyzed the lowest-radio-frequency pulses in the NANOGrav 12.5 yr data set over time, where broadening is the strongest, to investigate if variations in τ_d are detectable by our algorithm.

Note that the only currently published values for broadening measurements of PSR J1903+0327 come from Geiger & Lam (2022), where values were obtained by convolving Gaussians of fixed width with an exponential broadening function of variable width and comparing to the observed pulses (McKinnon 2014). We find values of τ_d which roughly agree with these results of ∼250 μs at 1200 MHz.

We created six summed profiles on which to deploy our CLEAN algorithm, with one profile corresponding to each year from 2012 to 2017 in our data set. We restricted the frequency band for each observation to 10 MHz centered at 1200 MHz to mitigate any additional broadening introduced

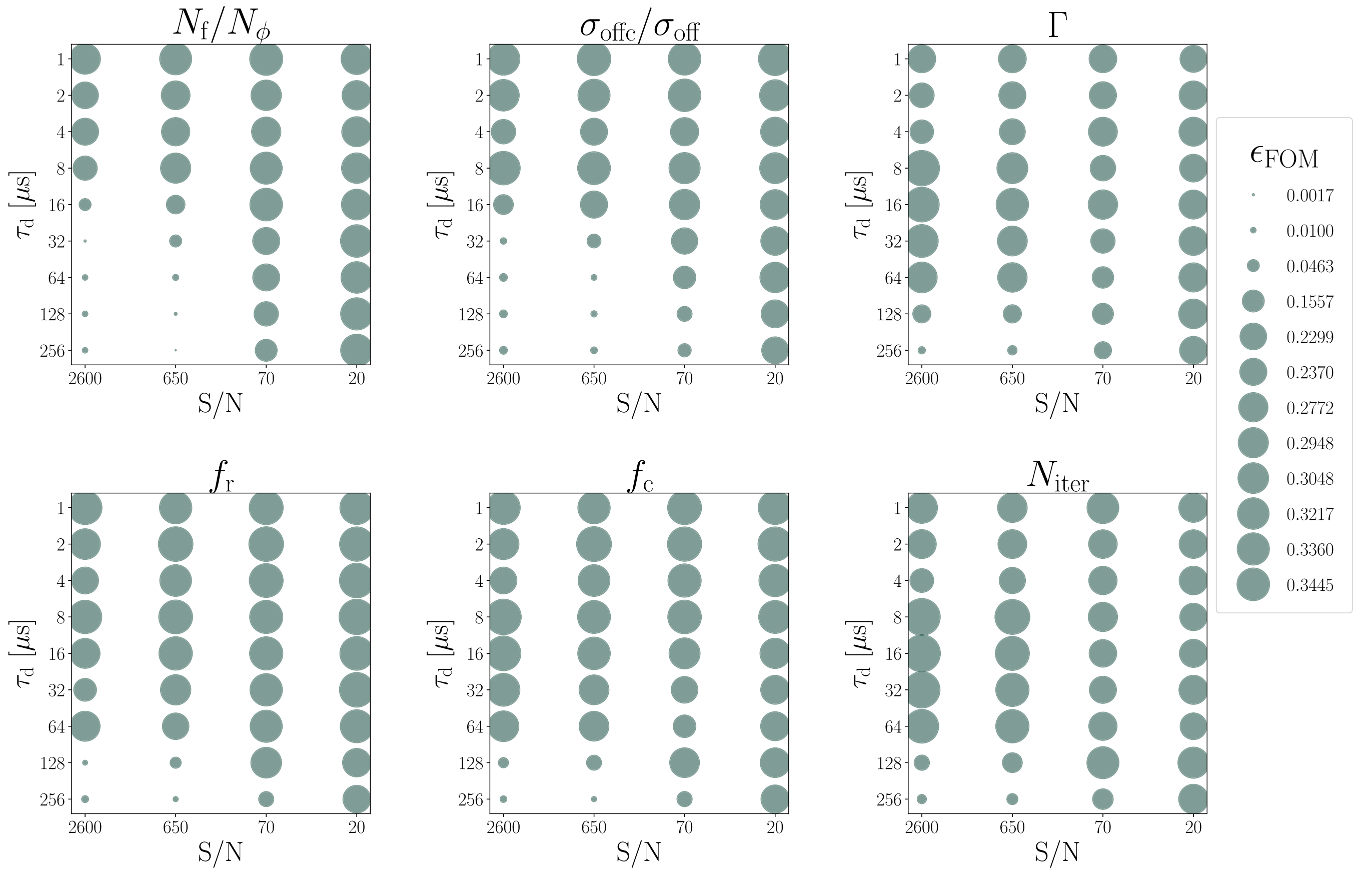

Figure 4. Overview of the relative performance for each FOM. Smaller circles indicate smaller biases or better performance of the FOMs on simulated data. These fractional averages were computed as described in Equation (13), with ϵ_i being composed of the τ_d returned by only one FOM instead of a weighted sum of all returned τ_d values. These values are labeled as $\epsilon_{\rm FOM}$ in this plot. In general, we see better performance for higher values of τ_d and S/N. Interestingly, the N_f/N_ϕ and $\sigma_{\rm offc}/\sigma_{\rm off}$ FOMs appear to perform better than the f_r and the Γ FOMs, which were highlighted in the Bhat et al. (2003) paper.

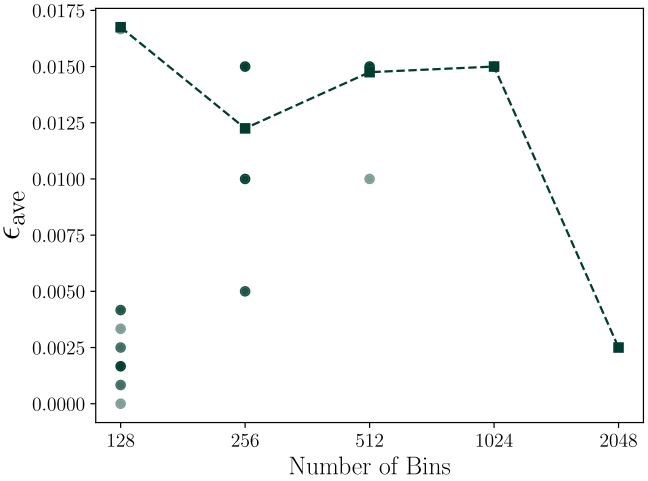

Figure 5. Results of parameterization runs with changing number of phase bins. The y-axis shows the fractional average relative bias size across 20 simulations for each bin value. The average recall error is denoted by the fully opaque green squares connected by the dashed line. The lighter circles indicate the recall error from each run. The average biases range within 1%.

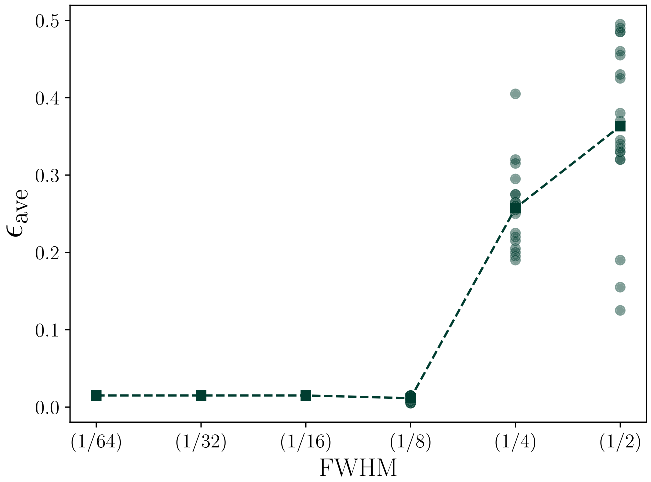

Figure 6. Results of parameterization runs with changing values of the FWHM of the intrinsic pulse, as a fraction of the number of phase bins. We see for FWHMs less than a quarter, the relative bias changes less than 1%, with the $\epsilon_{\rm ave}$ change still under 40% for the FWHM being half the number of phase bins. This effect has been seen with the CS approach as well.

when frequency-averaging the pulses together. Each summed profile consists of 12 monthly observations summed via cross-correlation. Cross-correlation is used to ensure the peaks of our profiles are properly aligned in time before they are summed, resulting in the highest possible S/N of the summed profile. This process was performed iteratively, with each new profile being cross-correlated with and then added to the summed profile.

The refractive timescale of PSR J1903+0327 is estimated to be between 1 and 2 yr (Geiger & Lam 2022). Therefore, summing across 1 yr of observations is consistent with the PBF remaining unchanged across this time span. To further increase

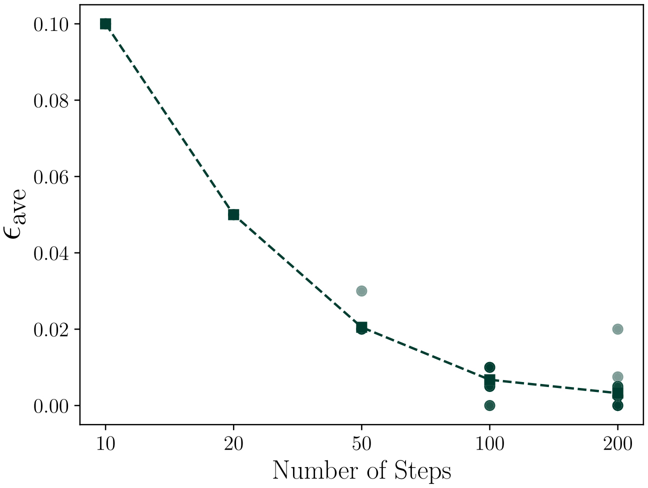

Figure 7. Results of parameterization runs with changing numbers of steps in the test $\tau_{\rm d}$ ranges. Although there is some range in the average errors returned, all values agree within ≈10%.

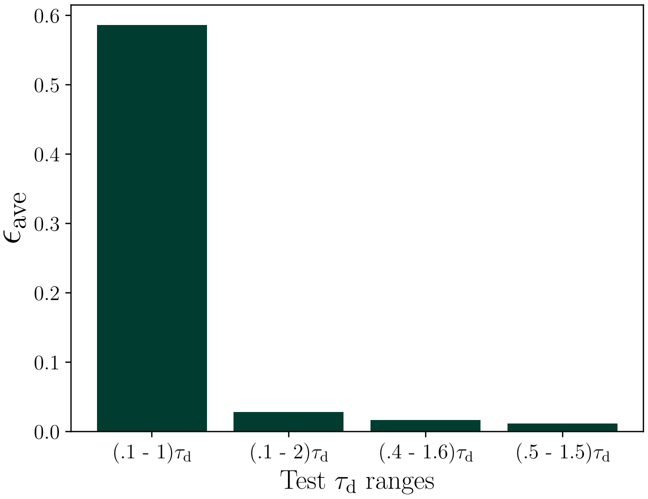

Figure 8. Results of parameterization runs with changing ranges of test $\tau_{\rm d}$ iterated over. We see that as long as the correct $\tau_{\rm d}$ is included within the range of test $\tau_{\rm d}$ iterated over, the range of test $\tau_{\rm d}$ was not a large contributor to the overall recall error.

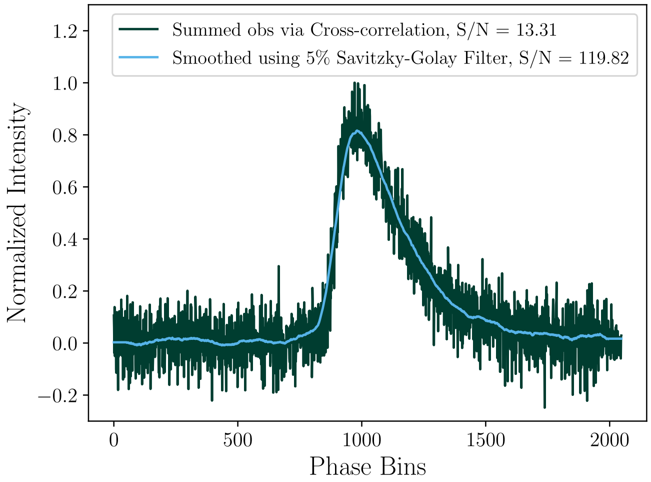

Figure 9. Summed and smoothed pulse profile for PSR J1903+0327 at 1200 MHz from NANOGrav observations during 2014.

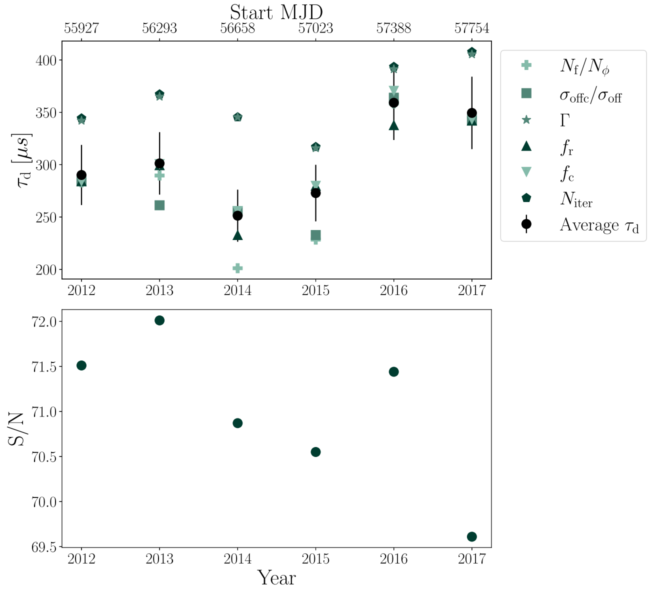

Figure 10. Time series for PSR J1903+0327 from 2012 to 2017 using NANOGrav data. This time series is constructed using a Savitzky–Golay filter with a window filter size chosen to obtain an S/N of around 70.

the S/N values for each profile, we used different Savitzky–Golay filters to smooth the resulting summed profile to the desired S/N level. In Figure 9, we see an example of this summed and smoothed pulse profile.

For our time-series analysis, we used two different filtering techniques, both employing a Savitzky–Golay filter using a polynomial of order zero to fit the profiles, using a filter window size necessary to achieve an S/N of 70 and using a filter window of 5% of the number of bins in pulse profile, in this case 2048 bins, to achieve a higher S/N. We chose to create a time series at two different levels of S/N to showcase the dependence of the algorithm's performance on S/N. We also chose to weight our FOMs using the weights found for $S/N = 70$ and $\tau_{\rm d} = 256\,\mu$s for both time series. We iterated through test $\tau_{\rm d}$ values ranging from 100 to 500 bins for each run, with a step size of one bin. We can see the results of these runs in Figures 10 and 11, where we converted our returned $\tau_{\rm d}$ values into units of microseconds.

We see what is expected in the FOMs for these time series: greater precision and more visibly apparent points of slope change for the high-S/N FOMs. Looking at Figures 12 and 13, we can see examples of how higher S/N results in sharper points of change in the FOMs, thus making choosing the correct $\tau_{\rm d}$ a more precise process. While this is true for all FOMs presented here, this difference can be seen most explicitly in the $\sigma_{\rm offc}/\sigma_{\rm off}$ FOM (panel (2)), where there is a noticeable location where the slope begins increasing in our larger-S/N FOMs versus our lower-S/N FOMs, where there is a more gradual increase in the slope of the $\sigma_{\rm offc}/\sigma_{\rm off}$ FOM, making the correct $\tau_{\rm d}$ more difficult to pinpoint. This increased sharpness of the points of change of our FOMs translated into greater accuracy and better agreement across our FOMs, which can be seen reflected in the tighter clusters around the average returned $\tau_{\rm d}$ values in our time series.

We also note some interesting results of this time-series analysis, particularly the dip in 2015, followed by a drastic increase the following year. However, coupled with the unusual scattering indices measured in Geiger & Lam (2022), it is clear

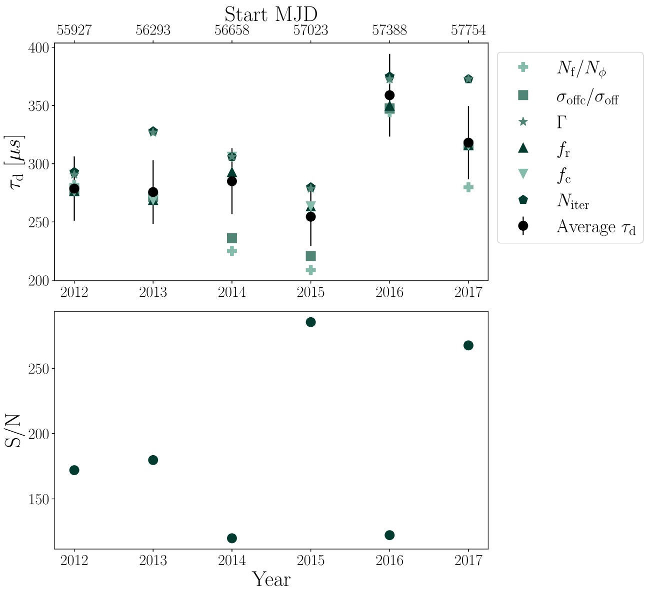

Figure 11. Time series for PSR J1903+0327 from 2012 to 2017 using NANOGrav data. This time series is constructed using a Savitzky–Golay filter with a window filter size of 5% of the length of the observation.

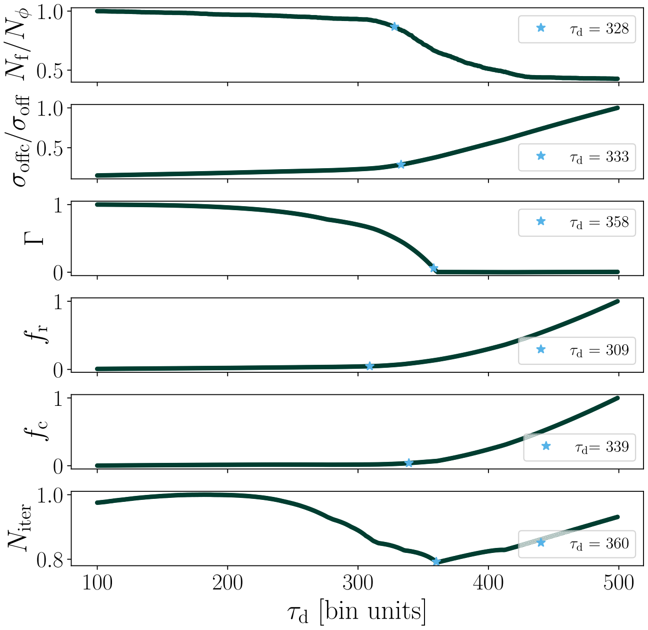

Figure 12. Normalized FOMs for PSR J1903+0327 at 1200 MHz from 2016 at S/N = 70. We see a tight grouping for the returned τ_d values for each FOM, with a mean value of $\tau_\mathrm{d} = 360.3$ μs and an error of 10% based on our simulation runs.

that an exponential PBF is not supported along this line of sight and a more complex model is necessary (A. Geiger et al. 2024, in preparation). Nonetheless, we have shown via this analysis that not only does our CLEAN algorithm perform as expected on observational radio data for pulsar J1903+0327 given our set of assumptions, but also that employment of this algorithm holds potential for scientific insight into the ever-changing ISM.

5. Future Work and Conclusions

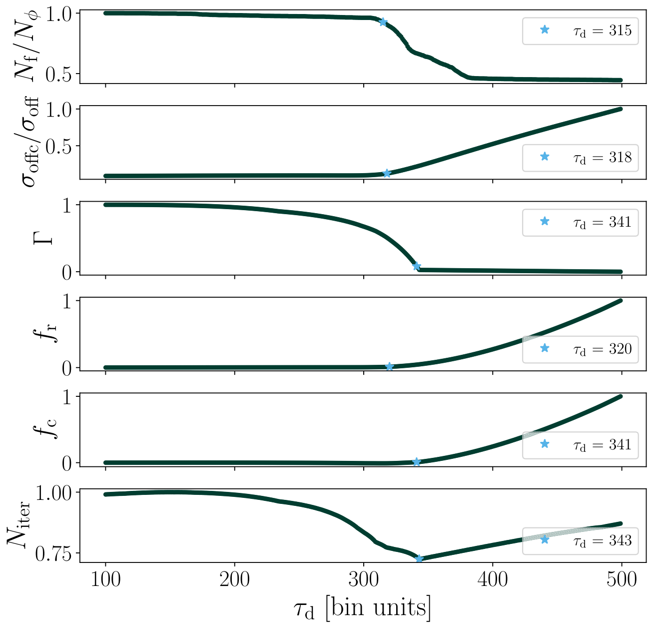

Figure 13. Normalized FOMs for PSR J1903+0327 at 1200 MHz from 2016 for high-S/N values achieved by using a Savitzky–Golay filter with a window size of 5% of the length of the observation. We see an even tighter grouping for the returned τ_d values for each FOM than compared to the lower-S/N FOM, with a mean value of $\tau_\mathrm{d} = 369.3$ μs and an error between 2% and 10% based on our simulation runs.

Within this work, we discussed our motivations, reimplemented CLEAN deconvolution as presented in Bhat et al. (2003), developed weights for our FOMs, discussed the results and products of our implementation of CLEAN, our parameterization work, and results on observational data of PSR J1903+0327. Through our parameterization work, we have concluded that our replicated CLEAN algorithm works as expected: the main factors that influence the recall of the algorithm are the S/N and τ_d of the pulse profile, and higher values of S/N and τ_d result in better recall. We have produced an algorithm that we can confidently deploy on larger sets of observational data within the parameterization bounds described in this work. To that end, we have presented a brief analysis of PSR J1903+0327 at two S/N levels and discussed our findings, showing that our methods prove to be effective on observational data from idealized radio pulsar observations, and can thus provide insight into the time dependence on pulse broadening timescales for similar pulsars via automatic deployment.

Significant future work is required to improve the robustness of this updated technique in estimating the scattering timescale for realistic PBFs and multicomponent pulse profiles, making it more applicable to a larger range of pulsars and providing greater insight beyond a simple measure of the scattering timescale. Moving forward, we aim to further develop our CLEAN algorithm into a broadly applicable tool, focusing on improving upon or removing the need for several simplifications used within this paper's method, primarily the assumption of intrinsically symmetric, single-peak pulse profiles. We will also deploy our algorithm on the data set used in Bhat et al. (2004), the follow-up to the original CLEAN method introductory paper, and on additional large-scale data sets (e.g., Stovall et al. 2015; Bilous et al. 2020). Using these data sets, we will use our CLEAN algorithm to provide measurements of τ_d across multiple frequencies along many lines of sight. This will give us greater insight into both the composition of the ISM and the intrinsic emission of radio pulsars.

Within this work, we have extensively tested our algorithm's performance on simulated pulses broadened using a thin-screen model of the ISM for our PBF. Future work will entail testing the effects of different PBFs, namely PBFs based on thick and uniform medium ISM models, on the performance of our algorithm. In addition, while our third derivative method for determining the intrinsic τ_{d} from our FOMs works well given high levels of S/N and large τ_{d} values, this may not hold for low τ_{d} values and low levels of S/N as the FOMs are not as smooth. Therefore, we will work on improving our automation efforts via the implementation of machine learning, thus allowing our recall rates to better reflect the performance of the algorithm.

We have greatly simplified radio pulsar emission by assuming symmetric Gaussian intrinsic pulses. However, perfectly symmetric pulses are uncommon in radio pulsars (e.g., Bilous et al. 2016). Should the intrinsic pulse be nonsymmetric, our Γ FOM will either be completely ineffective or lead to incorrect values of τ_{d} being chosen. Therefore, we must further probe the effects of nonsymmetry on our FOM, and develop new FOMs that do not rely on assumed symmetry.

Acknowledgments

O.Y. is supported by the National Science Foundation Graduate Research Fellowship under grant No. DGE-2139292. We gratefully acknowledge support received from NSF AAG award number 2009468, and NSF Physics Frontiers Center award number 2020265, which supports the NANOGrav project. We acknowledge Research Computing at the Rochester Institute of Technology for providing computational resources and support that have contributed to the research results reported in this publication.

ORCID iDs

Olivia Young https://orcid.org/0000-0002-0883-0688
Michael T. Lam https://orcid.org/0000-0003-0721-651X

References

Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2021, ApJL, 923, L22
Bhat, N. D. R., Cordes, J. M., Camilo, F., Nice, D. J., & Lorimer, D. R. 2004, ApJ, 605, 759
Bhat, N. D. R., Cordes, J. M., & Chatterjee, S. 2003, ApJ, 584, 782
Bilous, A. V., Bondonneau, L., Kondratiev, V. I., et al. 2020, A&A, 635, A75
Bilous, A. V., Kondratiev, V. I., Kramer, M., et al. 2016, A&A, 591, A134
Champion, D. J., Ransom, S. M., Lazarus, P., et al. 2008, Sci, 320, 1309
Cordes, J. M., Pidwerbetsky, A., & Lovelace, R. V. E. 1986, ApJ, 310, 737
Cordes, J. M., & Rickett, B. J. 1998, ApJ, 507, 846
Demorest, P. B. 2011, MNRAS, 416, 2821
Dolch, T., Stinebring, D. R., Jones, G., et al. 2021, ApJ, 913, 98
Geiger, A., & Lam, M. T. 2022, The Frequency-dependent scattering of pulsar J1903+0327 Memo Series No. 8, NANOGrav Technical Report
Högbom, J. A. 1974, A&AS, 15, 417
Jones, G., Cordes, J., Demorest, P. B., et al. 2013, in 2013 US National Committee of URSI National Radio Science Meeting (USNC-URSI NRSM) (IEEE), 1
Lam, M. T., 2017 PyPulse: PSRFITS handler, Astrophysics Source Code Library, ascl:1706.011
Lentati, L., Kerr, M., Dai, S., et al. 2017, MNRAS, 468, 1474
Lorimer, D. R., D'Amico, N., Lyne, A. G., et al. 1997, in Joint European and National Astronomical Meeting. 6th European and 3rd Hellenic Astronomical Conference, ed. J. D. Hadjidemetriou & J. H. Seiradakis, 236
Lorimer, D. R., & Kramer, M. 2004, Handbook of Pulsar Astronomy, Vol. 4 (Cambridge: Cambridge Univ. Press)
Manchester, R. N., Hobbs, G., Bailes, M., et al. 2013, PASA, 30, e017
Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, AJ, 129, 1993
McKinnon, M. M. 2014, PASP, 126, 476
Shannon, R. M., & Cordes, J. M. 2017, MNRAS, 464, 2075
Stinebring, D. 2013, CQGra, 30, 224006
Stovall, K., Ray, P. S., Blythe, J., et al. 2015, ApJ, 808, 156
Tsai, W. J., Simonetti, J. H., & Kavic, M. 2017, PASP, 129, 024301
Verbiest, J. P. W., Lentati, L., Hobbs, G., et al. 2016, MNRAS, 458, 1267
Williamson, I. P. 1972, MNRAS, 157, 55
Young, O. 2022, PhD thesis, Rochester Institute of Technology