Constraining the Selection-corrected Luminosity Function and Total Pulse Count for Radio Transients

Fengqiu Adam Dong[1] , Antonio Herrera-Martin[2,3] , Ingrid Stairs[1] , Radu V. Craiu[3] , Kathryn Crowter[1] ,
Gwendolyn M. Eadie[2,3] , Emmanuel Fonseca[4,5] , Deborah Good[6] , James W. Mckee[7,8] , Bradley W. Meyers[9] ,
Aaron B. Pearlman[10,11] , and David C. Stenning[12]

[1] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada; fengqiu.dong@gmail.com
[2] David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
[3] Department of Statistical Science, University of Toronto, Ontario Power Building, 700 University Avenue, 9th Floor, Toronto, ON, M5G 1Z5, Canada
[4] Department of Physics and Astronomy, West Virginia University, Morgantown, WV 26506-6315, USA
[5] Center for Gravitational Waves and Cosmology, Chestnut Ridge Research Building, Morgantown, WV 26505, USA
[6] Physics Department, University of Montana, Charles H Clapp Building, Rm 126, 32 Campus Drive, Missoula, MT 59812, USA
[7] E.A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[8] Centre of Excellence for Data Science, Artificial Intelligence and Modelling, University of Hull, Cottingham Road, Hull, HU6 7RX, UK
[9] International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia
[10] Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A 2T8, Canada
[11] Trottier Space Institute, McGill University, 3550 rue University, Montréal, QC H3A 2A7, Canada
[12] Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada

Abstract

Studying transient phenomena, such as individual pulses from pulsars, has garnered considerable attention in the era of astronomical big data. Of specific interest to this study are rotating radio transients (RRATs), nulling, and intermittent pulsars. This study introduces a new algorithm named LuNfit, tailored to correct the selection biases originating from the telescope and detection pipelines. Ultimately LuNfit estimates the intrinsic luminosity distribution and nulling fraction of the single pulses emitted by pulsars. LuNfit relies on Bayesian nested sampling so that the parameter space can be fully explored. Bayesian nested sampling also provides the additional benefit of simplifying model comparisons through the Bayes ratio. The robustness of LuNfit is shown through simulations and applying LuNfit onto pulsars with known nulling fractions. LuNfit is then applied to three RRATs, J0012 +5431, J1538+1523, and J2355+1523, extracting their intrinsic luminosity distribution and burst rates. We find that their nulling fractions are 0.4(2), 0.749(5), and 0.995(2), respectively. We further find that a log-normal distribution likely describes the single pulse luminosity distribution of J0012+5431 and J1538+1523, while the Bayes ratio for J2355+1523 slightly favors an exponential distribution. We show the conventional method of correcting selection effects by "scaling up" the missed fraction of radio transients can be unreliable when the mean luminosity of the source is faint relative to the telescope sensitivity. Finally, we discuss the limitations of the current implementation of LuNfit while also delving into potential enhancements that would enable LuNfit to be applied to sources with complex pulse morphologies.

Unified Astronomy Thesaurus concepts: Radio pulsars (1353); Radio transient sources (2008); Neutron stars (1108); Compact objects (288)

1. Introduction

Neutron stars result from the remnants of massive stars after they have undergone core-collapse supernova. Rapid rotation of neutron stars likely emerges due to angular momentum conservation following stellar collapse. The rapid rotation combined with beamed radiation creates a lighthouse effect and results in pulsed signals at observatories on Earth. The observed emission exhibits a frequency-dependent delay due to cold-plasma dispersion, imparted onto the radio signal as it traverses the interstellar medium. The dispersion is dependent on the integrated line-of-sight electron density and is inversely proportional to the square of the observing frequency; i.e., $\Delta t = \frac{1}{2.41 \times 10^{-4}}$ s DM pc cm^{-3}, $((\nu_1/\text{GHz})^2 - (\nu_2/\text{GHz})^2)$; (Lorimer & Kramer 2004), where Δt is the time delay, and ν_1 and ν_2 are the upper and lower observing frequencies. DM is

the dispersion measure. The conventional technique to detect pulsars is to first correct for the dispersion and then fold the observation on the pulsar period to build signal-to-noise ratio (S/N). Therefore, pulsar detection sensitivity can be significantly boosted by longer integration times. Pulsar survey strategies must strike a balance between integration time and sky coverage when using conventional single-dish radio telescopes (Manchester et al. 2001; Cordes et al. 2006; Stovall et al. 2014). Often, but not always, this leads to surveys focusing on the Galactic plane. A counterexample is the Green Bank North Celestial Cap survey, which surveyed the whole Northern Hemisphere (Stovall et al. 2014). Because of the large sky coverage, the survey only dwelled for 120 s per pointing. The contemporary pulsar astronomy landscape operates in a distinctly different parameter space, such as the Canadian Hydrogen Intensity Mapping Experiment (CHIME) and the Low-Frequency Array (LOFAR); (CHIME/FRB Collaboration et al. 2018; Sanidas et al. 2019). Their wide field of view allows for simultaneous significant sky coverage and long dwell times. With these new capabilities, intermittent pulsars,

nulling pulsars, and rotating radio transients (RRATs) become easier to find and characterize.

It has long been known that some pulsars do not emit on every rotation (Backer 1970). This phenomenon is called nulling, and up until recently the population of known nulling pulsars has been small compared to the total pulsar population, likely due to the aforementioned survey methods (Konar & Deka 2019). With access to instruments with wide fields of view, this may be changing. For example, Ng et al. (2020) showed that even long-known pulsars thought to be persistent emitters could exhibit nulling behavior. It is important to understand the nulling phenomenon as this can lead to better models for pulsar emission mechanisms (van Leeuwen et al. 2003). On the extreme end of the nulling pulsar spectrum are RRATs, highly intermittent pulsars that can only be detected via their single pulses. As a large portion of the pulsar population likely consists of RRATs (McLaughlin et al. 2006; Keane & McLaughlin 2011), understanding their intrinsic characteristics is vital to understanding the pulsar population. However, RRATs cannot be folded to increase S/N and thus may require long observations to detect appreciable amounts of single pulses.

A critical aspect of population analysis lacking in the current literature is probing the intrinsic luminosity function and burst rate of highly intermittent sources using only the detected pulsar single pulses. Many studies of RRATs only describe an observed burst rate (e.g., Karako-Argaman et al. 2015; Deneva et al. 2016; Good et al. 2021; Dong et al. 2023) rather than correcting for selection effects to probe the intrinsic values. There have been attempts to probe the intrinsic parameters for nulling pulsars (Ritchings 1976; Kaplan et al. 2018; Anumarlapudi et al. 2023) and also another similar class of astrophysical phenomena, repeating fast radio bursts (FRBs), luminous extragalactic radio bursts with similar burst characteristics to RRATs (Li et al. 2021). The most established method described in Ritchings (1976) relies on forming histogram bins and calculating $\Delta = ON - NF \times OFF$ for each histogram bin, where ON and OFF are the histogram bins of the integrated intensities during the on-pulse window and off-pulse window, respectively, and NF is the nulling fraction. One would then minimize $|\Delta|$ by changing NF to find the preferred nulling fraction. However, this method and the Gaussian mixture model in Kaplan et al. (2018) operate on the folded time series of the pulsar observations and necessarily requires two conditions: the first of these conditions is a period and a correctly identified pulse arrival phase. This is not always possible with RRATs due to their extreme sporadicity (i.e., Dong et al. 2023). The second requirement is recording the full-time series for the observation, even during times of no emission. This is problematic for some contemporary commensal instruments such as CHIME/FRB (CHIME/FRB Collaboration et al. 2018), which only save data segments around detected single pulses. While the method described in Li et al. (2021) looks at only single pulses, they account for the missed pulses by "scaling up" by the measure detection fraction. However, we show in Section 5.1 that this method can be unreliable, especially when the source has emitted many bursts below the observing telescope detection threshold. In an attempt to avoid the limitations of earlier works, we implement a new method, *Luminosity and N fit* (LuNfit), that utilizes Bayesian inference using the nested sampling algorithm (Skilling 2004) to constrain the intrinsic single-pulse

Table 1
Properties of the CHIME/Pulsar System

	CHIME/Pulsar
Receiver noise temperature	∼50 K
Frequency range	400–800 MHz
Number of beams	10 (Tracking)
Beam width (FWHM)	30'(400 MHz)–15'(800 MHz)
Time resolution	327.68 μs
Search frequency resolution	390.625 kHz
Coherent dedispersion	Yes

luminosity distribution and the burst rates simultaneously. LuNfit is designed to remove any systematic biases in the single-pulse luminosity distribution and burst rate by empirically measuring the selection effects of the telescope and detection pipeline.

This study is presented in the following way. Section 2 describes the CHIME instrument and the implementation of LuNfit. We discuss the likelihood and derivation of LuNfit. Section 2.1 describes the underlying luminosity distributions used. Section 2.2 details the priors used in LuNfit. Section 2.3 addresses the issue of discrete parameters. Section 2.4 details the CHIME/Pulsar backend, the detection of pulses, and how single-pulse characteristics are retrieved following the detection of pulses. Section 2.5 details the process of injecting simulated pulses into the CHIME/Pulsar detection pipeline to find the selection biases of the telescope. Section 3 details the simulations we performed to show the efficacy and robustness of LuNfit. Section 3.1 compares LuNfit results with the known nulling fractions of two pulsars, B1905+39 and J2044+4614. Section 3.2 describes the conversion to flux units. Section 4 applies LuNfit to three RRATs, J0012+54, J1528+2345, and J2355+1523. We present their intrinsic luminosity distributions and burst rates. Finally, Section 5 compares LuNfit and conventional methods of correcting selection biases and discusses LuNfit's potential future applications to pulsars, RRATs, and FRBs.

2. Methods

CHIME is a commensal transit radio telescope situated in British Columbia, Canada. It comprises three backend instruments: CHIME/cosmology, CHIME/FRB, and CHIME/Pulsar (CHIME/FRB Collaboration et al. 2018; CHIME/Pulsar Collaboration et al. 2021; CHIME Collaboration et al. 2022). Our study only utilizes the CHIME/Pulsar backend, which processes up to 10 simultaneous steerable tracking beams as sources transit the CHIME sky. The properties of CHIME/Pulsar are provided in Table 1. The transit nature of the CHIME telescope allows it to cover the whole northern sky every day, and while each observation is only around ∼10 minutes CHIME/Pulsar can achieve daily cadence on astrophysical sources. Every day CHIME/Pulsar can observe ∼500 individual sources. For this study, CHIME/Pulsar collects high-time resolution spectra in the form of Sigproc[13] style filterbank data. CHIME's sensitivity changes relative to the zenith angle of the source; therefore, to maintain consistency within each observation, we use only the data in which pulses from the nominal pulsar position would be visible through the entire frequency range.

---

[13] https://sigproc.sourceforge.net/

For this study, we are focused on RRATs, which are only detectable via single pulses. A Bernoulli process is a sequence of binary random variables, i.e., they take only values of 0 or 1. Each pulse from an RRAT can be described as one of these binary variables where 0 represents nondetection and 1 represents detection. While RRAT pulses have not been modeled as Bernoulli process previously, RRATs are often assumed to follow a Poisson distribution, the extreme case for the Bernoulli process (e.g., Chen et al. 2022; McKenna et al. 2024). During any observation of an RRAT, any radio telescope will detect only a portion of the emitted pulses due to limited sensitivity. This is called the "selection" of the telescope. Therefore, assuming the selection effects of the telescope are understood, we forward model the intrinsic luminosity function of the RRAT. For N pulses emitted by an RRAT, the likelihood is a series of Bernoulli trials for each pulse emitted. This can be written as

$$L(N, \phi_1 ... \phi_N) \propto \prod_{i=1}^{N} \phi_i^d (1 - \phi_i)^{1-d} \qquad (1)$$

where L is the likelihood, i is the pulse number, N is the total (but unknown) number of pulses emitted by the RRAT, ϕ_i is the probability of detecting pulse i, and $d = 1$ if a detection was made, or $d = 0$ otherwise. Removing the terms in the product that go to one (e.g., when $d = 1$, then the terms $(1 - \phi_i)^{1-1} = 1$), we can rewrite this as

$$L(N, \phi_1 ... \phi_N) \propto \left(\prod_{i=1}^{n} \phi_i \right) \left(\prod_{j=n+1}^{N} (1 - \phi_j) \right) \qquad (2)$$

where n is the number of detected pulses. The first product results from all the pulses detected by the telescope, and the second product results from all those not detected by the telescope. In practice, ϕ_i depends on the pulse's brightness and the intrinsic luminosity function of the RRAT. Therefore,

$$\phi_i(S_{i,\mathrm{det}}, \alpha) = P(\mathrm{det}|S_{i,\mathrm{det}}) \int P(S_{i,\mathrm{det}}|S_T, \sigma_d) P(S_T|\alpha) dS_T \qquad (3)$$

where $S_{i,\mathrm{det}}$ is the detected brightness of pulse i. S_T is the true intrinsic brightness as emitted by the RRAT. We emphasize the difference between S_T and S_{det}, where S_{det} is the detected brightness and comes with the errors due to the stochastic background noise of the telescope. S_T is the intrinsic brightness of the pulse emitted from the source without any noise applied to it. The latter, even if inherently low, can be amplified by random noise fluctuations, rendering it detectable and vice versa. Expressed differently, $S_{\mathrm{det}} = S_T + S_\sigma$, where S_σ is the stochastic noise of the data. Depending on the chosen $S_{i,\mathrm{det}}$ cutoff, this effect can be non-negligible. Figure 2 shows that this effect is most apparent at the turnoff locations, near detection fractions of 0 and 1. $P(\mathrm{det}|S_{i,\mathrm{det}})$ encodes all the selection biases of the telescope and detection pipeline. $P(S_T|\alpha)$ is the probability density of the intrinsic luminosity distribution of the RRAT, where α is a vector containing the parameters of the luminosity model; that is, $\alpha = [\mu, \sigma]$ or $\alpha = [k]$ for a log-normal or exponential distribution, respectively. This is discussed in further in Section 2.1. $P(S_{i,\mathrm{det}}|S_T, \sigma_d)$ designates the Gaussian detection error with σ_d being the known and fixed error. Error σ_d is measured empirically by performing injections, as discussed in

Section 2.5. Due to S_T being an unmeasurable parameter, we marginalize over it. A computational efficiency boost is achieved by realizing that $P(S_{i,\mathrm{det}}|S_T, \sigma_d)$ is a Gaussian and $P(S_{\mathrm{det}}|S_T, \sigma_d) = P(S_{\mathrm{det}} - S_T|0, \sigma_d)$, allowing the right-hand integral of Equation (3) to be written as a convolution.

For pulses that are not detected, we assume that the probability of nondetection is the same for every pulse. That is,

$$1 - \phi_j(\alpha) = 1 - \int P(\mathrm{det}|S_T) P(S_T|\alpha) dS_T. \qquad (4)$$

This assumption must be made as there is no way to know the brightness of a pulse that is not detected. Finally, as the pulses are assumed to be independent, the order does not matter. The independence assumption does not consider refractive scintillation or the clustering of RRAT pulses. Refractive scintillation can change the overall shape of $P(S_T|\alpha)$ and occurs on timescales of weeks (Narayan 1992). Therefore, when applying LuNfit to data that span longer than a few weeks, we point out that the fitted luminosity function will be the intrinsic luminosity function modulated by refractive scintillation. However, carefully accounting for the effects of refractive scintillation with different classes of luminosity functions is out of the scope of this study. We note that in this study, the timescales for observations in Section 4 span months, and therefore the luminosity function that has been fit is likely modulated by refractive scintillation. We do not make any attempts to remove these effects. Despite this, as seen in Section 4, it is unlikely that the shape of the luminosity function is much different from a log-normal or exponential distribution. Any clustering of RRAT pulses should not cause a problem for LuNfit as there is no dependence on each pulse's arrival time. We assume that all pulses from each cluster conform to a global luminosity distribution and do not consider the case where each cluster possesses an independent luminosity distribution. Thus, the likelihood is multiplied by $\binom{N}{n}$ and becomes

$$L(\alpha, N|n, S_{\mathrm{det}}) = \binom{N}{n} \left(1 - \int P(\mathrm{det}|S_T) P(S_T|\alpha) dS_T \right)^{N-n}$$
$$\times \prod_{i=1}^{n} P(\mathrm{det}|S_{i,\mathrm{det}}) \int P(S_{i,\mathrm{det}}|S_T, \sigma_d) P(S_T|\alpha) dS_T. \qquad (5)$$

For this analysis, we take a Bayesian approach utilizing nested sampling (Skilling 2004) through the dynesty package (Speagle 2020).[14] Bayesian nested sampling is a method that breaks up the posterior into many nested "slices," assigns weights and volumes to each slice, and recombines them again to form the posterior and evidence. It is a distinct way to approach Bayesian analysis compared to Markov Chain Monte Carlo and possesses multiple benefits. For example, nested sampling can sample the whole parameter space and avoid problems with multimodal distributions. By extension, obtaining the model evidence (marginal likelihood) for model comparisons through the Bayes ratio is trivial. Both these features are crucial for LuNfit.

We use the detected signal-to-noise ratio (S/N) for $S_{i,\mathrm{det}}$. The strict definitions are given in Section 2.4. Nevertheless, other

---

[14] https://dynesty.readthedocs.io/en/stable/

Table 2
Prior Types and Parameters

Parameter	Prior Type	p_1	p_2
μ	Gaussian	0	4
σ	Inverse gamma	1.938	1
k	Inverse gamma	1	1
N	Uniform	n	$N_{\rm rot} = (T_{\rm obs}/{\rm period})$

brightness metrics, such as detected flux density or fluence, can be employed instead. Details of changing to flux density units are given in Section 3.2.

2.1. Underlying Luminosity Distributions

One caveat in this formulation is the luminosity distribution chosen to describe the detected single pulses, where choosing the incorrect model can significantly alter the luminosity and nulling characteristics of the source in question. However, as we employ a Bayesian nested sampling approach, we use the Bayes ratio to differentiate between different models. In Section 3, we demonstrate with simulations that the Bayes ratio is robust in general. For this reason, we have selected two models to describe pulsar emission characteristics. The first model is the log-normal distribution parameterized by the location, μ, and scale, σ, described by

$$P(S_T | \mu, \sigma) = \frac{1}{S_T \sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln S_T - \mu)^2}{2\sigma^2}\right). \quad (6)$$

The second model is the exponential distribution parameterized by k, given by:

$$P(S_T | k) = k \exp(-k S_T). \quad (7)$$

We chose these two models because they have demonstrated their ability to explain various neutron star single-pulse emission phenomena, such as RRAT pulses (Burke-Spolaor et al. 2011; Meyers et al. 2018, 2019; Mickaliger et al. 2018; Good et al. 2021; Dong et al. 2023), giant pulses (Karuppusamy et al. 2010; Bera & Chengalur 2019; McKee et al. 2019), and regular single pulses (Burke-Spolaor et al. 2012; Liu et al. 2015; Bilous et al. 2022). Repeating FRBs have shown double-peaked Gaussian distributions that could also be considered in future studies (Li et al. 2021) .

2.2. Priors

We assign general priors for both the luminosity function parameters and the number of pulses. These priors are given in Table 2. Unlike Markov Chain Monte Carlo methods that do not necessarily require normalized priors, we need the Bayes ratio for model comparisons and thus need robust evidence estimates. The priors are chosen according to expected behavior. However, they do not restrict unexpected behavior. To achieve this, all priors do not have strict limits unless those limits are unphysical, such as in the case of N. All values on the priors have been chosen based on simulations of pulsars for CHIME, where the underlying unit is S/N. Users of LuNfit should adjust these values if they are using flux, fluence, or other brightness metrics. For μ, we choose a Gaussian prior

described by the following equation

$$P(\mu) = \frac{1}{p_2 \sqrt{2\pi}} \exp\left(-\frac{(\mu - p_1)^2}{2p_2^2}\right) \quad (8)$$

such that the prior contains $\sim68\%$ of the probability mass between -4 and 4. This ensures little bias toward particularly faint or bright pulsars. Note that μ has units of $\ln(\mathrm{S/N})$, such that -4 corresponds to $\mathrm{S/N} \approx 0.02$ and 4 corresponds to $\mathrm{S/N} \approx 54.60$. Therefore, we allow a wide range of median values. If the user knows that the pulsar of interest is particularly bright/faint relative to the telescope sensitivity, the prior on μ should be adjusted to reflect that prior knowledge.

For σ and k we choose an inverse gamma distribution as the prior. They are described by

$$P(\alpha) = \frac{p_2^{p_1}}{\Gamma(p_1)} (1/\alpha)^{p_1+1} \exp(p_2/\alpha). \quad (9)$$

These priors ensure support for σ and k across the viable parameter range between 0 and ∞. However, we choose p_1 and p_2 such that for σ $\sim90\%$ of the probability mass is between 0 and 2, and for k $\sim90\%$ of the probability mass is between 0 and 10. The ranges represent the most likely parameter space for detectable pulsars. Since σ is unitless, we believe that the priors chosen for σ should be near universal for all pulsars. We note that the prior on σ did not need to be altered in any set of simulated parameters. Although k is also unitless, k determines the mean of an exponential distribution. Like μ, the chosen prior for k can differ if the pulsar is particularly bright/faint relative to the telescope sensitivity.

The prior for N is determined by the number of detections made and the amount of time we have observed the source. We assign a uniform distribution between n and $N_{\rm rot} = \lfloor T_{\rm obs}/P \rfloor$ where $\lfloor \rfloor$ is the floor function, $T_{\rm obs}$ is the amount of observation time, and P is the pulsar period. If P is unknown, we set $N_{\rm rot} = \lfloor T_{\rm obs}/1\,\mathrm{ms} \rfloor$.

2.3. Discrete N

Discrete parameters are nontrivial to implement in Bayesian nested sampling. As N is necessarily discrete, we implement a pragmatic approach approximating its discreteness. Dynesty draws samples from the inverse cumulative distribution function (CDF). Therefore, to ensure discrete values of N are drawn, we use the `scipy.stats.randint.ppf` function.[15] This applies the following CDF:

$$\mathrm{CDF} = \frac{\lfloor N \rfloor - n + 1}{N_{\rm rot} - n}. \quad (10)$$

While this implementation may cause errors if n and N are $\mathcal{O}(1)$, in the case of this study, both n and N and their associated errors are much greater than 1. Therefore, we do not expect significant effects on the inference of N.

2.4. CHIME and Detection of Pulses

Here, we describe how we obtain $S_{i,\mathrm{det}}$ for each of the detected pulses. To identify single pulses from known RRATs

---

[15] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.randint.html

and pulsars, we first search the CHIME/Pulsar filterbank data using the CHIME/Pulsar Single-pulse PIPEline (CHIPSPIPE; Dong et al. 2023). Once we detect a pulse, we extract 2 s of data around the time of arrival (TOA) and downsample it by a factor of 3, achieving a time resolution of approximately 1 ms. Next, we mask the channels contaminated by radio frequency interference (RFI), dedisperse the data to the nominal DM of the pulsar, and average the data in frequency, producing a single time series per pulse. To remove the underlying baseline, we remove the pulse from the data and fit polynomials up to 10th order. For each fit, we calculate a reduced χ^2 value and accept the fit where the reduced χ^2 is closest to 1. This process is important as the CHIME/Pulsar baseline can be highly variable. Once the baseline is fit, we subtract it from the frequency collapsed time series, and the remaining background noise is described by the standard deviation of X_n, the baseline-subtracted, pulse removed time series (σ_{noise}). To determine the amplitude of the pulse, we subtract the fitted baseline from the original 2 s chunk of data. Then, we perform a maximum likelihood fit of a Gaussian pulse given by the equation:

$$ F(X) = A \exp\left(-\frac{1}{2}\frac{(X - X_{\mathrm{loc}})^2}{W^2}\right) + B \qquad (11) $$

where X_{loc} represents the location of the Gaussian peak, X is the frequency collapsed time series, A is the peak amplitude, W is the Gaussian width, and B accounts for any remaining DC offset in the time series. Throughout our study, unless otherwise specified, we use $\mathrm{S/N} = S_{\mathrm{det}} = \frac{A}{\sigma_{\mathrm{noise}}}$. We perform the Gaussian pulse fit on each pulse detected by CHIPSPIPE. After fitting, we discard all $S_{\mathrm{det}} < 2$ to ensure that every pulse is a good fit to the Gaussian. This fiducial cutoff is also reflected in the selection effects that are measured.

2.5. Injections

In this section, we describe the process of obtaining $P(\det|S_{i,\mathrm{det}})$, $P(\det|S_T)$, and $P(S_{i,\mathrm{det}}|S_T, \sigma_d)$, with the primary goal of empirically determining the retrieval probability (or detection fraction) at different S_T and $S_{i,\mathrm{det}}$. These values are obtained empirically since the complexity of CHIME/Pulsar and CHIPSPIPE makes it impossible to derive them analytically. We inject simulated Gaussian pulses into real CHIME/Pulsar data. The simulated pulses are broadband with flat spectral indices and widths equal to the mean of the observed widths of the source in question. CHIPSPIPE is not sensitive to width changes for regular slow pulsars or RRATs, so injecting with one width per pulsar is sufficient. Figure 1 shows the effects of width on pulse retrieval for CHIPSPIPE. We find that as long as most pulse widths remain above 10 ms then the width effect of CHIPSPIPE is minimal. Discussions of situations where one must include width effects in the likelihood are provided in more detail in Section 5.3. Since there are hundreds of hours of CHIME/Pulsar observations for each source spread across hundreds of observations, we spread out the injections over multiple observations to perform about 20,000 injections per pulsar. Spreading the injections across multiple observation epochs allows for an accurate description of day-to-day RFI variations.

For a given injection we find the amplitude of the injected pulse in the following way. We subtract the baseline by fitting a polynomial around each injection TOA, and we measure σ_{noise} in the same way that is described in Section 2.4. One can only
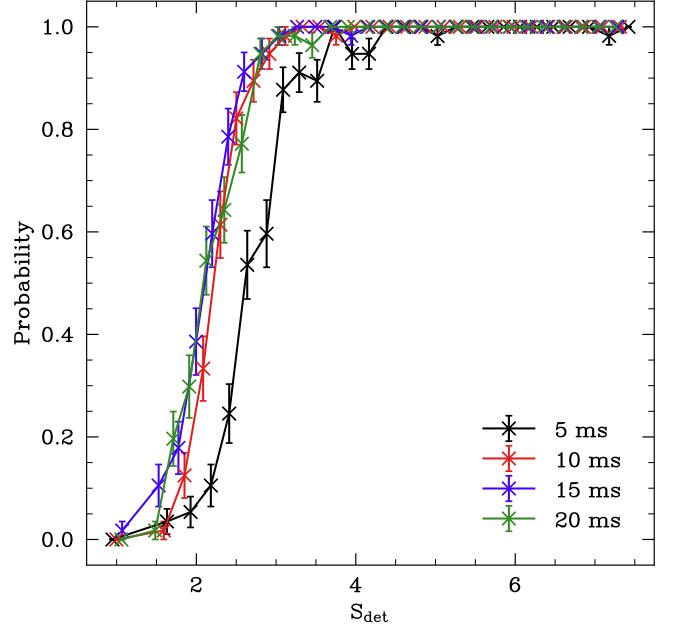

Figure 1. The width effects of CHIPSPIPE. The plot shows the probability of retrieval for different S_{det} at varying widths. For regular slow pulsars, the Gaussian widths are generally greater than 10 ms. Thus, they are not affected much by CHIPSPIPE selection effects. For sources that exhibit high levels of pulse-to-pulse width variability, one needs to model the width along with S_{det} in LuNfit. For the current iteration of LuNfit, we only chose pulsars with Gaussian widths larger than 5 ms to avoid the sensitivity drop-off.

control S_T and not S_{det} when injecting. Thus, we inject the desired S_T based on the σ_{noise}. For each S_T, we inject multiple copies at different TOAs in different observation files. The injected "observations" are saved and processed using the same pipelines for real pulsar data. An example of the recovery rate is shown in Figure 2, and binomial errors are assumed for each data point $\left(\sqrt{\frac{p(1-p)}{m}}\right)$, where p is the measured recovery fraction and m is the total number of injections at a specific S_T. The low S_T values observed here are due to a matter of S/N definition, and Figure 3 illustrates the conversion between CHIPSPIPE/PRESTO detection significance and S_T.

We assume that $P(S_{i,\mathrm{det}}|S_T, \sigma_d)$ is Gaussian, and as S_T is a controlled parameter, σ_d is the only unknown. To measure σ_d, we fit each injected pulse as described in Section 2.4. Therefore for each group of injected S_T we obtain a group of detected $S_{i,\mathrm{det}}$. We then take the standard deviation of the retrieved $S_{i,\mathrm{det}}$ for each of the three brightest values of S_T to use for σ_d. Since $P(S_{i,\mathrm{det}}|S_T, \sigma_d)$ is dominated by the stochastic noise in the data, we assume σ_d is constant across all injected S_T.

Two disparate methods are provided to measure $P(\det|S_{i,\mathrm{det}})$. The first approach involves measuring $S_{i,\mathrm{det}}$ directly from the injected "observations." For each injected pulse, we measure $S_{i,\mathrm{det}}$ by the method outlined in Section 2.4. We bin them and calculate the recovery fraction at each bin by comparing them with the detected injections. This approach is the most accurate but may not be feasible for systems like CHIME/FRB, where one cannot go back through the data stream to calculate S_{det} for nondetections. However, we injected into filterbank data, and even nondetections can be fitted as we know the data segment where the injections were placed. This is shown by red circles in the right panel of Figure 2.

In cases where the first approach is not feasible, a second approach is provided, where one can forward model based on
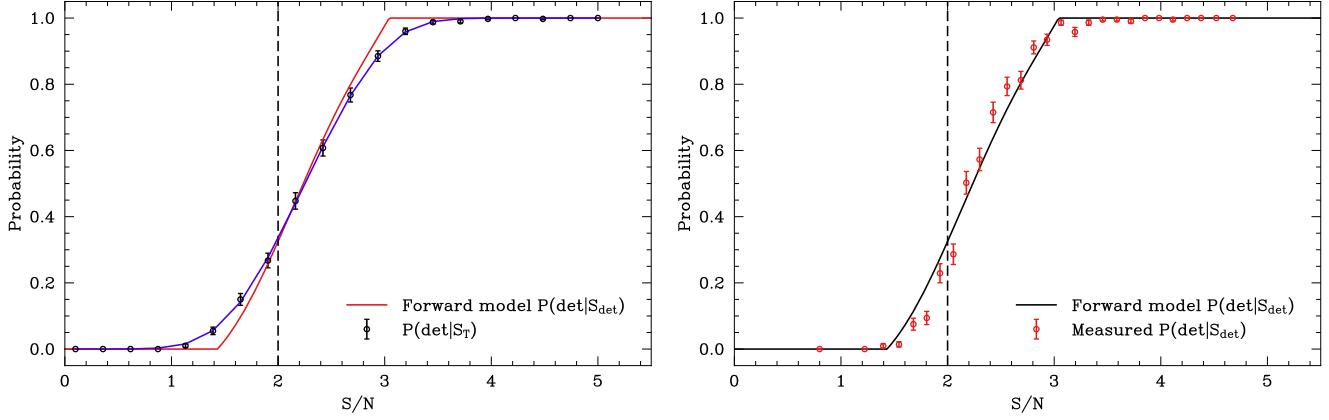
Figure 2. The selection effects of CHIPSPIPE. This plot shows the probability of detection given $S_{i,\mathrm{det}}$ or S_T. The left shows the fitted $P(\det|S_T)$ compared to the forward modeled $P(\det|S_{i,\mathrm{det}})$. The right shows how the forward modeled $P(\det|S_{i,\mathrm{det}})$ compares to the measured values. This study uses the LuNfit implementation. We show the full selection effects here, but when performing LuNfit, we cut off the selection at $S_{i,\mathrm{det}} = 2$ to ensure robust pulses and retrieval probability measurements. The dotted line indicates this cut. The blue line is a best-fit polynomial for $P(\det|S_T)$.
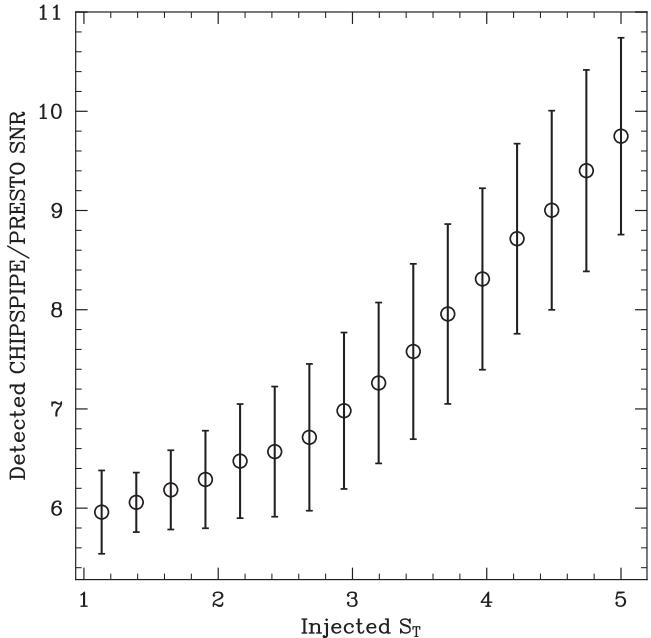
Figure 3. The detected CHIPSPIPE single-pulse search S/N vs. the injected S_T value. Not all injected pulses are detected for detections of $S_T < \sim 3$. Thus, the CHIPSPIPE S/N will be biased high as higher S/N pulses are being preferentially detected. This causes the flattening of the dependence at low S_T.

$P(S_{i,\mathrm{det}}|S_T, \sigma_d)$ and $P(\det|S_T)$. It is found that for CHIPSPIPE, $P(\det|S_{i,\mathrm{det}})$ can be well fit by a piecewise nth-order polynomial

$$P(\det|S_{\mathrm{det}}) = \begin{cases} 0 & S_d < S_{d0} \\ c_0 + c_1 x_1 + c2 x_2^2 ... + c_n x_n^n & S_{d0} < S_d < S_{d1} \\ 1 & S_d > S_{d1} \end{cases}$$ (12)

where S_{d0} is determined by when $P(\det|S_{\mathrm{det}})$ first reaches 0 and S_{d1} is when $P(\det|S_{\mathrm{det}})$ first reaches 1, enforcing the function to be continuous

$$P(\det|S_T) = \int P(\det|S_{\mathrm{det}}) P(S_{\mathrm{det}}|S_T) dS_{\mathrm{det}}.$$ (13)

Equation (13) shows the relationship between the measured probabilities, $P(\det|S_T)$, and the derived property $P(\det|S_{\mathrm{det}})$.

Table 3
Table of Results for J0012+5431, J1538+2345, and J2355+1523

Pulsar	J0012+5431	J1538+2345	J2355+1523
No. of detections	194	6272	453
μ	$-0.8(2)$	1.48(3)	0.1(3)
σ	0.62(5)	1.01(2)	0.8(1)
$N_{\log n}$	88,000(36,000)	9400(200)	2300(1900)
k	2.1(1)	0.141(2)	0.62(3)
$N_{\exp}$	39,030(8700)	8910(70)	2010(160)
μ_{flux}	$-4.0(2)$	$-1.94(3)$	$-3.1(6)$
σ_{flux}	0.62(5)	1.01(2)	0.8(1)
k_{flux}	51(5)	4.31(6)	16(1)
Observation time (s)	442,410	129,196	352,933
Period (s)	3.025	3.449	0.913
Nulling fraction	0.4(2)	0.749(5)	0.995(3)
$\ln Z_{\mathrm{ln}} - \ln Z_{\exp}$	0.7(3)	299.3(4)	$-0.6(5)$

Note. While both exponential and log-normal fit parameters are provided, two of the three RRATs favor the log-normal model. Whereas J0012+5431 only weakly favors the log-normal distribution, J1538+2345 has a strong preference for the log-normal distribution. The only RRAT to favor the exponential model is J2355+1523, but it is a weak Bayes ratio with large errors. In either case, both distributions for J2355+1523 strongly suggest a high nulling fraction. All the uncertainties provided are the 68th percentile confidence intervals.

Forward modeling can be done by using the user's preferred fitting method. In this case, we show a maximum likelihood fit for $P(\det|S_{\mathrm{det}})$ in Figure 2. We perform many maximum likelihood fits with the polynomials of Equation (12) ascending from first order to tenth order. The polynomial with the closest reduced χ_r^2 to 1 is accepted. Therefore, the polynomial order can change for each set of injections depending on the best fit.

3. Simulations

To assess the limitations of LuNfit, we simulate various sets of pulsars with different luminosity parameters and numbers of detected pulses. For each set of luminosity parameters, we simulate a pulse, add Gaussian noise based on the measured $P(S_{i,\mathrm{det}}|S_T, \sigma_d)$, and then determine whether it will be detected according to the measured $P(\det|S_{i,\mathrm{det}})$. This process is repeated until the desired number of pulses is obtained. We then employ

LuNfit to characterize each set of simulations. In total, 20 different sets of parameters are simulated for each luminosity distribution. We simulate one realization for each set of α parameters for the following scenarios:

1. 150 detections, which represents the minimum required for a reliable fit for LuNfit
2. 500 detections, a reasonable number typically observed for the most prolific RRATs discovered by CHIME (Dong et al. 2023)
3. 1000 detections, approximately the number of detections made by the Five-hundred-meter Aperture Spherical Telescope (FAST) for the most prolific FRBs (Li et al. 2022; Zhou et al. 2022; Zhang et al. 2023)
4. 5000 detections, the best-case scenario that can be achieved for the most prolific RRATs.

These simulation results are presented in Appendix Figures 10 and 11. Note that all simulations of the log-normal luminosity model were made with $\sigma = 0.5$. In summary, we recover the correct parameters in almost all situations, with the exponential model requiring fewer detections for the same error region compared to the log-normal model. This is due to the fact that the log-normal distribution is characterized by two parameters (as opposed to one for an exponential distribution), which allows more flexibility when modeling the data. The recovered parameters are generally more accurate when there are more detected pulses and/or the median luminosity is higher. The true N will be lower for a higher median luminosity, and LuNfit will not need to account for as many missed pulses, leading to smaller error regions. Conversely, for lower median luminosities, LuNfit accounts for more missed bursts, and the error regions are larger. This characteristic is universal across luminosity models. These simulations allow us to understand the limitations of LuNfit and provide valuable insights into its performance under various scenarios.

We compare the exponential and log-normal luminosity distributions via the Bayes ratio to see if LuNFit can correctly identify the correct true underlying distribution. We simulate four scenarios for both the log-normal and exponential distributions: 150, 500, 1000, and 5000 detections (that is, the true underlying distribution is log-normal or exponential). We simulate 20 parameter sets for each scenario. We vary μ from –0.5 to 1 for the log-normal model while keeping $\sigma = 0.5$. For the exponential distribution, we vary k from 0.5 to 2. Appendix Figures 12 and 13 show the Bayes ratios between the log-normal and exponential models for all of our simulations. LuNfit can retrieve the correct luminosity distribution for most cases. In general, when a pulsar follows an exponential distribution, LuNfit can always choose the correct model. On the other hand, with only 150 detections, the Bayes ratio is always low for the log-normal distribution. The Bayes ratio for log-normal distributions increases as more detections are made or μ increases. The reason for the μ dependence is that log-normal distributions approach exponential distributions for low values of μ. Therefore, it becomes more difficult to parse the two distributions apart.

3.1. Validation on Known Pulsars

We validate our method by performing LuNfit on two pulsars with known nulling fractions, where the nulling fraction is given by NF $= 1 - \frac{N}{N_{\rm rot}}$ and $N_{\rm rot} = \frac{T_{\rm obs}}{P}$ is the total number of rotations that the observation time allows. This is shown in

Figure 4. B1905+39 ($P = 1.236$ s) is a bright, slow isolated pulsar and not known to exhibit any nulling. Indeed, we find no evidence of nulling in the LuNfit results given in Figure 4. We used one CHIME/Pulsar observation on MJD 59,161, totaling 871 s, and detected 473 pulses. When LuNfit is applied to this pulsar, it accurately indicates a nulling fraction of $-0.02(5)$, consistent with 0 and confirming its non-nulling nature. In this calculation, we have lifted the prior on N and thus allow NF < 0 to show the efficacy of LuNfit. This is shown in Figures 4(a) and (b). In Figure 4(a), we remove the constraints on the prior for N to show that LuNfit converges on the correct value even when not constrained. The constraints on N are then reapplied in Figure 4(b).

J2044+4614 ($P = 1.393$ s), on the other hand, is a known nulling pulsar, and its nulling properties were previously discovered in Ng et al. (2020). We used 39 CHIME/Pulsar observations from MJD 58,808 to MJD 58,579 and detected 585 bursts with 34,130 s of observations. We show the LuNfit results in Figure 4. To compare our results with previously documented methods, we use the nulling measurement method discussed in Ritchings (1976). The LuNfit derived nulling fraction is 0.4(2) and agrees with the Ritchings (1976) method value of ~ 0.41. The two methods are shown in Figure 4(c). In addition, the nulling fraction as measured by Ng et al. (2020) is $>9\%$, consistent with our findings and shown as the shaded region in Figure 4(c).

3.2. Conversion to Flux Density Units

Conversion from the S/N-based $S_{\rm det}$ to Jansky-based units can be done by applying the single-pulse radiometer equation (e.g., Good et al. 2021),

$$S_{\rm peak} = S/N \frac{T_{\rm sys}}{G\sqrt{\Delta\nu \Delta t\, n_p}} \qquad (14)$$

to each detection. Then one would repeat the analysis detailed in Section 2 with $S_{\rm peak}$ in place of $S_{\rm det}$, being careful to apply the same radiometer conversion factor to both the $P(\det|S_{i,\det})$ and $P(S_{i,\det}|S_T, \sigma_d)$. We show the results of this in Table 3. Only the luminosity function is altered in this change of variables, and thus the result for N does not change in Table 3. We caution, however, that due to CHIME/Pulsar's poor flux density calibration issues (Good et al. 2021; Dong et al. 2023), these values should only be taken as a lower limit.

An alternative method for converting the flux is to transform the resultant distribution. We can directly transform the units of the luminosity distribution to flux density by

$$P(S/N)dS/N = P(S_{\rm peak})dS_{\rm peak}. \qquad (15)$$

Here, $P(S/N)$ is the probability distribution of the S/N, and $P(S_{\rm peak})$ is the probability distribution of the peak flux density. Rearranging and applying Equation (14) gives

$$P(S_{\rm peak}) = P(S/N)\left(\frac{T_{\rm sys}}{G\sqrt{\Delta\nu \Delta t n_p}}\right)^{-1}. \qquad (16)$$

4. Results

We present our findings concerning the constrained luminosity function, the intrinsic nulling fractions, and the total pulse number for three RRATs discovered in Karako-Argaman et al. (2015)

(a) B1905+39 without nulling fraction cut off

(b) B1905+39 with nulling fraction cut off at 0

(c) J2044+4614
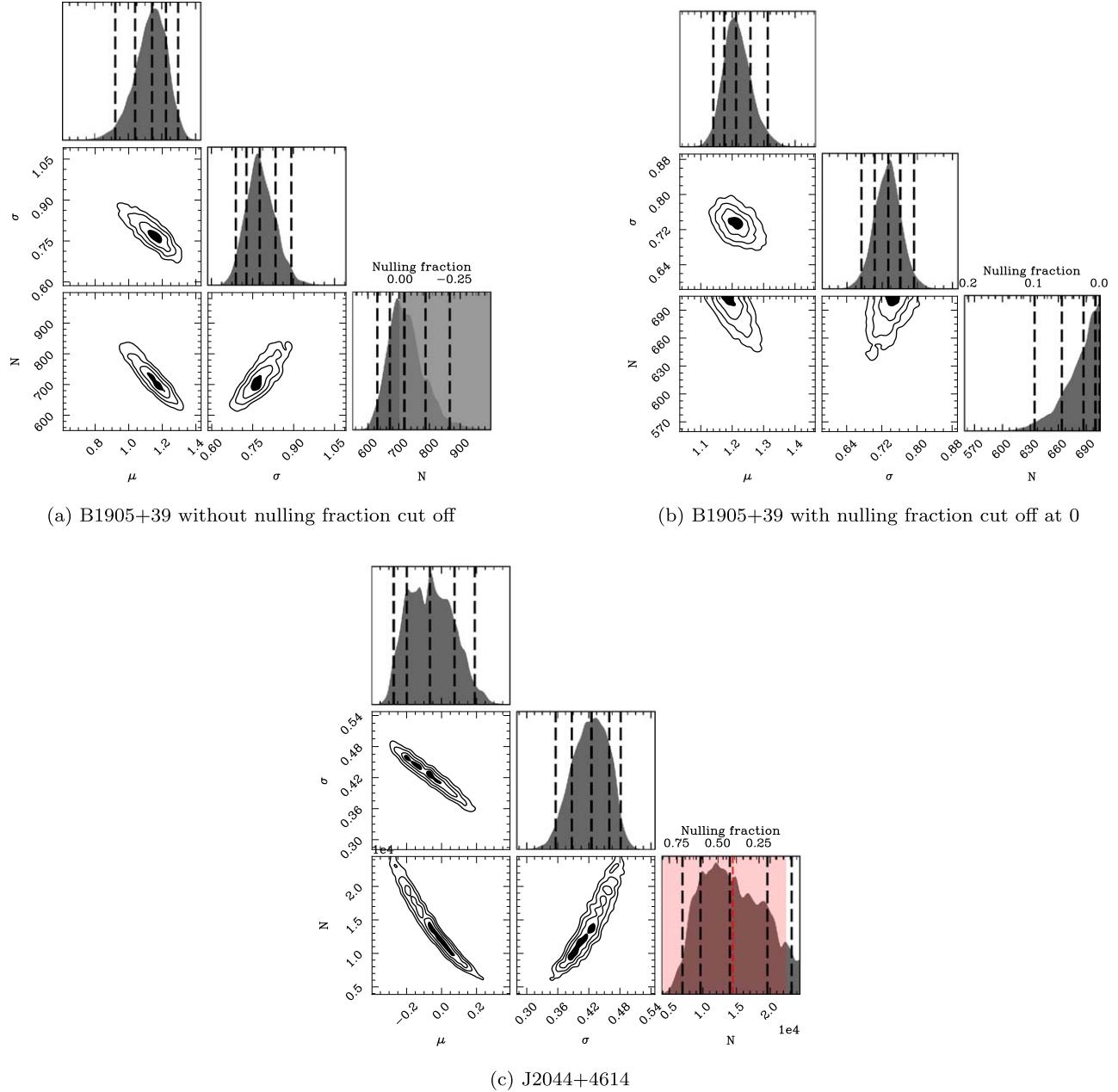
Figure 4. (a) LuNfit for B1905+39: as expected, the nulling fraction peaks at ∼0. The shaded region is forbidden due to negative nulling fractions. We show this region to demonstrate that even when physical constraints are not set, LuNfit converges on reasonable values. (b) LuNfit for B1905+39 again, with the physical limits for nulling fraction. (c) LuNfit for J2044+4614. The predicted nulling fraction according to the method outlined in Ritchings (1976) is ∼0.41. This is shown by the red dashed line. The red-shaded region shows the allowed nulling fractions provided by Ng et al. (2020). The black dashed lines show the 68th percentile and 95th percentile confidence intervals and the median.

and Dong et al. (2023). They are J0012+5431, J1538+2345, and J2355+1523. Detailed results are summarized in Table 3, while visual representations of the `dynesty` nested sampling fits can be found in Figures 5, 6, and 7.

Upon examination, a notable trend emerges: the fitted exponential distribution consistently exhibits a higher median S_T than the log-normal distribution, encompassing all pulsars outlined in Table 3. This distinction consequently results in a diminished total pulse count, N, for the exponential distribution.

J0012+5431 is an RRAT first discovered in Dong et al. (2023). As a result, we had an abundant amount of CHIME/Pulsar data available for this analysis. The LuNfit results for

J0012+5431 are presented in Figure 5. While it initially appears to be extremely sporadic (∼1.6 pulses per hour), LuNfit shows that this source is actually very faint and potentially less sporadic than initially thought. Consequently, CHIME/Pulsar appears to be detecting only the most intensely luminous tail of the distribution. We folded multiple observations to find signs of continuous emission without success. However, this is in line with expectations, considering the source's faintness, prolonged rotational period, and the brevity of CHIME/Pulsar observation windows. The long rotational period dictates that within the brief ∼10 minutes of a CHIME/Pulsar observation, accumulating a substantial number of rotations to bolster the S/N significantly is unattainable. This
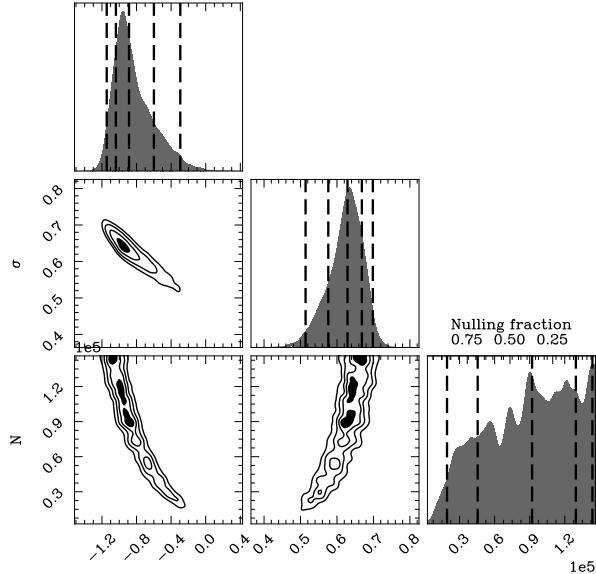
(a)
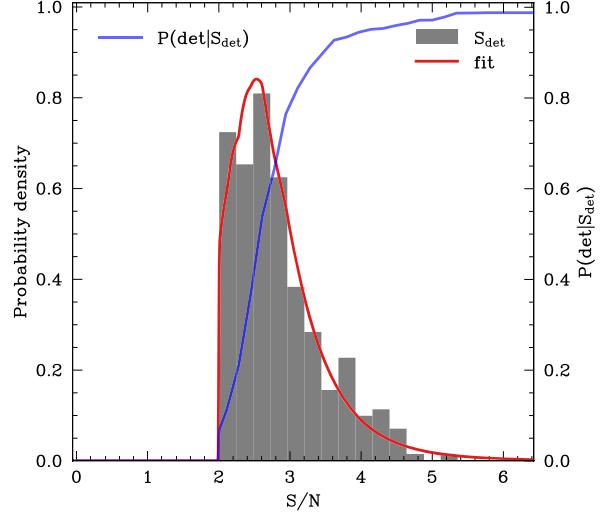
(b)
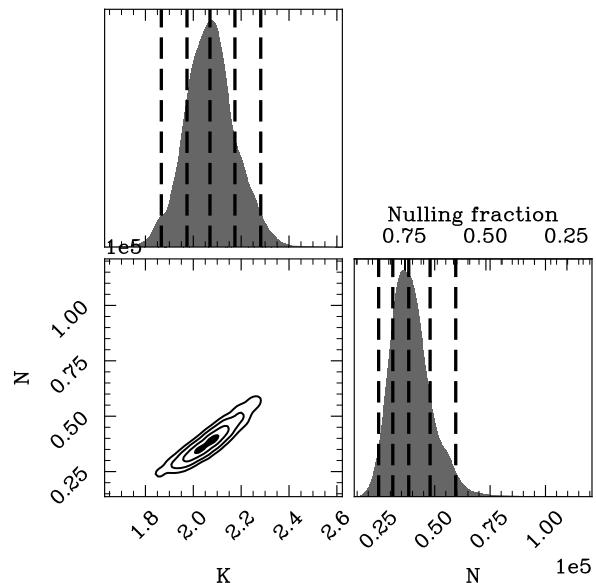
(c)
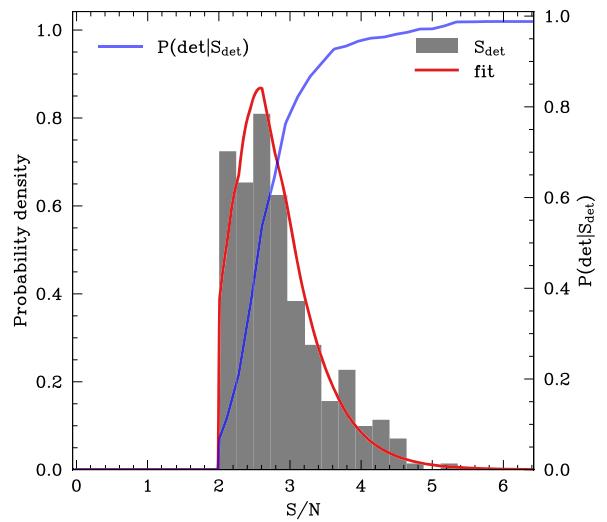
(d)

Figure 5. Corner plot and fit for J0012+5431. The log of the Bayes ratio is 0.7(3). Panels (a) and (b) show the log-normal model, and (c) and (d) show the exponential model. Figures (a) and (b) show the LuNfit corner plots for J0012+5431. Figures (b) and (d) show the LuNfit fits as applied to the detections of J0012+5431. The black dashed lines show the 68 percentile, 95 percentile, and the median. We also provide the selection effect for this source as the blue line.

suggests that follow-up with a more sensitive instrument with longer integration times such as FAST or the Green Bank Telescope (GBT) could confirm our findings. With the Bayes ratio only slightly leaning in favor of the log-normal model, we present LuNfit's results for both the log-normal and exponential models in Figure 5.

J1538+2345 is an RRAT first discovered in Karako-Argaman et al. (2015). For this RRAT, LuNfit heavily favors the log-normal distribution. Thus, only that is provided in the LuNfit results in Figure 6. This source is extremely prolific for an RRAT with a nulling fraction of 0.749(5). As many telescopes have observed this source, we compare our results to

that of the observational burst rates reported by the LOFAR (Karako-Argaman et al. 2015), the GBT (Karako-Argaman et al. 2015), and the FAST (Lu et al. 2019) who reported observational nulling fractions of 0.94, 0.93, and 0.71, respectively. The FAST observations align the best with our selection-corrected nulling fraction of 0.749(5). This is likely due to the extreme sensitivity of FAST, enabling it to probe the intrinsic properties of J1538+2345. We obtained a much higher number of observations than Lu et al. (2019). Therefore, we likely have a more accurate long-term average. The slight discrepancy between the LuNfit results and the FAST results is likely due to the length of observations. The
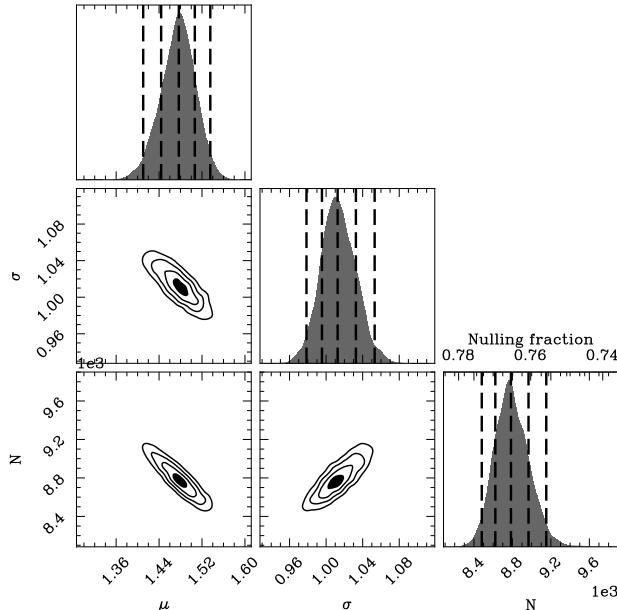
Figure 6. Corner plot and fit for J1538+2345. As J1538+2345 highly prefers the log-normal model, only that is shown. The left panel shows the log-normal constraints on J1538+2345, and the right panel shows the fit given the mean values. The black dashed lines show the 68 percentile, 95 percentile, and the median.

study by Lu et al. (2019) was conducted with only four observations, each lasting 30 minutes. The initial two observations yielded significantly elevated levels of emitted pulses from J1538+2345, which likely contributed to the lower nulling fractions reported by FAST.

J2355+1523 is an RRAT first discovered in Dong et al. (2023). Among the RRATs in that study, it was one of the most prolific sources identified, making it an ideal candidate for LuNfit. The LuNfit results are shown in Figure 7. We find that the Bayes ratio marginally favors the exponential distribution. Thus, we provide LuNfit results for both models in Figure 7. In either case, the source is bright, and the predicted nulling fraction is close to 1.

5. Discussion

We have shown through this study that LuNfit is an effective tool for probing the intrinsic luminosity function and burst rate of RRATs. This has been validated through simulation and comparison with pulsars of known nulling fractions. We then applied LuNfit to three known RRATs, finding their intrinsic luminosity functions and nulling fractions. Below, we provide a detailed discussion of LuNfit's limitations, how it compares with other techniques, and its use cases. Furthermore, we provide a discussion on how LuNfit can be adapted to be used for exotic radio transients such as FRBs and radio magnetars.

5.1. Comparisons with Other Methods

The methods in the literature for correcting for the selection effects of single pulses involve estimating the true burst count by scaling up by the lost fraction, e.g., for FRB121102 (Li et al. 2021). This method involves sampling the selection effects like LuNfit. Then, they place each detection into a histogram bin. The histogram bin is divided by the detection fraction to recover the "true count." While this method can be effective when the telescope's sensitivity accurately captures the "turn-over" of the luminosity function, it becomes unreliable in cases where the source is dim, the telescope lacks sensitivity, or a combination of both.
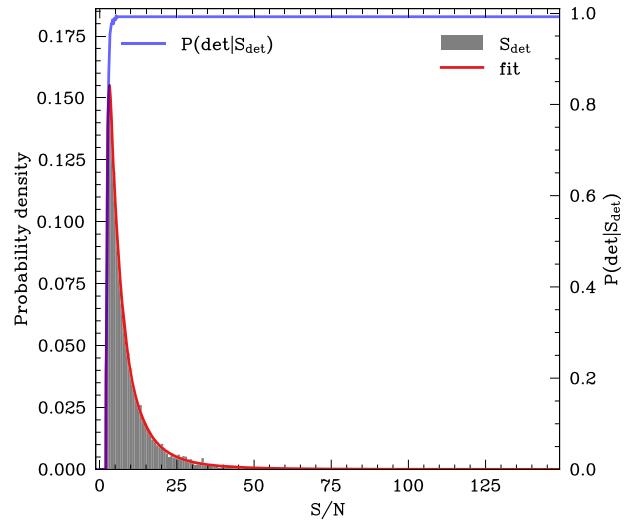
Three factors contribute to the limitations of the Li et al. (2021) approach. First, accurately measuring the selection effects becomes extremely challenging when the detection fraction is low. As the detection fraction decreases, probing the selection function requires exponentially more computing resources. To address this issue in our analysis, we cut off our selection function at a minimum value of $S_{det} = 2\grave{a}$. Second, even if the selection effects are well measured, in parameter spaces where the detection fraction is effectively zero, the simple correction used in conventional methods leads to nonsensical results, as the corrected values go to infinity. Lastly, the Li et al. (2021) correction requires binning of the detected brightnesses, losing information in the process.

To illustrate the performance of LuNfit compared to the Li et al. (2021) method outlined, we provide a comparison in Figures 8 and 9. For the Li et al. (2021) method, we correct each histogram bin by dividing by the detection fraction of that bin. We perform a maximum likelihood fit to the scaled bin heights to get the luminosity function parameters. We find that LuNfit exhibits much greater accuracy and precision compared to conventional methods in all cases, especially when the true N is much higher than the detected n. In summary, LuNfit offers an improved approach to handling selection effects accurately and can provide more reliable results, especially in cases where conventional methods may struggle due to low detection fractions.

5.2. Use for Intermittent Pulsars and Comparisons with Other Tools

LuNfit has the potential to measure the nulling fraction of many intermittent pulsars and RRATs using the vast observation capabilities of CHIME. While the sample analyzed here is too small to say anything about the population, it does show that there can exist a large range of nulling behavior between 0 and 1 nulling fractions. If raw filterbank data were saved, then the Gaussian mixture model (Kaplan et al. 2018) or the Ritchings (1976) method also can provide comparable results to LuNfit for measurements of the nulling fraction. However,
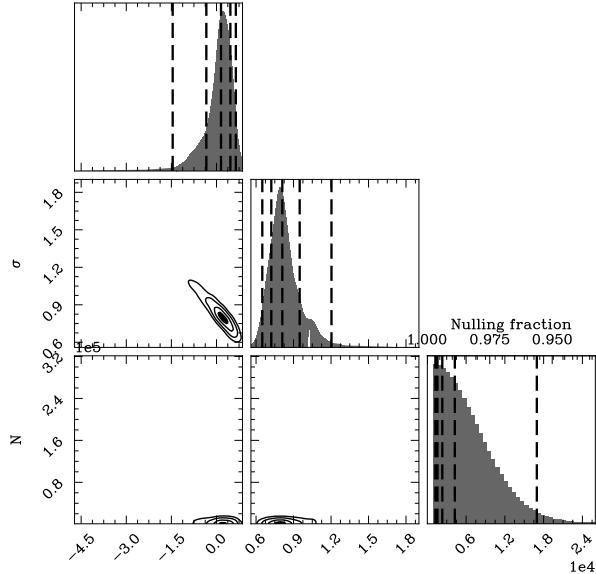
(a)
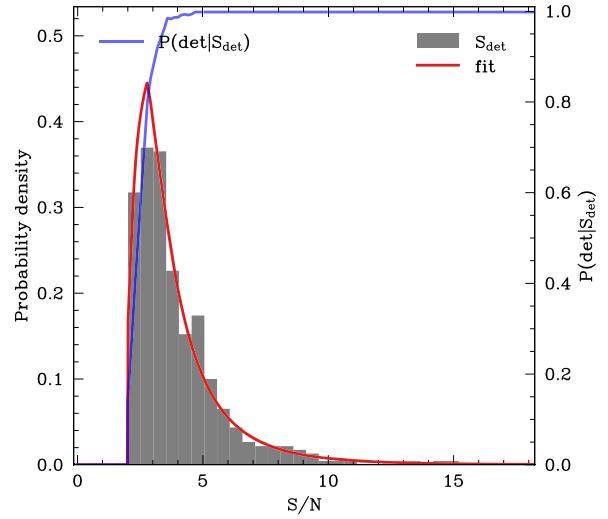
(b)
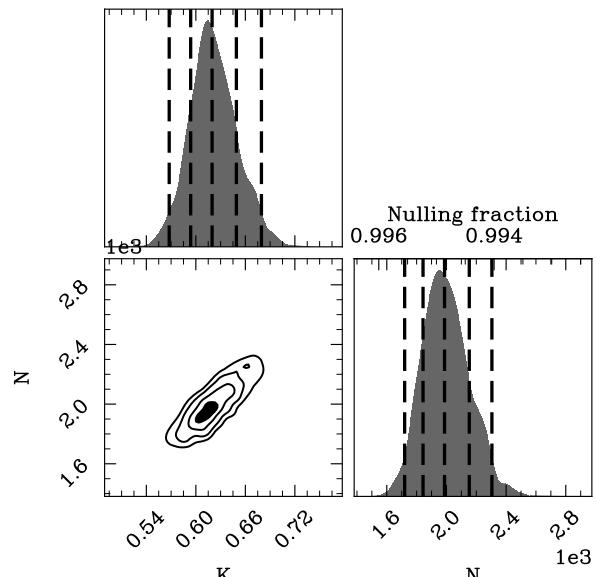
(c)
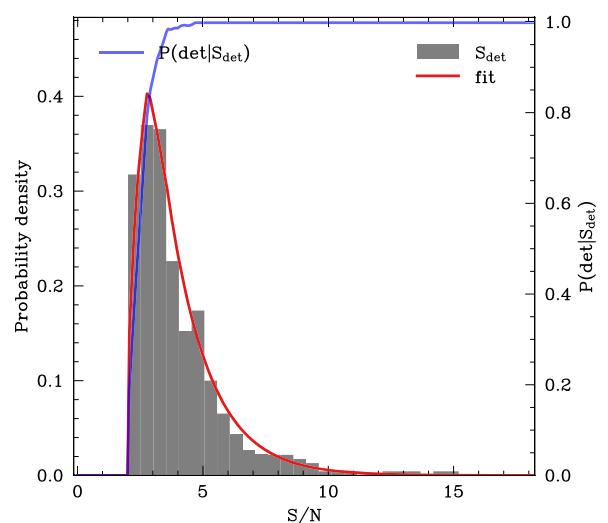
(d)

Figure 7. Fit and corner plots for J2355+1523. The log of the Bayes ratio is 2.7(4). Panels (a) and (b) show the log-normal model, while (c) and (d) show the exponential model. Figures (a) and (b) show the LuNfit corner plots for J2355+1523. Figures (b) and (d) show the LuNfit fits as applied to the detections of J2355 +1523. The black dashed lines show the 68 percentile, 95 percentile, and the median. We also provide the selection effect for this source as the blue line.

LuNfit has four benefits over the aforementioned tools. First, LuNfit can provide the single-pulse luminosity function of intermittent pulsars. Although, arguably, the Gaussian mixture model can also provide a similar measure, this is not possible with the method from Ritchings (1976). Second, as transient astronomy enters an era of big data, many observatories and programs such as CHIME (CHIME/FRB Collaboration et al. 2018), UTMOST (Farah et al. 2019), CRAFT (Macquart et al. 2010), and MeerTRAP (Rajwade et al. 2022) are not saving all observations, rather, only the segments that contain identified pulses by their real-time pipelines. This poses a problem for the Gaussian mixture model and the Ritchings (1976) method as they require data measurements during nondetections. LuNfit avoids this problem by only requiring a measurement of the selection function, which has already been implemented in some experiments like CHIME/FRB (Merryfield et al. 2023). Third, the Gaussian mixture method and the Ritchings (1976) method require identifying the pulse phase where the pulsar emits. If the source is extremely intermittent, where there are only a few single pulses per observation, a few pulses spread across many observations (i.e., some observations contain no detectable emission), or the source is very faint such that even when folded, no appreciable emission is seen, the analysis becomes near impossible with both the Gaussian mixture method and the Ritchings (1976) method. This is not a problem for LuNfit as long as enough single-pulse detections were made
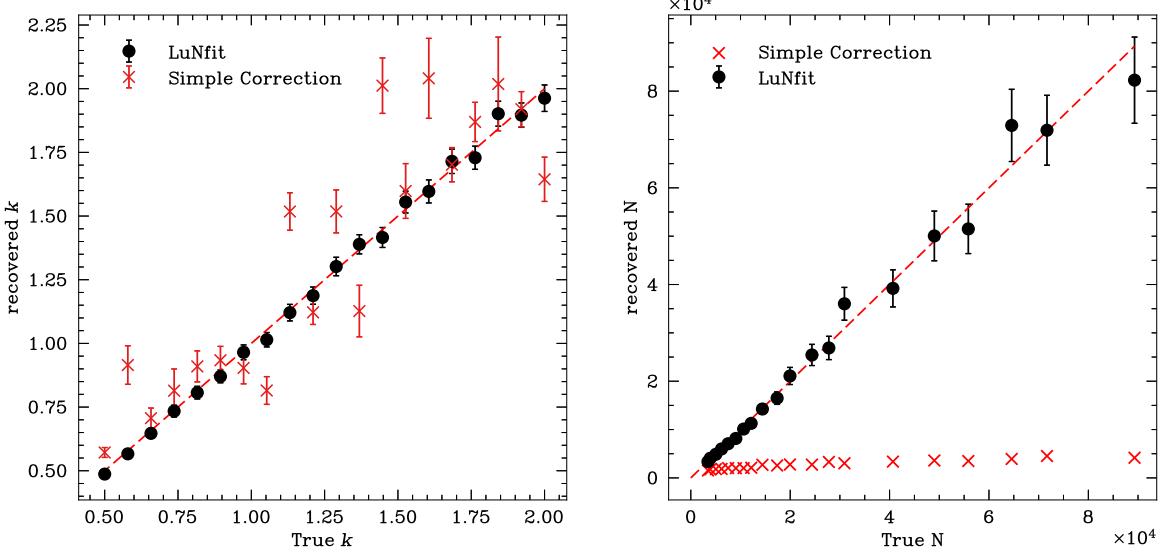
Figure 8. Comparison between a simple correction using the selection function and LuNfit for an exponentially generated distribution with 1000 detections. The "simple correction" is the method described in Li et al. (2021), and a full description is provided in Section 5.1.
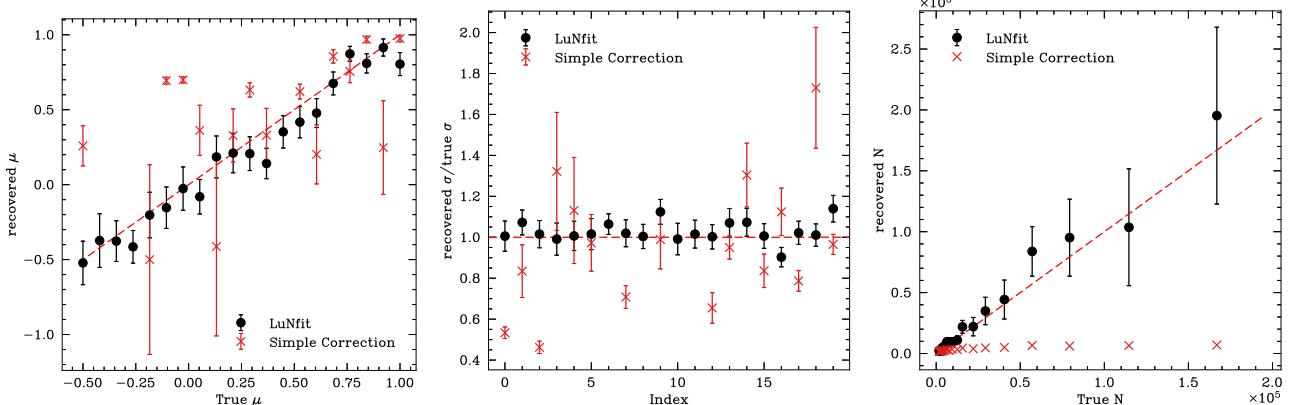
Figure 9. Comparison between a simple correction using the selection function and LuNfit for a log-normal generated distribution with 1000 detections. The red dashed line is the identity line. The "simple correction" is the method described in Li et al. (2021), and a full description is provided in Section 5.1. All simulations are performed with a true σ of 0.5.

on the higher luminosity tail of the distribution. This is because LuNfit does not require pulse phase information, only whether or not detections were made. Finally, LuNfit gives the ability to select between different physically driven models. In this study, we focused on the exponential and log-normal distributions. However, other models, such as the double Gaussian, can be used. LuNfit is based on nested sampling, so the Bayes ratio can always be used to select the preferred model.

5.3. Potential for Repeating FRB Studies and Limitations

We have outlined here the application of LuNfit for RRATs and intermittent slow pulsars. LuNfit can also constrain other radio transients, such as repeating FRBs. Tools like the Gaussian mixture model (Kaplan et al. 2018) and the Ritchings (1976) method cannot be used on repeating FRBs due to the lack of strict periodicity. In our current implementation of LuNfit, we do not include the width of the pulses in the analysis. This decision is based on the observation that for slow pulsars, the width variations are generally small enough to ignore their impact on the selection function. However, this assumption may not hold for repeating FRBs that exhibit

drastic width variability (Pleunis et al. 2021). Therefore, LuNfit can only be applied to repeating FRBs that show minimal width variation. Future iterations of LuNfit will incorporate width dependence to address this issue. As more parameters are introduced, the computational complexity of the analysis increases. Dealing with selection effects in multiple parameters becomes more resource intensive. Therefore, optimizations will be required to retrieve selection effects while including width variations effectively.

6. Conclusion and Future Work

In this study, we demonstrate the importance of accounting for selection biases inherent to the telescope and detection pipeline. This allows the accurate determination of the intrinsic pulse rate and luminosity function of a slow pulsar or RRAT. To achieve this, we created an analysis framework named LuNfit utilizing Bayesian nested sampling, leveraging the `dynesty` package.

We present detailed simulations and validation procedures for LuNfit in Section 2. This involved many sets of simulations of fake pulsars and characterizing the nulling fraction of two

pulsars, B1905+39 and J2044+4614, where the nulling fraction is known. As a conclusive application of LuNfit, we apply it to three known RRATs, successfully ascertaining their intrinsic luminosity functions and burst rates. We show that the LuNfit nulling fractions for J1538+2345 align most closely with the FAST observations. We argue that this is likely due to the impressive sensitivity of FAST, enabling them to probe the intrinsic properties of J1538+2345 without much selection bias. The nulling fraction measured by FAST is likely biased slightly high due to short observations. Notably, our findings in Section 5.1 highlight the limitations of conventional techniques in capturing the intrinsic luminosity distribution and burst rate when compared with LuNfit.

Looking ahead, future work includes improving LuNfit by incorporating width dependence and more model types, such as the single and double Gaussian. This extension will render LuNfit more versatile, enabling its effective utilization for sources emitting bursts with complex and varying morphologies. This enhancement will allow LuNfit to be applied to all repeating FRBs, pulsars, and radio magnetars. Additionally, harnessing LuNfit's current capabilities, we intend to quantify the nulling fraction for numerous intermittent pulsars and RRATs using CHIME/Pulsar observations. This will provide a systematic understanding of the nulling slow pulsar population. Finally, we plan to make the codebase publicly available with the next iteration of LuNfit in the near future.

Appendix
Simulations

Figures 10–13 show simulations demonstrating the recovery capabilities of LuNfit.

(a) 150 detections

(b) 150 detections

(c) 500 detections

(d) 500 detections

(e) 1000 detections

(f) 1000 detections

(g) 5000 detections

(h) 5000 detections
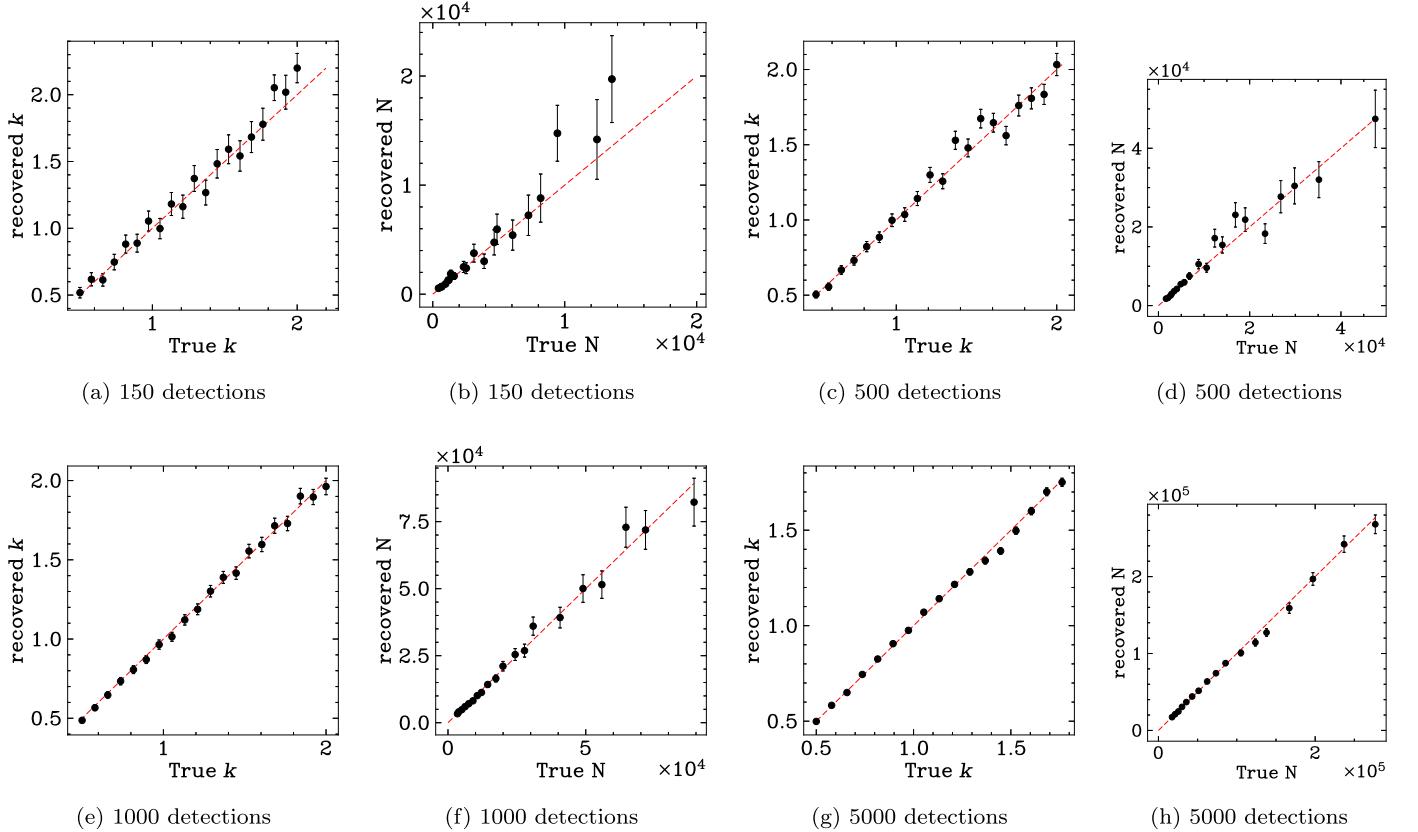
Figure 10. Simulations and recovery capabilities of LuNfit for the exponential distribution. The axes show the LuNfit recovered quantities against the true simulated quantities. N and k are the parameters constrained for an exponential luminosity distribution. The red dashed line is the identify line. All points should lie on this line for a perfect fitting algorithm. This figure shows the recovered k (N) value given a simulated (true) k (N) value. As expected for high true N, the error bars increase as this corresponds to an intrinsically fainter pulsar. In other words, LuNfit needs to predict more missed pulses for higher true N values. For an exponential distribution, a fainter pulsar corresponds to a larger k. Therefore, we similarly find that the errors for k increase for larger k. LuNfit performs significantly better for an exponential distribution compared to a log-normal distribution due to the smaller parameter space.
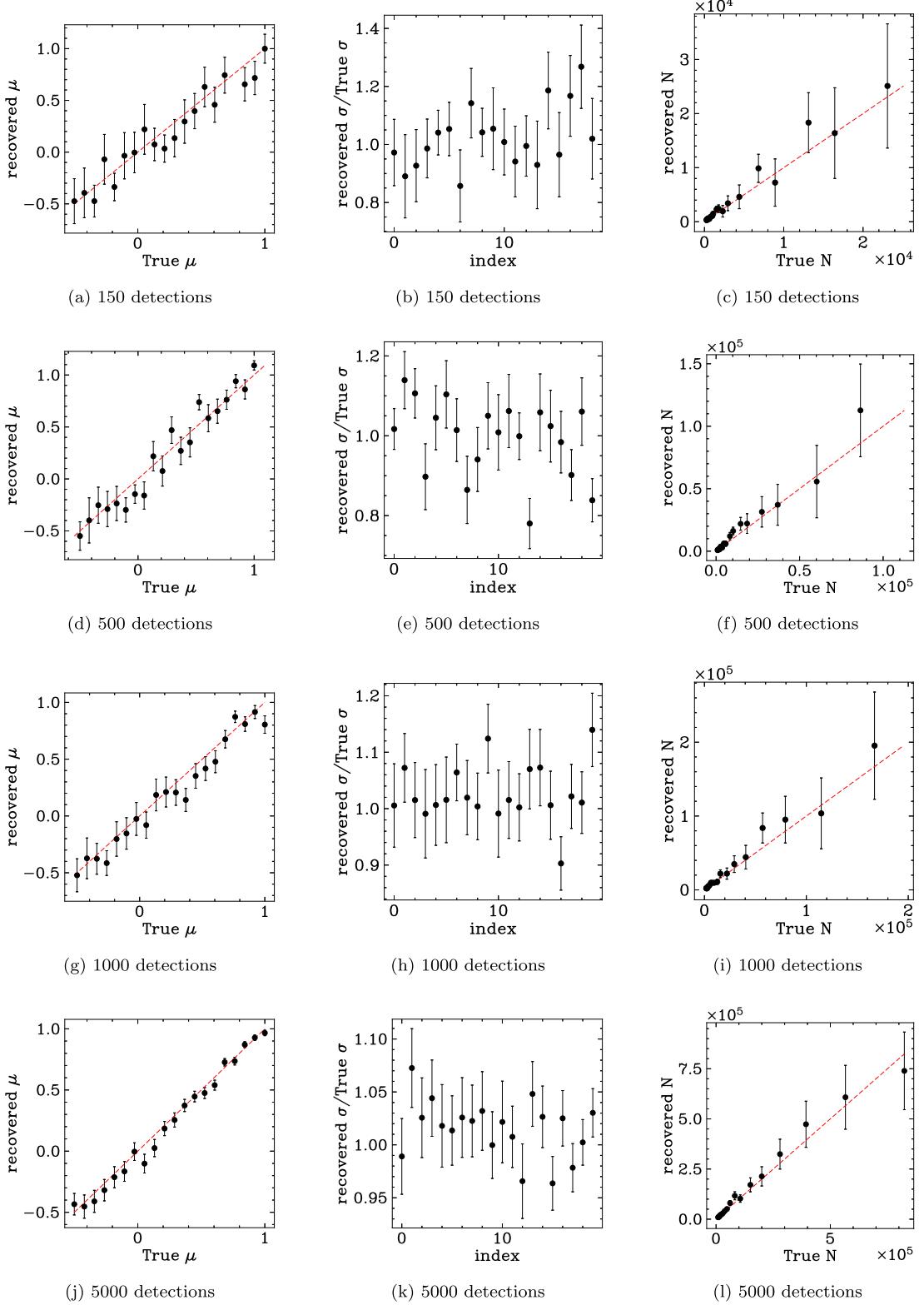
Figure 11. Simulations and recovery capabilities of LuNfit for the log-normal distribution. The axes show the LuNfit recovered quantities against the true simulated quantities. Terms μ, σ, and N are the parameters constrained for a log-normal luminosity distribution. All simulations shown here have a σ of 0.5. Again, the red dashed line is the identity line. All points should fall on this line for a perfect fitting algorithm. Like the exponential distribution shown in Figure 10, the fit is worse for high true N and low true μ as this indicates a fainter pulsar. The uncertainties of the log-normal luminosity model are markedly larger than that of the exponential luminosity model. This is because of the larger parameter space to explore. All simulations were performed with $\sigma = 0.5$; thus, all data points in the middle column should equal 1 (within uncertainty). This figure shows that in most cases LuNfit is able to retrieve the correct σ value.

(a) 150 detections

(b) 500 detections

(c) 1000 detections

(d) 5000 detections
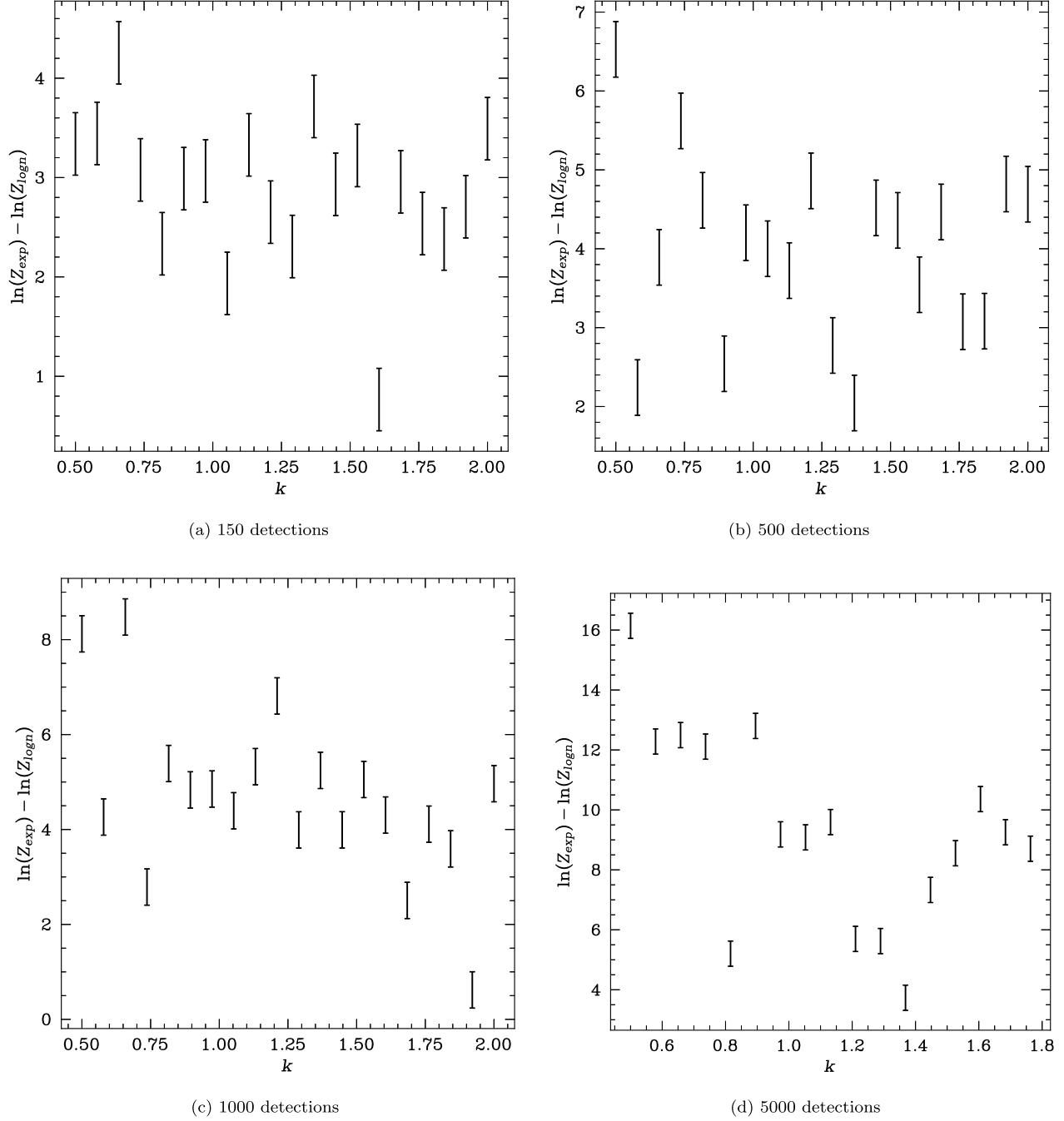
Figure 12. Bayes ratio for a simulated exponential distribution. The Bayes ratio is given in favor of the exponential distribution for four sets of exponential luminosities distributions. That is, positive values indicate a preference for the exponential distribution. There are 150 (a), 500 (b), 1000 (c), and 5000 (d) detected pulses in each simulation. If the true underlying distribution is exponential, LuNfit always correctly identifies this, as shown by the absence of negative values. In general, the Bayes ratio is larger for low k and a larger number of detections.

(a) 150 detections

(b) 500 detections


(c) 1000 detections

(d) 5000 detections
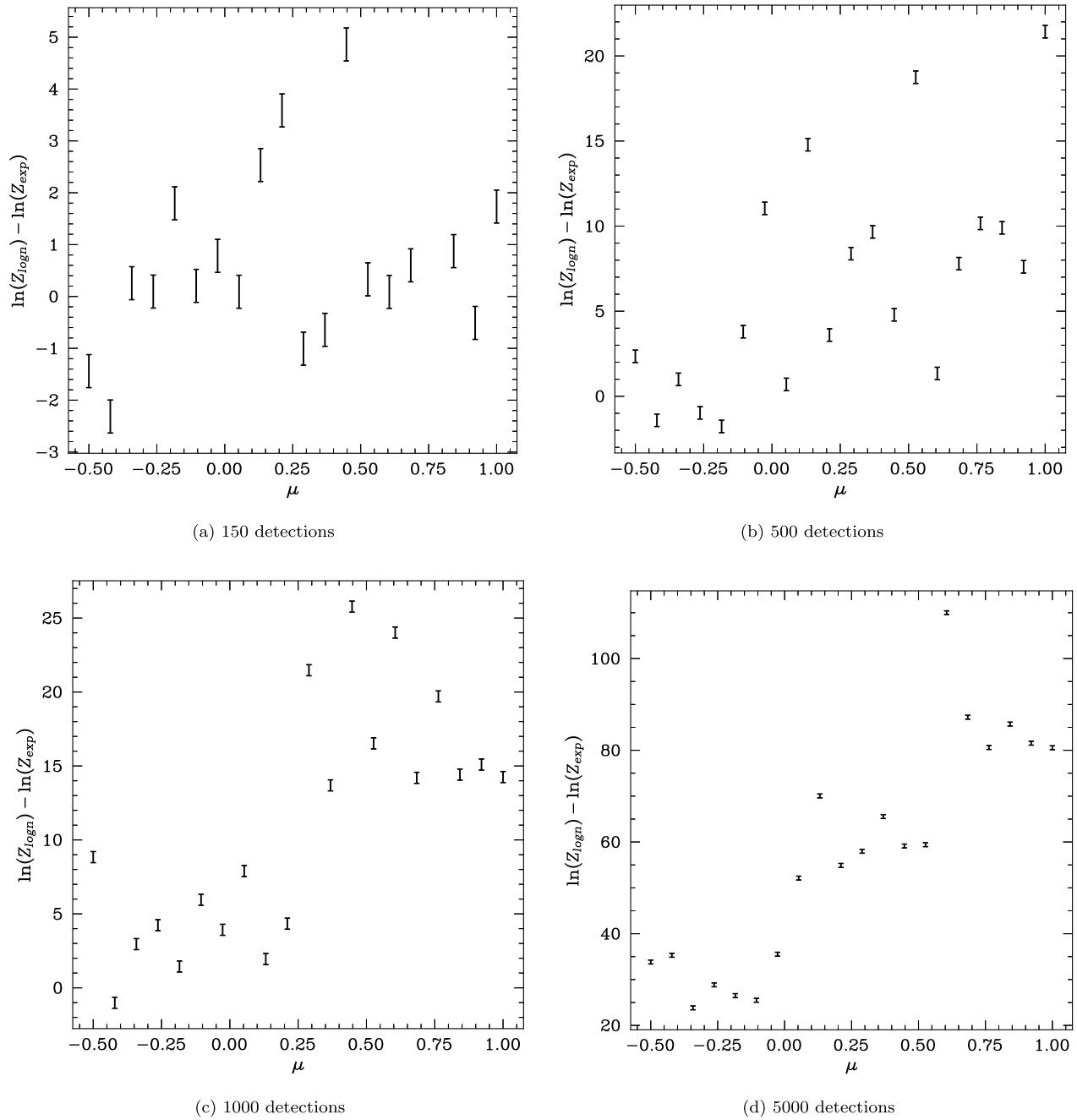
Figure 13. Bayes ratio for simulated log-normal distributions. The Bayes ratio is given in favor of the log-normal distribution for four sets of log-normal luminosities. That is, positive values indicate a preference for a log-normal distribution. There are 150 (a), 500 (b), 1000 (c), and 5000 (d) detected pulses in each simulation. The σ for these simulations was set to 0.5. The effects of a brighter median luminosity (higher μ) are much more pronounced for the log-normal distribution than the exponential distribution. Higher μ values are clearly positively correlated with higher Bayes ratios. This is due to LuNfit being able to completely fit the "turnover" in log-normal distributions. Additionally, the log-normal tends toward an exponential for low μ values. Once again, a larger number of detections aid the differentiation between log-normal and exponential distributions.

ORCID iDs

Fengqiu Adam Dong ⓘ https://orcid.org/0000-0003-4098-5222
Antonio Herrera-Martin ⓘ https://orcid.org/0000-0002-3654-4662
Ingrid Stairs ⓘ https://orcid.org/0000-0001-9784-8670
Radu V. Craiu ⓘ https://orcid.org/0000-0002-1348-8063
Kathryn Crowter ⓘ https://orcid.org/0000-0002-1529-5169

Gwendolyn M. Eadie ⓘ https://orcid.org/0000-0003-3734-8177
Emmanuel Fonseca ⓘ https://orcid.org/0000-0001-8384-5049
Deborah Good ⓘ https://orcid.org/0000-0003-1884-348X
James W. Mckee ⓘ https://orcid.org/0000-0002-2885-8485
Bradley W. Meyers ⓘ https://orcid.org/0000-0001-8845-1225
Aaron B. Pearlman ⓘ https://orcid.org/0000-0002-8912-0732
David C. Stenning ⓘ https://orcid.org/0000-0002-9761-4353

References

Anumarlapudi, A., Swiggum, J. K., Kaplan, D. L., & Fichtenbauer, T. D. J. 2023, ApJ, 948, 32

Backer, D. C. 1970, Natur, 228, 42

Bera, A., & Chengalur, J. N. 2019, MNRAS, 490, L12

Bilous, A. V., Grießmeier, J. M., Pennucci, T., et al. 2022, A&A, 658, A143

Burke-Spolaor, S., Bailes, M., Johnston, S., et al. 2011, MNRAS, 416, 2465

Burke-Spolaor, S., Johnston, S., Bailes, M., et al. 2012, MNRAS, 423, 1351

Chen, J. L., Wen, Z. G., Yuan, J. P., et al. 2022, ApJ, 934, 24

CHIME Collaboration, Amiri, M., Bandura, K., et al. 2018, ApJ, 863, 48

CHIME Collaboration, Amiri, M., Bandura, K., et al. 2021, ApJS, 255, 5

CHIME Collaboration, Amiri, M., Bandura, K., et al. 2022, ApJS, 261, 29

Cordes, J. M., Freire, P. C. C., Lorimer, D. R., et al. 2006, ApJ, 637, 446

Deneva, J. S., Stovall, K., McLaughlin, M. A., et al. 2016, ApJ, 821, 10

Dong, F. A., Crowter, K., Meyers, B. W., et al. 2023, MNRAS, 524, 5132

Farah, W., Flynn, C., Bailes, M., et al. 2019, MNRAS, 488, 2989

Good, D. C., Andersen, B. C., Chawla, P., et al. 2021, ApJ, 922, 43

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357

Kaplan, D. L., Swiggum, J. K., Fichtenbauer, T. D. J., & Vallisneri, M. 2018, ApJ, 855, 14

Karako-Argaman, C., Kaspi, V. M., Lynch, R. S., et al. 2015, ApJ, 809, 67

Karuppusamy, R., Stappers, B. W., & van Straten, W. 2010, A&A, 515, A36

Keane, E. F., & McLaughlin, M. A. 2011, BASI, 39, 333

Konar, S., & Deka, U. 2019, JApA, 40, 42

Li, D., Wang, P., Zhu, W. W., et al. 2021, Natur, 598, 267

Li, D., Wang, P., Zhu, W. W., et al. 2022, Natur, 601, E1

Liu, K., Karuppusamy, R., Lee, K. J., et al. 2015, MNRAS, 449, 1158

Lorimer, D. R., & Kramer, M. 2004, Handbook of Pulsar Astronomy, Vol. 4 (Cambridge: Cambridge Univ. Press)

Lu, J., Peng, B., Liu, K., et al. 2019, SCPMA, 62, 959503

Macquart, J.-P., Bailes, M., Bhat, N. D. R., et al. 2010, PASA, 27, 272

Manchester, R. N., Lyne, A. G., Camilo, F., et al. 2001, MNRAS, 328, 17

McKee, J. W., Stappers, B. W., Bassa, C. G., et al. 2019, MNRAS, 483, 4784

McKenna, D. J., Keane, E. F., Gallagher, P. T., & McCauley, J. 2024, MNRAS, 527, 4397

McLaughlin, M. A., Lyne, A. G., Lorimer, D. R., et al. 2006, Natur, 439, 817

Merryfield, M., Tendulkar, S. P., Shin, K., et al. 2023, AJ, 165, 152

Meyers, B. W., Tremblay, S. E., Bhat, N. D. R., et al. 2018, ApJ, 869, 134

Meyers, B. W., Tremblay, S. E., Bhat, N. D. R., et al. 2019, PASA, 36, e034

Mickaliger, M. B., McEwen, A. E., McLaughlin, M. A., & Lorimer, D. R. 2018, MNRAS, 479, 5413

Narayan, R. 1992, RSPTA, 341, 151

Ng, C., Wu, B., Ma, M., et al. 2020, ApJ, 903, 81

Pleunis, Z., Good, D. C., Kaspi, V. M., et al. 2021, ApJ, 923, 1

Rajwade, K. M., Bezuidenhout, M. C., Caleb, M., et al. 2022, MNRAS, 514, 1961

Ransom, S. M. 2001, PhD thesis, Harvard Univ.

Ritchings, R. T. 1976, MNRAS, 176, 249

Sanidas, S., Cooper, S., Bassa, C. G., et al. 2019, A&A, 626, A104

Skilling, J. 2004, in AIP Conf. Ser. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint (Melville, NY: AIP), 395

Speagle, J. S. 2020, MNRAS, 493, 3132

Stovall, K., Lynch, R. S., Ransom, S. M., et al. 2014, ApJ, 791, 67

van Leeuwen, A. G. J., Stappers, B. W., Ramachandran, R., & Rankin, J. M. 2003, A&A, 399, 223

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261

Zhang, Y.-K., Li, D., Zhang, B., et al. 2023, ApJ, 955, 142

Zhou, D. J., Han, J. L., Zhang, B., et al. 2022, RAA, 22, 124001