The Astrophysical Journal, 928:67 (11pp), 2022 March 20 https://doi.org/10.3847/1538-4357/ac54ae

A Detection of Red Noise in PSR J1824–2452A and Projections for PSR B1937+21 Using NICER X-Ray Timing Data

Jeffrey S. Hazboun[1], Jack Crump[2], Andrea N. Lommen[2], Sergio Montano[2], Samantha J. H. Berry[3], Jesse Zeldes[2], Elizabeth Teng[2,4], Paul S. Ray[5], Matthew Kerr[5], Zaven Arzoumanian[6], Slavko Bogdanov[7], Julia Deneva[8], Natalia Lewandowska[9], Craig B. Markwardt[6], Scott Ransom[10], Teruaki Enoto[11], Kent S. Wood[12], Keith C. Gendreau[6], David A. Howe[13,14], and Aditya Parthasarathy[15]

[1] University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011, USA; hazboun@uw.edu
[2] Haverford College, 370 Lancaster Ave., Haverford, PA 19041, USA
[3] Bryn Mawr College, 101 North Merion Ave., Bryn Mawr, PA 19010, USA
[4] Department of Physics and Astronomy, Northwestern University, 1800 Sherman Ave., Evanston, IL 60201, USA
[5] U.S. Naval Research Laboratory, Washington, DC 20375, USA
[6] Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[7] Columbia Astrophysics Laboratory, Columbia University, 550 W. 120th St., New York, NY 10027, USA
[8] George Mason University, resident at Naval Research Laboratory, Washington, DC 20375, USA
[9] Department of Physics and Astronomy, Swarthmore College, Swarthmore, PA 19081, USA
[10] NRAO, 520 Edgemont Rd., Charlottesville, VA 22903, USA
[11] Department of Astronomy, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan
[12] Praxis, resident at the Naval Research Laboratory, Washington, DC 20375, USA
[13] Time and Frequency Division, NIST, Boulder, CO 80305, USA
[14] Department of Physics, University of Colorado, Boulder, CO 80309, USA
[15] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany
Received 2021 December 3; revised 2022 February 10; accepted 2022 February 12; published 2022 March 28

Abstract

We have used X-ray data from the Neutron Star Interior Composition Explorer (NICER) to search for long-timescale temporal correlations ("red noise") in the pulse times of arrival (TOAs) from the millisecond pulsars PSR J1824−2452A and PSR B1937+21. These data more closely track intrinsic noise because X-rays are unaffected by the radio-frequency-dependent propagation effects of the interstellar medium. Our search yields strong evidence (natural log Bayes factor of 9.634 ± 0.016) for red noise in PSR J1824−2452A, but the search is inconclusive for PSR B1937+21. In the interest of future X-ray missions, we devise and implement a method to simulate longer and higher-precision X-ray data sets to determine the timing baseline necessary to detect red noise. We find that the red noise in PSR B1937+21 can be reliably detected in a 5 yr mission with a TOA error of 2 μs and an observing cadence of 20 observations per month compared to the 5 μs TOA error and 11 observations per month that NICER currently achieves in PSR B1937+21. We investigate detecting red noise in PSR B1937+21 with other combinations of observing cadences and TOA errors. We also find that time-correlated red noise commensurate with an injected stochastic gravitational-wave background having an amplitude of $A_{\rm GWB} = 2 \times 10^{-15}$ and spectral index of timing residuals of $\gamma_{\rm GWB} = 13/3$ can be detected in a pulsar with similar TOA precision to PSR B1937+21. This is with no additional red noise in a 10 yr mission that observes the pulsar 15 times per month and has an average TOA error of 1 μs.

Unified Astronomy Thesaurus concepts: Pulsars (1306); Millisecond pulsars (1062); Red noise (1956); X-ray astronomy (1810); Pulsar timing method (1305)

1. Introduction

Rotation-powered pulsars, particularly millisecond pulsars (MSPs), are extraordinarily stable rotators, but there is noise in the arrival time of the pulses. Pulsar timing noise is the subject of much research, both to understand the pulsar emission mechanism and for the sake of using pulsars as clocks, e.g., to make a gravitational-wave detector using a pulsar timing array (PTA; Hobbs et al. 2010; Cordes & Shannon 2012; Arzoumanian et al. 2016; Lentati et al. 2016; Perera et al. 2019). The noise we observe likely comes from both intrinsic and extrinsic effects, but what fraction of the noise is intrinsic to the pulsar is difficult to disentangle from the other noise sources (Lam et al. 2016; Lentati et al. 2016; Lam et al. 2017; Hazboun et al. 2020; Goncharov et al. 2021b). Intrinsic noise could come from jitter in the location of the emission mechanism, i.e., the beam could be nonstationary on the pulsar (Lam et al. 2016), or it could come from various changes in the interior of the neutron star, effectively changing the moment of inertia in a stochastic manner (Cordes & Shannon 2010; Melatos & Link 2014). Extrinsic noise could come from the intervening medium, i.e., the interstellar medium (ISM), from measurement noise, or from gravitational waves. In the case of X-ray observations, the ISM noise is essentially nonexistent, since its effects decrease significantly with increases in the observation frequency (Stinebring 2013). The Neutron Star Interior Composition Explorer (NICER) gives us a unique chance to perform high-precision, long-term timing of pulsars in the X-ray band, free from ISM noise. The detection of red noise that we document in this paper is significant because it is a detection of red noise where we know that none of the red noise is due to the ISM.

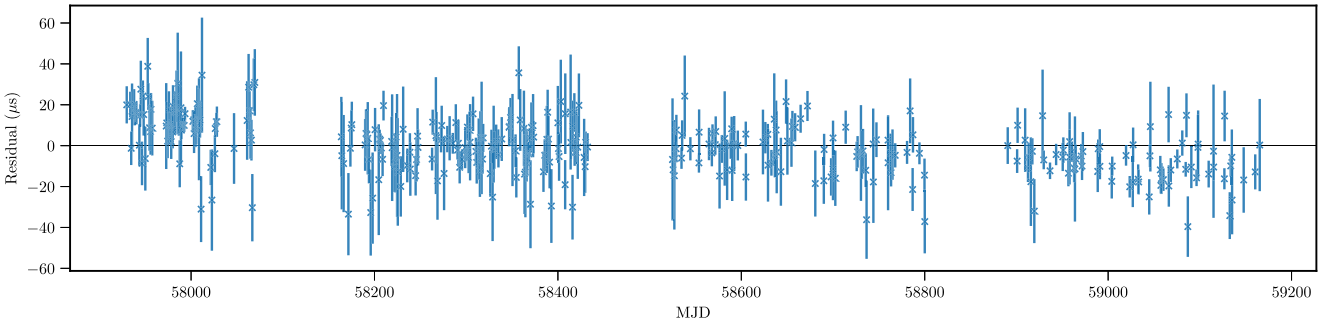

Figure 1. NICER TOA residuals for PSR J1824−2452A using the $\ddot{f}$ fit discussed in Section 4.

In this paper, we have applied existing techniques to X-ray data that have traditionally been used in radio data (see Arzoumanian et al. 2018a; Alam et al. 2020a; Hazboun et al. 2020, and references therein). We use existing software, including the Enhanced Numerical Toolbox Enabling a Robust PulsaR Inference SuitE (ENTERPRISE; Ellis et al. 2017) and a Python timing package, Pint Is Not TEMPO3 (PINT; Luo et al. 2021), to search for red noise in X-ray times of arrival (TOAs).

In Section 2, we discuss how we process NICER data, briefly reviewing the various filters we use to select events more likely to come from the pulsar. In Section 3, we introduce the various types of noise to illustrate why X-ray data are important, how they are accounted for or measured in our model, and how we construct our model. In Section 4, we describe the Bayesian analysis of our data sets, leading to our measurement of red noise using X-ray timing. In Section 5, we also discuss the analysis of an extension of the NICER mission, as well as the prospects of future high-sensitivity X-ray missions, using our newly developed data simulation tools. In Section 6, we discuss our results and summarize our investigation.

2. Observations and Data Processing

The primary instrument on NICER is the X-ray Timing Instrument with 52 coaligned X-ray concentrators, each paired with an active detector that records event time and pulse height information for each detected photon (Gendreau et al. 2017).

In addition to the photons of interest, the detectors on NICER are sensitive to cosmic rays, trapped particles, optical light, and other radiation that can produce photon-like events. We filter out spurious events and periods of high background using a variety of criteria, as explained in Deneva et al. (2019). The photons filtered out would increase the uncertainty of TOAs but would not add any type of time-correlated signal.

The measured event times have a precision of about 40 ns and are referenced to UTC with an accuracy better than 100 ns after the fine timing bias calibrations are applied in the standard level 1 processing.

In our processing, we construct one pulse TOA per NICER `ObsID`, which means one TOA per UTC day, typically with hundreds to thousands of seconds of exposure. Our data for PSR J1824−2452A include 337 TOAs taken between 2017 June 25 and 2020 November 12 with a mean TOA error of 11.4 μs and a mean observing cadence of 8.2 observations per month.

Our data for PSR B1937+21 include 466 TOAs taken between 2017 June 28 and 2020 November 18 with a mean TOA error of 5.1 μs and a mean observing cadence of 11.3 observations per month.

Figures 1 and 2 show the residuals (TOA minus the deterministic timing model used to model the TOAs) and errors for PSR J1824−2452A and PSR B1937+21, respectively. Timing models based on radio data were refit using NICER data to obtain pulsar ephemerides based solely on the X-ray data. In observations of PSR J1824−2452A, we filtered out photons outside of the 1–5.5 keV range, while in observations of PSR B1937+21, we filtered out photons outside of the 1.15–5.55 keV range. The TOAs for each of these pulsars are calculated using maximum-likelihood fits to an analytic pulse template. More complete details of the NICER timing accuracy and TOA generation procedure can be found in Section 2 of Deneva et al. (2019).

3. Methods

Understanding the different noise sources is crucial to our understanding of pulsars themselves and also the key to our ability to use them to detect gravitational waves (Arzoumanian et al. 2018a).

We divide noise in pulsar timing residuals into two categories, white and red, which is a reference to the spectrum of the fluctuations. The spectral density of white noise is independent of frequency, while red noise has a larger amplitude at lower frequency. To the eye, red noise in a time series appears as if the data wander over long timescales.

We adopt our model for this work from PTAs, which use the data from many pulsars to observe the stochastic background of gravitational waves from supermassive binary black holes (Lentati et al. 2016; Perera et al. 2019). This signal manifests in the timing data as red noise that is additionally spatially correlated, i.e., dependent upon the observation angle between pulsar lines of sight from Earth (Hellings & Downs 1983). The standard likelihood for gravitational-wave analysis with PTAs is well documented in the literature (Demorest et al. 2013; Lentati et al. 2013; van Haasteren & Levin 2013; Lentati et al. 2014; van Haasteren & Vallisneri 2015; Arzoumanian et al. 2016). Here we forgo the spatially correlated part of the signal model and search only for generic red noise within a single pulsar data set. The covariance matrix is built from the various white-noise components discussed in Section 3.1, and the red noise is modeled using a Gaussian process with a Fourier basis separately parameterized for each pulsar, as discussed in Section 3.2. The model includes a linearized timing model where the timing model parameters are marginalized over during the analysis. The linearization uses a design matrix built from the first derivatives of the best-fit timing model (Blandford et al. 1984; Hazboun et al. 2019b). The design matrix used here is the same used for radio-based pulsar timing, except that it does not include any of the chromatic (radio-frequency-dependent) components of the timing model.

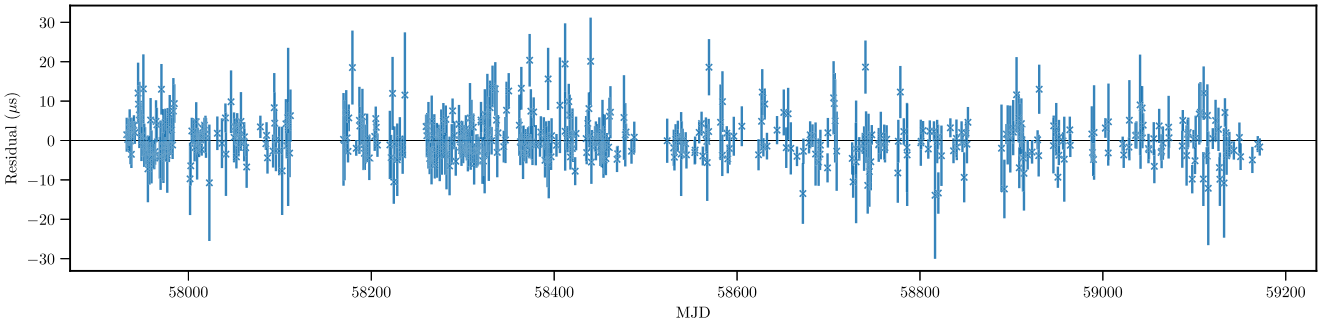

Figure 2. NICER TOA residuals for PSR B1937+21 using the $\ddot{f}$ fit discussed in Section 4.

3.1. White Noise

Additional noise parameters are often used to augment pulsar timing data errors. Here we discuss EFAC and EQUAD,[16] which were used in our analysis (see Lam et al. 2016, and references therein for further details.) In EQUAD, Q is an error added in quadrature to TOA errors that models short-timescale errors from diffractive scintillation, similar propagation effects, and pulse jitter. EFAC augments the TOA errors as a multiplicative factor, F, and models underestimates in TOA errors from low signal-to-noise ratio (S/N) TOAs, as well as allowing timing model fits to achieve reduced χ^2 fits of one. Therefore, the most generic model for the white noise in NICER data is

$$\sigma^2_{\rm total} = F^2\sigma^2_{\rm S/N} + Q^2, \tag{1}$$

where $\sigma_{\rm total}$ is the total white-noise value used to construct the covariance matrix. We will see in Section 4.1 that the additional parameters Q and F were in fact not favored in a Bayesian model selection with NICER data for PSR J1824−2452A.

3.2. Red Noise

Red noise can be categorized into chromatic and achromatic processes, where "chromatic" here refers to a dependence on the frequency of the pulsed light from the neutron star.

In radio observations of pulsars, the largest chromatic red noise comes from the bulk movement of the ISM across the line of sight to a pulsar, which causes changes in the integrated electron column density, known as the dispersion measure (DM), and hence the dispersion of the pulses (Shannon & Cordes 2017; Arzoumanian et al. 2018b). The DM effects scale as $1/\nu^2$, where ν is the electromagnetic frequency. Higher-order effects (e.g., $1/\nu^4$) can also stem from scattering of the radio pulses through the ISM and secondary frequency-dependent dispersion (Hemberger & Stinebring 2008).

Achromatic red noise can be the result of many different processes. For example, magnetospheric state switching can affect the neutron star rotation itself (Lyne et al. 2010). Some MSPs can have asteroid belts, which can cause perturbations in their orbits that appear as achromatic red noise in the residuals (Shannon et al. 2013).

Understanding the nature of red noise is crucial to building accurate pulsar timing models (Coles et al. 2011; Stinebring 2013). Furthermore, gravitational-wave detection using pulsars, i.e., pulling a very weak signal out of noisy data, requires an accurate description of the noise (Lam et al. 2016, 2017; Alam et al. 2020a).

Various scaling laws using optimal statistics to build S/Ns, ρ, have been developed extensively in the literature for understanding when/if a pulsar data set is sufficient for the detection of red noise. In Section 5, we will use a single pulsar version of the ρ expression in Chamberlin et al. (2015),

$$\rho = \left(2T\int_{f_L}^{f_H} df \frac{b^2 f^{-2\gamma}}{(bf^{-\gamma} + 2\sigma^2\Delta t)^2}\right)^{1/2}, \tag{2}$$

for detecting power-law red noise. Here γ is the power-law spectral index, Δt is the time between observations (1/cadence), the length of the data set is given by T, the white noise is given by σ, and b is defined using both the amplitude and spectral index of the red noise as the following:

$$b \equiv \frac{A^2_{\rm RN}}{12\pi^2}\left(\frac{1}{f_{\rm yr}}\right)^{3-\gamma}, \tag{3}$$

where $f_{\rm yr}$ is the reference frequency, here chosen to be 1 yr^{-1}.

The X-ray observations can play a key role in the understanding of these types of noise for the following reason: there is no chromatic noise in X-ray observations. Radio waves are coherently scattered in the ISM with a strength depending on the inverse square of the frequency, strongly affecting the observed light through dispersive delays, scintillation, multipath propagation, etc. The X-rays are (photoelectrically) absorbed or incoherently (Compton) scattered, resulting only in simple, static attenuation of the light (Longair 2011). Therefore, observing in X-ray data eliminates a significant category of noise. Further, in many pulsars, the chromatic red noise measured is orders of magnitude larger than the various achromatic red-noise types (Perera et al. 2019). Finally, mismodeled chromatic red noise can appear as achromatic noise (Cordes et al. 2016), so eliminating chromatic noise from the data set will allow for a proper characterization of achromatic noise.

Since the achromatic red noise searched for in pulsar data is a stochastic signal, Gaussian process regression (Rasmussen & Williams 2006) is often used as the main tool for modeling. Here a normal kernel process was used in the Fourier domain with a power-law model

$$P_{\rm RN} = \frac{A^2_{\rm RN}}{12\pi^2}\left(\frac{f}{f_{\rm yr}}\right)^{-\gamma} {\rm yr}^3 \tag{4}$$

[16] ECORR refers to noise that is correlated between TOAs within a particular observing epoch across radio frequencies but not correlated between epochs (Arzoumanian et al. 2014; Lam et al. 2016). Because ECORR is correlated across frequencies and necessitates multiple TOAs from a single observing epoch, it is not required in our analysis.

for the power spectral density (PSD), a technique used often by PTAs (van Haasteren & Levin 2013; Lentati et al. 2016). There are n frequencies considered ranging from $(1/T, n/T)$, where T is the time span of the data set. In longer data sets, n is usually 30, but with these much shorter data sets, we used $n = 10$, since the red noise will only manifest in the lowest few frequencies. The PSD is given in units of TOA residual power, yr^3. The amplitude, A_{RN}, is unitless and referenced to a frequency of 1 yr^{-1}. The prior used in our search is log-uniform $(10^{-20}, 10^{-11})$. The spectral index, γ, prior was uniform (0, 7), where $\gamma = 0$ is equivalent to white noise. Both are standard for pulsar noise searches (Alam et al. 2020a, 2020b).

In addition, we undertook a search for excess noise power more generically using a free spectral model. Such a search is unrestricted by any PSD model and creates a posterior for the amplitudes for all frequencies searched (Arzoumanian et al. 2018b). The priors for these parameters were log-uniform in the range $(10^{-10}, 10^{-4})$ s.

3.3. Software

The models were built using the ENTERPRISE Python package developed for full PTA analyses (Ellis et al. 2017). Two different techniques were used for sampling the likelihood. We used a standard Markov Chain Monte Carlo sampler, PTMCMCSAMPLER, for doing parameter estimation and model selection, while a nested sampler, `Dynesty`, was used in order to calculate the evidence for the red-noise likelihoods (Ellis & van Haasteren 2017; Speagle 2020). This evidence calculation allowed us to calculate the Bayes factor with respect to the base noise model, in this case, a model that only includes the TOA errors. The purely Python-based timing package, PINT, was used to fit pulsar timing models to our TOAs, as well as for adding noise to our data simulations (Luo et al. 2021).

The noise was simulated using a set of routines built explicitly for adding noise into these X-ray data sets with PINT and is available on GitHub.[17] The routines are based on a set of publicly available routines in a Python wrapper package for the pulsar timing software TEMPO2 (Hobbs et al. 2006), called LIBSTEMPO (Vallisneri 2015). The TOAs were zeroed out to match the deterministic timing model before adding noise. EFAC was added by inflating TOA errors before adding Gaussian-distributed noise, while EQUAD was added as a separate set of Gaussian-distributed perturbations. Red noise was added by pulling Fourier coefficients across 30 frequencies that were Gaussian-distributed around the chosen power-law values.

Using this software, we injected various known values of power-law red noise and checked that our recovery with ENTERPRISE was within the expected ranges, given the sensitivity of the data set.

4. Red-noise Search

We undertook a series of analyses on the NICER X-ray TOAs from PSR B1937+21 and PSR J1824−2452A using ENTERPRISE. Prior to these analyses, the X-ray TOAs from each of these pulsars were fit in PINT using an ephemeris initially based on radio TOAs but refit with only NICER data. This is an important milestone, as both the PSR B1937+21 and PSR J1824−2452A NICER data sets are now mature enough to be independently fit. Following the analysis in Deneva et al. (2019), the TOAs were fit with the second derivative of the spin frequency, $\ddot{f}$, which effectively subtracts a cubic-order polynomial fit from the data. A number of different physical processes can result in a significant detection of this parameter, including acceleration of the pulsar system, braking of the pulsar's spin from electromagnetic interaction with its surroundings, or the presence of a long-period binary companion (Liu et al. 2018, and references therein).

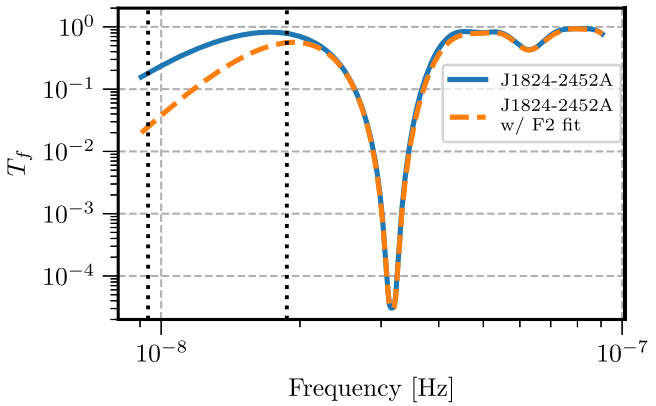

Figure 3. Transmission function for PSR J1824−2452A calculated for NICER data using `hasasia` (Hazboun et al. 2019a) with and without a fit for $\ddot{f}$. Here T_f shows the proportion of power transmitted through the timing model fit as a function of frequency. The two vertical dotted lines show $1/T$ and $2/T$, where T is the time span of the data set. The $\ddot{f}$ fit removes a substantial amount of additional power at a frequency of the inverse time span of the data.

A fit for $\ddot{f}$ removes power from the data at frequencies proportional to the inverse time span of the data (see Figure 3) and hence can absorb significant power from any red-noise process (Blandford et al. 1984; Hazboun et al. 2019b). Just as in Deneva et al. (2019), $\ddot{f}$ was removed from the pulsar ephemeris model in the search for the red noise, i.e., $\ddot{f}$ was removed from the timing model design matrix used in the linearized timing model marginalization of the analysis. This would leave power in the residuals that was initially removed by the $\ddot{f}$ fit. This fitting and then removing process ensures that any parameters covariant to $\ddot{f}$ (such as $\dot{f}$, the quadratic spin-down term) are within the linear regime for the red-noise analysis, since a linearized timing model is used to marginalize over the timing model in our search. The parameters for both pulsars from this type of fit are presented in Table 1.

As a check, the search for red noise was also carried out using pulsar models with no initial fit for $\ddot{f}$, i.e., the fit initially included only f and $\dot{f}$. This leads to a slightly different set of parameters, but using this fit led to equivalent results, showing that the linear timing model marginalization was effective at bridging the gap for the purposes of these red-noise analyses. Lastly, a third set of analyses were carried out where $\ddot{f}$ was included in the red-noise search. In the case of PSR J1824−2452A, these showed much less support for red noise, as expected. See Figure 3 as an illustration of how much power this parameter pulls from low frequencies.

Equation (2) in Johnston & Galloway (1999) allows us to calculate our expected braking index given a fit for f, $\dot{f}$, and $\ddot{f}$ for PSR J1824−2452A. If we assume that PSR J1824−2452A has $\ddot{f}$, then calculation yields a breaking index of ∼2400, which is not physically likely.

[17] https://github.com/Hazboun6/pta_sim/blob/master/pta_sim/pint_sim.py

Table 1
Timing Model for PSR J1824−2452A and PSR B1937+21

Parameter[a]	PSR J1824−2452A	PSR B1937+21
Solar system ephemeris	DE438	DE438
TT realization	TT(BIPM2019)	TT(BIPM2019)
Barycentric timescale	TDB	TDB
Start (MJD)	57,929.8	57,932.5
Finish (MJD)	59,165.3	59,171.5
DILATEFREQ	N	N
Number of TOAs	337	466
Right ascension (J2000) (hh:mm:ss.s)	**18:24:32.0077(2)**	**19:39:38.56133(6)**
decl. (J2000) (dd:mm:ss)	**−24:52:10.99(6)**	**21:34:59.126(2)**
Proper motion in R.A. (mas yr^{-1})	**−0.7(6)**	**−0.10(11)**
Proper motion in decl. (mas yr^{-1})	**20(13)**	**−0.51(17)**
Annual parallax (mas)	**−0.6(14)**	**1.9(9)**
Epoch of position (MJD)	56,999.9998000000000000	55,321.0000000000000000
F0 (s^{-1})	**27.405534870665(6)**	**641.928221244395(4)**
F1 (s^{-2})	**$-1.735216(3)\times10^{-13}$**	**$-4.33071(2)\times10^{-14}$**
Epoch (MJD)	58,547.5718813381904546	58,552.0580235874009410
TZRMJD	56,974.6086015081282523	55,800.9121432622426041
TZRSITE	ncyobs	gbt
TZRFRQ	1302.344971	812.187012

Note.
[a] Parameters in bold are those that were fit.

4.1. White-noise Model Selection

A hypermodel framework (Hee et al. 2016) was used to investigate whether any additional white-noise parameters, as discussed in Section 3.1, are needed in order to inflate the TOA errors of the NICER data. The hypermodel framework allows on-the-fly Bayesian model selection by using a hyperlikelihood built of the various models under investigation. These are combined, along with a parameter that chooses the specific model for which to evaluate the likelihood. This type of analysis was carried out recently in Goncharov et al. (2021a) for in-depth noise model selection on pulsars. An exhaustive analysis of models was done that included all combinations of EQUAD, EFAC, and power-law-modeled red noise for both pulsar data sets. The results for both pulsars are reported in Table 2. The table gives the proportion of samples in a given model, normalized to the number of samples spent in the most favored model. The odds ratios can be read off by comparing two numbers in the same column, e.g., the odds ratio of model C to model D in PSR J1824−2452A is 1:0.77. Entries of zero mean that a model was never visited. In the case of PSR J1824−2452A, the preferred model was model C, without any additional white-noise parameters but with a power-law red-noise model. For PSR B1937+21, the most preferred model was model B, which included only EQUAD and not EFAC or red noise. In both pulsars, models containing EFAC were highly disfavored, evidence that this parameter is not necessary for these X-ray data sets. Models containing EQUAD were only moderately favored or disfavored compared to models with red noise only or red noise plus EQUAD. As the red noise becomes more distinguishable in these data sets, the preference for EQUAD might become clearer.

Table 2
Sample Fractions for the Noise Model Selection in PSR J1824−2452A and PSR B1937+21

				Sample Fraction	
Model	RN	EFAC	EQUAD	J1824−2452A	B1937+21
A		✓		0	0.01
B			✓	9×10^{-5}	**1**
C	✓			**1**	0.54
D	✓		✓	0.77	0.73
E		✓	✓	0	0.02
F	✓	✓		0.06	0.01
G	✓	✓	✓	0.05	0.02

Note. A check mark designates that a given element was included in the model. The fractions are normalized by the most favored model, bolded for the two pulsars. The odds ratio can be read off by comparing two numbers from the same column, i.e., the odds ratio of model B to model C for PSR B1937+21 is 1:0.54.

4.2. Red-noise Analysis and Evidence Calculation

We proceeded to run the analyses in PSR B1937+21 and PSR J1824−2452A with the most preferred model for each pulsar found in Section 4.1. In addition to the power-law model for the PSD of the red noise used above, we also used a free spectral model where the amplitudes of the PSD at each of 10 frequencies are free parameters and are not restricted to a power-law model. The parameters for the power-law search are shown in Figures 4 and 5 and are compared to red-noise values from longer time span radio timing data. In the case of PSR B1937+21, the posteriors are not very informative, except to set an upper limit on the noise in the X-ray data. Note that the radio data parameters are close to the sensitivity threshold

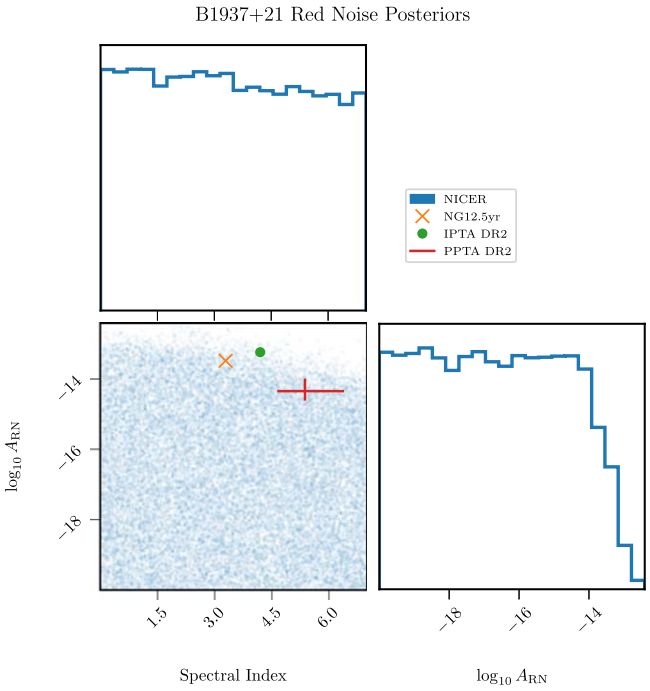

Figure 4. Bayesian two-dimensional posterior for the power-law red-noise model in the PSR B1937+21 data set. The X-ray data are currently fairly uninformative; however, the upper extent of the posteriors is nearing the measured values in Perera et al. (2019), Alam et al. (2020a), and Goncharov et al. (2021a).

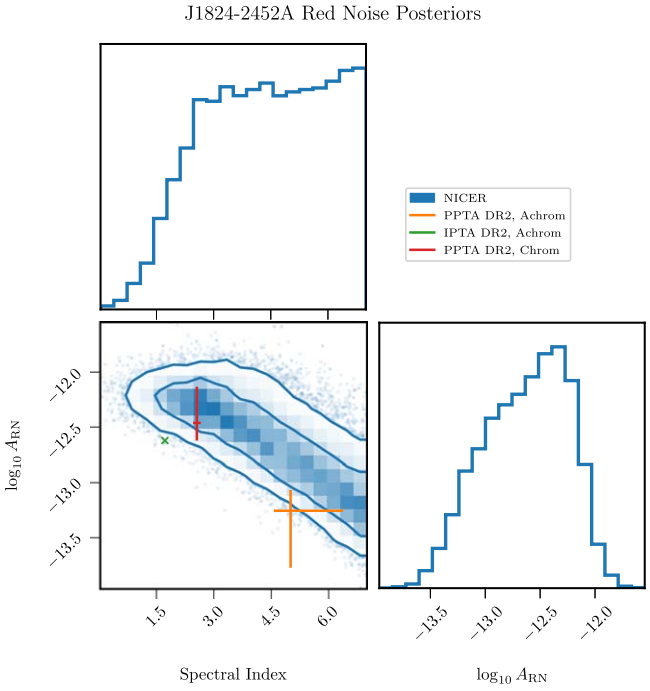

Figure 5. Bayesian two-dimensional posterior for the power-law red-noise model in the PSR J1824−2452A data set. While the amplitude posterior from the X-ray data shows a significant detection, the spectral index is fairly unconstrained. Values from other PTA data sets (Perera et al. 2019; Alam et al. 2020a; Goncharov et al. 2021a) are also shown, along with errors when available. Note that the spectral index/amplitudes for the PPTA chromatic and achromatic process are both in fair agreement with the NICER posterior.

of the NICER data set. The power-law red-noise posterior for PSR J1824−2452A shows a significant detection of red noise that is in broad agreement with various radio data results. The

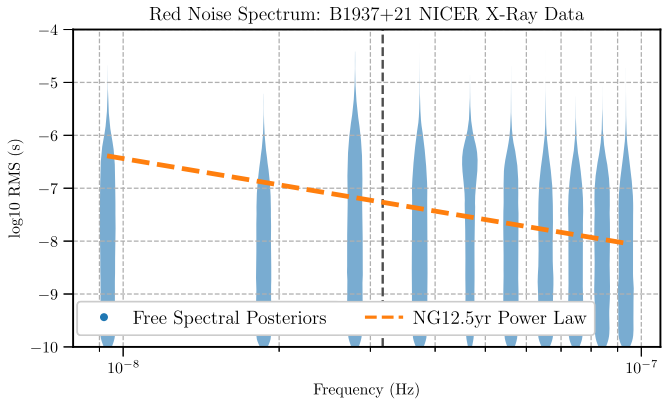

Figure 6. Free spectral analysis posteriors from PSR B1937+21 NICER data. Each of the violin plots shows the posterior probability that parameterizes the amplitude of noise in the respective frequency bin. The dashed orange line shows the best-fit power-law noise from Alam et al. (2020a).

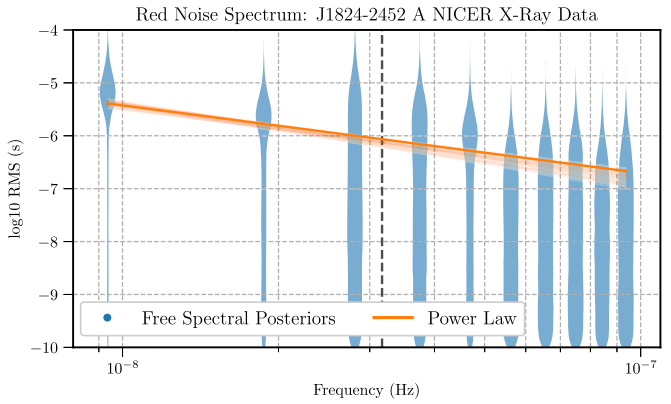

Figure 7. Free spectral analysis posteriors from PSR J1824−2452A NICER data. Each of the violin plots shows the posterior probability that parameterizes the amplitude of noise in the respective frequency bin. The thin tails extending to the minimum rms values in the two lowest-frequency bins represent significant detections of power in those bins. The solid orange lines show the two-dimensional maximum a posteriori value (bold) and a number of other realizations of the power law from the X-ray data analysis.

parameter posteriors for the free spectral model are shown in Figures 6 and 7, along with a fiducial power law for comparison. We see strong detections in the lowest two frequencies for PSR J1824–2452A.

In order to quantify the significance of both of these models, Bayesian evidence calculations were carried out using the nested sampling Python package `Dynesty`. This was done for models that included power-law red-noise models and were compared against the models without red noise. The most favored models from the model selection analysis (see Table 2) were used as the base models. In the case of PSR J1824−2452A, since red noise was the only preferred additional noise element, the noise-only model included TOA errors plus a linear timing model perturbation. In the case of PSR B1937+21, the base model had TOA errors, a linear timing model perturbation, and EQUAD. The Bayes factors calculated are given in Table 3.

We find a Bayes factor very close to 1 for the red-noise search in PSR B1937+21; therefore, we cannot claim or refute the presence of red noise in the NICER data. We find strong evidence for a power-law red-noise model in PSR J1824−2452A with a natural log Bayes factor of 9.634 ± 0.016.

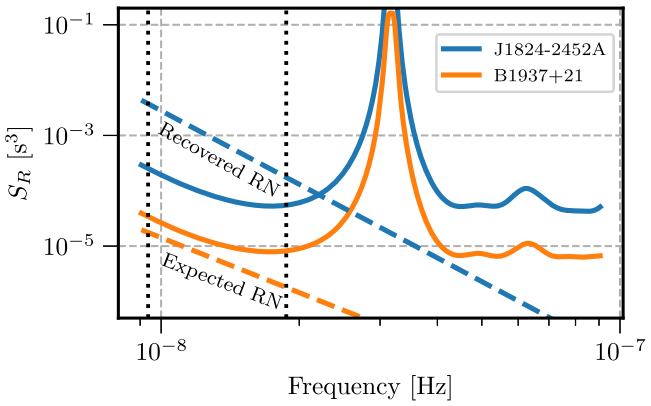

Figure 8. Comparison of pulsar sensitivity curves with power-law red noise. The solid curves show the PSD of the white noise in the residuals for the two NICER data sets. The curves were constructed analytically with `hasasia` using the preferred white-noise model from Section 4.1 and the X-ray timing models for these pulsars. The dashed lines show the best-fit power-law models for red noise from our analysis for PSR J1824−2452A and from Alam et al. (2020a) for PSR B1937+21. The red-noise power is larger than the white-noise-only S_R for the lowest two frequencies in PSR J1824−2452A but does not rise above the curve, even in the lowest frequency of PSR B1937+21. The two vertical dotted lines show $1/T$ and $2/T$, where T is the time span of the data set.

Table 3
Bayes Factors for the Presence of Red Noise in NICER Data

PSR	ln(BF)
B1937+21	-0.31 ± 0.01
J1824−2452A	9.634 ± 0.016

Note. The data for PSR J1824−2452A show a clear detection of achromatic red noise, while the data for PSR B1937+21 show no strong evidence for the presence or absence of red noise.

The theoretical noise PSD, S_R, often referred to as the sensitivity curve or just the sensitivity, of individual pulsars is well understood in the context of pulsar timing model fits and various sources of white noise (Hazboun et al. 2019b). These data sets are strongly affected by the short time span, especially when looking for steep red noise. The fit for the spin-down parameters pulls power out at the lowest frequencies important for the detection of red noise, while the astrometric fit leaves a broad peak at 1 yr^{-1} (Blandford et al. 1984). Figure 8 shows the sensitivities of both of these pulsar data sets made using `hasasia` (Hazboun et al. 2019a), along with red-noise PSDs drawn with the parameters retrieved in this analysis (in the case of PSR J1824−2452A) or from longer radio data sets (in the case of PSR B1937+21.) It is obvious from the comparisons that we would expect to detect the much larger amplitude process in PSR J1824−2452A but not the red noise in PSR B1937+21, since the radio timing parameters put the red noise below the sensitivity of this pulsar.

5. Future Prospects: Simulated X-Ray Data

Future X-ray data sets may well be able to detect red noise in PSR B1937+21 and other pulsars, as we did in PSR J1824−2452A. We ran simulations to determine what specifications (cadence and TOA error) would allow a mission to make such a detection. These simulations allow us to assess when a future X-ray mission, like the proposed Spectroscopic Time-Resolving Observatory for Broadband Energy X-rays (STROBE-X), will detect red noise in PSR B1937+21 using X-ray data alone.

5.1. Simulations

To simulate residuals, we extended the NICER 3.5 yr data set that was integrated by ObsID (data that were collected on the same UTC day) forward in time and matched the current observing cadence as closely as possible using the following algorithm. First, we took the difference between the Modified Julian Dates (MJDs) of two adjacent TOAs chosen at random within the NICER data set. We added this difference to the MJD of the most recent TOA in the growing data set to create the MJD for the next TOA. We repeated this process until the data set created after the end of the NICER data reaches our desired length (e.g., 5 yr). The result was an extended data set with a similar observing cadence but nonidentical spacing in time. We also scaled the observing cadence to create data with varying cadences. This allowed us to create data sets with different observing cadences from the NICER data but with realistic variation. We simulated errors on these new points by taking the error of a randomly chosen point in the actual NICER data and making that the error of a TOA in the simulated data set. We also scaled these errors by a multiplicative factor determined by the ratio between the desired error and the mean error of the NICER data in order to test data with smaller or larger error values. Once we have created the simulated TOAs, we use the PINT scripts mentioned in Section 3.3 to subtract the calculated residual from each TOA so that the simulated TOAs fit the timing model.

5.2. Noise Models in PSR B1937+21

The most conservative model for detecting red noise in PSR B1937+21 is one that has red noise, EQUAD, and EFAC all injected, as that adds the most noise into the data. Conversely, the most optimistic model is a model that only has red noise without EQUAD or EFAC. For the purposes of determining if red noise will be detected in PSR B1937+21 at a given observing cadence, TOA error, and mission length, any other noise model would fall between these two models in terms of our ability to make a detection of red noise. Given the expensive nature of these simulations, these two cases were adopted to bracket our understanding of NICER's ability to detect red noise in PSR B1937+21 in the near future.

5.3. Detectability of Red Noise in Future X-Ray Data

We injected red noise into the simulated data sets at the level currently observed in the NANOGrav 12.5 yr data set ($\gamma = 3.387$ and $\log_{10} A_{\rm RN} = -13.46$; Alam et al. 2020a). We conducted the analysis in the same manner as for the existing NICER data, described in Section 3.3, this time searching for a signal that we had injected. Figure 9 shows the posterior probabilities for the red-noise amplitude recovered in our analysis of simulated NICER data for two different mission lengths (5 yr as the solid blue line and 15 yr as the dashed orange line). The dotted gray line shows the prior, while the solid black line shows the injected value.

In Figure 9, we see the difference in the character of the posteriors for an upper limit (solid) and a detection (dashed), as well as the difference that a data set three times as long makes in our ability to detect red noise. The 15 yr data set has a

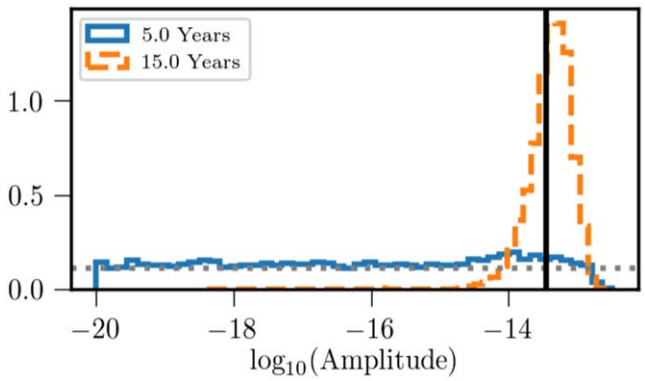

Figure 9. Posterior of recovered $\log_{10} A_{\rm RN}$ values for 5 yr of simulated NICER data on PSR B1937+21 with a cadence of 10 observations per month and an error value of 5 μs (solid blue). This is close to the real NICER data cadence and error value. Also shown is the posterior of recovered amplitude values for 15 yr of simulated NICER data on PSR B1937+21 (dashed orange). The dotted gray line represents the priors, and the solid black line represents the injected log of the red-noise amplitude of $\log_{10} A_{\rm RN} = -13.46$. At 5 yr, there is no detection of red noise, though we can see an upper limit beginning to be set on the posterior. At 15 yr, there is a clear detection of the injected red noise.

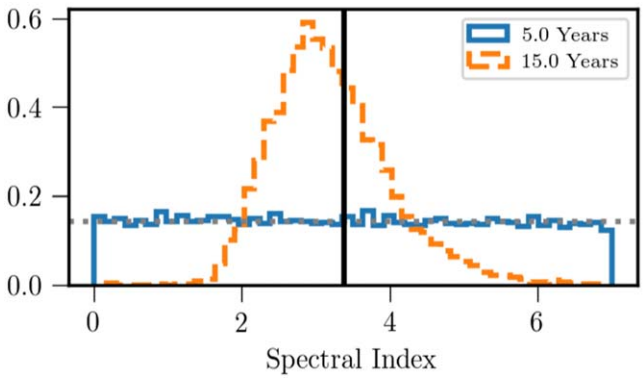

Figure 10. Posterior of recovered spectral indices for 5 yr of simulated NICER data on PSR B1937+21 (solid blue) and 15 yr of simulated NICER data on PSR B1937+21 (dashed orange). The dotted gray line represents the priors, and the solid black line represents the injected spectral index of $\gamma = 3.387$. At 15 yr, the posterior is localized, indicating a more significant detection, whereas at 5 yr, the posterior is similar to the priors.

narrow posterior, tightly localized around the injected red-noise value of $\log_{10} A_{\rm RN} = -13.46$. Hence, we were able to successfully recover the injected value. By contrast, the 5 yr posterior resembles the prior, except in the high-amplitude region that is ruled out by the data. While we are able to set an upper limit, this particular analysis is not a detection.

Figure 10 shows the posterior probabilities for the power-law spectral index recovered in our analysis of simulated NICER data of two different lengths (5 yr as the solid blue line and 15 yr as the dashed orange line). The dotted gray line shows the prior, while the solid black line shows the injected value.

Again, we see that the longer data set in Figure 10 has a clear high probability density region in the posterior, with the most likely value settling in a range near $\gamma = 3$. This is near the injected spectral index of $\gamma = 3.387$. By contrast, in the analysis of the shorter data set, the posterior still closely resembles the priors.

5.4. Detecting Red Noise in PSR B1937+21

Figure 11 reveals the strategies that a future 5 yr X-ray mission (like STROBE-X) could use for detecting red noise in PSR B1937+21. In particular, it shows the direct exchange that such a mission could take between TOA error and observing cadence and still detect red noise. For example, observations that yield a 2 μs TOA error could be made 20 times per month and detect red noise. Alternatively, a 3 μs TOA error would require observations more frequently: 30 observations per month.

In order to achieve a smaller uncertainty in TOAs than NICER, STROBE-X would need to have a larger effective collecting area than NICER, observe each pulsar for longer, or both. For example, NICER is achieving roughly 5 μs precision on PSR B1937+21 using ObsID TOAs. In order to achieve 2.5 μs precision, STROBE-X would need to have four times the effective collecting area of NICER, i.e., 7600 cm^2 instead of 1900 cm^2, or observe for four times as long (NICER observes each TOA for hundreds or even thousands of seconds; Deneva et al. 2019; Markwardt et al. 2021).

For a longer 10 yr X-ray mission, Figure 11 shows that red noise in PSR B1937+21 will be widely detectable at a variety of observing cadences and TOA errors. Specifically, if NICER continues to observe PSR B1937+21 for a total of 10 yr, it would be expected to detect red noise in the pulsar. This is shown in the region of the graph in Figure 11 closest to the intersection of the blue lines. These represent the current NICER cadence and TOA error, and the purple color is indicating a high likelihood of significant detection of red noise.

While the simulated data set used to create Figure 11 is injected with both EQUAD and EFAC according to Section 3.1, the simulated data set used to create Figure 12 only has white noise commensurate with the TOA errors and no EQUAD or EFAC. Accordingly, Figure 11 shows a more conservative analysis, as it contains a larger amount of noise than Figure 12. Therefore, the likelihood of making a significant detection with the same cadence and TOA error is higher in Figure 12 than in Figure 11.

In order to compare our numerical simulations with the theoretical S/N given in Equation (2), we need to relate ρ to the Bayes factors we have been calculating. The Laplace approximation (MacKay 2002; Romano & Cornish 2017), $2\ln\mathcal{B} \approx \rho^2 + 2\ln\left(\frac{\Delta V_1 / V_1}{\Delta V_0 / V_0}\right)$, where $\Delta V_{\mathcal{M}}$ is the characteristic spread of the likelihood around the maximum and $V_{\mathcal{M}}$ is the total parameter space volume, serves as a crude relation between the Bayes factor and the S/N. The second term on the right-hand side is a negative term that encodes the Occam penalty for the use of too many parameters. A rough approximation can be made by assuming the second term is negligible. Our detection threshold of a Bayes factor of 100 then corresponds to $\rho \approx 3$.

The white lines in Figures 11–13 show where $\rho = 3$ for the injected red-noise parameters. We would expect that along the white line, there is a 50% chance of detecting the injected red noise. Accordingly, the left panel of Figure 11 is more pessimistic than Equation (2) would indicate, while the right panel of Figures 11 and 12 are more optimistic.

5.5. Detecting Temporal Correlations Induced by the Gravitational-wave *B*ackground

We used our existing infrastructure to ask a similar question about the low-amplitude noise that might arise from a stochastic gravitational-wave background (GWB) rather than about high-

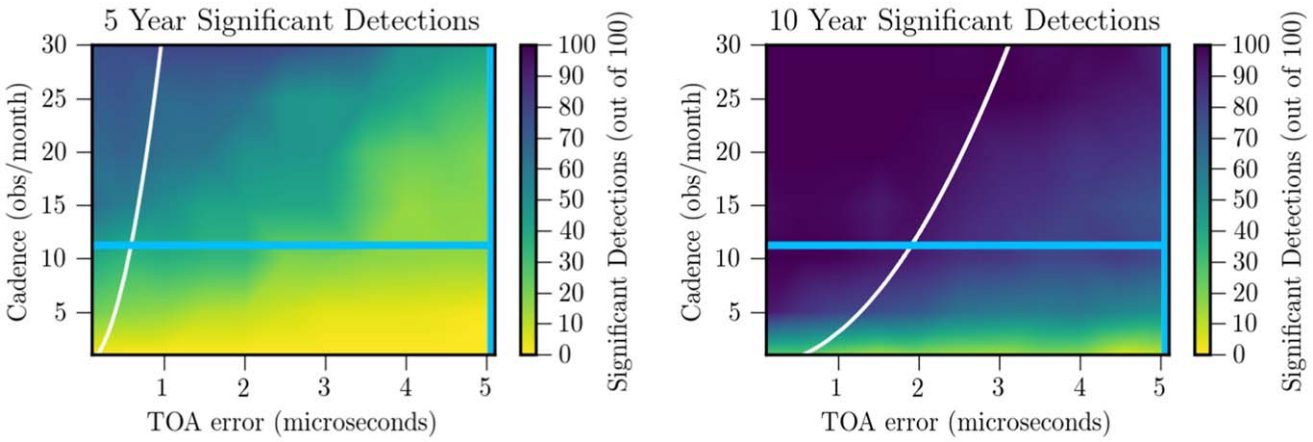

Figure 11. Number of significant red-noise recoveries as a function of TOA error and cadence for a 5 (left panel) and 10 (right panel) yr mission with injected values of $\log_{10}A_{\rm RN} = -13.46$ and $\gamma = 3.387$. The noise model was used on simulated data constructed from NICER data following the procedure in Section 5.1. At least 100 different iterations of simulated data and noise injected following the procedures in Sections 3.1 and 3.2 using a model with red noise, EQUAD, and EFAC were run at 77 points on the graph. The color of the graph, interpolated from the 77 points, indicates how many of those simulations resulted in a Bayes factor larger than 100 and a recovered $\log_{10}A_{\rm RN}$ within 1 of the injected value. The light blue lines indicate the existing TOA error and cadence for the 3.5 yr NICER data. The white line indicates where there is an S/N of 3 according to Equation (2). These figures show what the mission requirements for the next X-ray mission (such as STROBE-X) would need to be in order to detect red noise in PSR B1937+21 (see text).

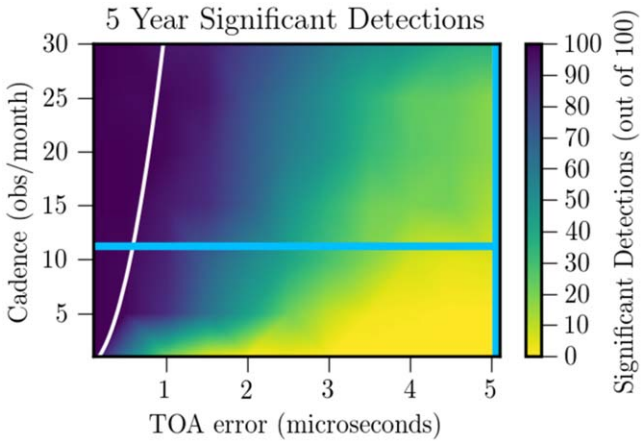

Figure 12. Number of significant red-noise recoveries as a function of TOA error and cadence for a 5 yr mission with injected values of $\log_{10}A_{\rm RN} = -13.46$ and $\gamma = 3.387$. The noise-fitting model was used on simulated data constructed from NICER data following the procedure in Section 5.1. At least 100 different iterations of simulated data and noise injected following the procedure in Section 3.2 using a model with only red noise were run at 77 points on the graph. The color of the graph, interpolated from the 77 points, indicates how many of those simulations resulted in a Bayes factor larger than 100 and a recovered $\log_{10}A_{\rm RN}$ within 1 of the injected value. The light blue lines indicate the existing TOA error and cadence for the 3.5 yr NICER data. The white line indicates where there is an S/N of 3 according to Equation (2).

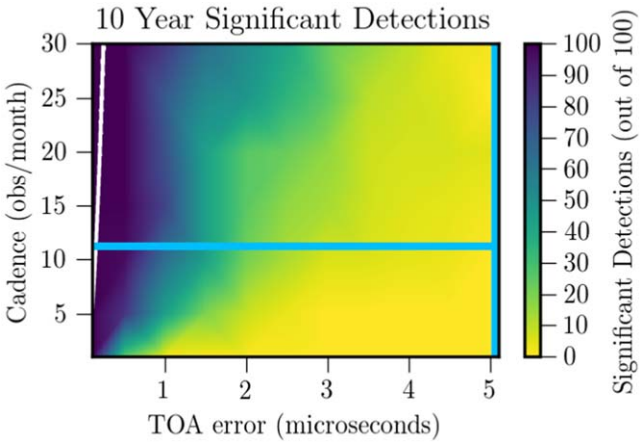

Figure 13. Number of times that the injected GWB-induced red noise was recovered as a function of TOA error and cadence for a 10 yr mission with injected values of $\log_{10}A_{\rm GWB} = -14.699$ and $\gamma_{\rm GWB} = 4.333$. The noise-fitting model was used on simulated data constructed from NICER data following the procedure in Section 5.1. At least 100 different iterations of simulated data and noise injected following the procedure in Section 3.2 with a model that only includes red noise were run at 77 points on the graph. The color of the graph, interpolated from the 77 points, indicates how many of those simulations resulted in a Bayes factor larger than 100 and a recovered $\log_{10}A_{\rm RN}$ within 1 of the injected value. The light blue lines indicate the existing TOA error and cadence for the 3.5 yr NICER data. The white line indicates where there is an S/N of 3 according to Equation (2). For this graph, the GWB was injected instead of the red noise, and no additional white noise was injected. The red noise will overpower the GWB in PSR B1937+21 in this time frame, so this is not a realistic depiction of how long it will take to detect the GWB in PSR B1937+21.

amplitude red noise. The former cannot be detected with a single pulsar, requiring the detection of spatial correlations across many lines of sight. Using the latest values from Arzoumanian et al. (2018a), we injected $\log_{10}A_{\rm GWB} = -14.699$ and $\gamma_{\rm GWB} = 4.333$ into the simulated data for PSR B1937+21 in place of red noise without injecting EQUAD or EFAC with a simulated 10 yr mission length. The results are shown in Figure 13. It is important to note that the GWB was injected in place of the red noise, not in addition to it, and that no additional white noise was injected. Because the red noise in PSR B1937+21 is strong, and our best model for PSR B1937+21 shown in Section 4.1 does include EQUAD, this is not a realistic representation of a detection of GWB-induced red noise in actual PSR B1937+21 data. However, this analysis is useful in determining mission requirements for detecting GWB-induced red noise in a pulsar like PSR B1937+21 that only has the GWB as noise in addition to the NICER TOA errors. This would be the most optimistic scenario for the detection of the GWB-induced red noise in a pulsar using X-ray data. In such a pulsar, a future X-ray mission would need to have an observing cadence close to once a day with a TOA error under 2.5 μs or a TOA error under 1.0 μs and a more infrequent observing cadence to detect the GWB-induced red noise with a 10 yr mission.

6. Discussion and Summary

Using NICER X-ray data alone, we detect red noise in PSR J1824−2452A with $\log_{10} A_{\rm RN} = -12.60^{+0.36}_{-0.46}$ and a spectral index of $\gamma = 4.41^{+1.81}_{-1.83}$, in agreement with radio observations. The X-ray observations are free from the influence of ISM that impacts radio data, making this detection significant, as it is a detection of red noise that is known not to be caused by the ISM. While the red noise in PSR B1937+21 was below the threshold of a detection, our analysis of simulated future data will help design future X-ray missions and inform when such a detection will take place. The detection of achromatic red noise in PSR B1937+21 is well known in the pulsar timing community (Kaspi et al. 1994; Lentati et al. 2016; Alam et al. 2020a, 2020b). As the NICER data set becomes more sensitive in the next few years, it will be extremely interesting to see how much red noise is detected and how much of the red noise seen in radio timing data is achromatic. These studies will be bolstered by the substantial amount of data from gamma-ray timing of pulsars (Kerr et al. 2015) by the Fermi Gamma-ray Telescope and the ongoing effort to search for the GWB in Fermi data (Kerr et al. 2022, submitted).

As has been the case in the pulsar timing community for the last decade, an analysis of noise models that yields the PSD of the noise, such as we have done, here replaces the σ_z analysis outlined in Matsakis et al. (1997) and earlier attempts at using second-order structure functions (Cordes & Downs 1985) that were previously developed to quantify the stability of a pulsar using time domain methods. The clock community and the pulsar timing community are both interested in the stability of pulsars, and they each have a different language to describe them, but we believe this analysis will appeal to both. Both communities speak the language of PSD, and ultimately, that is what both communities would like to describe their clocks, i.e., what is the PSD of the noise?

If one wishes to relate the PSD, e.g., our results, to σ_z, Matsakis et al. (1997) gave a relationship between the spectral index of a power-law PSD in the residuals, $S_R \propto f^{-\gamma}$, $0 < \gamma < 6$ and their statistic, $\sigma_z^2 \propto \tau^{\gamma-3}$. Our results for PSR J1824−2452A of $\gamma \sim 3$ would give a flat dependence on τ, similar to the Allan deviation. Allan statistics have an exact correspondence to PSD power-law models for spectral index, $0 < \gamma < 4$; $\gamma = 3$ is formally a flicker-FM noise type. A low-frequency cutoff filter can be employed, if necessary, for the case $2 < \gamma < 6$ (Makdissi et al. 2010). Future analysis will implement Allan statistical treatments so that standard clock characterizations of NICER data can be readily used by clock analysts (Howe 2006; Howe & Schlossberger 2022)

Future projects will improve on the NICER data set in both sensitivity and quantity of data. We have used our analysis of the NICER data to make recommendations for future missions. STROBE-X will observe in the 0.2–12 keV band using the X-ray Concentrator Array with a planned collecting area of 21,760 $\rm cm^2$ at 1.5 keV (Ray et al. 2019). This increase in collecting area alone will provide a marked improvement over NICER's observations. By conducting tests with longer time spans and smaller uncertainties, we aim to provide thresholds for these future experiments in X-ray timing.

J.S.H. wishes to thank Jim Cordes for useful discussions. The NANOGrav project receives support from the National Science Foundation (NSF) as Physics Frontier Center award No. 1430284. Portions of this work performed at NRL were supported by NASA.

ORCID iDs

Jeffrey S. Hazboun https://orcid.org/0000-0003-2742-3321
Andrea N. Lommen https://orcid.org/0000-0003-4137-7536
Elizabeth Teng https://orcid.org/0000-0003-0420-2067
Paul S. Ray https://orcid.org/0000-0002-5297-5278
Matthew Kerr https://orcid.org/0000-0002-0893-4073
Slavko Bogdanov https://orcid.org/0000-0002-9870-2742
Julia Deneva https://orcid.org/0000-0003-1226-0793
Natalia Lewandowska https://orcid.org/0000-0003-0771-6581
Craig B. Markwardt https://orcid.org/0000-0001-9803-3879
Scott Ransom https://orcid.org/0000-0001-5799-9714
Teruaki Enoto https://orcid.org/0000-0003-1244-3100
Keith C. Gendreau https://orcid.org/0000-0001-7115-2819
Aditya Parthasarathy https://orcid.org/0000-0002-4140-5616

References

Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2020a, ApJS, 252, 1
Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2020b, ApJS, 252, 5
Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2018a, ApJ, 859, 47
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2014, ApJ, 794, 141
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2016, ApJ, 821, 13
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2018b, ApJS, 235, 37
Blandford, R., Narayan, R., & Romani, R. W. 1984, JApA, 5, 369
Chamberlin, S. J., Creighton, J. D. E., Siemens, X., et al. 2015, PhRvD, 91, 044048
Coles, W., Hobbs, G., Champion, D. J., Manchester, R. N., & Verbiest, J. P. W. 2011, MNRAS, 418, 561
Cordes, J. M., & Downs, G. S. 1985, ApJS, 59, 343
Cordes, J. M., & Shannon, R. M. 2010, arXiv:1010.3785
Cordes, J. M., & Shannon, R. M. 2012, ApJ, 750, 89
Cordes, J. M., Shannon, R. M., & Stinebring, D. R. 2016, ApJ, 817, 16
Demorest, P. B., Ferdman, R. D., Gonzalez, M. E., et al. 2013, ApJ, 762, 94
Deneva, J. S., Ray, P. S., Lommen, A., et al. 2019, ApJ, 874, 160
Ellis, J., Taylor, S., Baker, P., & Vallisneri, M. 2017, ENTERPRISE v3.3.1, Zenodo, doi: 10.5281zenodo.4059815
Ellis, J., & van Haasteren, R. 2017, PTMCMCSampler: Official Release v2.0, Zenodo, doi:10.5281/zenodo.1037579jellis18/
Gendreau, K., Arzoumanian, Z. & NICER Team 2017, AAS Meeting, 229, 309.03
Goncharov, B., Reardon, D. J., Shannon, R. M., et al. 2021a, MNRAS, 502, 478
Goncharov, B., Shannon, R. M., Reardon, D. J., et al. 2021b, ApJL, 917, L19
Hazboun, J., Romano, J., & Smith, T. 2019a, JOSS, 4, 1775
Hazboun, J. S., Romano, J. D., & Smith, T. L. 2019b, PhRvD, 100, 104028
Hazboun, J. S., Simon, J., Taylor, S. R., et al. 2020, ApJ, 890, 108
Hee, S., Handley, W. J., Hobson, M. P., & Lasenby, A. N. 2016, MNRAS, 455, 2461
Hellings, R. W., & Downs, G. S. 1983, ApJL, 265, L39
Hemberger, D. A., & Stinebring, D. R. 2008, ApJL, 674, L37
Hobbs, G., Archibald, A., Arzoumanian, Z., et al. 2010, CQGra, 27, 084013
Hobbs, G. B., Edwards, R. T., & Manchester, R. N. 2006, MNRAS, 369, 655
Howe, D. A., & Schlossberger, N. 2022, IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, 69, 468
Howe, D. A. 2006, Metro, 43, S322
Johnston, S., & Galloway, D. 1999, MNRAS, 306, L50
Kaspi, V. M., Taylor, J. H., & Ryba, M. F. 1994, ApJ, 428, 713
Kerr, M., Ray, P. S., Johnston, S., Shannon, R. M., & Camilo, F. 2015, ApJ, 814, 128
Lam, M. T., Cordes, J. M., Chatterjee, S., et al. 2016, ApJ, 819, 155
Lam, M. T., Cordes, J. M., Chatterjee, S., et al. 2017, ApJ, 834, 35
Lentati, L., Alexander, P., Hobson, M. P., et al. 2013, PhRvD, 87, 104021
Lentati, L., Hobson, M. P., & Alexander, P. 2014, MNRAS, 444, 3863
Lentati, L., Shannon, R. M., Coles, W. A., et al. 2016, MNRAS, 458, 2161
Liu, X. J., Bassa, C. G., & Stappers, B. W. 2018, MNRAS, 478, 2359
Longair, M. S. 2011, High Energy Astrophysics (Cambridge: Cambridge Univ. Press)

Luo, J., Ransom, S., Demorest, P., et al. 2021, ApJ, 911, 45
Lyne, A., Hobbs, G., Kramer, M., Stairs, I., & Stappers, B. 2010, Sci, 329, 408
MacKay, D. J. C. 2002, Information Theory, Inference & Learning Algorithms (New York: Cambridge Univ. Press)
Makdissi, A., Vernotte, F., & Clercq, E. D. 2010, ITUFF, 57, 1011
Markwardt, C., Gendreau, K., Arzoumanian, Z., Corcoran, M., & Ray, P. 2021, NICER Mission Guide, NASA, https://heasarc.gsfc.nasa.gov/docs/nicer/mission_guide/
Matsakis, D. N., Taylor, J. H., & Eubanks, T. M. 1997, A&A, 326, 924
Melatos, A., & Link, B. 2014, MNRAS, 437, 21
Perera, B. B. P., DeCesar, M. E., Demorest, P. B., et al. 2019, MNRAS, 490, 4666
Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press), 4
Ray, P. S., Arzoumanian, Z., Ballantyne, D., et al. 2019, arXiv:1903.03035
Romano, J. D., & Cornish, N. J. 2017, LRR, 20, 2
Shannon, R. M., & Cordes, J. M. 2017, MNRAS, 464, 2075
Shannon, R. M., Cordes, J. M., Metcalfe, T. S., et al. 2013, ApJ, 766, 5
Speagle, J. S. 2020, MNRAS, 493, 3132
Stinebring, D. 2013, CQGra, 30, 224006
Vallisneri, M. 2015, libstempo: Python wrapper for Tempo2, Astrophysics Source Code Library, ascl:2002.017
van Haasteren, R., & Levin, Y. 2013, MNRAS, 428, 1147
van Haasteren, R., & Vallisneri, M. 2015, MNRAS, 446, 1170