The Astrophysical Journal, 973:94 (20pp), 2024 October 1 https://doi.org/10.3847/1538-4357/ad656e

Reducing Instrumental Errors in Parkes Pulsar Timing Array Data

Axl F. Rogers[1], Willem van Straten[2], Sergei Gulyaev[1], Aditya Parthasarathy[3], George Hobbs[4], Zu-Cheng Chen[5,6], Yi Feng[7], Boris Goncharov[8,9], Agastya Kapur[4], Xiaojin Liu[10], Daniel Reardon[11,12], Christopher J. Russell[13], and Andrew Zic[14]

[1] Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand
[2] Manly Astrophysics, 15/41-42 East Esplanade, Manly, NSW 2095, Australia
[3] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[4] CSIRO Astronomy & Space Science, Australia Telescope National Facility, P.O. Box 76, Epping, NSW, 1710, Australia
[5] Department of Physics and Synergetic Innovation Center for Quantum Effects and Applications, Hunan Normal University, Changsha, Hunan 410081, People's Republic of China
[6] Institute of Interdisciplinary Studies, Hunan Normal University, Changsha, Hunan 410081, People's Republic of China
[7] Research Center for Astronomical Computing, Zhejiang Laboratory, Hangzhou 311100, People's Republic of China
[8] Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy
[9] INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi, Italy
[10] Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, People's Republic of China
[11] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, Victoria 3122, Australia
[12] OzGrav: The Australian Research Council Centre of Excellence for Gravitational Wave Discovery, Hawthorn VIC 3122, Australia
[13] CSIRO Scientific Computing, Australian Technology Park, Locked Bag 9013, Alexandria, NSW 1435, Australia
[14] Department of Physics and Astronomy and MQ Research Centre in Astronomy, Astrophysics and Astrophotonics, macquarie University, NSW 2109, Australia
Received 2024 April 20; revised 2024 July 16; accepted 2024 July 17; published 2024 September 23

Abstract

This paper demonstrates the impact of state-of-the-art instrumental calibration techniques on the precision of arrival times obtained from 9.6 yr of observations of millisecond pulsars using the Murriyang 64 m CSIRO Parkes Radio Telescope. Our study focuses on 21 cm observations of 25 high-priority pulsars that are regularly observed as part of the Parkes Pulsar Timing Array project, including those predicted to be the most susceptible to calibration errors. We employ measurement equation template matching (METM) for instrumental calibration and matrix template matching (MTM) for arrival time estimation, resulting in significantly improved timing residuals with up to a sixfold reduction in white noise compared to arrival times estimated using scalar template matching and conventional calibration based on the ideal feed assumption. The median relative reduction in white noise is 33%, and the maximum absolute reduction is 4.5 μs. For PSR J0437−4715, METM and MTM reduce the best-fit power-law amplitude (2.7σ) and spectral index (1.7σ) of the red noise in the arrival time residuals, which can be tentatively interpreted as mitigation of $1/f$ noise due to otherwise unmodeled steps in polarimetric response. These findings demonstrate the potential to directly enhance the sensitivity of pulsar timing array experiments through more accurate methods of instrumental calibration and arrival time estimation.

Unified Astronomy Thesaurus concepts: Millisecond pulsars (1062); Gravitational waves (678); Polarimetry (1278)

1. Introduction

Pulsar timing arrays (PTAs) are invaluable tools for detecting spatially correlated signal fluctuations at low frequencies, spanning from nHz to μHz (Sazhin 1978; Detweiler 1979; Hellings & Downs 1983). Within this frequency band, the dominant anticipated signal is the stochastic gravitational-wave background (GWB) generated by a cosmic population of inspiralling supermassive black hole binaries (Sesana et al. 2004; Burke-Spolaor et al. 2019). Additional speculative sources of gravitational waves (GWs) in the nHz range encompass cosmic strings (Siemens et al. 2007; Blanco-Pillado et al. 2018), phase transitions (Caprini et al. 2010; Kobakhidze et al. 2017; Xue et al. 2021), and a primordial GWB originating from quantum fluctuations of the gravitational field during the early Universe, amplified by inflation (Grishchuk 1975; Lasky et al. 2016).

The GWB is expected to manifest as a common red-noise process, characterized by a similar spectral signature, in all pulsars within the array (Phinney 2001). The detection of spatially correlated signals in the times of arrival (ToAs) of multiple pulsars (Rajagopal & Romani 1995), with the quadrupolar signature initially proposed by Hellings & Downs (HD correlation; 1983), provides compelling evidence of a stochastic GWB detection (Agazie et al. 2023a). Over the past decade, PTA collaborations have steadily improved the sensitivity of their data sets for GW searches, progressively reducing the upper limits on the stochastic GWB amplitude (van Haasteren et al. 2011; Demorest et al. 2013; Shannon et al. 2013a, 2015; Arzoumanian et al. 2014, 2016, 2018; Lentati et al. 2015; Verbiest et al. 2016; Perera et al. 2018; Alam et al. 2021a, 2021b), and working toward identifying individual GW sources (Yardley et al. 2010; Zhu et al. 2014; Babak et al. 2016; Aggarwal et al. 2019).

Beyond GWs, various astrophysical processes can introduce red noise that is unique to each pulsar (Coles et al. 2011; van Haasteren & Levin 2013; Lentati et al. 2014). These include intrinsic spin noise (Shannon & Cordes 2010; Melatos & Link 2014; Lam et al. 2017), magnetospheric torque variations (Lyne et al. 2010), variable dispersion (Keith et al. 2013; Jones et al. 2017), and multipath propagation effects in the interstellar medium (ISM; Hemberger & Stinebring 2008; Cordes &

Shannon 2010; Dolch et al. 2021), and the presence of undetected objects in orbit around pulsars (Shannon et al. 2013b).

The International Pulsar Timing Array (IPTA) collaboration, composed of various PTA projects, is dedicated to the pursuit of a common goal: the detection of nHz frequency GWs (Hobbs et al. 2010; Verbiest et al. 2016; Perera et al. 2019). Currently, the IPTA encompasses four PTA members: the European Pulsar Timing Array (EPTA; Kramer & Champion 2013; Desvignes et al. 2016), the Indian Pulsar Timing Array (Joshi et al. 2018; Susobhanan et al. 2020), the North American Nanohertz Observatory for Gravitational-Waves (NANOGrav; McLaughlin 2013; Arzoumanian et al. 2018; Cordes & McLaughlin 2019; Ransom et al. 2019), and the Parkes Pulsar Timing Array (PPTA; Manchester et al. 2013; Reardon et al. 2016; Kerr et al. 2020). Furthermore, efforts are underway to establish a PTA project in China (Chinese Pulsar Timing Array, hereafter CPTA), leveraging the Five-hundred-meter Aperture Spherical Telescope (Lee 2016; Hobbs et al. 2019). Other scientific initiatives, such as the MeerKAT PTA in South Africa (Bailes et al. 2016), and the Canadian Hydrogen Intensity Mapping Experiment pulsar collaboration in Canada (Ng 2018), are poised to contribute to the IPTA's collaborative endeavors. An essential addition to this list is the γ-ray PTA (Fermi PTA), which offers an independent probe of the GWB and stands as the sole means of confirming radio PTA results (Ajello et al. 2022).

In 2020 and 2021, significant advancements were made in our understanding of PTAs. The NANOGrav 12.5 yr data set (NG12.5; Arzoumanian et al. 2020), EPTA second data release (EPTA DR2; Chen et al. 2021), and PPTA second data release (PPTA DR2; Goncharov et al. 2021) all revealed a common uncorrelated red-noise (CURN) process within their data sets, marking a crucial discovery. However, these studies failed to conclusively confirm or refute the presence of hypothetical HD correlations.

This revelation gained further support through an analysis of the second data release from IPTA (Perera et al. 2019), which consolidated historical data from EPTA, NANOGrav, and PPTA, and confirmed the existence of a CURN process (Antoniadis et al. 2022). However, an independent analysis of PPTA DR2 by Goncharov et al. (2021) shed light on the potential misinterpretation of noise without a statistically identical spectrum between pulsars as a common red process.

PTAs demonstrate sensitivity not only to the quadrupolar correlation of GWs (Taylor et al. 2016; Burke-Spolaor et al. 2019) but also to various other correlated signals. These can include monopolar correlation due to terrestrial time standards errors (Hobbs et al. 2012, 2020), and dipolar correlation stemming from errors in the solar system ephemeris model (Champion et al. 2010; Caballero et al. 2018). Incorrect modeling of these sources may introduce red noise into timing residuals, potentially compromising GW detection sensitivity (Tiburzi et al. 2016). While NG12.5 ruled out monopolar and dipolar spatially correlated signals (Arzoumanian et al. 2020), the analyses of timing residuals in NG12.5, EPTA DR2, and PPTA DR2 do not yield statistical support for a quadrupolar spatial correlation (Arzoumanian et al. 2020; Chen et al. 2021; Goncharov et al. 2021).

The most recent GWB search papers by the EPTA (Antoniadis et al. 2023a), NANOGrav (Agazie et al. 2023b), and PPTA (Reardon et al. 2023) present analyses of their latest data releases: EPTA DR2 (Antoniadis et al. 2023b), NANOGrav 15 yr data set (Agazie et al. 2023a), and PPTA third data release (PPTA DR3; Zic et al. 2023). These papers unveil evidence for an HD-correlated GWB with varying levels of significance, with estimated probabilities of false alarm rates at 3σ, 3σ–4σ, and 2σ, respectively. A comparison of these studies showcases consistent measurements of nHz GWB parameters, even with diverse data modeling approaches, demonstrating agreement within 1σ. The coherence of pulsar noise parameters in the majority of analyzed pulsars and the standardization of noise models reconcile modeling disparities, refining constraints on GWB amplitude and HD correlations. This advancement provides a robust foundation for IPTA's Data Release 3 by extending data sets to encompass additional pulsars (Agazie et al. 2024). Concurrently, CPTA also reported similar findings on the HD-correlated GWB (Xu et al. 2023), aligning with the broader consensus from other PTA projects.

The sensitivity of PTA experiments is founded upon the accuracy and precision with which arrival times can be estimated. Therefore, it remains important to study and quantify the extent to which PTA sensitivity may be limited by unmodeled instrumental artifacts and calibration errors. Instrumental polarization can distort pulse profiles, leading to correlated errors in the ToAs of each pulsar that can mimic signals associated with a stochastic GWB (van Straten 2013; Lentati et al. 2016). Consequently, polarimetric calibration is crucial for minimizing systematic timing errors (van Straten 2006, hereafter S06; Guillemot et al. 2023), which are most readily observed as dramatic variations of arrival time residuals as a function of the parallactic angle (e.g., Figure 1 of van Straten 2013). Various methods, including enhanced arrival time estimation (Hotan et al. 2005; S06), instrumental calibration (Jenet & Anderson 1998; van Straten 2004), and radio frequency interference (RFI) mitigation (Lazarus et al. 2016, 2020; Reardon 2021), have been developed to quantify and mitigate sources of systematic error, thereby enhancing the precision and accuracy of ToA estimates.

For most pulsars, modest instrumental distortion can induce systematic timing errors of the order of 100 ns (van Straten 2013), significantly hindering efforts to detect the stochastic GWB (Jenet et al. 2005). In this study, we investigate the instrumental distortion of arrival times estimated for the 25 high-priority millisecond pulsars (MSPs) regularly observed for the PPTA project, which includes pulsars with a great potential for improvement by addressing calibration errors (van Straten 2013).

The paper's structure is as follows: we describe the data set in Section 2, detail our methods in Section 3, present our results in Section 4, and conclude with a discussion of our findings and prospects for future research in Section 5.

2. Observations

The observations analyzed in this study were carried out using the Murriyang 64 m radio telescope at Parkes. Our analysis focuses on a subset of the data from PPTA DR2, which provides precise arrival times for 26 ms pulsars. (PSR J1732−5049 was excluded from this research as it was recently removed from the PPTA's high-priority list.) DR2 spans a 14 yr period, with an observational cadence of approximately 3 weeks, and includes observations from various radio frequency bands (10, 20, 40/50 cm) and backend instruments (CASPSR, CPSR2, PDFBs, WBCORR;

Table 1
Observational Characteristics of 25 High-priority PPTA Pulsars, Including J2000.0 Coordinates (JNAME), Spin Period (P) in Milliseconds, Observed ToAs ($O_{\rm ToA}$), Remaining ToAs after Outlier Rejection ($N_{\rm ToA}$), Observation Time Span (in Years), and Corresponding Modified Julian Date Range

Pulsar (JNAME)	P (ms)	$O_{\rm ToA}$	$N_{\rm ToA}$	Span (yr)	MJD Range (start–finish)
PSR J0437−4715	5.76	882	700	8.210	55409–58933
PSR J0613−0200	3.06	238	221	8.773	55427–58752
PSR J0711−6830	5.49	349	333	9.565	55427–58932
PSR J1017−7156	2.34	353	333	8.764	55472–58749
PSR J1022+1001	16.45	241	222	8.172	55426–58410
PSR J1024−0719	5.16	138	127	9.098	55426–58749
PSR J1045−4509	7.47	187	177	9.051	55444–58749
PSR J1125−6014	2.63	150	141	8.865	55694–58932
PSR J1446−4701	2.19	170	145	8.913	55677–58932
PSR J1545−4550	3.58	133	123	6.621	56513–58931
PSR J1600−3053	3.60	189	174	9.551	55444–58932
PSR J1603−7202	14.84	222	212	9.095	55427–58749
PSR J1643−1224	4.63	167	155	9.057	55445–58753
PSR J1713+0747	4.57	256	246	9.059	55444–58753
PSR J1730−2304	8.12	186	165	9.598	55427–58932
PSR J1744−1134	4.07	253	237	9.057	55445–58753
PSR J1824−2452A	3.05	87	80	8.923	55493–58752
PSR J1832−0836	2.72	78	72	6.804	56447–58932
PSR J1857+0943	5.36	138	127	9.515	55457–58932
PSR J1909−3744	2.95	374	344	8.926	55444–58752
PSR J1939+2134	1.56	153	142	7.823	55472–58329
PSR J2124−3358	4.93	227	221	9.474	55472–58933
PSR J2129−5721	3.73	255	238	9.516	55457–58933
PSR J2145−0750	16.05	224	211	9.474	55471–58932
PSR J2241−5236	2.19	399	329	9.644	55410–58933

Note. Note that, for PSR J1022+1001, ToAs from August 15 to 30 each year were excluded in consideration of solar conjunction.

Kerr et al. 2020). Notably, DR2 yielded the lowest rms timing residuals for each pulsar up to the point of our analysis, which precedes PPTA DR3.

The selected subset of PPTA DR2 comprises observations conducted between MJDs 55409 and 58933 (from 2010 August 1 to 2020 March 25) at 1400 MHz (corresponding to the 20 cm band observations) employing the H–OH and 21 cm Multi-beam receivers and CASPSR backend. Although CASPSR was also used to observe in the 10 cm (∼3100 MHz) and 40/50 cm (∼700 MHz) bands, our analysis focuses solely on the 20 cm band observations due to the limited availability of 10 cm data and the significant RFI contamination in the 40/50 cm band (Parthasarathy et al. 2019). Please refer to Table 1 for more details about the observations of each pulsar.

3. Methods

The following sections describe the two methods of polarimetric calibration and the two methods of arrival time estimation that are compared in this work.

3.1. Polarimetric Calibration

To establish a baseline for comparing the impact of calibration techniques, data are calibrated using an estimate of the polarimetric response of the observing system based on an approximation known as the *ideal feed assumption* (IFA; Manchester et al. 2013). The IFA calibration model includes the assumptions that the receptors are perfectly orthogonally polarized and that the reference source (e.g., a pulsed noise diode coupled to the receptors) is 100% linearly polarized, illuminating both receptors equally and in-phase (Caleb et al. 2019). The IFA is an incomplete description of the instrumental response, and, for some systems, calibration based on the IFA results in significant systematic variations of the total intensity profile and arrival time distortions. Therefore, for comparison in this study, the data are also calibrated using measurement equation template matching (METM; van Straten 2013). METM uses a single well-calibrated observation of a pulsar with a high signal-to-noise ratio (S/N) as a polarized reference source, one or more uncalibrated observations of the same pulsar, and (optionally) observations of an amplitude-modulated reference source.

For this study, PSR J0437−4715 is used as the polarized reference source, and its template profile is derived from multiple measurement equation modeling (MEM) solutions. MEM uses uncalibrated observations of a pulsar observed at multiple parallactic angles and an amplitude-modulated reference source. We employ the polarimetric calibration pipeline (PSRPL)[15] to perform both MEM and METM on multiple data sets. Following RFI excision using MEERGUARD,[16] the Meertime extension of COASTGUARD[17] (Lazarus et al. 2016, 2020), we generate 5 minute subintegrations of PSR J0437−4715 and 2 minute subintegrations of noise diode observations. Separate calibrator models are produced for each data set.

During the MEM stage, Equation (19) of Britton (2000) is used to parameterize the unknown response of the nonideal receiver, and the degenerate model parameters described in Appendix B of van Straten (2004) are set to zero. The degenerate model parameters include the difference in the receptor ellipticities, δ_χ, which mixes Stokes I and Stokes V, and the rotation of the receiver about the line of sight, σ_θ, which mixes Stokes Q and Stokes U. By setting these parameters to zero, the receptors are assumed to have equal (and opposite) ellipticities, and the rotation of the receiver about the line of sight is assumed to be zero. The unknown polarization of the nonideal noise diode signal is parameterized by the three components of the normalized Stokes polarization vector, $\hat{\boldsymbol{C}} = (\hat{C}_1, \hat{C}_2, \hat{C}_3) = (C_1, C_2, C_3)/C_0$, where C_0 is the total intensity of the noise diode, and, for linearly polarized receptors, C_1, C_2, and C_3 correspond to Stokes Q, U, and V of the noise diode.

From a total of 66 best-fit MEM solutions, the 52 most robust solutions with a weighted mean reduced χ^2 (averaged over all frequency channels) less than 1.05 and greater than 0.85 are selected. The selected solutions are then ranked based on the product of several range-normalized attributes. Given the minimum and maximum values, $x_{\rm min}$ and $x_{\rm max}$, of some attribute x, and the value of that attribute derived from the ith solution x_i, the (dimensionless) range-normalized attribute

$$\hat{x}_i = \frac{x_i - x_{\rm min}}{x_{\rm max} - x_{\rm min}} \quad (1)$$

lies on the interval [0, 1]. This normalization gives equal weight to each of the attributes included in the rank metric, which is a function of the integration length of the session, the

[15] http://psrchive.sourceforge.net/manuals/psrpl/
[16] https://github.com/danielreardon/MeerGuard
[17] https://github.com/plazar/coast_guard

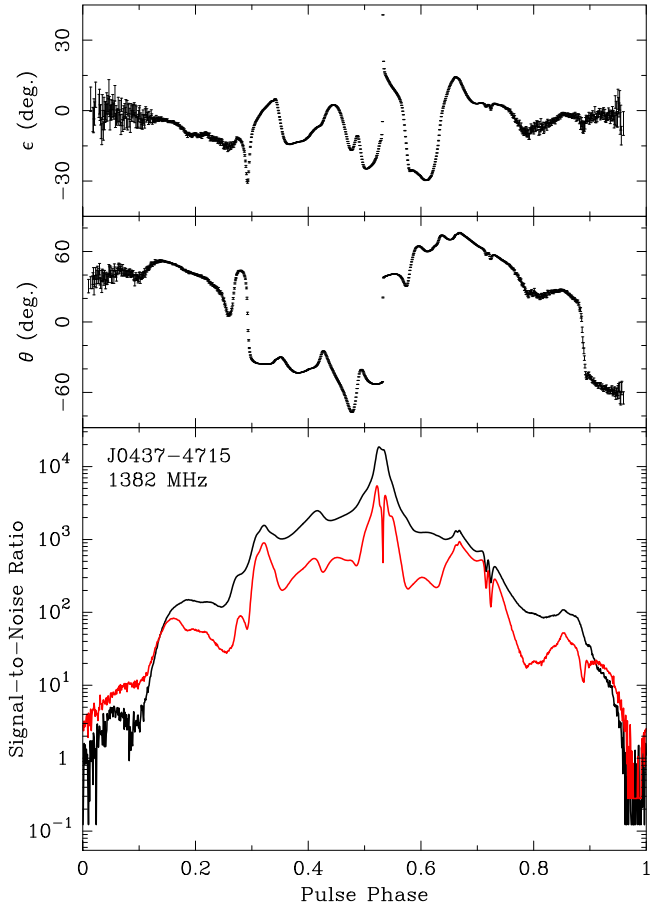


Figure 1. Average polarization of PSR J0437−4715, plotted as a function of pulse phase using polar coordinates: orientation, θ, ellipticity, ϵ, and polarized intensity plotted in red below the total intensity. Flux densities are normalized by the standard deviation of the off-pulse total intensity phase bins. Calibrated using MEM and used as the template for METM, this profile is integrated from 160 hr of observations that span 9.6 yr and 400 MHz of bandwidth centered at 1382 MHz.

S/N of the integrated average profile, the weighted mean reduced χ^2, the median uncertainties of the estimated values of δ_θ (the difference in receptor orientations) and σ_χ (the receptor ellipticities), and the fraction of the band that was lost to RFI. Based on these attributes, the MEM solution derived from the observing session recorded on 2014 April 15 is ranked as the best, and the calibrated total profile observed on this day is selected as the reference profile.

The 51 other MEM-calibrated total profiles are matched to this reference profile using matrix template matching (MTM; S06), then the reference profile and matched totals are integrated to form a template profile with an integration length of 160 hr, shown in Figure 1. Between pulse phase ∼0.9 and 0.13, the polarized flux appears to be greater than the total intensity, which is not physically possible. This artifact of imperfect baseline removal may indicate that radiation from this pulsar is received at all times; therefore, there is no off-pulse region of the pulse phase, and each of the four Stokes parameters is offset by an arbitrary amount. This baseline artifact does not affect arrival time estimates, which are calculated using only the nonzero spin harmonics of the Fourier transform of each pulse profile. Although PSR J0437−4715 is the brightest pulsar with the highest precision in the PPTA, it is also exceptionally susceptible to calibration errors, primarily

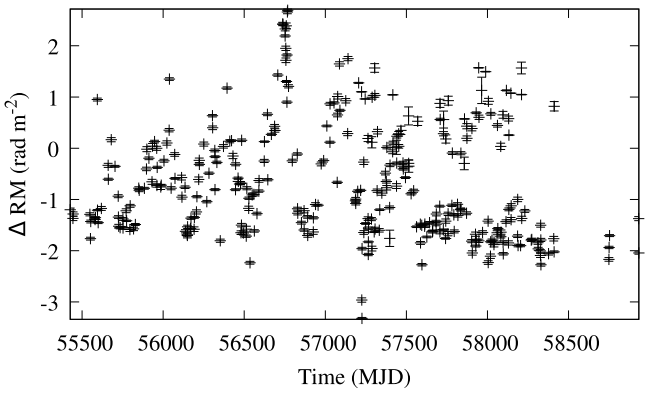


Figure 2. Temporal variation of average daily ionospheric contribution to Faraday rotation, measured with respect to the ionosphere on 2014 April 15 (MJD 56762). Most of the error bars, which denote the 1σ uncertainty of the RM estimate, are too small to be distinguished from the data point.

owing to the transition between orthogonally polarized modes in the middle of its main pulse.

This template profile is used as the polarized reference source in the following stage of METM analysis, during which Equation (19) of Britton (2000) is used to parameterize the receiver, and all model parameters are varied. A total of 296 robust METM solutions are selected by placing a lower limit on the fraction of the band that was lost to RFI and by applying Tukey's fence thresholds (see Section 3.2) to various attributes, including the median (over all frequency channels) reduced χ^2 and the median uncertainties of the estimated values of δ_θ, σ_χ, $\hat{C}_1$, and $\hat{C}_2$.

In this analysis, a variation of ionospheric Faraday rotation is not included in the METM model; therefore, for each METM solution, the variation of the best-fit estimate of σ_θ with radio frequency is used to derive an estimate of the average change in ionospheric Faraday rotation on that day, relative to 2014 April 15. After excluding outliers, the ionospheric Faraday rotation measure differences (ΔRM shown in Figure 2) vary between -3.3 and $+2.7$ rad m^{-2} over the 9.6 yr spanned by robust METM solutions. The predominantly day-to-night variation of ionospheric total electron content is observed as annual variations in the derived ionospheric ΔRM estimates owing to the annual drift between solar time and the sidereal times at which PSR J0437−4715 is observed. The observed peak in ionospheric ΔRM around MJD 56750 (2014 March/April) is near the peak in solar magnetic activity cycle 24.

These estimates of ionospheric RM are used to correct the estimates of σ_θ in each METM solution.

To reduce the noise in the estimated model parameters and interpolate across gaps in both time and radio frequency, seven of the METM model parameters (σ_θ, σ_χ, δ_θ, δ_χ, $\hat{C}_1$, $\hat{C}_2$, and $\hat{C}_3$) are smoothed using two-dimensional penalized splines (Eilers & Marx 1996; Grimstad et al. 2015). The optimal smoothing factor is determined using an iterative search algorithm. For each trial smoothing factor, the average goodness-of-fit is evaluated over four iterations of Monte Carlo cross-validation, also known as repeated random subsampling validation. On each of the four iterations, the smoothing spline is fit to a randomly selected half of the parameter estimates, and the other half of the estimates are used to validate the goodness-of-fit of the spline. The smoothing splines fit to the best estimates of $\hat{\boldsymbol{C}}$ are shown in Figure 3. We speculate that the apparent overpolarization of the noise diode, which starts around MJD

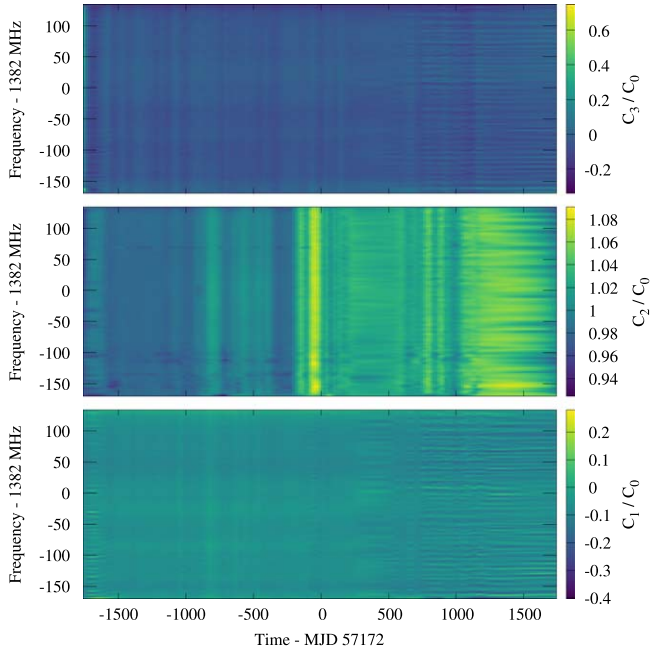


Figure 3. Two-dimensional smoothing splines fit to the best estimates of the normalized polarization vector of the artificial noise source; each component is plotted as a function of time and frequency.

57000, is of an unknown instrumental origin. In future work, it might prove useful to model instrumental impurity using a depolarizing Mueller matrix (Lu & Chipman 1996) and include this model in the calibration solution.

Given an observation of the artificial noise source, the components of $\hat{\boldsymbol{C}}$ predicted by the smoothing splines are used to derive estimates of the absolute gain G, differential gain γ, and differential phase ϕ of the instrument as described in Ord et al. (2004).[18] These are combined with the values of σ_θ, σ_χ, δ_θ, and δ_χ predicted by the associated smoothing splines to fully describe the polarimetric response of the instrument at any epoch and radio frequency spanned by the splines.

The critical steps detailed above are outlined here:

1. Prepare data.
   (a) Collect observations of the polarized reference source and its calibrators to create MEM sessions, ensuring that each session has good parallactic angle coverage and sufficient S/N for an accurate analysis.
2. Produce calibrator models.
   (a) Produce MEM calibrator model solutions.
   (b) Create template archives for METM analysis.
   (c) Produce METM calibrator model solutions.
   (d) Correct the ionospheric Faraday rotation.
   (e) Produce spline-smoothed METM solutions.
3. Calibrate pulsar data.
   (a) Apply calibration solutions to pulsar observations.

After completing these steps, arrival times are computed as described in the following section. A more detailed outline of the PSRCHIVE commands used for both polarimetric calibration and arrival time estimation is provided in Appendix C.

[18] Only $\hat{\boldsymbol{C}}$ is required to determine G, γ, and ϕ because the METM model is configured to include the reference source in the signal chain after the front-end component described by σ_θ, σ_χ, δ_θ, and δ_χ.

Table 2
Relative Arrival Time Uncertainties for Each Pulsar

Pulsar (JNAME)	τ_β (ns)	σ_τ (ns)	τ_β/σ_τ	$\hat{\sigma}_\varphi$	$\hat{\sigma}_{\tilde{\varphi}}$
PSR J0437−4715	205	116	1.77	0.82	1.43
PSR J0613−0200	59	1018	0.06	0.95	1.49
PSR J0711−6830	73	979	0.07	0.89	1.53
PSR J1017−7156	74	635	0.12	0.92	1.58
PSR J1022+1001	278	1555	0.18	0.73	1.67
PSR J1024−0719	33	1023	0.03	0.73	2.20
PSR J1045−4509	338	2570	0.13	0.87	1.50
PSR J1125−6014	8	1510	0.005	0.94	1.30
PSR J1446−4701	85	1359	0.08	0.96	1.33
PSR J1545−4550	66	808	0.08	0.84	1.74
PSR J1600−3053	119	572	0.21	0.88	1.39
PSR J1603−7202	143	1316	0.11	0.84	1.56
PSR J1643−1224	269	2248	0.12	0.92	1.39
PSR J1713+0747	5	287	0.02	0.86	1.57
PSR J1730−2304	196	1322	0.15	0.74	1.69
PSR J1744−1134	108	385	0.28	1.59	6.98
PSR J1824−2452A	20	2628	0.008	0.84	4.01
PSR J1832−0836	17	563	0.03	0.95	1.44
PSR J1857+0943	121	1208	0.10	0.92	1.43
PSR J1909−3744	22	152	0.15	0.83	1.52
PSR J1939+2134	44	586	0.08	0.92	1.49
PSR J2124−3358	115	2551	0.05	0.85	1.45
PSR J2129−5721	225	964	0.23	1.08	1.61
PSR J2145−0750	147	995	0.15	0.95	1.45
PSR J2241−5236	21	334	0.06	0.98	1.36

Note. Columns, from left to right, include J2000.0 coordinates (JNAME), predicted timing error for a 1% calibration error (τ_β), PPTA DR2 timing precision (σ_τ, Kerr et al. 2020), fraction of PPTA DR2 timing residuals that could be due to predicted calibration error (τ_β/σ_τ, van Straten 2013), theoretical relative ToA uncertainty between MTM and STM template-matching algorithms ($\hat{\sigma}_\varphi$), and the ratio between uncertainties in arrival times derived from the invariant interval and total intensity ($\hat{\sigma}_{\tilde{\varphi}}$). For PSR J0437−4715, PPTA DR2 timing precision is achieved by timing the invariant profile (Britton 2000).

3.2. Arrival Time Estimation and Analysis

Arrival times are estimated using both conventional scalar template matching (STM; Taylor 1992; T92) and MTM (S06). MTM quadruples the number of observational constraints while introducing only 6 degrees of freedom. For 23 out of 25 pulsars in our analysis, arrival time estimates derived from the polarization profile (using MTM) are expected to have greater precision than those derived from the total intensity profile alone (using STM), as indicated by the theoretical relative ToA uncertainty between MTM and STM ($\hat{\sigma}_\varphi$ in Table 2). Table 2 also shows the predicted timing error for a 1% calibration error (τ_β; van Straten 2013) and the timing precision achieved by PPTA DR2 (σ_τ; Kerr et al. 2020). For a given MSP, the potential significance of systematic timing errors due to inaccurate instrumental calibration is characterized by τ_β/σ_τ (column (4)).

The predicted values of relative uncertainty $\hat{\sigma}_\varphi$ are based on analysis of spectral content in the phase-resolved average profiles of all four Stokes parameters and additional (white) radiometer noise. However, over sufficiently long timescales, pulse arrival time residuals exhibit (red) timing noise, and it is necessary to model and remove this red noise before the white-noise content of residuals can be quantified and compared. In principle, the systematic timing errors induced by polarization

Table 3
Prior Ranges on White-noise and Red-noise Model Parameters, Comprising the Dimensionless Base-10 Logarithms of the Error Scale Factor $\hat{E}_{f,10} = \log_{10}(E_f)$, the Error Added in Quadrature $\hat{E}_{q,10} = \log_{10}(E_q/\mathrm{s})$, and the Red-noise Amplitude $\hat{A}_{\mathrm{red},10} = \log_{10}(A_{\mathrm{red}}/\mathrm{yr}^{3/2})$; and the Power-law Spectral Index β

Parameter	Prior Range	Type
$\hat{E}_{f,10} = \log_{10}(E_f)$	$(-1, 1)$	log-uniform
$\hat{E}_{q,10} = \log_{10}(E_q/\mathrm{s})$	$(-9, -5)$	log-uniform
$\hat{A}_{\mathrm{red},10} = \log_{10}(A_{\mathrm{red}}/\mathrm{yr}^{3/2})$	$(-18, -10)$	log-uniform
β	$(0, 7)$	log-uniform

distortions could also induce red noise, such as the $1/f$ noise produced by unmodeled steps in an instrumental response. Therefore, it is also interesting to quantify and compare the red-noise content of pulsar timing residuals.

In this study, we analyze both the white- and red-noise components of timing residuals using TEMPO2 (Edwards et al. 2006; Hobbs et al. 2006) and TEMPONEST (Lentati et al. 2014).[19] TEMPO2 is employed to fit the timing model to the observed ToAs by minimizing the timing residuals. TEMPONEST utilizes the multimodal nested sampling algorithm MULTINEST (Feroz et al. 2009)[20] to explore the parameter space of the nonlinear pulsar timing model. Simultaneously, it determines a red-noise model and two time-independent white-noise modifiers:

1. *Error scale factor.* This accounts for any errors that are proportional to the estimated uncertainty (including potential miscalibrated radiometer noise in the system). It modifies each ToA uncertainty by a constant scale factor E_f.
2. *Error added in quadrature.* This compensates for additional white noise by adding a constant E_q in quadrature to each ToA uncertainty (Lentati et al. 2014).

The total ToA uncertainty is therefore obtained by adjusting the uncertainty σ_τ as follows:

$$\sigma'_\tau = \sqrt{E_q^2 + E_f^2 \sigma_\tau^2}. \tag{2}$$

Error scale factor (EFAC) and error added in quadrature (EQUAD) values are typically applied to all ToAs in a pulsar timing data set and adjusted iteratively until the fitted model's reduced χ^2 reaches unity (Shannon et al. 2014).

TEMPONEST models red noise using a power-law spectrum characterized by an amplitude (A_{red}) and spectral index β:

$$P_r(f) = \frac{A_{\mathrm{red}}^2}{12\pi^2}\left(\frac{f}{f_{\mathrm{yr}}}\right)^{-\beta}. \tag{3}$$

Here, f_{yr} is a reference frequency of 1 cycle yr^{-1}, and the amplitude A_{red} is in units of $\mathrm{yr}^{3/2}$ (Lentati et al. 2014; Parthasarathy et al. 2019). The prior ranges for the red noise and additional white-noise parameters used in our analysis are detailed in Table 3. Note that TEMPONEST is configured to sample the spectrum using 120 Fourier coefficients.

Outliers in pulsar timing measurements bias both timing model and noise model parameter estimates and reduce the accuracy of estimated parameter uncertainties (Vallisneri & van Haasteren 2017). To address this challenge, we employ robust and automated outlier removal techniques. An initial outlier rejection involves removing any instances of ToAs with errors equal to 0 and performing a 5σ outlier removal on both the relative error for each residual (residual divided by uncertainty) and ToA goodness-of-fit. We then run TEMPO2 with the initial (slightly corrupted) TEMPONEST red-noise model, producing whitened residuals that are corrupted by some remaining outliers. These remaining outliers are removed using Tukey's fence (Tukey 1977), a robust statistical method for outlier detection and removal (Morello et al. 2019), which is used to enhance timing accuracy. This approach typically flags only a small portion (around 5%) of the data as outliers and has been shown to improve timing accuracy by a factor of 2 for many pulsars (Lower et al. 2020). Tukey's fence defines a "reasonable" range based on the interquartile range (IQR), which is the difference between the 75th percentile (Q3) and the 25th percentile (Q1) of the data set. The range extends from Q1 $- q^*$IQR to Q3 $+ q^*$IQR, where q is a parameter. While a common choice for q is 1.5, it can be adjusted according to specific analysis requirements, determining the stringency of the outlier cutoff. We used a less stringent cutoff value of $q = 2$ for our analysis.

Following the application of Tukey's fence to each MSP, we update the best-fit model by rerunning TEMPONEST to generate a red-noise model that is not corrupted by outliers. PSR J0437 −4715 and PSR J2241−5721 are exceptions. An additional 5σ removal was performed on the relative error for each residual to catch a few outliers missed by Tukey's fence for these pulsars. A final run of TEMPO2 with the final TempoNest red-noise model produces whitened residuals for each calibration and ToA estimation method combination. By comparing the red- and white-noise model parameters, we aim to evaluate whether advanced techniques like METM and MTM, which better account for instrumental calibration errors, can enhance the experimental sensitivity of PTAs over extended periods compared to conventional methods like IFA and STM.

4. Results

In this section, we present the results of our analysis, comparing advanced polarimetric calibration (METM; van Straten 2013) and arrival time estimation (MTM; S06) methods with conventional approaches (IFA and STM). We applied these techniques to 25 high-priority pulsars regularly observed as part of the PPTA project, resulting in four distinct data sets (IFA-MTM, METM-MTM, IFA-STM, and METM-STM).

The quality of the arrival time estimation procedure is summarized in Figure 4, which presents a heat map of the median arrival time goodness-of-fit for each pulsar. Goodness-of-fit is characterized by the reduced χ^2 of the fit between the observed pulse profile and the template pulse profile. Large values of the reduced χ^2 may indicate that the template pulse profile is not an accurate representation of the average pulse profile, or that the observed pulse profiles vary by more than what would be expected due to radiometer noise alone; e.g., owing to additional pulsar self-noise known as jitter (Cordes & Downs 1985; Osłowski et al. 2011).

The results of modeling the arrival time estimates are presented in Table 4, which lists, for each pulsar and each combination of methods, the number of ToAs, the weighted

[19] https://github.com/LindleyLentati/TempoNest
[20] https://github.com/farhanferoz/MultiNest

Figure 4. The reduced χ^2 heat map illustrates the median goodness-of-fit over all arrival times obtained for each pulsar. PSR J0437−4715 is excluded owing to exceptionally large values of 2.34, 7.77 2.04, and 2.61 for the IFA-MTM, IFA-STM, METM-MTM, and METM-STM methods, respectively.

standard deviations of the postfit residuals output by TEMPO2, and the noise model parameters obtained through TEMPONEST.

The parameters that characterize the white noise of each pulsar are summarized in Table 5, which compares the minimum, median, and maximum values of E_f, E_q, and the uncertainty-weighted standard deviation of the whitened (red noise removed) postfit timing residuals, $\sigma_{\tau,\mathrm{wh}}$, for the four data sets. Notably, IFA-STM has the largest median values of E_f and E_q; therefore, it serves as the baseline for model comparison in this study.

Table 6 compares the best-fit E_q estimates of the IFA-STM and METM-MTM data sets for those seven pulsars with statistically significant differences.

The IFA-STM and METM-MTM data sets are further compared in Table 7, which lists $\sigma_{\tau,\mathrm{wh}}$ for each pulsar. For each pulsar, column (4) lists the white-noise quotient

$$Q = \frac{\sigma_{\tau,\mathrm{wh}}\{\mathrm{METM} - \mathrm{MTM}\}}{\sigma_{\tau,\mathrm{wh}}\{\mathrm{IFA} - \mathrm{STM}\}} \quad (4)$$

that is used to define the percentage improvement, $(1 - Q) \times 100\%$. METM-MTM yields significant reductions in $\sigma_{\tau,\mathrm{wh}}$; a similar result was found for three pulsars observed with the Nançay Radio Telescope (Guillemot et al. 2023). For the MSPs in our sample, the median reduction in white noise is 33%, and the maximum reduction of 85% (an impressive factor of 6.6) is observed for PSR J0437−4715. For this pulsar, we plot the noise model parameter distributions derived from each data set in Figure 5 and compare the amplitude spectra of the residuals in Figure 6.

Finally, Figure 7 depicts the improvement in timing precision for each pulsar by comparing the $\sigma_{\tau,\mathrm{wh}}$ values from PPTA DR2 with those derived in this study using METM-MTM and IFA-STM. As discussed in more detail in Section 5.7, this visual comparison should be treated as indicative because there are significant differences between the 21 cm data included in the PPTA DR2 analysis and the subset of data analyzed in this work.

5. Discussion

Comparing METM and MTM with conventional methods, we observe a significant reduction in the white noise in pulse arrival times across all pulsars in our sample. This is evident from the significantly smaller uncertainty-weighted standard deviations of whitened postfit timing residuals for METM-MTM compared to IFA-STM (up to 6.6 times smaller; see Tables 7 and 4). For the majority of PPTA pulsars, the reduction in white noise achieved with METM-MTM exceeds the predicted relative ToA uncertainty between MTM and STM algorithms ($\hat{\sigma}_\varphi$, see Table 2). This includes pulsars like PSR J1744−1134 and PSR J2129−5721, for which MTM is predicted to perform worse than STM, owing to the multiple correlation between the phase shift and unknown Jones matrix model parameters. These results indicate that MTM is able to mitigate the impact of polarization calibration errors, which contribute additional white noise to STM-derived arrival times.

Further supporting this interpretation, the smallest improvements are achieved for pulsars with low susceptibility to calibration error as characterized by τ_β (see Table 2), such as PSR J1125−6014, PSR J1713+0747, PSR J1909+3744, PSR J1939−2134, PSR J1824−2452A, and PSR J2241−5236 (see Table 7). Similarly, for pulsars that are highly susceptible to calibration error, such as PSR J0437−4715, PSR J1022+1001, PSR J1045−4509, and PSR J1643−1224, METM-MTM significantly reduces $\sigma_{\tau,\mathrm{wh}}$. The largest reduction in $\sigma_{\tau,\mathrm{wh}}$, with a quadrature difference of 4.5 μs between IFA-STM

Table 4
Noise Statistics and Noise Model Parameters for Each Timing Data Set

Pulsar	Method	$N_{\rm toas}$	$\overline{\sigma_\tau}$ μs	$\overline{\sigma_{\tau,\rm wh}}$ μs	$\hat{E}_{f,10}$	$\hat{E}_{q,10}$	$\hat{A}_{\rm red,10}$	β
PSR J0437−4715	METM-MTM	700	0.584	0.100	0.60(5)	−7.02(2)	−13.48(7)	3.0(4)
	IFA-MTM	700	0.653	0.103	0.64(5)	−7.02(2)	−13.47(7)	3.1(4)
	METM-STM	700	0.680	0.231	0.6(1)	−6.64(1)	−13.5(1)	3.1(7)
	IFA-STM	700	0.947	0.663	0.3(4)	−6.17(1)	−13.1(1)	2.0(5)
PSR J0613−0200	METM-MTM	221	0.661	0.583	0.04(3)	−7.0(4)	−15(1)	4(2)
	IFA-MTM	221	0.648	0.583	0.04(3)	−7.0(4)	−16(1)	4(2)
	METM-STM	221	1.111	1.083	0.22(6)	−6.3(3)	−16(1)	3(2)
	IFA-STM	221	1.200	1.117	0.22(7)	−6.3(3)	−16(1)	3(2)
PSR J0711−6830	METM-MTM	333	0.816	0.754	−0.02(2)	−6.6(3)	−14(1)	2(2)
	IFA-MTM	333	0.805	0.722	−0.01(2)	−6.8(3)	−13.8(8)	2(2)
	METM-STM	333	0.925	0.842	0.02(2)	−6.7(4)	−13.6(9)	2(2)
	IFA-STM	333	1.069	0.972	0.04(2)	−6.6(3)	−14.0(7)	4(2)
PSR J1017−7156	METM-MTM	333	1.704	0.223	−0.09(2)	−7.4(4)	−12.88(6)	2.7(4)
	IFA-MTM	333	1.726	0.221	−0.09(2)	−7.4(4)	−12.89(5)	2.7(4)
	METM-STM	333	1.751	0.294	0.14(2)	−7.3(4)	−12.90(6)	2.9(5)
	IFA-STM	333	1.800	0.312	0.15(2)	−7.2(4)	−12.90(6)	2.9(5)
PSR J1022+1001	METM-MTM	215	0.857	0.833	0.00(5)	−6.19(4)	−16(1)	3(2)
	IFA-MTM	215	0.875	0.864	0.01(5)	−6.17(4)	−16(1)	3(2)
	METM-STM	215	1.341	1.281	0.07(4)	−6.00(4)	−15(1)	3(2)
	IFA-STM	215	1.619	1.619	0.04(5)	−5.88(4)	−16(1)	3(2)
PSR J1024−0719	METM-MTM	127	0.739	0.654	0.05(3)	−6.9(4)	−15(1)	4(2)
	IFA-MTM	127	0.787	0.721	0.07(3)	−6.8(4)	−15(1)	4(2)
	METM-STM	127	1.088	1.078	0.11(3)	−6.9(5)	−16(1)	3(2)
	IFA-STM	127	1.095	1.081	0.12(3)	−6.9(5)	−16(1)	3(2)
PSR J1045−4509	METM-MTM	177	4.025	1.280	−0.07(5)	−6.5(5)	−12.46(9)	2.6(5)
	IFA-MTM	177	4.689	1.311	−0.06(4)	−6.5(4)	−12.44(9)	2.6(5)
	METM-STM	177	7.010	4.364	−0.3(3)	−5.34(4)	−12.5(2)	2.7(9)
	IFA-STM	177	5.739	4.702	0.3(1)	−5.4(1)	−12.5(4)	2(1)
PSR J1125−6014	METM-MTM	141	1.939	0.833	0.0(1)	−6.06(5)	−12.9(1)	2.8(7)
	IFA-MTM	141	1.919	0.836	0.0(1)	−6.07(5)	−12.8(1)	2.6(6)
	METM-STM	141	2.030	0.844	0.0(1)	−6.08(5)	−12.8(1)	2.8(7)
	IFA-STM	141	1.993	0.833	0.0(1)	−6.08(5)	−12.8(1)	2.7(6)
PSR J1446−4701	METM-MTM	145	1.747	1.318	−0.01(7)	−6.00(9)	−15(1)	3(2)
	IFA-MTM	145	1.605	1.345	0.04(6)	−6.0(1)	−15(1)	3(2)
	METM-STM	145	2.132	1.978	0.34(7)	−5.8(1)	−16(1)	3(2)
	IFA-STM	145	2.247	2.227	0.33(8)	−5.8(1)	−16(1)	3(2)
PSR J1545−4550	METM-MTM	123	0.594	0.429	−0.1(1)	−6.9(4)	−13.3(2)	3(1)
	IFA-MTM	123	0.603	0.434	0.06(7)	−6.9(4)	−13.3(2)	3(1)
	METM-STM	123	1.068	0.776	0.13(9)	−6.6(4)	−13.2(3)	3(1)
	IFA-STM	123	1.052	0.785	0.12(9)	−6.5(4)	−13.1(3)	3(1)
PSR J1600−3053	METM-MTM	174	1.828	0.230	−0.04(6)	−7.0(4)	−13.3(2)	2.9(6)
	IFA-MTM	174	1.748	0.221	0.05(5)	−7.2(4)	−13.2(1)	2.7(5)
	METM-STM	174	1.757	0.311	0.08(6)	−7.1(4)	−14.0(5)	5(1)
	IFA-STM	174	2.100	0.341	0.10(7)	−6.9(3)	−14.1(5)	5(1)
PSR J1603−3053	METM-MTM	212	0.783	0.497	0.01(3)	−7.2(4)	−13.3(2)	2.6(8)
	IFA-MTM	212	0.819	0.504	0.01(3)	−7.2(4)	−13.3(2)	2.8(9)
	METM-STM	212	0.877	0.593	0.00(3)	−7.1(4)	−13.2(2)	2.6(8)
	IFA-STM	212	0.961	0.747	0.06(3)	−6.8(4)	−13.4(3)	3(1)
PSR J1643−1224	METM-MTM	155	2.481	0.452	−0.12(6)	−6.8(4)	−12.65(7)	2.0(3)
	IFA-MTM	155	2.153	0.438	−0.14(7)	−6.7(4)	−12.64(6)	1.9(3)
	METM-STM	155	2.282	1.087	0.26(4)	−6.6(4)	−12.8(2)	2.6(7)
	IFA-STM	155	2.133	1.039	0.24(4)	−6.6(4)	−12.8(2)	2.4(7)
PSR J1713+0747	METM-MTM	246	0.234	0.201	−0.08(5)	−6.76(3)	−13.6(1)	1.5(6)
	IFA-MTM	246	0.232	0.204	−0.07(6)	−6.75(3)	−13.6(1)	1.5(6)
	METM-STM	246	0.252	0.216	−0.04(6)	−6.72(3)	−13.6(1)	1.6(6)
	IFA-STM	246	0.247	0.215	−0.01(6)	−6.72(3)	−13.6(1)	1.5(6)

Table 4
(Continued)

Pulsar	Method	$N_{\rm toas}$	$\frac{\sigma_\tau}{\mu s}$	$\frac{\sigma_{\tau,\rm wh}}{\mu s}$	$\hat{E}_{f,10}$	$\hat{E}_{q,10}$	$\hat{A}_{\rm red,10}$	β
PSR J1730−2304	METM-MTM	165	0.935	0.668	0.11(3)	−7.1(4)	−13.7(4)	3(1)
	IFA-MTM	165	0.865	0.664	0.11(3)	−7.2(4)	−13.7(4)	3(1)
	METM-STM	165	1.460	0.963	0.11(3)	−6.9(5)	−13.5(4)	3(1)
	IFA-STM	165	1.488	1.171	0.14(4)	−6.4(4)	−13.7(8)	4(1)
PSR J1744−1134	METM-MTM	237	0.426	0.306	0.02(3)	−6.8(2)	−13.5(4)	2(1)
	IFA-MTM	237	0.438	0.311	0.04(3)	−6.9(2)	−13.5(4)	2(1)
	METM-STM	237	0.442	0.301	−0.09(4)	−6.63(5)	−13.6(4)	2(1)
	IFA-STM	237	1.034	0.682	−0.13(6)	−6.21(3)	−14.5(6)	5(2)
PSR J1824−2452A	METM-MTM	80	16.467	0.537	0.1(4)	−6.2(2)	−12.4(1)	3.5(6)
	IFA-MTM	80	15.780	0.507	0.0(4)	−6.2(2)	−12.3(1)	3.3(7)
	METM-STM	80	16.445	0.551	0.2(3)	−6.3(2)	−12.4(1)	3.6(6)
	IFA-STM	80	16.148	0.510	−0.1(4)	−6.2(2)	−12.4(1)	3.5(7)
PSR J1832−0836	METM-MTM	72	2.485	0.597	−0.02(6)	−6.8(4)	−14.1(4)	6(1)
	IFA-MTM	72	2.066	0.645	0.03(5)	−6.8(4)	−14.0(4)	5(1)
	METM-STM	72	1.605	0.929	0.1(1)	−6.4(4)	−13.9(5)	5(1)
	IFA-STM	72	1.593	0.809	0.0(2)	−6.3(4)	−13.0(5)	3(2)
PSR J1857+0943	METM-MTM	127	1.288	0.573	−0.06(5)	−6.8(4)	−13.9(4)	4(1)
	IFA-MTM	127	1.255	0.597	−0.04(5)	−6.8(4)	−13.8(4)	4(1)
	METM-STM	127	1.335	0.849	0.02(6)	−6.6(4)	−13.8(4)	4(1)
	IFA-STM	127	1.532	0.889	0.04(6)	−6.5(4)	−13.9(4)	4(1)
PSR J1909−3744	METM-MTM	344	0.518	0.200	−0.10(6)	−6.73(3)	−14.0(2)	3.7(8)
	IFA-MTM	344	0.512	0.204	−0.11(7)	−6.72(3)	−14.0(2)	3.7(8)
	METM-STM	344	0.530	0.204	−0.08(6)	−6.72(3)	−14.0(3)	3.6(9)
	IFA-STM	344	0.540	0.204	−0.13(7)	−6.72(3)	−14.0(3)	3.8(9)
PSR J1939+2134	METM-MTM	142	1.121	0.101	0.4(2)	−7.0(2)	−12.94(6)	2.5(3)
	IFA-MTM	142	1.126	0.101	0.4(2)	−7.0(2)	−12.95(6)	2.5(3)
	METM-STM	142	1.117	0.104	0.1(3)	−6.93(7)	−12.98(6)	2.7(3)
	IFA-STM	142	1.173	0.104	0.6(1)	−7.3(3)	−12.96(6)	2.5(3)
PSR J2124−3358	METM-MTM	221	1.341	1.320	−0.01(2)	−6.9(5)	−16(1)	3(2)
	IFA-MTM	221	1.275	1.272	−0.02(2)	−6.9(5)	−16(1)	3(2)
	METM-STM	221	2.936	2.983	0.37(2)	−6.4(5)	−16(1)	3(2)
	IFA-STM	221	2.920	2.886	0.38(3)	−6.4(5)	−16(1)	3(2)
PSR J2129−5721	METM-MTM	238	1.131	0.914	0.01(3)	−6.7(4)	−14.7(6)	5(1)
	IFA-MTM	238	1.162	0.911	0.00(3)	−6.7(4)	−14.4(6)	5(1)
	METM-STM	238	1.301	1.151	0.17(2)	−6.9(4)	−15(1)	4(2)
	IFA-STM	238	1.336	1.194	0.17(2)	−6.9(4)	−14.4(7)	5(2)
PSR J2145−0750	METM-MTM	211	0.592	0.531	0.00(3)	−6.51(7)	−16(1)	4(2)
	IFA-MTM	211	0.581	0.531	0.02(3)	−6.51(7)	−15(1)	4(2)
	METM-STM	211	0.832	0.602	0.00(3)	−6.41(6)	−14.6(8)	4(2)
	IFA-STM	211	0.810	0.684	0.03(4)	−6.32(7)	−15(1)	4(2)
PSR J2241−5236	METM-MTM	329	0.251	0.214	0.14(4)	−6.80(6)	−13.9(4)	3(1)
	IFA-MTM	329	0.255	0.217	0.13(5)	−6.80(6)	−14.0(5)	3(1)
	METM-STM	329	0.254	0.215	0.15(3)	−6.87(6)	−14.0(3)	3(1)
	IFA-STM	329	0.269	0.229	0.14(4)	−6.80(5)	−13.9(3)	3(1)

Note. Columns, from left to right, include each pulsar's J2000.0 coordinates (JNAME); the calibration-ToA estimation method; the number of ToAs, $N_{\rm toas}$; the uncertainty-weighted standard deviation of the postfit timing residuals σ_τ, and whitened timing residuals $\sigma_{\tau,\rm wh}$; and the maximum likelihood estimates of the noise model parameters, comprising the dimensionless base-10 logarithms of the EFAC $\hat{E}_{f,10} = \log_{10}(E_f)$, the EQUAD $\hat{E}_{q,10} = \log_{10}(E_q/\rm s)$, and the red-noise amplitude $\hat{A}_{\rm red,10} = \log_{10}(A_{\rm red}/\rm yr^{3/2})$; and the power-law spectral index β. Values in parentheses are the 1σ uncertainty in the last digit quoted.

and METM-MTM data sets, is observed for PSR J1045−4509. This pulsar also has the greatest susceptibility to calibration error τ_β.

The whitened residuals include a contribution from the pulsar-intrinsic jitter, which is not separately accounted for in our noise analysis. The white noise induced by jitter is statistically independent of the radiometer noise that dominates the error of each arrival time estimate; therefore, jitter increases E_q. The jitter noise varies with pulse phase, and when the subpulse structures that cause jitter are broader than the pulse

Table 5
Minimum, Median, and Maximum Values of the Error Scale Factor E_f, Error Added in Quadrature E_q, and Uncertainty-weighted Standard Deviation of the Whitened Postfit Timing Residuals $\sigma_{\tau,\mathrm{wh}}$, for Each Combination of Calibration and Arrival Time Estimation Methods

Model	E_f			E_q (μs)			$\sigma_{\tau,\mathrm{wh}}$ (μs)		
	Min	Med	Max	Min	Med	Max	Min	Med	Max
METM-MTM	0.77	1.02	4.02	0.04	0.16	0.99	0.10	0.54	1.32
IFA-MTM	0.72	1.07	4.34	0.04	0.15	0.95	0.10	0.51	1.35
METM-STM	0.53	1.27	4.21	0.05	0.23	4.61	0.10	0.84	4.36
IFA-STM	0.75	1.30	4.13	0.05	0.33	3.85	0.10	0.81	4.70

Table 6
Pulsars with Statistically Significant Different Estimates of Error Added in Quadrature, E_q

Pulsar (JNAME)	E_q (μs)		ΔE_q (μs)
	IFA-STM	METM-MTM	
PSR J0437−4715	0.67(2)	0.095(4)	0.67(2)
PSR J1022+1001	1.3(1)	0.64(7)	1.1(1)
PSR J1045−4509	4(1)	0.3(3)	4(1)
PSR J1446−4701	1.7(4)	1.0(2)	1.4(4)
PSR J1713+0747	0.19(1)	0.17(1)	0.08(2)
PSR J1744−1134	0.62(4)	0.15(6)	0.60(7)
PSR J2145−0750	0.48(8)	0.31(5)	0.37(9)

Note. The best-fit estimates of E_q derived from the IFA-STM and METM-MTM data sets are shown in columns (2) and (3), and the quadrature differences between them are listed in column (4). Values in parentheses are the 1σ uncertainty in the last digit quoted.

Table 7
Comparison of Uncertainty-weighted Standard Deviations of Whitened Postfit Timing Residuals, $\sigma_{\tau,\mathrm{wh}}$

Pulsar (JNAME)	$\sigma_{\tau,\mathrm{wh}}$ (ns)		Q
	IFA-STM	METM-MTM	
PSR J0437−4715	663	100	0.15
PSR J0613−0200	1117	583	0.52
PSR J0711−6830	972	754	0.78
PSR J1017−7156	312	223	0.71
PSR J1022+1001	1630	839	0.51
PSR J1024−0719	1081	654	0.60
PSR J1045−4509	4702	1280	0.27
PSR J1125−6014	833	833	1.00
PSR J1446−4701	2227	1318	0.59
PSR J1545−4550	785	429	0.55
PSR J1600−3053	341	230	0.67
PSR J1603−7202	747	497	0.67
PSR J1643−1224	1039	452	0.43
PSR J1713+0747	215	201	0.93
PSR J1730−2304	1171	668	0.57
PSR J1744−1134	682	306	0.45
PSR J1824−2452A	510	537	1.05
PSR J1832−0836	809	597	0.74
PSR J1857+0943	889	573	0.64
PSR J1909−3744	204	200	0.98
PSR J1939+2134	104	101	0.97
PSR J2124−3358	2886	1320	0.46
PSR J2129−5721	1194	914	0.77
PSR J2145−0750	684	513	0.75
PSR J2241−5236	229	214	0.93

Note. Estimates of $\sigma_{\tau,\mathrm{wh}}$ for the IFA-STM and METM-MTM methods are listed in columns (2) and (3), and the white-noise quotient, Q, defined by Equation (4) is listed in column (4).

phase bins used to resolve the average profile, the jitter noise is correlated between phase bins. That is, jitter causes the noise in each phase bin to no longer be independent and identically distributed, a fundamental assumption on which the STM and MTM algorithms are based. Therefore, jitter inflates the reduced χ^2 of each template-matching fit and causes an underestimation of the arrival time error, which in turn increases E_f. In summary, the pulsar-intrinsic jitter impacts the time-of-arrival goodness-of-fit, the EFAC E_f, and the EQUAD E_q. These useful metrics are discussed in the following three sections.

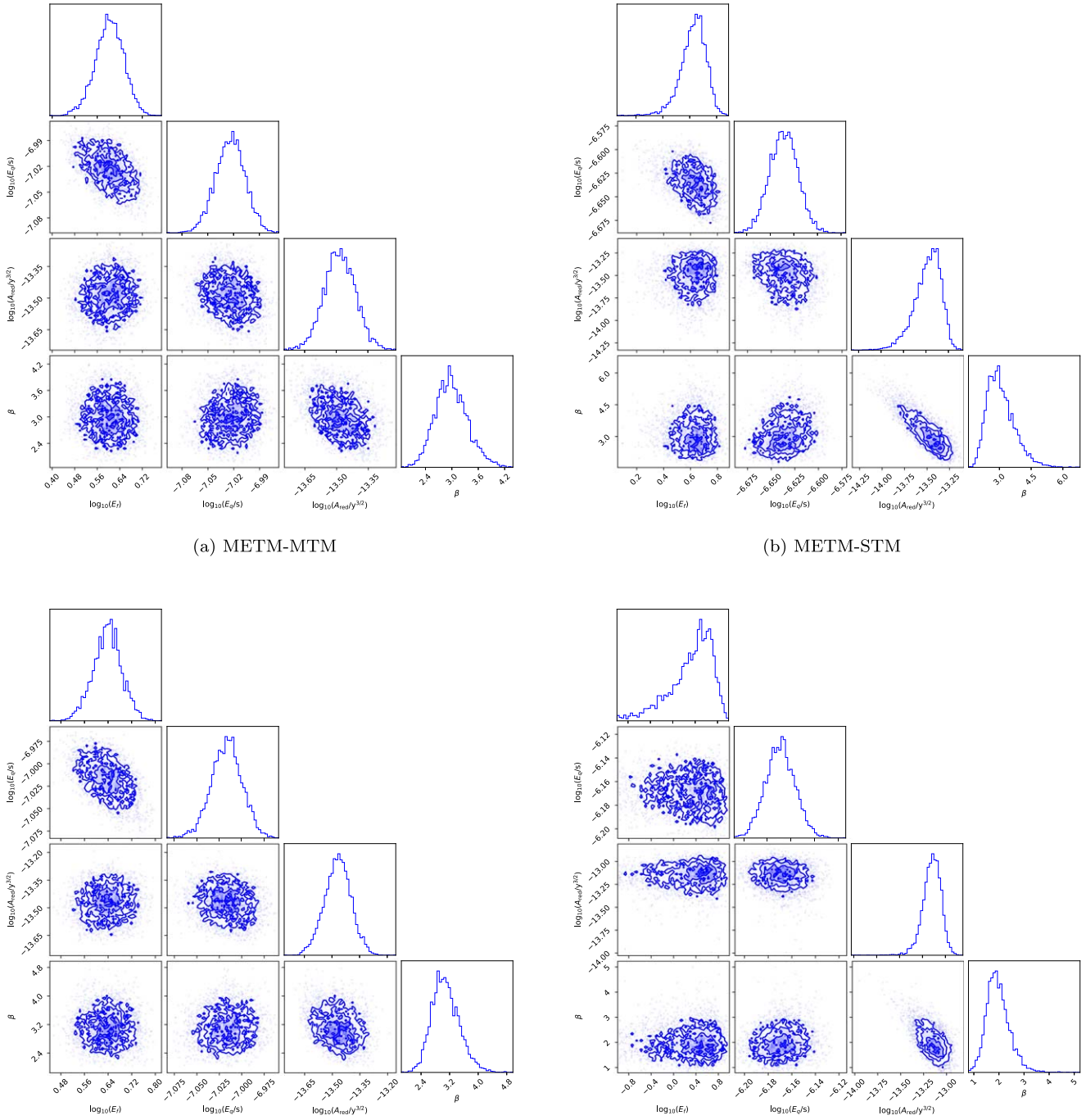
5.1. Time-of-arrival Goodness-of-fit

The template-matching reduced χ^2 heat map (see Figure 4) presents the median goodness-of-fit for arrival times derived from each pulsar in our study. To increase the dynamic range of the color scale in this figure, PSR J0437−4715 is excluded because it exhibits the greatest jitter relative to timing uncertainty, leading to exceptionally large values of reduced χ^2.

Calibration errors are expected to increase in the reduced χ^2 of each ToA estimate. Therefore, MTM is expected to yield lower reduced χ^2 values than yielded by STM. This is observed for several of the pulsars in our data set; most notably, PSR J1022+1001, renowned for its temporal profile variations (Kramer et al. 1999; Ramachandran & Kramer 2003; Hotan et al. 2004), has one of the two highest reduced χ^2 values for IFA-STM. This high value is significantly reduced when using MTM. Similarly, for PSR J1939+2134 and PSR J1713+0747, MTM yields lower reduced χ^2 values than STM. Notable exceptions include PSR J1824−2452A and PSR J2124−3358, for which MTM increases the median reduced χ^2. This warrants further explanation, which is currently only speculative. These pulsars may demonstrate greater intrinsic variability in Stokes Q, U, and V than in Stokes I, possibly due to switching between orthogonally polarized modes. The variability might also originate in the ISM or ionosphere; for example, uncorrected Faraday rotation variations (e.g., Yan et al. 2011) could lead to bandwidth depolarization, a transformation that cannot be represented by a Jones matrix and therefore cannot be modeled by MTM. A more quantitative investigation is beyond the scope of this paper, and presents an opportunity for future research.

5.2. Error Scale Factor

When ToA uncertainties accurately reflect the white-noise content of arrival time estimates, E_f is expected to be close to unity (i.e., $\log_{10}(E_f) \sim 0$). Values of $E_f > 1$ may suggest that the algorithm for arrival time estimation systematically underestimates ToA uncertainty. Alternatively, it could result from additional noise sources with amplitudes that are



(a) METM-MTM

(b) METM-STM

(c) IFA-MTM

(d) IFA-STM

Figure 5. Noise model parameter distributions for PSR J0437−4715. The two-dimensional posterior surfaces and one-dimensional marginal distributions for red- and white-noise parameters are derived from the MULTINEST chains generated by TEMPONEST during joint parameter space exploration (Lentati et al. 2014).

proportional to that of the radiometer noise in the integrated pulse profile.

The interpretation of E_f aligns with two main results. First, E_f appears to indicate that STM underestimates the uncertainty, as evidenced by the median E_f around 1.3 for STM compared to around 1.0 for MTM (see Table 5). However, for some pulsars, E_f is dominated by jitter; for instance, PSR J0437−4715 has the highest E_f value of approximately 4 for each model (for IFA-STM, E_f is $\sim 2 \pm 3$, which is poorly constrained and consistent with 4.) This pulsar displays the lowest METM-MTM whitened noise ($\sigma_{\tau,\mathrm{wh}} \sim 100$ ns) in our data set; it also has the highest level of self-noise relative to the radiometer noise (Osłowski et al. 2011; Parthasarathy et al. 2021).

Although jitter dominates E_f for some pulsars, the error factor remains a valuable diagnostic measure. For example, our initial analysis of E_f highlighted errors in both the phase gradient shift (PGS) algorithm (the default STM algorithm when estimating arrival times using PSRCHIVE tools like PAT) and the PSRCHIVE implementation of MTM. After correcting these errors, as described in Appendix A, there were no statistically significant differences between the E_f values

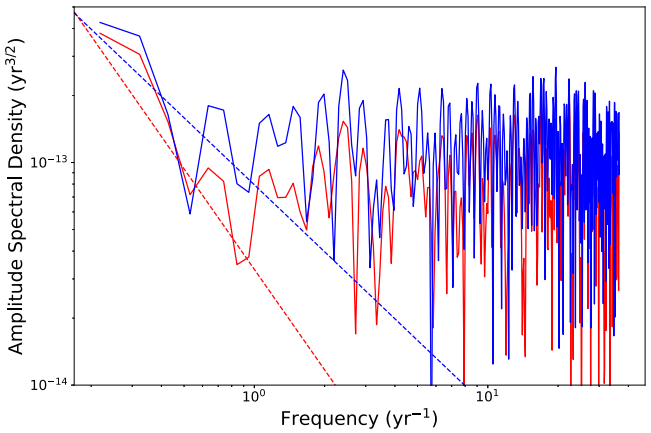


Figure 6. Amplitude spectra of the postfit residuals of PSR J0437−4715 for the METM-MTM (red) and IFA-STM (blue) data sets. The best-fit red-noise models output by TEMPONEST for each data set are indicated by dashed lines with matching colors.

derived from IFA-calibrated and METM-calibrated data for a given pulsar and arrival time estimation algorithm.

5.3. Error Added in Quadrature

EQUAD, E_q, characterizes any additional sources of unaccounted white noise with amplitudes that are uncorrelated with that of the radiometer noise estimated from the off-pulse baseline. When ToA uncertainties accurately reflect the noise content of arrival time estimates, $E_q \to 0$ and $\log_{10}(E_q) \to -\infty$; therefore, larger E_q values indicate greater unaccounted uncertainty in ToA measurements.

Out of all of the E_q estimates listed in Table 4, we compare those derived from the IFA-STM and METM-MTM data sets and list the seven pulsars with statistically significant differences in Table 6. Most notably, METM-MTM has reduced E_q for PSR J1045−4509 by an order of magnitude. This pulsar also has the greatest reduction in $\sigma_{\tau,\mathrm{wh}}$ and the largest susceptibility to calibration error as characterized by τ_β. It is also interesting to find PSR J1744−1134 in this subset. For this pulsar, the theoretical uncertainty for MTM-derived arrival times is predicted to be 1.6 times greater than that of STM-derived arrival times; it is one of only two pulsars for which $\hat{\sigma}_\varphi > 1$ (see Table 2). However, PSR J1744−1134 also has the second-largest susceptibility to calibration error relative to timing precision ($\tau_\beta/\sigma_\tau \sim 0.28$ in Table 2).

Across the entire set of pulsars, the median E_q for IFA-STM is 330 ns, which is approximately two times the median E_q of 160 ns for METM-MTM (see Table 5), and there is moderate correlation between ΔE_q and τ_β (Pearson correlation coefficient $r \sim 0.58$). There is also very high correlation between ΔE_q and the differences in the weighted standard deviations of the whitened residuals, $\Delta\sigma_{\tau,\mathrm{wh}}$ ($r \sim 0.92$), and moderate correlation between τ_β and $\Delta\sigma_{\tau,\mathrm{wh}}$ ($r \sim 0.63$). However, it should be noted that the correlations between these variables are reduced when PSR J1045−4509 is omitted from the data set. In this case, the correlation between ΔE_q and τ_β is low ($r \sim 0.37$), the correlation between ΔE_q and $\Delta\sigma_{\tau,\mathrm{wh}}$ is moderate ($r \sim 0.60$), and the correlation between $\Delta\sigma_{\tau,\mathrm{wh}}$ and τ_β is low ($r \sim 0.45$).

As both E_q and $\sigma_{\tau,\mathrm{wh}}$ are significantly reduced by METM-MTM and both are correlated with τ_β, we conclude that polarization calibration errors contribute significantly to the additional white noise in pulsar timing data.

5.4. Red Noise

For most of the PPTA pulsars, the results in Table 4 show no significant differences in red-noise model parameters across the four data sets considered. The IFA-STM and METM-STM data sets for PSR J1600−3053 and the IFA-STM data set for PSR J1744−1134 appear to have marginally smaller red-noise amplitudes and marginally steeper spectra; however, the noise model parameter distribution plots included in Appendix B (Figures 8 and 9) clearly show that, in these three cases, the red-noise amplitude and spectral index are both highly covariant and poorly constrained. Therefore, these results are omitted from further consideration.

In contrast, for PSR J0437−4715, the best-fit red-noise model for the IFA-STM data set has a significantly larger amplitude (2.7σ difference) and marginally smaller spectral index (1.7σ difference) than those derived from the other three data sets. Figure 5 verifies that the red-noise model parameters are well constrained and only moderately covariant in the IFA-STM and METM-STM data sets. Both the larger red-noise amplitude and smaller spectral index ($\beta = 2$) derived from the IFA-STM data set are consistent with the presence of additional $1/f$ noise in the amplitude spectrum of these data. It is equally likely that additional white noise in the IFA-STM data biases the red-noise model parameters.

Figure 6 plots the amplitude spectra of the METM-MTM and IFA-STM postfit residuals for PSR J0437−4715 (computed before subtracting the best-fit red-noise model). The amplitude spectral density is the square root of the power spectral density (PSD); therefore, a PSD characterized by a power law with a spectral index of β, as defined in Equation (3), corresponds to a power-law amplitude spectrum with a spectral index of $\beta' = \beta/2$. The steeper power-law spectrum with slope $\beta' = 3/2$ that best fits the METM-MTM data is either buried under additional red noise with slope $\beta' = 1$ in the IFA-STM data or buried under additional white noise at key frequencies around per year that constrain the slope of the spectrum.

It may be possible to differentiate between these two equally probable interpretations by analyzing data that span more time and radio frequencies (e.g., Zic et al. 2023), or by searching for the interpulsar correlated systematic timing error induced by polarimetric distortion (van Straten 2013). This systematic error is a function of the polarized emission from each pulsar, and is independent of the angular separation between them. Therefore, instrumental distortion adversely impacts the sensitivity of a PTA experiment to all moments (monopolar, dipolar, and quadrupolar) in a multipole expansion of residual timing delays.

5.5. Impact of METM

In the interest of better understanding which of the two techniques—METM or MTM—has the greatest impact on pulsar timing experiments, this section compares four statistical measures (ToA goodness-of-fit, EFAC, EQUAD, and weighted standard deviation of whitened residuals) for the two calibration methods (METM and IFA), and the following section compares these quantities for the two arrival time estimation methods (MTM and STM).

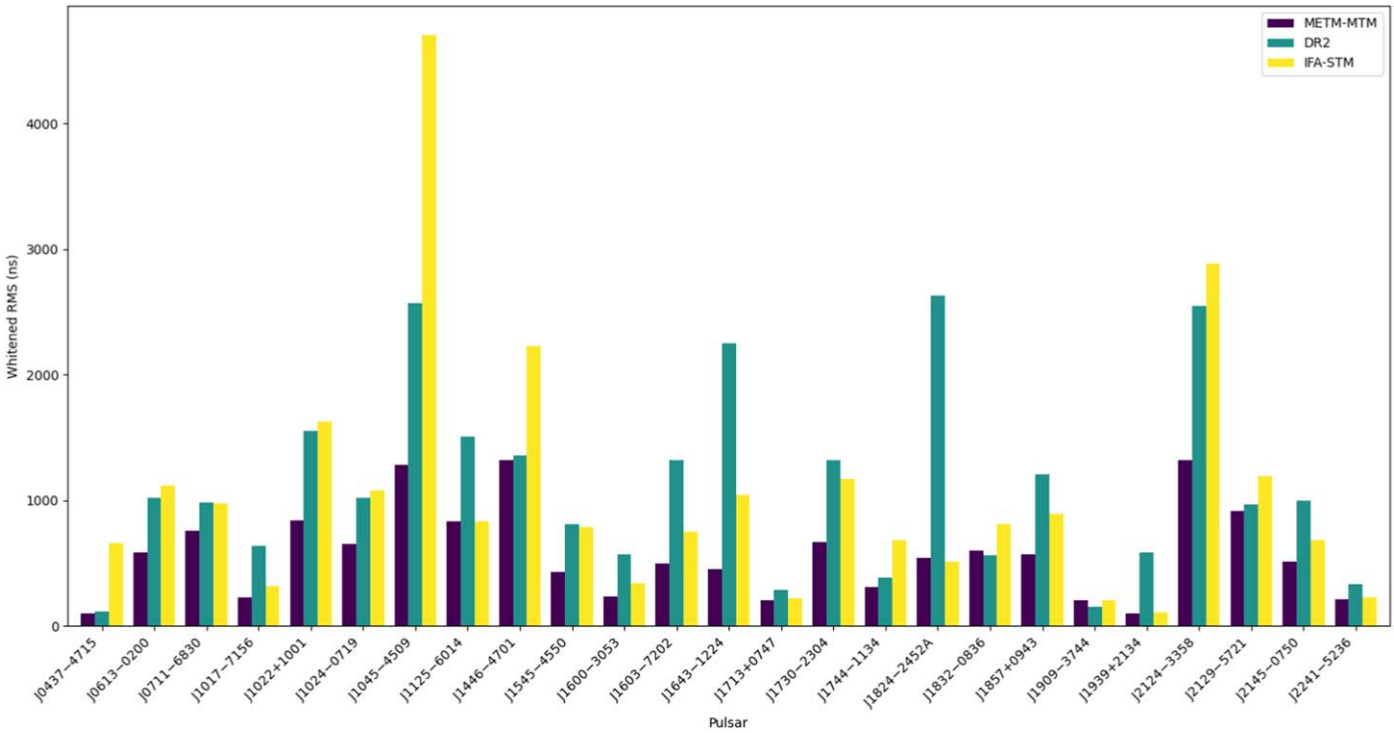


Figure 7. Comparison of DR2, METM-MTM, and IFA-STM $\sigma_{\tau,\mathrm{wh}}$ for all pulsars. Each bar represents the $\sigma_{\tau,\mathrm{wh}}$ value obtained using METM-MTM (blue), DR2 (green), and IFA-STM (yellow) for a specific pulsar. Note that DR2 uses the invariant interval (Britton 2000) for PSR J0437−4715.

The information in Figure 4 and its caption can be used to compare the goodness-of-fit of arrival times computed after calibration using either METM or IFA. First, when comparing IFA-STM with METM-STM (columns (3) and (5)), four pulsars stand out as having exceptionally high values of median reduced χ^2 in the IFA-STM data set: PSR J0437−4715, PSR J1022+1001, PSR J1713+0747, and PSR J1939+2134. For each of these pulsars, METM significantly improves the STM goodness-of-fit. When comparing IFA-MTM with METM-MTM (columns (2) and (4)), a similar reduction in reduced χ^2 is observed only for PSR J0437−4715. That is, METM generally has little impact on ToA goodness-of-fit when MTM is used.

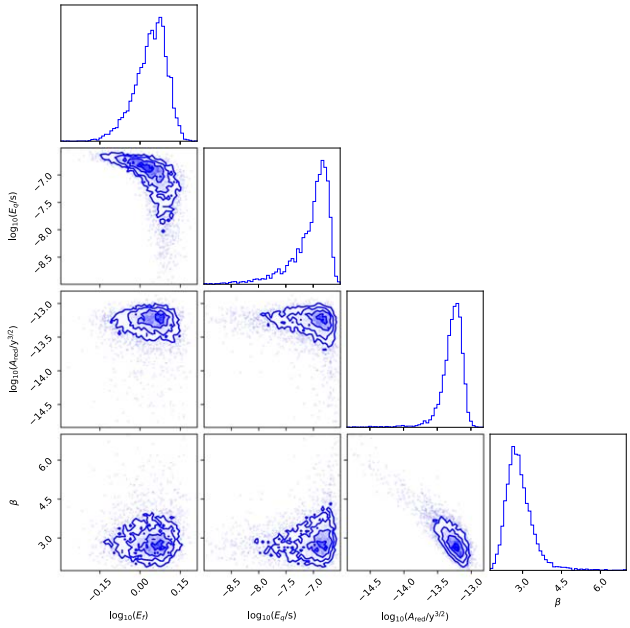
In Table 5, the comparison of the median values of E_f for METM-STM and IFA-STM (bottom two rows) and the median values of E_f for METM-MTM and IFA-MTM (top two rows) shows that METM only marginally reduces the EFAC for both template-matching techniques. For arrival times derived using STM, the median value of E_q drops by 30% from 0.33 (IFA) to 0.23 μs (METM). For MTM-derived ToAs, there is negligible difference in the median E_q for IFA and METM. In summary, METM generally has minimal impact on white-noise model parameters when MTM is used.

Although METM significantly reduces the $\sigma_{\tau,\mathrm{wh}}$ of STM-derived arrival times for a number of pulsars in Table 4 (e.g., PSR J0437−4715, PSR J1022+1001, PSR J1603−3053, and PSR J1744−1134), there are also some pulsars for which METM increases $\sigma_{\tau,\mathrm{wh}}$(e.g., PSR J1832−0836 and PSR J2124−3358). Consequently, the median values of $\sigma_{\tau,\mathrm{wh}}$ listed in Table 5 show that, with respect to IFA, METM slightly increases the median value of $\sigma_{\tau,\mathrm{wh}}$ by about 30 ns for arrival times derived using either template-matching method. When comparing $\sigma_{\tau,\mathrm{wh}}$ of MTM-derived arrival times of each pulsar, the differences between METM and IFA are very small.

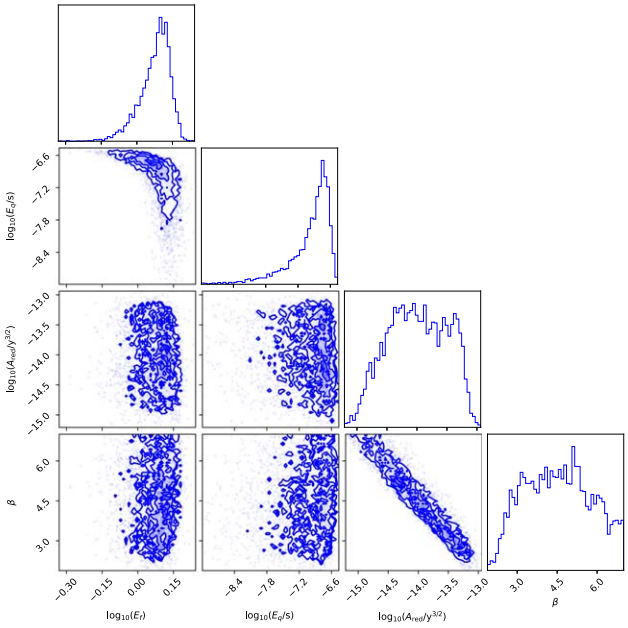
In summary, METM has negligible impact on ToAs derived using MTM, and only minimal impact on ToAs derived using STM. METM improves the STM goodness-of-fit and reduces the $\sigma_{\tau,\mathrm{wh}}$ of STM-derived arrival times for a small number of pulsars. For the STM-derived arrival times of all pulsars, METM reduces the median value of E_q by 100 ns, but also increases the median value of $\sigma_{\tau,\mathrm{wh}}$ by 30 ns.

5.6. Impact of MTM

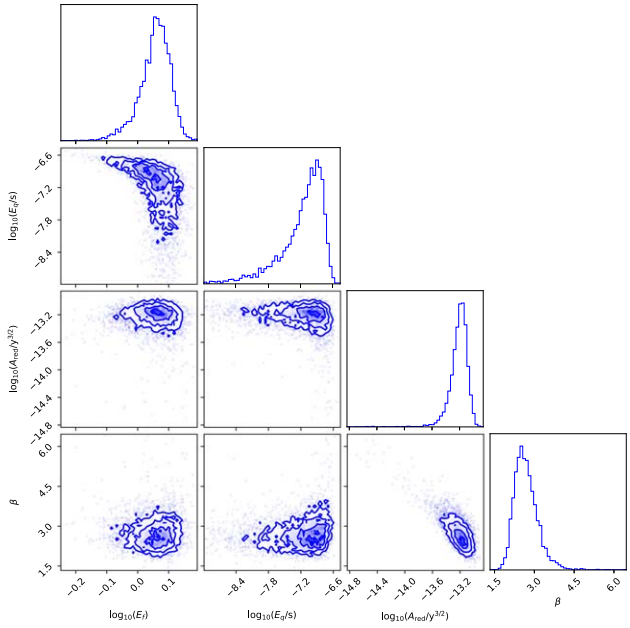
As done in the previous section, the information in Figure 4 (and its caption) is used to compare the goodness-of-fit of arrival times computed using either MTM or STM. We start by comparing IFA-MTM with IFA-STM (columns (2) and (3)). For each of the four previously identified pulsars with exceptionally high values of IFA-STM median reduced χ^2 (PSR J0437−4715, PSR J1022+1001, PSR J1713+0747, and PSR J1939+2134), MTM significantly improves the goodness-of-fit. For PSR J0437−4715 and PSR J1022+1001, the improvement due to MTM is greater than the improvement due to METM. For PSR J1939+2134, the opposite is observed. When comparing METM-MTM with METM-STM (columns (4) and (5)), the reduced χ^2 values for MTM and STM are similar (difference less than 0.05%) for most pulsars except PSR J0437−4715, PSR J1022+1001, and PSR J2241−5236, for which MTM achieves a better fit than STM; and PSR J1824−2452A, PSR J1939+2134, and PSR J2124−3358, for which STM achieves a better fit than MTM. Overall, the differences between MTM and STM ToA goodness-of-fit are greatest in the IFA-calibrated data set.



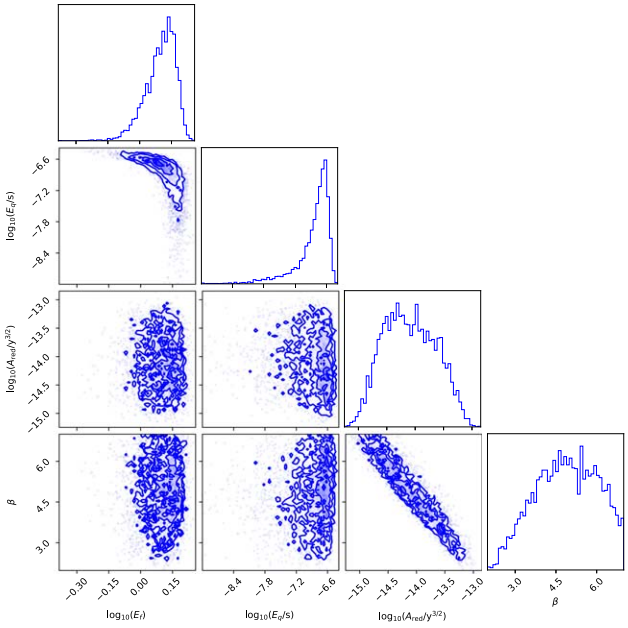
(a) METM-MTM



(b) METM-STM



(c) IFA-MTM



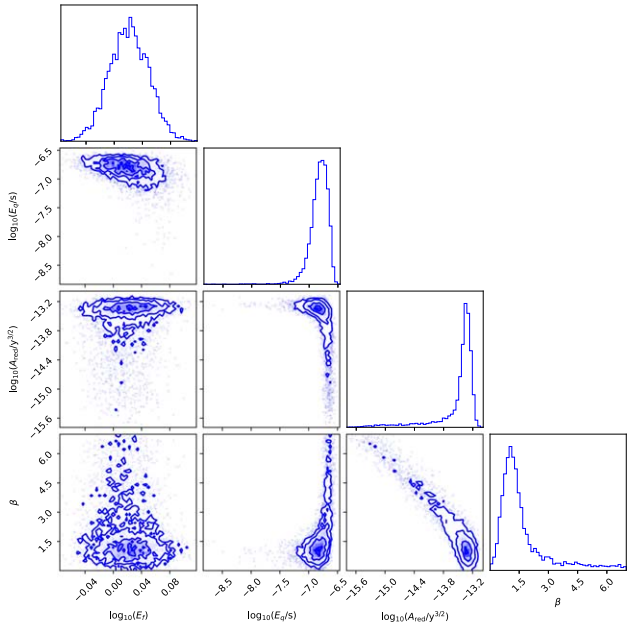
(d) IFA-STM

Figure 8. Noise model parameter distributions for PSR J1600−3053. See Figure 5 for information.

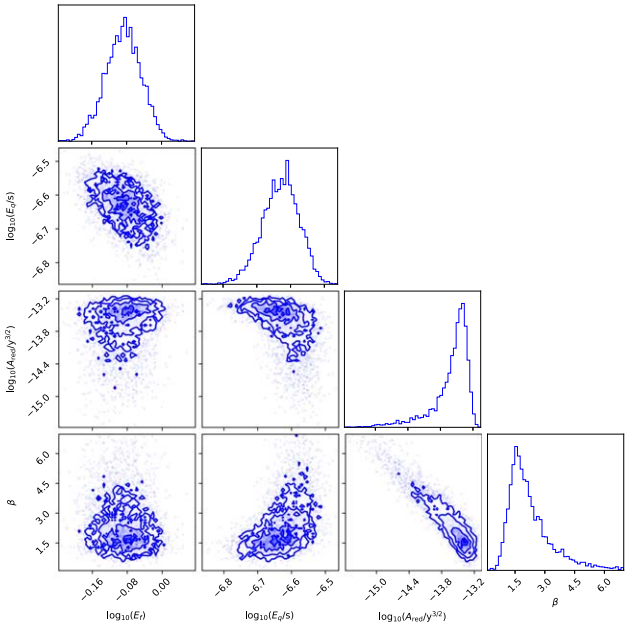
In Table 5, the comparison of the median values of E_f for METM-MTM and METM-STM (rows (1) and (3)) and the median values of E_f for IFA-MTM and IFA-STM (rows (2) and (4)) shows that MTM significantly reduces the EFAC for both calibration techniques. For data calibrated using METM, the median value of E_q drops by 30% from 0.23 (STM) to 0.16 μs (MTM). For IFA-calibrated data, the median value of E_q drops by 55% from 0.33 (STM) to 0.15 μs (MTM). In general, MTM significantly reduces the median values of both white-noise model parameters regardless of which calibration technique is used. MTM also significantly reduces the maximum values of E_q for both calibration methods (from 3.85 to 0.95 μs for IFA; and from 4.61 to 0.99 μs for METM).

A similar trend is observed in the weighted standard deviations of the whitened residuals. For both calibration techniques, MTM reduces the median value of $\sigma_{\tau,\mathrm{wh}}$ by 0.3 μs and the maximum value of $\sigma_{\tau,\mathrm{wh}}$ by 3 μs.

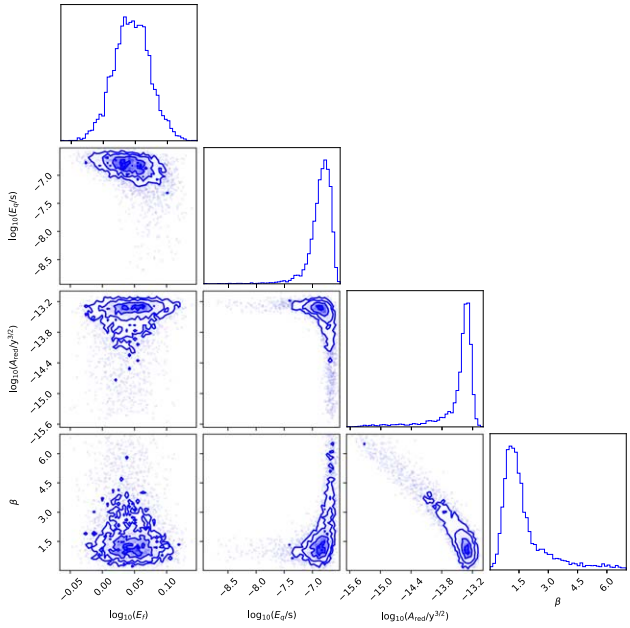
In summary, MTM has significant impact on ToAs derived from data using either calibration technique. For data calibrated using either IFA or METM, MTM improves the ToA goodness-of-fit, reduces the median E_f, and reduces both the median and maximum values of E_q and $\sigma_{\tau,\mathrm{wh}}$.



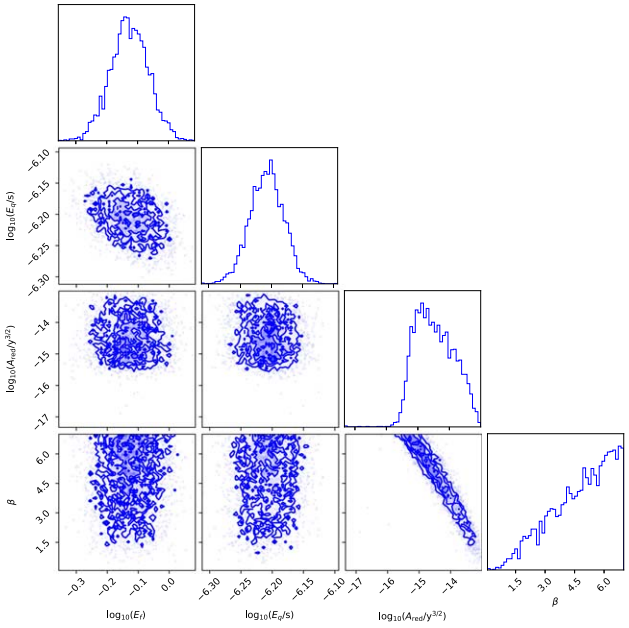
(a) METM-MTM



(b) METM-STM



(c) IFA-MTM



(d) IFA-STM

Figure 9. Noise model parameter distributions for PSR J1744−1134. See Figure 5 for information.

5.7. Comparison with PPTA DR2

The values of $\sigma_{\tau,\mathrm{wh}}$ derived in this work are compared with those reported for PPTA DR2 in Figure 7. Although this visual comparison may usefully indicate those pulsars for which improved techniques can be expected to have the greatest impact, some caution and consideration are required when interpreting this plot. First, the PPTA DR2 20 cm data span longer periods of time and include observations made using a variety of instrumental backends, each of which may require the inclusion of jumps to model unknown time delays. In contrast, our analysis focuses on only one backend. Second, the PPTA DR2 data were calibrated using MEM (van Straten 2004), and therefore, the IFA-STM results presented in this work are expected to have greater values of $\sigma_{\tau,\mathrm{wh}}$ than those reported for PPTA DR2. This is particularly true for PSR J1045−4509 and PSR J1446−4701. Furthermore, for PPTA DR2, PSR J0437−4715 was timed using the polarimetric invariant profile (Britton 2000), which greatly mitigates the impact of polarization calibration errors (van Straten et al. 2001). Consequently, for this pulsar, $\sigma_{\tau,\mathrm{wh}}$ for IFA-STM is significantly greater than that of PPTA DR2. Keeping in mind

the above caveats, Figure 7 shows that METM-MTM yields a better timing precision than PPTA DR2 for 23 of the 25 pulsars studied. The two exceptions are PSR J1832−0836 and PSR J1909−3744, pulsars that have low susceptibility to calibration error as characterized by τ_β.

6. Conclusion

Compared to conventional approaches for polarization calibration and arrival time estimation, the combination of METM (van Straten 2013) and MTM (S06) significantly reduces the white noise in pulse arrival times. For PSR J0437−4715, METM and MTM reduce the best-fit amplitude of the red noise in the timing residuals, either by mitigating additional red noise or by increasing the accuracy of red-noise model parameter estimates by reducing the bias on these parameters due to white noise.

In this work, we evaluated the impact of METM and MTM on the timing precision using four different quantities: the ToA goodness-of-fit, EFAC E_f, EQUAD E_q, and the weighted standard deviation of the whitened postfit residuals $\sigma_{\tau,\mathrm{wh}}$. We found that, with respect to the baseline IFA-STM data set, both METM and MTM improve the ToA goodness-of-fit; however, METM generally has little impact when MTM is used. A similar pattern is observed in Figure 16 of Guillemot et al. (2023).

Whereas METM only marginally reduces the median EFAC, MTM significantly decreases the median value of E_f. In general, this indicates that the STM algorithm used in this work typically underestimates arrival time uncertainty; however, for PSR J0437−4715, E_f is dominated by jitter for both template-matching techniques. We also found that both METM and MTM significantly reduce the EQUAD; however, the median value of E_q differs very little between METM and IFA when MTM is used. In contrast, MTM significantly decreases both the median and maximum values of E_q for both METM and IFA calibration methods. Finally, METM was shown to marginally increase the median value of $\sigma_{\tau,\mathrm{wh}}$, whereas MTM significantly decreases both the median and maximum values of $\sigma_{\tau,\mathrm{wh}}$. (As expected, changes in E_q and $\sigma_{\tau,\mathrm{wh}}$ are highly correlated.)

In summary, MTM significantly improves the arrival time precision, regardless of the calibration technique used. This indicates that, as long as the method of calibration is sufficiently accurate to minimize bandwidth depolarization, MTM is able to model any residual calibration errors and mitigate their impact on arrival time estimates. Therefore, we recommend that MTM should be used for most pulsars in every PTA experiment. Depending on the instrument, achieving sufficiently well-calibrated observations may necessitate the use of either MEM or METM.

An increased arrival time precision has the potential to enhance our ability to detect errors in solar system ephemerides (Vallisneri et al. 2020) and terrestrial time models (Hobbs et al. 2020), facilitate new or more accurate measurements of pulsar properties (e.g., van Straten 2013), and increase PTA sensitivity to the stochastic GWB. Therefore, MTM should be adopted and utilized for all future IPTA data releases.

Acknowledgments

We are grateful to Bill Coles for helpful advice and to the anonymous referee for constructive criticism that greatly improved the manuscript. We also extend our thanks to the collaborators for their significant contributions to this work. The Parkes radio telescope is part of the Australia Telescope National Facility (https://ror.org/05qajvd42), which the Australian Government funds for operation as a National Facility managed by CSIRO. We acknowledge the Wiradjuri people as the Traditional Owners of the Observatory site.

Appendix A
Arrival Time Uncertainty Corrections

A.1. Scalar Template-matching Correction

The PSRCHIVE software used for this study includes two different algorithms for estimating the phase shift between a high-S/N template profile and an observed profile by cross-correlation in the Fourier domain (T92). The PGS algorithm uses the Van Wijngaarden–Dekker–Brent method (Brent 1973) to find the phase shift τ that minimizes an objective χ^2 merit function, then uses the curvature, $\partial^2\chi^2/\partial\tau^2$, to calculate the theoretical uncertainty of τ. The Fourier domain with Markov Chain Monte Carlo method minimizes the same objective merit function using the Levenberg–Marquardt algorithm, then optionally uses Markov Chain Monte Carlo to sample the distribution of τ and calculate its uncertainty. The MTM algorithm also uses the curvature of χ^2 to compute arrival time uncertainty; therefore, in this work, we experimentally compare MTM and STM using the PGS implementation.

Both MTM and PGS are expected to underestimate the uncertainty at low S/N, as demonstrated for the PGS algorithm through simulations (Appendix A of Hotan et al. 2005), mathematical proof (Appendix B of Arzoumanian et al. 2015), and comparative analysis of experimental data (Wang et al. 2022). However, the original implementation of PGS also overestimates the arrival time uncertainty when the fit between the template and observation is poor. This unexpected result is most obvious when analyzing the PSR J0437−4715 timing residuals. For arrival times estimated using the original PGS implementation, the best-fit estimate of the E_f noise model parameter increased from ∼1.8 for IFA-calibrated data to ∼5.9 for METM-calibrated data.

The apparent increase in uncertainty is an artifact of the original PGS implementation, which computes the formal error of the phase shift based on the incorrect assumption that the reduced χ^2 is unity. This is equivalent to assuming that the noise in the postfit residual profile is equivalent to the radiometer noise in each harmonic (σ in Equations (A6) through (A11) of T92). This assumption breaks down when the observed profile is not a good match to the template, as is the case when the total intensity is significantly distorted by residual calibration errors.

Poor template-matching fits in the IFA-calibrated data are expected to increase χ^2; however, owing to the incorrect definition of σ in the original PGS implementation, they also artificially inflate the derived arrival time uncertainty (σ_τ, defined by Equation (A10) of T92). In contrast, METM calibration significantly reduces distortions to the total intensity profile, and thereby improves the STM fit and reduces the arrival time uncertainties yielded by PGS. Consequently, a larger value of E_f is required to account for things like the pulsar-intrinsic jitter in the METM-calibrated data. Although it is smaller, a similar inflation of E_f was also observed for

PSR J1022+1001, which is also highly susceptible to polarization calibration errors (van Straten 2013).

The erroneous assumption that the reduced χ^2 equals unity was made optional in the PSRCHIVE software (on 2023 July 4), and this assumption was disabled before reproducing the results and analysis presented in this paper.

A.2. MTM Correction

During our initial analysis of the best-fit noise model parameters produced by TEMPONEST, we found that all arrival times estimated using MTM had a median E_f of 0.7, indicating that this algorithm systematically *overestimates* the arrival time uncertainty by a factor of approximately $\sqrt{2}$. To better understand the origin of this erroneous scale factor, we revisited both the derivation and the implementation of the equations that define the uncertainty of the best-fit MTM phase shift, first presented in Section 3.2 of S06. Here, the covariance matrix that defines the formal uncertainties of the MTM model parameters is given by $\boldsymbol{C} = \boldsymbol{\alpha}^{-1}$, where $\boldsymbol{\alpha}$ is the curvature matrix defined by Equation (14) of S06. This relationship between $\boldsymbol{\alpha}$ and $\boldsymbol{C}$ follows Equation (15.5.15) of Numerical Recipes (Press et al. 1992; hereafter NR), and the definition of $\boldsymbol{\alpha}$ is based on Equations (15.5.8) and (15.5.11) of NR. (In particular, following the discussion in Section 15.5 of NR, the term containing a second derivative has been dropped.) Compared to Equation (15.5.11) of NR, Equation (14) of S06 includes an extra factor of 2; however, this factor of 2 was missing in the PSRCHIVE implementation of the calculation. Erroneously dividing $\boldsymbol{\alpha}$ by 2 is equivalent to multiplying $\boldsymbol{C}$ by 2 and inflating the MTM arrival time uncertainty by $\sqrt{2}$.

A similar factor of 2 error appeared in the PSRCHIVE calculation of the gradient vector, $\boldsymbol{\beta}$, which is also defined by Equation (15.5.8) of NR. Therefore, for future reference, a complete derivation of both $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ is provided here, beginning with the merit function defined by Equation (10) of S06

$$\chi^2 = \sum_{m=1}^{N/2} \sum_{k=0}^{3} |S_{m,k} - \mathrm{tr}[\boldsymbol{\sigma}_k \, \boldsymbol{\rho}'_m]|^2 \varsigma_k^{-2}. \tag{A1}$$

In this equation, $S_{m,k}$ are the complex-valued Fourier transforms of the average pulse profiles of the observed Stokes parameters, as a function of pulsar spin harmonic m and Stokes parameter index k; $\boldsymbol{\rho}'_m$ is the model coherency matrix for harmonic m, $\boldsymbol{\sigma}_k$ are the Hermitian basis matrices, and tr[$\mathbf{A}$] is the trace of matrix $\boldsymbol{A}$. Define the observed coherency matrix at harmonic m

$$\boldsymbol{\rho}_m = \frac{1}{2} \sum_{k=0}^{3} S_{m,k} \boldsymbol{\sigma}_k, \tag{A2}$$

such that $S_{m,k} = \mathrm{tr}[\boldsymbol{\sigma}_k \, \boldsymbol{\rho}_m]$, then use the linearity of the matrix trace and the transitivity of matrix multiplication to express the difference between observed and model Stokes parameters for harmonic m

$$D_{m,k} = \mathrm{tr}[\boldsymbol{\sigma}_k (\boldsymbol{\rho}_m - \boldsymbol{\rho}'_m)]. \tag{A3}$$

Assuming that the noise in each Stokes parameter is equal, let $\varsigma_k = \varsigma$. The first partial derivative of χ^2 with respect to model parameter η_r is then

$$\frac{\partial \chi^2}{\partial \eta_r} = -\frac{2}{\varsigma^2} \sum_{m=1}^{N/2} \sum_{k=0}^{3} \mathrm{Re} \left[D_{m,k} \mathrm{tr} \left(\boldsymbol{\sigma}_k \frac{\partial \boldsymbol{\rho}'_m}{\partial \eta_r} \right)^* \right]. \tag{A4}$$

Because $\mathrm{tr}(\boldsymbol{\sigma}_k \, \boldsymbol{\rho})^* = \mathrm{tr}(\boldsymbol{\sigma}_k \, \boldsymbol{\rho}^\dagger)$ and both the trace of a matrix and the real part of a complex number are linear,

$$\frac{\partial \chi^2}{\partial \eta_r} = -\frac{2}{\varsigma^2} \sum_{m=1}^{N/2} \mathrm{Re} \, \mathrm{tr} \left[\sum_{k=0}^{3} D_{m,k} \boldsymbol{\sigma}_k \frac{\partial \boldsymbol{\rho}'^\dagger_m}{\partial \eta_r} \right]. \tag{A5}$$

Combining the definitions of the coherency matrix and the Stokes parameters yields

$$\boldsymbol{\rho} = \frac{1}{2} \sum_{k=0}^{3} \mathrm{tr}(\boldsymbol{\sigma}_k \, \boldsymbol{\rho}) \boldsymbol{\sigma}_k; \tag{A6}$$

therefore,

$$\sum_{k=0}^{3} D_{m,k} \boldsymbol{\sigma}_k = 2(\boldsymbol{\rho}_m - \boldsymbol{\rho}'_m) \tag{A7}$$

and

$$\beta_r \equiv -\frac{1}{2} \frac{\partial \chi^2}{\partial \eta_r} = \frac{2}{\varsigma^2} \sum_{m=1}^{N/2} \mathrm{Re} \, \mathrm{tr} \left[(\boldsymbol{\rho}_m - \boldsymbol{\rho}'_m) \frac{\partial \boldsymbol{\rho}'^\dagger_m}{\partial \eta_r} \right]. \tag{A8}$$

Taking the partial derivative of $-\beta_r$ with respect to model parameter η_s (and dropping the term containing a second derivative) yields

$$\alpha_{rs} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial \eta_r \partial \eta_s} = \frac{2}{\varsigma^2} \sum_{m=1}^{N/2} \mathrm{Re} \, \mathrm{tr} \left[\frac{\partial \boldsymbol{\rho}'^\dagger_m}{\partial \eta_r} \frac{\partial \boldsymbol{\rho}'_m}{\partial \eta_s} \right], \tag{A9}$$

which is equivalent to Equation (14) of S06. Both β_r and α_{rs} include a factor of 2 that was missing in the PSRCHIVE adaptation of the Levenberg–Marquardt algorithm to complex-valued matrices. This error was corrected (on 2023 July 9) before reproducing the results and analysis presented in this paper.

Appendix B
Rejected Red-noise Models

When comparing the four methods applied to each pulsar in Table 4, there are slight differences in the best-fit red-noise model parameters that mostly fall within the estimated uncertainties. Exceptions include three data sets (PSR J1600−3053 IFA-STM and METM-STM, and PSR J1744−1134 IFA-STM) that exhibit marginally smaller red-noise amplitudes and steeper spectra; however, the following plots show that these parameters are highly covariant and poorly constrained in these three cases. Consequently, these minor differences are not further considered.

Appendix C
Calibration Steps

To help others integrate MEM, METM, and MTM into existing pipelines, this appendix provides a more detailed description of the commands used for polarimetric calibration.

C.1. Produce MEM Solutions

Inputs are as follows:

1. uncalibrated observations of the reference pulsar (e.g., PSR J0437−4715 at Parkes), divided into long sessions, where each session spans at most 1 day (horizon to horizon) and includes a minimum of 2 hr of observations; and
2. uncalibrated observations of the square-wave amplitude-modulated noise diode (CAL) observed prior to each pulsar observation.

Steps are as follows:
Assuming that the CAL observations are listed in a single flat ASCII file named `database.txt` (as produced by pac) and that the pulsar observations for a single session are listed in a file named `session.ls`, the following steps are performed for each session:

A. Calibrate the data using the IFA, and produce a time-integrated total named `choose.ar`, which is later passed to the polarization calibration modeling (pcm) program for use when choosing the phase bins to include in the MEM fit.
```
pac -P -O pac_out -d database.txt -M session.ls
psradd -T -o choose.ar pac_out/*.calibP
```
B. Run pcm in MEM mode.
```
pcm $args -c choose.ar -d database.txt -M session.ls
```

where `$args` includes the following command-line options:

`-m bri00e19` uses Equation (19) of Britton (2000) to model the instrumental response,
`-k` assumes that the receptors have equal ellipticities,
`-Q` models the noise diode as coupled after the orthomode transducer,
`-s` normalizes Stokes parameters by the phase-integrated invariant interval,
`-n 64` uses 64 phase bins as model constraints,
`-a 0` disables the phase-alignment check,
`-K 3.0` rejects outliers when computing CAL levels,
`-step 3.0` detects and models steps in instrumental response,
`-X 2.0` masks channels with $\chi^2/N_{\rm free} > 2.0$,
`-N` does not unload calibrated data files.

Outputs are as follows:

1. `pcm.fits` is the best-fit parameters that describe the model of the instrumental response and the polarization of the noise diode (CAL).
2. `total.ar` is the calibrated time-integrated full-polarization average pulse profile for this session.

The `pcm.fits` files should be reviewed, and solutions with any obvious model-fitting problems, errors, or poorly constrained model parameters should be discarded.

C.2. Create METM Template Profile

Inputs are as follows:

1. The `total.ar` files were produced for each MEM session.

Steps are as follows:

A. Choose the best MEM solution (criteria described in the paper), and rename its `total.ar` file to `chosen.ar`.
B. Model and eliminate any temporal variations (e.g., changes in instrumental response, ionospheric Faraday rotation, etc.) before further integrating in time. This is for each `total.ar`
```
pcm -t8 -n 128 -S chosen.ar -out total.mtm total.ar
```
C. Integrate the calibrated `total.calib` files output by pcm and the `chosen.ar` template.
```
psradd -T -o pulsar.std chosen.ar */total.calib
-j "weight snr fscrunch=1" -phath 0.003.
```

Outputs are as follows:

1. `pulsar.std` is the template to be used for METM.

C.3. Produce METM Solutions

Inputs are as follows:

1. `pulsar.std` as the high-S/N, well-calibrated, full-polarization average profile of the reference pulsar;
2. uncalibrated observations of the reference pulsar (e.g., PSR J0437−4715 at Parkes), divided into short sessions that include a minimum of 1 hr of observations; and
3. uncalibrated observations of the square-wave amplitude-modulated noise diode (CAL) observed prior to each pulsar observation.

Steps are as follows:
Assuming that the CAL observations are listed in a single flat ASCII file named `database.txt` (as produced by pac) and that the pulsar observations for a single short session are listed in a file named `session.ls`, the following step is performed for each session,

```
pcm -S pulsar.std -d database.txt -M session.ls $args
```

where `$args` includes the following command-line options:

`-m bri00e19` uses Equation (19) of Britton (2000) to model the instrumental response,
`-Q` models the noise diode as coupled after the orthomode transducer,
`-s` normalizes Stokes parameters by the phase-integrated invariant interval,
`-n 200` uses 200 harmonics as model constraints,
`-K 3.0` rejects outliers when computing CAL levels,
`-step 3.0` detects and models steps in instrumental response,
`-X 2.0` masks channels with $\chi^2/N_{\rm free} > 2.0$.

Outputs are as follows:

1. `pcm.fits` is the best-fit parameters that describe the model of the instrumental response and the polarization of the noise diode (CAL); and
2. `total.ar` is the calibrated time-integrated full-polarization average pulse profile for this session.

The `pcm.fits` files should be reviewed, and solutions with any obvious model-fitting problems, errors, or poorly constrained model parameters should be discarded.

C.4. Correct Ionospheric Faraday Rotation

Inputs are as follows:

1. All of the solutions contained in `pcm.fits` files were produced using METM.

Steps are as follows: Estimate the ionospheric RM contribution, and subtract it from each session.

```
pcmrm pcm.fits.
```

Outputs are as follows:

1. `pcm.rms` is the `pcm.fits` solution with σ_θ corrected for ionospheric Faraday rotation.

C.5. Smooth METM Solutions

Inputs are as follows:

1. `file.ls` is the listing of the `pcm.rmc` files corrected for ionospheric Faraday rotation using `pcmrm`.

Steps are as follows: Spline-smooth the METM solutions using Monte Carlo cross-validation.

```
smint -p1 -cross -cross-m 4 -cross-f 0.5
-cross-iqr 0 -M files.ls.
```

The following command-line options are specified:

`-p1` is the initial guess for the penalized spline smoothing factor.
`-cross` find the optimal spline smoothing factor using Monte Carlo cross-validation.
`-cross-m 4` is the number of cross-validation iterations.
`-cross-f 0.5` is the fraction of data used to validate the best-fit spline on each iteration.
`-cross-iqr 0` disables outlier excision using the uncertainty-weighted IQR.
`-M files.ls` is the list of `pcm.fits` files (corrected for ionospheric Faraday rotation).

Outputs are as follows:

1. `smint.fits` is the spline-smoothed calibrator solution,
2. `cal_stokes_fit*.eps`
3. `pcal_fit_*.eps.`

The postscript files plot the spectrum for each of the smoothed model parameters; each page shows the data extracted from one of the input `pcm.fits` files (black points with error bars) and the spline fit to the data at this epoch (red line).

C.6. Calibrate the Pulsar Data

Inputs are as follows:

1. `metm_database.txt` is the database of spline-smoothed METM calibrator solutions.
2. `cal_database.txt` is the database of CAL files.
3. `fluxcal_database.txt` is the database of FLUXCAL files.
4. `uncalibrated.ls` is a list of the pulsar observations to be calibrated.

Steps are as follows: Perform METM calibration.

```
pac $args -M uncalibrated.ls/ -d cal_data-
base.txt -d fluxcal_database.txt -d metm_-
database.txt
```

where `$args` includes the following command-line options:

`-K 3.0` rejects outliers when computing calibrator levels,
`-g` frequency-averages the data to match the number of channels of the calibrator,
`-m b` uses only calibrators observed before the pulsar,
`-S` uses the complete reception model,
`-e cmetm` is the extension added to output file names.

Outputs are as follows:

1. calibrated pulsar observations with a new extension, `*.cmetm.`

C.7. Produce Arrival Time Estimates

Inputs are as follows:

1. Calibrated pulsar observations are listed in a text file named `calibrated.ls.`
2. There is a well-calibrated template profile (e.g., the calibrated observation with the highest S/N).

Steps are as follows:

A. Time- and frequency-average the profile data.
```
pam -TF -e TF -M calibrated.ls
```
B. Produce the MTM arrival time estimates in a `.tim` file.
```
pat -Fpcs pulsar.std -f tempo2 -C gof *.
TF > mtm.tim
```

Outputs are as follows:

1. `mtm.tim` is a TEMPO2-formatted file of arrival time estimates produced by MTM.

ORCID iDs

Axl F. Rogers https://orcid.org/0009-0004-7606-9343
Willem van Straten https://orcid.org/0000-0003-2519-7375
Sergei Gulyaev https://orcid.org/0000-0003-0186-5551
Aditya Parthasarathy https://orcid.org/0000-0002-4140-5616
George Hobbs https://orcid.org/0000-0003-1502-100X
Zu-Cheng Chen https://orcid.org/0000-0001-7016-9934
Yi Feng https://orcid.org/0000-0002-0475-7479
Boris Goncharov https://orcid.org/0000-0003-3189-5807
Agastya Kapur https://orcid.org/0009-0001-5071-0962
Xiaojin Liu https://orcid.org/0000-0002-2187-4087
Daniel Reardon https://orcid.org/0000-0002-2035-4688
Christopher J. Russell https://orcid.org/0000-0002-1942-7296
Andrew Zic https://orcid.org/0000-0002-9583-2947

References

Agazie, G., Alam, M. F., Anumarlapudi, A., et al. 2023a, ApJL, 951, L9
Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023b, ApJL, 951, L8
Agazie, G., Antoniadis, A., Anumarlapudi, A., et al. 2024, ApJ, 966, 105
Aggarwal, K., Arzoumanian, Z., Baker, P. T., et al. 2019, ApJ, 880, 116
Ajello, M., Atwood, W. B., Baldini, L., et al. 2022, Sci, 376, 521
Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2021a, ApJS, 252, 5
Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2021b, ApJS, 252, 4
Antoniadis, J., Arumugam, P., Arumugam, S., et al. 2023a, A&A, 678, A50

Antoniadis, J., Arzoumanian, Z., Babak, S., et al. 2022, MNRAS, 510, 4873
Antoniadis, J., Babak, S., Bak Nielsen, A. S., et al. 2023b, A&A, 678, A48
Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2020, ApJL, 905, L34
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2014, ApJ, 794, 141
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2015, ApJ, 813, 65
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2016, ApJ, 821, 13
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2018, ApJS, 235, 37
Babak, S., Petiteau, A., Sesana, A., et al. 2016, MNRAS, 455, 1665
Bailes, M., Barr, E., Bhat, N. D. R., et al. 2016, in Proc. of MeerKAT Science: On the Pathway to the SKA, Vol. 277, 11
Blanco-Pillado, J. J., Olum, K. D., & Siemens, X. 2018, PhLB, 778, 392
Brent, R. P. 1973, Algorithms for Minimization Without Derivatives (1st ed.; Englewood Cliffs, NJ: Prentice-Hall)
Britton, M. C. 2000, ApJ, 532, 1240
Burke-Spolaor, S., Taylor, S. R., Charisi, M., et al. 2019, A&ARv, 27, 5
Caballero, R. N., Guo, Y. J., Lee, K. J., et al. 2018, MNRAS, 481, 5501
Caleb, M., van Straten, W., Keane, E. F., et al. 2019, MNRAS, 487, 1191
Caprini, C., Durrer, R., & Siemens, X. 2010, PhRvD, 82, 063511
Champion, D. J., Hobbs, G. B., Manchester, R. N., et al. 2010, ApJL, 720, L201
Chen, S., Caballero, R. N., Guo, Y. J., et al. 2021, MNRAS, 508, 4970
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., & McLaughlin, M. A. 2019, BAAS, 51, 447
Cordes, J. M., & Shannon, R. M. 2010, arXiv:1010.3785
Demorest, P. B., Ferdman, R. D., Gonzalez, M. E., et al. 2013, ApJ, 762, 94
Desvignes, G., Caballero, R. N., Lentati, L., et al. 2016, MNRAS, 458, 3341
Detweiler, S. 1979, ApJ, 234, 1100
Dolch, T., Stinebring, D. R., Jones, G., et al. 2021, ApJ, 913, 98
Edwards, R. T., Hobbs, G. B., & Manchester, R. N. 2006, MNRAS, 372, 1549
Eilers, P. H. C., & Marx, B. D. 1996, StaSc, 11, 89
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Goncharov, B., Shannon, R. M., Reardon, D. J., et al. 2021, ApJL, 917, L19
Grimstad, B., et al. 2015, SPLINTER:A Library for Multivariate Function Approximation with Splines, http://github.com/bgrimstad/splinter
Grishchuk, L. P. 1975, JETP, 40, 409
Guillemot, L., Cognard, I., van Straten, W., Theureau, G., & Gérard, E. 2023, A&A, 678, A79
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., Coles, W., Manchester, R. N., et al. 2012, MNRAS, 427, 2780
Hobbs, G., Dai, S., Manchester, R. N., et al. 2019, RAA, 19, 020
Hobbs, G. B., Edwards, R. T., & Manchester, R. N. 2006, MNRAS, 369, 655
Hobbs, G., Guo, L., Caballero, R. N., et al. 2020, MNRAS, 491, 5951
Hotan, A. W., Bailes, M., & Ord, S. M. 2004, MNRAS, 355, 941
Hotan, A. W., Bailes, M., & Ord, S. M. 2005, MNRAS, 362, 1267
Jenet, F. A., & Anderson, S. B. 1998, PASP, 110, 1467
Jenet, F. A., Hobbs, G. B., Lee, K. J., & Manchester, R. N. 2005, ApJL, 625, L123
Jones, M. L., McLaughlin, M. A., Lam, M. T., et al. 2017, ApJ, 841, 125
Joshi, B. C., Arumugasamy, P., Bagchi, M., et al. 2018, JApA, 39, 51
Keith, M. J., Coles, W., Shannon, R. M., et al. 2013, MNRAS, 429, 2161
Kerr, M., Reardon, D. J., Hobbs, G., et al. 2020, PASA, 37, e020
Kobakhidze, A., Lagger, C., Manning, A., & Yue, J. 2017, EPJC, 77, 570
Kramer, M., & Champion, D. J. 2013, CQGra, 30, 224009
Kramer, M., Xilouris, K. M., Camilo, F., et al. 1999, ApJ, 520, 324
Lam, M. T., Cordes, J. M., Chatterjee, S., et al. 2017, ApJ, 834, 35
Lasky, P. D., Mingarelli, C. M. F., Smith, T. L., et al. 2016, PhRvX, 6, 011035
Lazarus, P., Karuppusamy, R., Graikou, E., et al. 2016, MNRAS, 458, 868
Lazarus, P., Karuppusamy, R., Graikou, E., et al. 2020, CoastGuard: Automated Timing Data Reduction Pipeline, Astrophysics Source Code Library, ascl:2003.008
Lee, K. J. 2016, in ASP Conf. Ser. 502, Frontiers in Radio Astronomy and FAST Early Sciences Symp., ed. L. Qain & D. Li (San Francisco, CA: ASP), 19
Lentati, L., Alexander, P., Hobson, M. P., et al. 2014, MNRAS, 437, 3004
Lentati, L., Shannon, R. M., Coles, W. A., et al. 2016, MNRAS, 458, 2161
Lentati, L., Taylor, S. R., Mingarelli, C. M. F., et al. 2015, MNRAS, 453, 2576
Lower, M. E., Bailes, M., Shannon, R. M., et al. 2020, MNRAS, 494, 228
Lu, S., & Chipman, R. A. 1996, JOSAA, 13, 1106
Lyne, A., Hobbs, G., Kramer, M., Stairs, I., & Stappers, B. 2010, Sci, 329, 408
Manchester, R. N., Hobbs, G., Bailes, M., et al. 2013, PASA, 30, e017
McLaughlin, M. A. 2013, CQGra, 30, 224008
Melatos, A., & Link, B. 2014, MNRAS, 437, 21
Morello, V., Barr, E. D., Cooper, S., et al. 2019, MNRAS, 483, 3673
Ng, C. 2018, in IAU Symp. 337, Pulsar Astrophysics the Next Fifty Years, ed. P. Weltevrede et al. (Cambridge: Cambridge Univ. Press), 179
Ord, S. M., van Straten, W., Hotan, A. W., & Bailes, M. 2004, MNRAS, 352, 804
Osłowski, S., van Straten, W., Hobbs, G. B., Bailes, M., & Demorest, P. 2011, MNRAS, 418, 1258
Parthasarathy, A., Bailes, M., Shannon, R. M., et al. 2021, MNRAS, 502, 407
Parthasarathy, A., Shannon, R. M., Johnston, S., et al. 2019, MNRAS, 489, 3810
Perera, B. B. P., DeCesar, M. E., Demorest, P. B., et al. 2019, MNRAS, 490, 4666
Perera, B. B. P., Stappers, B. W., Babak, S., et al. 2018, MNRAS, 478, 218
Phinney, E. S. 2001, arXiv:astro-ph/0108028
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, in Numerical Recipes in C. The Art of Scientific Computing, ed. L. Cowles & A. Harvey (Cambridge: Cambridge Univ. Press)
Rajagopal, M., & Romani, R. W. 1995, ApJ, 446, 543
Ramachandran, R., & Kramer, M. 2003, A&A, 407, 1085
Ransom, S., Brazier, A., Chatterjee, S., et al. 2019, BAAS, 51, 195
Reardon, D. J. 2021, MeerGuard, https://github.com/danielreardon/MeerGuard
Reardon, D. J., Hobbs, G., Coles, W., et al. 2016, MNRAS, 455, 1751
Reardon, D. J., Zic, A., Shannon, R. M., et al. 2023, ApJL, 951, L6
Sazhin, M. V. 1978, SvA, 22, 36
Sesana, A., Haardt, F., Madau, P., & Volonteri, M. 2004, ApJ, 611, 623
Shannon, R. M., & Cordes, J. M. 2010, ApJ, 725, 1607
Shannon, R. M., Cordes, J. M., Metcalfe, T. S., et al. 2013b, ApJ, 766, 5
Shannon, R. M., Oslowski, S., Dai, S., et al. 2014, MNRAS, 443, 1463
Shannon, R. M., Ravi, V., Coles, W. A., et al. 2013a, Sci, 342, 334
Shannon, R. M., Ravi, V., Lentati, L. T., et al. 2015, Sci, 349, 1522
Siemens, X., Mandic, V., & Creighton, J. 2007, PhRvL, 98, 111101
Susobhanan, A., Gopakumar, A., Hobbs, G., & Taylor, S. R. 2020, PhRvD, 101, 043022
Taylor, J. H. 1992, RSPTA, 341, 117
Taylor, S. R., Vallisneri, M., Ellis, J. A., et al. 2016, ApJL, 819, L6
Tiburzi, C., Hobbs, G., Kerr, M., et al. 2016, MNRAS, 455, 4339
Tukey, J. W. 1977, Exploratory Data Analysis (Reading, MA: Addison-Wesley)
Vallisneri, M., Taylor, S. R., Simon, J., et al. 2020, ApJ, 893, 112
Vallisneri, M., & van Haasteren, R. 2017, MNRAS, 466, 4954
van Haasteren, R., & Levin, Y. 2013, MNRAS, 428, 1147
van Haasteren, R., Levin, Y., Janssen, G. H., et al. 2011, MNRAS, 414, 3117
van Straten, W. 2004, ApJS, 152, 129
van Straten, W. 2006, ApJ, 642, 1004
van Straten, W. 2013, ApJS, 204, 13
van Straten, W., Bailes, M., Britton, M., et al. 2001, Natur, 412, 158
Verbiest, J. P. W., Lentati, L., Hobbs, G., et al. 2016, MNRAS, 458, 1267
Wang, J., Shaifullah, G. M., Verbiest, J. P. W., et al. 2022, A&A, 658, A181
Xu, H., Chen, S., Guo, Y., et al. 2023, RAA, 23, 075024
Xue, X., Bian, L., Shu, J., et al. 2021, PhRvL, 127, 251303
Yan, W. M., Manchester, R. N., Hobbs, G., et al. 2011, Ap&SS, 335, 485
Yardley, D. R. B., Hobbs, G. B., Jenet, F. A., et al. 2010, MNRAS, 407, 669
Zhu, X. J., Hobbs, G., Wen, L., et al. 2014, MNRAS, 444, 3709
Zic, A., Reardon, D. J., Kapur, A., et al. 2023, PASA, 40, e049