



# Timing stability of millisecond pulsars and prospects for gravitational-wave detection

J. P. W. Verbiest,<sup>1,2,3\*</sup> M. Bailes,<sup>1</sup> W. A. Coles,<sup>4</sup> G. B. Hobbs,<sup>2</sup> W. van Straten,<sup>1</sup> D. J. Champion,<sup>2</sup> F. A. Jenet,<sup>5</sup> R. N. Manchester,<sup>2</sup> N. D. R. Bhat,<sup>1</sup> J. M. Sarkissian,<sup>2</sup> D. Yardley,<sup>2,6</sup> S. Burke-Spolaor,<sup>1,2</sup> A. W. Hotan<sup>7</sup> and X. P. You<sup>8</sup>

<sup>1</sup>Swinburne University of Technology, Centre for Astrophysics and Supercomputing Mail H39, PO Box 218, VIC 3122, Australia

<sup>2</sup>Australia Telescope National Facility – CSIRO, PO Box 76, Epping, NSW 1710, Australia

<sup>3</sup>Department of Physics, West Virginia University, PO Box 6315, WV 26506, USA

<sup>4</sup>Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, USA

<sup>5</sup>CGWA, University of Texas, Brownsville, TX 78520, USA

<sup>6</sup>Sydney Institute for Astronomy, School of Physics A29, The University of Sydney, NSW 2006, Australia

<sup>7</sup>Curtin Institute for Radio Astronomy, Curtin University of Technology, Bentley, WA 6102, Australia

<sup>8</sup>School of Physical Science and Technology, Southwest University, 2 Tiansheng Road, Chongqing 400715, China

Accepted 2009 August 6. Received 2009 August 3; in original form 2009 June 19

## ABSTRACT

The analysis of high-precision timing observations of an array of  $\sim 20$  millisecond pulsars (a so-called ‘timing array’) may ultimately result in the detection of a stochastic gravitational-wave background. The feasibility of such a detection and the required duration of this type of experiment are determined by the achievable rms of the timing residuals and the timing stability of the pulsars involved. We present results of the first long-term, high-precision timing campaign on a large sample of millisecond pulsars used in gravitational-wave detection projects. We show that the timing residuals of most pulsars in our sample do not contain significant low-frequency noise that could limit the use of these pulsars for decade-long gravitational-wave detection efforts. For our most precisely timed pulsars, intrinsic instabilities of the pulsars or the observing system are shown to contribute to timing irregularities on a 5-year time-scale below the 100 ns level. Based on those results, realistic sensitivity curves for planned and ongoing timing array efforts are determined. We conclude that prospects for detection of a gravitational-wave background through pulsar timing array efforts within 5 years to a decade are good.

**Key words:** gravitational waves – pulsars: general.

## 1 INTRODUCTION

The rotational behaviour of pulsars has long been known to be predictable, especially in the case of millisecond pulsars (MSPs). Current models suggest that such pulsars have been spun up by accretion from their binary companion star to periods of several milliseconds, making them spin much faster than the more numerous younger pulsars, which typically have periods of about a second. The rotational stability of MSPs is generally three to four orders of magnitude better than that of normal pulsars, and on time-scales of several years it has been shown that some MSPs have a timing stability comparable to the most precise atomic clocks (Matsakis, Taylor & Eubanks 1997). This timing stability is most

clearly quantified through the technique of pulsar timing, which compares arrival times of pulses to a model describing the pulsar, its binary orbit and the interstellar medium (ISM) between the pulsar and the Earth (as detailed by Edwards, Hobbs & Manchester 2006). This technique has enabled determination of physical parameters at outstanding levels of precision, such as the orbital characteristics of binary star systems (e.g. van Straten et al. 2001), the masses of pulsars and their companions (e.g. Jacoby et al. 2005; Nice 2006) and the turbulent character of the ISM (e.g. You et al. 2007). The strong gravitational fields of pulsars in binary systems have also enabled stringent tests of general relativity (GR) and alternative theories of gravity, as described by, for example, Kramer et al. (2006) and Bhat, Bailes & Verbiest (2008). Finally, pulsars have provided the first evidence that gravitational waves (GWs) exist at levels predicted by GR (Taylor & Weisberg 1982) and have placed the strongest limit yet on the existence of a background of GWs in the Galaxy (Jenet et al. 2006).

\*E-mail: Joris.Verbiest@mail.wvu.edu

Sazhin (1978) was the first to investigate the potential effect of GWs on the times-of-arrival (TOAs) of pulsar pulses and to conclude that direct detection of GWs could be possible through pulsar timing. Subsequent analyses and theoretical predictions for astronomical sources of GWs have determined that a stochastic gravitational-wave background (GWB) from binary black holes in the cores of galaxies is the most likely signal to be detectable. As summarized in Jenet et al. (2006), the energy density of such a GWB per unit logarithmic frequency interval can be expressed as

$$\Omega_{\text{gw}}(f) = \frac{2}{3} \frac{\pi^2}{H_0^2} A^2 \frac{f^{2\alpha+2}}{f_{\text{ref}}^{2\alpha}}, \quad (1)$$

where  $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$  is the Hubble constant,  $f$  is the GW frequency,  $f_{\text{ref}} = (1 \text{ yr})^{-1}$ ,  $A$  is the dimensionless amplitude of the GWB and  $\alpha$  is the spectral index of the GWB. The one-sided power spectrum of the effect of such a GWB on pulsar timing residuals is given by

$$P(f) = \frac{1}{12\pi^2 f^3} h_c(f)^2, \quad (2)$$

where  $h_c$  is the characteristic strain spectrum, defined as

$$h_c(f) = A \left( \frac{f}{f_{\text{ref}}} \right)^\alpha. \quad (3)$$

Jenet et al. (2006) also summarized the characteristics and expected ranges for various GWBs of interest. Most importantly, for a GWB created by supermassive black hole mergers,  $\alpha = -2/3$  and  $A$  is predicted to be between  $10^{-15}$  and  $10^{-14}$  (Jaffe & Backer 2003; Wyithe & Loeb 2003). Sesana, Vecchio & Colacino (2008) expanded upon these analyses and showed that the actual GWB spectrum strongly depends on the merger history, with a variety of spectral indices possible. They concluded, however, that  $\alpha = -2/3$  was a reasonable approximation for practical purposes. For a background of GWs that were formed in the early Universe,  $\alpha \approx -1$  and the amplitude range predicted by Grishchuk (2005) is  $A = 10^{-17} - 10^{-15}$ , but standard models (e.g. Boyle & Buonanno 2008) predict much lower amplitudes. A third GWB that may be detected by pulsar timing arrays (PTAs) is formed by cosmic strings (Damour & Vilenkin 2005; Caldwell, Battye & Shellard 1996), with predicted amplitudes between  $10^{-16}$  and  $10^{-14}$ , and  $\alpha = -7/6$  (Maggiore 2000).

Hellings & Downs (1983) first investigated the correlations that arise between timing data of different pulsars due to the presence of a stochastic and isotropic GWB in the Galaxy. They demonstrated that the GWs cause a quadrupolar correlation between the timing residuals of different pulsars. Romani (1989) and Foster & Backer (1990) expanded this analysis and introduced the concept of a PTA, in which an ensemble of pulsars is timed and their residuals correlated with each other. The PTA concept uses the quadrupolar correlation signature first derived by Hellings & Downs (1983) to separate the effect of a GW from all other contributions to the residuals, such as intrinsic pulsar timing irregularities, clock errors, ISM effects and Solar system ephemeris errors. Alternatively, the correlation signature for non-Einsteinian GWs (as recently derived by Lee, Jenet & Price 2008) could be used.

The PTA concept was more rigorously explored by Jenet et al. (2005) who first determined the sensitivity of PTA experiments to backgrounds of GWs (equation 12 of Jenet et al. 2005). Their analysis showed that the sensitivity of a PTA depends on four main parameters: the number of pulsars, the data span ( $T$ ), the rms of the timing residuals ( $\sigma$  henceforth) and the number of observations in each of the pulsar timing data sets ( $N_{\text{TOA}}$ ). They further determined

that, for a PTA consisting of weekly observations of 20 MSPs, all with a timing rms of 100 ns, a 5-year observational campaign would be required to make an  $\sim 3\sigma$  detection of a GWB with  $\alpha = -2/3$  and  $A = 10^{-15}$ . It follows from equation (12) of Jenet et al. (2005) that the lowest amplitude of a GWB from supermassive black hole mergers to which a PTA is sensitive, scales as

$$A_{\min, \text{GWB}} \propto \frac{\sigma}{T^{5/3} \sqrt{N_{\text{TOA}}}}. \quad (4)$$

Depending on the achievable rms residual of MSPs, an alternative PTA could therefore achieve the same results through timing of 20 MSPs on a biweekly basis for 10 years with an rms of close to 300 ns. This raises two questions related to the potential of PTAs to detect a GWB. First, down to which precision can MSPs be timed ( $\sigma_{\min}$ ), and secondly can a low residual rms be maintained over long campaigns (i.e. does  $\sigma/T^{5/3}$  decrease with time)? In the context of this second question, we will henceforth use the term ‘timing stability’ when referring to the potential of an MSP timing data set to maintain a constant, preferably low rms residual at all time-scales up to the time-span of a PTA project, which is typically envisaged to be 5 years or longer.

It has been shown for a few pulsars that timing with a residual rms of a few hundred nanoseconds is possible for campaigns lasting a few years. Specifically, Hotan, Bailes & Ord (2006) presented a timing rms of 200 ns over 2 years of timing on PSRs J1713+0747 and J1939+2134 (PSR B1937+21) and 300 ns over 2 years of timing on PSR J1909–3744; Splaver et al. (2005) reported an rms of 180 ns on 6 years of timing PSR J1713+0747, and Verbiest et al. (2008) timed PSR J0437–4715 at 200 ns over 10 years. Similar results for PSRs J0437–4715 and J1939+2134 were obtained by Hobbs et al. (2009) over 5 years of timing. It has, however, not been demonstrated thus far that MSPs can be timed with an rms residual of  $\leq 100$  ns over 5 years or more.

The second question – whether a low rms residual can be maintained over 10 years or longer – also remains unanswered. Kaspi, Taylor & Ryba (1994) detected excess low-frequency noise in PSR J1939+2134; Splaver et al. (2005) presented apparent instabilities in long-term timing of PSR J1713+0747 and Verbiest et al. (2008) noted correlations in the timing residuals of PSR J0437–4715, but apart from these no long-term timing of MSPs has been presented to date. Given the low rms residual reported on all three sources, it is unclear how strongly the reported non-Gaussian noise would affect the use of these pulsars in a GWB-detection effort.

In this article, we present the first high-precision long-term timing results for a sample of 20 MSPs. The source selection, observing systems and data analysis methods are described in Section 2. Our updated timing models and residual plots for all pulsars in our sample are also presented in Section 2, allowing the reader a fundamental inspection of the reliability of our timing results. In Section 3, we perform a stability analysis of the timing data, with the dual purpose of identifying low-frequency noise in any of our timing data and of assessing the potential impact of such noise on the use of pulsars in a timing array. In Section 4, we outline a new way of quantifying different components of the pulsar timing rms. Through this analysis, we separate the levels of receiver noise, noise with a dependency on observing frequency and any temporal instabilities, providing a bound on the residual rms that might be achievable on a 5-year time-scale. We apply this analysis to three of our most precisely timed pulsars. In Section 5, we calculate sensitivity curves for ongoing and planned PTAs. These sensitivity curves take into account the inhomogeneous character of a realistic array (i.e. the fact that the rms will differ between pulsars) and assume a bound on

residual rms as determined in Section 4. In Section 6, we summarize our findings.

## 2 OBSERVATIONS AND DATA REDUCTION

### 2.1 Sample selection

The data presented in this paper have been collated from two pulsar timing programmes at the Parkes radio telescope. The oldest of these commenced during the Parkes 70-cm MSP survey (Bailes et al. 1994), aiming to properly characterize the astrometric and binary parameters of the MSPs found in the survey. Initial timing results from this campaign were published by Bell et al. (1997) and Toscano et al. (1999). The bright millisecond pulsars PSRs J1713+0747 and B1937+21 (both discovered earlier at Arecibo) were also included in this programme. A few years later, as new discoveries were made in the Swinburne intermediate-latitude survey (Edwards et al. 2001), these pulsars were also added, resulting in a total of 16 MSPs that were regularly timed by 2006. Improved timing solutions for these 16 pulsars were presented by Hotan et al. (2006) and Ord et al. (2006).

Besides the projects described above, the Parkes Pulsar Timing Array (PPTA; Manchester 2008) project commenced more regular timing observations of these pulsars in late 2004, expanding the number of MSPs to 20 (listed in Table 1) and adding regular monitoring at a low observing frequency (685 MHz) to allow correction for variations of the ISM electron density. A detailed analysis of these low-frequency observations and ISM effects was recently presented by You et al. (2007), and an analysis of the combined data on PSR J0437–4715 was published by Verbiest et al.

(2008). For this pulsar, we will use the timing results presented in that publication; for all other pulsars, we will present our improved timing models in Section 2.4.

### 2.2 Observing systems

Unless otherwise stated, the data presented were obtained at the Parkes 64-m radio telescope, at a wavelength of 20 cm. Two receivers were used: the H-OH receiver and the 20 cm multibeam receiver (Staveley-Smith et al. 1996). Over the last 5 years, observations at 685 MHz were taken with the 10/50 cm coaxial receiver for all pulsars; however, they were only used directly in the final timing analysis of PSR J0613–0200, whose profile displays a sharp spike at this frequency if coherent dedispersion is applied. For PSRs J1045–4509, J1909–3744 and J1939+2134, the 685 MHz observations were used to model and remove the effects of temporal variations in interstellar dispersion delays, and hence are included indirectly in the timing analysis. For all other pulsars, any such variations were below the level of our sensitivity.

Three different observing systems were used. First, the Caltech Fast Pulsar Timing Machine (FPTM; Sandhu et al. 1997; Sandhu 2001), between 1994 and 2001 November. This is an autocorrelation spectrometer with a total bandwidth of up to 256 MHz. Secondly, the 256 MHz bandwidth analogue filterbank (FB) was used in 2002 and 2003. Finally, the Caltech–Parkes–Swinburne Recorder 2 (CPSR2; Hotan et al. 2006) was used from 2002 November onwards. The CPSR2 is a baseband data recorder with two 64 MHz bandwidth observing bands (one usually centred at 1341 MHz, the other at 1405 MHz) and phase-coherent dispersion removal occurring in near real time.

**Table 1.** Pulsars in our sample.

Pulsar name	Discovery	Previous timing solution <sup>a</sup>	Pulse period (ms)	Orbital period (d)	Dispersion measure (cm <sup>-3</sup> pc)
J0437–4715	Johnston et al. (1993)	1, 2	5.8	5.7	2.6
J0613–0200	Lorimer et al. (1995)	3	3.1	1.2	38.8
J0711–6830	Bailes et al. (1997)	3, 4	5.5	–	18.4
J1022+1001	Camilo et al. (1996)	3	16.5	7.8	10.3
J1024–0719	Bailes et al. (1997)	3	5.2	–	6.5
J1045–4509	Bailes et al. (1994)	3	7.5	4.1	58.2
J1600–3053	Ord et al. (2006)	5	3.6	14.3	52.3
J1603–7202	Lorimer et al. (1996)	3	14.8	6.3	38.0
J1643–1224	Lorimer et al. (1995)	4	4.6	147.0	62.4
J1713+0747	Foster, Wolszczan & Camilo (1993)	3, 6	4.6	67.8	16.0
J1730–2304	Lorimer et al. (1995)	4	8.1	–	9.6
J1732–5049	Edwards & Bailes (2001)	7	5.3	5.3	56.8
J1744–1134	Bailes et al. (1997)	3	4.1	–	3.1
B1821–24; J1824–2452	Lyne et al. (1987)	8, 10	3.1	–	120.5
B1855+09; J1857+0943	Segelstein et al. (1986)	3, 9	5.4	12.3	13.3
J1909–3744	Jacoby et al. (2003)	3, 11	2.9	1.5	10.4
B1937+21; J1939+2134	Backer et al. (1982)	3, 9	1.6	–	71.0
J2124–3358	Bailes et al. (1997)	3	4.9	–	4.6
J2129–5721	Lorimer et al. (1996)	3	3.7	6.6	31.9
J2145–0750	Bailes et al. (1994)	3, 12	16.1	6.8	9.0

*Note.* Column 2 gives the reference for the discovery paper, while Column 3 provides references to recent or important publications on timing of the sources. For the three pulsars with original B1950 names, these names are given besides the J2000.0 names.

<sup>a</sup>References: (1) Verbiest et al. (2008); (2) van Straten et al. (2001); (3) Hotan et al. (2006); (4) Toscano et al. (1999); (5) Ord et al. (2006); (6) Splaver et al. (2005); (7) Edwards & Bailes (2001); (8) Hobbs et al. (2004); (9) Kaspi et al. (1994); (10) Cognard & Backer (2004); (11) Jacoby et al. (2005); (12) Löhmer et al. (2004).

### 2.3 Arrival time determination

The processing applied differs for data from different observing systems. The FPTM data were calibrated using a real-time system to produce either two or four Stokes parameters which were later combined into Stokes I. The FB data were produced from a search system with no polarimetric calibration possible. This system produced Stokes I profiles after folding 1-bit data. Data from both of these systems were integrated in frequency and time to produce a single profile for each observation. These observations were  $\sim 25$  min in duration. For CPSR2 data, in order to minimize the effects of aliasing and spectral leakage, 12.5 per cent of each edge of the bandpass was removed. To remove the worst radio frequency interference, any frequency channel with power more than  $4\sigma$  in excess of the local median was also removed ('local' was defined as the nearest 21 channels and the standard deviation  $\sigma$  was determined iteratively). The CPSR2 also operated a total power monitor on microsecond time-scales, which removed most impulsive interference.

The CPSR2 data were next integrated for 5 min and calibrated for differential gain and phase to correct for possible asymmetries in the receiver hardware, if calibrator observations were available. (Especially in the years directly following the CPSR2 commissioning, observations of a pulsating noise source, needed for polarimetric calibration, were not part of the standard observing schedule.) Subsequently the data were integrated for the duration of the observation, which was typically 32 min for PSRs J2124–3358, J1939+2134 and J1857+0943 and 64 min for all other pulsars. In the case of PSR J1643–1224, the integration time was 32 min until 2005 December and 64 min from 2006 onwards. Finally, the CPSR2 data were integrated in frequency and the Stokes parameters were combined into total power. CPSR2 data that did not have calibrator observations available were processed identically, except for the calibration step. While for some pulsars (like PSR J0437–4715) these uncalibrated data are provably of inferior quality (see e.g. van Straten 2006), in our case this is largely outweighed by the improved statistics of the larger number of TOAs and by the extended timing baseline these observations provided. We therefore include both calibrated and uncalibrated observations in our data sets.

Pulse TOAs were determined through cross-correlation of the total intensity profiles thus obtained with pulsar and frequency-dependent template profiles. These template profiles were created through addition of a large number of observations and were phase-aligned for both CPSR2 observing bands. As there were only few high signal-to-noise ratio (S/N) observations obtained with the FPTM and FB backends for most pulsars, these data were timed against standards created with the CPSR2 backend. This may affect the reliability of their derived TOA errors. For this reason, we have evaluated the underestimation of TOA errors for each backend separately, as explained in the next section. While the TOA errors were generally determined through the standard Fourier phase gradient method, the Gaussian interpolation method produced more accurate estimates for pulsars with low S/Ns (Hotan, Bailes & Ord 2005) – specifically for PSRs J0613–0200, J2129–5721, J1732–5049, J2124–3358 and J1045–4509. The PSRCHIVE software package (Hotan, van Straten & Manchester 2004) was used to perform all of the processing described above.

### 2.4 Timing results

The TEMPO2 software package (Hobbs, Edwards & Manchester 2006) was used to calculate the residuals from the TOAs and initial

timing solutions (Table 1). In order to account for the unknown instrumental delays and pulsar-dependent differences in observing setup, arbitrary phase offsets were introduced between the data from different backends. Where available, data at an observing frequency of 685 MHz were included in an initial fit to inspect visually for the presence of dispersion measure (DM) variations. In the case of PSRs J1045–4509, J1909–3744, J1939+2134 and J0437–4715, such variations were significant and dealt with in the timing software through a method similar to that presented by You *et al.* (2007). The average DM values presented in Tables 3–5 were determined from the 20 cm data exclusively. The uncertainties in these DM values do not take into account possible pulse shape differences between the profiles at these slightly varying frequencies. We updated all the pulsar ephemerides to use International Atomic Time [implemented as TT(TAI) in TEMPO2] and the DE405 Solar system ephemeris (Standish 2004).

In order to correct for any underestimation of TOA uncertainties resulting from (amongst others) the application of CPSR2-based template profiles to the FB and FPTM data (as mentioned in Section 2.3) and to allow comparison of our timing model parameters to those published elsewhere, the TOA uncertainties were multiplied by error factors (so-called 'EFACs') that are dependent on the pulsar and observing system. Specifically, this part of the analysis was performed as follows. First, the timing data from each observing system were pre-whitened by fitting harmonically related sine/cosine pairs if required. Next, the TOA uncertainties were multiplied by an EFAC value that produced a reduced  $\chi^2$  value of unity for that pre-whitened subset of the data. Because of the potential non-Gaussian noise in the data, application of these backend-specific EFACs does not necessarily result in a reduced  $\chi^2$  value of unity for the entire, recombined data set. To account for such non-Gaussian noise in the data, a 'global' EFAC was applied to the entire data set, making the reduced  $\chi^2$  after pre-whitening equal to unity and increasing the parameter uncertainties reported in the timing models appropriately. As mentioned, the pre-whitening method was based on fitting of sine/cosine pairs to the data, according to the following formula described by Martin (2001) (and replicated in Hobbs *et al.* 2006):

$$\Delta R = \sum_{k=1}^{n_H} A_{\sin,k} \sin(k\omega_{\text{pw}}\Delta t) + A_{\cos,k} \cos(k\omega_{\text{pw}}\Delta t),$$

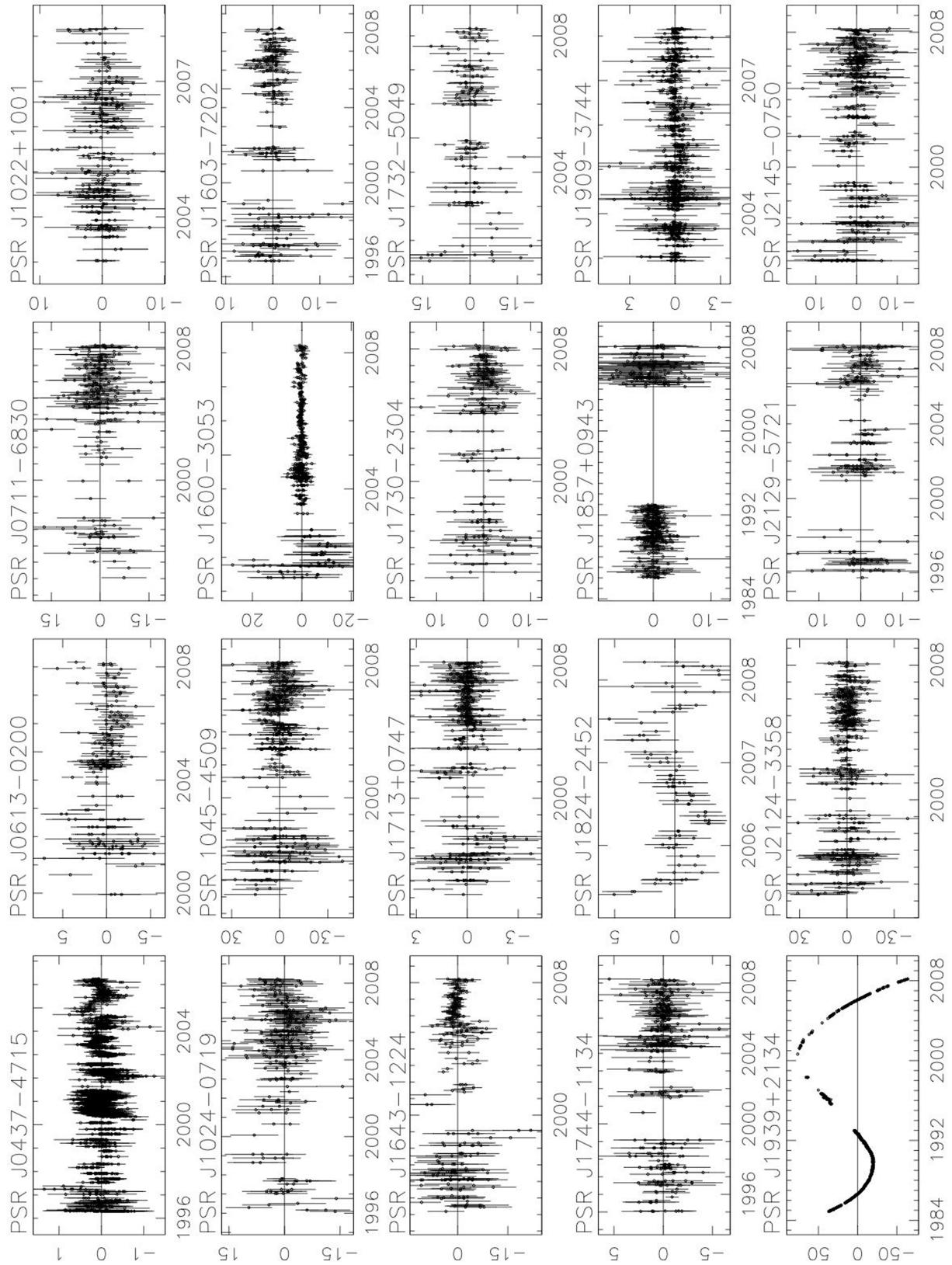
where  $k$  runs over all sine/cosine pairs,  $n_H$  is the total number of harmonically related pairs fitted,  $A_{\cos,k}$  and  $A_{\sin,k}$  are, respectively, the amplitude of the  $k$  th cosine and sine waves and  $\omega_{\text{pw}}$  is the fundamental frequency derived from

$$\omega_{\text{pw}} = \frac{2\pi}{T(1 + 4/n_H)},$$

with  $T$  being the length of the data set. If pre-whitening terms were included in the final fit, we provide the values for  $\omega_{\text{pw}}$ ,  $A_{\cos,k}$  and  $A_{\sin,k}$  as part of our timing model.

Because the potential non-Gaussian noise present in the data is the subject of our investigations in the remaining sections of this paper, the pre-whitening terms as well as the global EFACs were not included in any subsequent analysis. The residuals plotted in Fig. 1 and the parameters presented in Table 2 therefore do not include pre-whitening terms or global EFACs.

The system-specific EFACs were generally less than two, with the only major outliers being the CPSR2 data of PSR J1939+2134 with an EFAC of 5.27 and the 32-min CPSR2 integrations (pre-2004 CPSR2 data) of PSR J1643–1224, which have an EFAC of 4.9. In the former case, this large EFAC may be caused by incomplete pre-whitening, as the non-Gaussian noise is badly modelled



**Figure 1.** Timing residuals of the 20 pulsars in our sample. Scaling on the x-axis is in years and on the y-axis in  $\mu\text{s}$ . For PSRs J1857+0943 and J1939+2134, these plots include the Arecibo data made publicly available by Kaspi et al. (1994); all other data are from the Parkes telescope, as described in Section 2. Sudden changes in white noise levels are due to changes in the pulsar backend setup (see Section 2 for more details).

**Table 2.** Summary of the timing results, sorted in the order of decreasing rms residual.

Pulsar name	rms ( $\mu\text{s}$ )	T (yr)	$N_{\text{TOA}}$
J1909–3744	0.166	5.2	893
J1713+0747	0.198	14.0	392
J0437–4715	0.199	9.9	2847
J1744–1134	0.617	13.2	342
J1939+2134	0.679	12.5	168
J1600–3053	1.12	6.8	477
J0613–0200	1.52	8.2	190
J1824–2452	1.63	2.8	89
J1022+1001	1.63	5.1	260
J2145–0750	1.88	13.8	377
J1643–1224	1.94	14.0	241
J1603–7202	1.98	12.4	212
J2129–5721	2.20	12.5	179
J1730–2304	2.52	14.0	180
J1857+0943	2.92	3.9	106
J1732–5049	3.23	6.8	129
J0711–6830	3.24	14.2	227
J2124–3358	4.01	13.8	416
J1024–0719	4.17	12.1	269
J1045–4509	6.70	14.1	401

*Note.* The columns present the pulsar name, the rms timing residual (without pre-whitening), the length of the data set and the number of TOAs. For PSRs J1939+2134 and J1857+0943, this table only contains the Parkes data (see Sections 2.4 and 3 for details).

by polynomials or sine/cosine pairs. The underestimation of PSR J1643–1224 TOA uncertainties is likely caused by the low S/N of these observations, which causes the Fourier phase gradient method to underestimate TOA errors (as previously reported by Hotan et al. 2005). We note that the EFAC for the 64 min integrations is much lower, at 2.5. In deriving the timing models, the global EFAC was at most 1.1 and for more than half of our sources less than 1.05.

The fact that most of our EFAC values are close to unity and show little variation with backend suggests that the parameter and error estimates are fairly robust. In order to account for the different sensitivity of the backends used and to limit effects of scintillation on our timing, we opt for a weighted analysis. It is therefore important to consider the impact of the TOA errors and applied EFACs on the different parts of this analysis. Given that for most pulsars the EFACs applied to the different backends are nearly equal, the resulting timing models will be little affected by these EFAC values. The reported uncertainties on the timing model parameters will be affected but will be comparable to previous publications, since our analysis method is similar. A full error analysis (as suggested by Verbiest et al. 2008) is needed to provide any more reliable parameter uncertainties. Since the focus of this paper is on the overall timing stability and implications for PTA science, we defer such error analysis (and the interpretation of any previously unpublished parameters in our timing models) to a later paper. We have, however, investigated the effect of weighting and EFACs on the timing stability analysis (Section 3), but have not uncovered any unexpected deviations beyond statistical noise. We therefore conclude that the weighting and applied EFACs do not invalidate our analysis.

A summary of the lengths of the data sets and the achieved rms residual can be found in Table 2, highlighting the superior residual

rms of PSRs J1909–3744, J0437–4715 and J1713+0747 when compared to other pulsars. The timing residuals for our data sets are presented in Fig. 1 and the timing models are presented in Tables 3–5, where  $2\sigma$  errors are given, in accordance with previous practice. We encourage observers to use the improved models when observing. We also note that all but a few of the parameters in our timing models are consistent with those published previously.

### 3 PULSAR TIMING STABILITY

In Section 1, we demonstrated that one of two vital questions relating to the potential of PTAs to detect a GWB is whether a low residual rms can be maintained over long time-spans (a property we refer to as ‘timing stability’). Effectively, this question breaks down into two parts: to what degree of significance low-frequency noise is present in our pulsar timing data and how any such low-frequency noise can be expected to affect sensitivity to a GWB. In order to answer this question fully, a spectral-analysis-based investigation of pulsar timing residuals, that includes identification and modelling of potential non-Gaussian noise sources, would be required. Because of various pulsar timing-specific issues such as clustering of data, large gaps in data sampling and large variations in error-bar size, however, standard spectral analysis methods fail to provide reliable power spectra of pulsar timing data. We therefore use the alternative approach provided by the  $\sigma_z$  statistic, as described by Matsakis et al. (1997). A brief explanation of this statistic along with a presentation of the  $\sigma_z$  values of our data is presented in Section 3.1, and a discussion of these results in terms of PTA science is provided in Section 3.2.

#### 3.1 $\sigma_z$ stability analysis

Originally proposed by Matsakis et al. (1997), the  $\sigma_z$  statistic is defined as

$$\sigma_z(\tau) = \frac{\tau^2}{2\sqrt{5}} \langle c_3^2 \rangle^{1/2},$$

where  $\langle \rangle$  denotes the average over subsets of the data,  $c_3$  is determined from a fit of the polynomial

$$c_0 + c_1(t - t_0) + c_2(t - t_0)^2 + c_3(t - t_0)^3$$

to the timing residuals for each subset and  $\tau$  is the length of the subsets of the data. In order for the  $\sigma_z$  values to be independent of each other, we use  $\tau = T, T/2, T/4, T/8, \dots$  only. The interpretation of this statistic in terms of power spectra deserves some attention. As presented by Matsakis et al. (1997), a power spectrum with spectral index  $\beta$

$$P(v) \propto f^\beta$$

would translate into a  $\sigma_z$  curve

$$\sigma_z(\tau) \propto \tau^\mu,$$

where the spectral indices are related as

$$\mu = \begin{cases} -\frac{1}{2}(\beta + 3) & \text{if } \beta < 1 \\ -2 & \text{otherwise.} \end{cases} \quad (5)$$

Equation (5) implies that spectra have different slopes in a  $\sigma_z$  graph than in a power spectrum. Along with the  $\sigma_z$  graphs for our data sets, Fig. 2 provides some examples of spectra for guidance: lines with a slope of  $-3/2$  (dotted lines in Fig. 2) represent spectrally white data ( $\beta = 0$  into equation 5 gives  $\mu = -3/2$ ) and a GWB with a spectral index  $\alpha = -2/3$  in the gravitational strain spectrum

**Table 3.** Timing parameters for the single pulsars PSRs J0711–6830, J1024–0719, J1730–2304, J1744–1134, J1824–2452, J1939+2134 and J2124–3358. Numbers in brackets give twice the formal standard deviation in the last digit quoted. Note that these parameters are determined using TEMPO2, which uses the International Celestial Reference System and Barycentric Coordinate Time. As a result, this timing model must be modified before being used with an observing system that inputs TEMPO format parameters (see Hobbs et al. 2006 for more information).

	Fit and data set parameters			
Pulsar name	J0711–6830	J1024–0719	J1730–2304	J1744–1134
MJD range	49 373.6–54 546.4	50 117.5–54 544.6	49 421.9–54 544.8	49 729.1–54 546.9
Number of TOAs	227	269	180	342
rms timing residual ( $\mu\text{s}$ )	3.24	3.80	2.52	0.617
Reference epoch for $P, \alpha, \delta$ and DM determination	49 800	53 000	53 300	53 742
	Measured quantities			
Right ascension, $\alpha$ (J2000.0)	07:11:54.22579(15)	10:24:38.68846(3)	17:30:21.6611(3)	17:44:29.403209(4)
Declination, $\delta$ (J2000.0)	−68:30:47.5989(7)	−07:19:19.1700(10)	−23:04:31.29(8)	−11:34:54.6606(2)
Proper motion in $\alpha, \mu_\alpha \cos \delta$ (mas yr <sup>−1</sup> )	−15.55(8)	−35.3(2)	20.27(6)	18.804(15)
Proper motion in $\delta, \mu_\delta$ (mas yr <sup>−1</sup> )	14.23(7)	−48.2(3)	—	−9.40(6)
Annual parallax, $\pi$ (mas)	—	—	—	2.4(2)
DM (cm <sup>−3</sup> pc)	18.408(4)	6.486(3)	9.617(2)	3.1380(6)
Pulse frequency, $\nu$ (Hz)	182.117 234 869 347(4)	193.715 683 568 44(13)	123.110 287 192 301(2)	245.426 119 748 3027(5)
Pulse frequency derivative, $\dot{\nu}$ (10 <sup>−16</sup> s <sup>−2</sup> )	−4.944 06(15)	−6.95(3)	−3.059 06(10)	−5.381 88(4)
	Pre-whitening terms			
Fundamental wave frequency, $\omega_{\text{pw}}$ (yr <sup>−1</sup> )	—	0.103 68	—	—
Amplitude of wave 1 cosine and sine terms, $A_{\cos,1}; A_{\sin,1}$ (10 <sup>−4</sup> s)	—	2(13); 4.7(21)	—	—
	Fit and data set parameters			
Pulsar name	J1824–2452	J1939+2134	J2124–3358	
MJD range	53 518.8–54 544.9	46 024.8–54 526.9	49 489.9–54 528.9	
Number of TOAs	89	180	416	
rms timing residual ( $\mu\text{s}$ )	0.986	0.354	4.03	
Reference epoch for $P, \alpha, \delta$ and DM determination	54 219	52 601	53 174	
	Measured quantities			
Right ascension, $\alpha$ (J2000.0)	18:24:32.00796(2)	19:39:38.561297(2)	21:24:43.85347(3)	
Declination, $\delta$ (J2000.0)	−24:52:10.824(6)	+21:34:59.12950(4)	−33:58:44.6667(7)	
Proper motion in $\alpha, \mu_\alpha \cos \delta$ (mas yr <sup>−1</sup> )	—	0.072(2)	−14.12(13)	
Proper motion in $\delta, \mu_\delta$ (mas yr <sup>−1</sup> )	−9(5)	−0.415(3)	−50.34(25)	
Annual parallax, $\pi$ (mas)	—	0.13(13)	3.1(11)	
DM (cm <sup>−3</sup> pc)	120.502(2)	71.0227(9)	4.601(3)	
Pulse frequency, $\nu$ (Hz)	327.405 594 6921(6)	641.928 233 642(12)	202.793 893 879 496(2)	
Pulse frequency derivative, $\dot{\nu}$ (10 <sup>−16</sup> s <sup>−2</sup> )	−1736.5(3)	−429.1(6)	−8.4597(2)	
	Pre-whitening terms			
Fundamental wave frequency, $\omega_{\text{pw}}$ (yr <sup>−1</sup> )	0.447 34	0.149 96	—	
Amplitudes of cosine and sine terms (10 <sup>−4</sup> s):				
Wave 1: $A_{\cos,1}; A_{\sin,1}$	−20(6); 2.1(14)	286(41); −413(60)	—	
Wave 2: $A_{\cos,2}; A_{\sin,2}$	—	30(5); 84(12)	—	
Wave 3: $A_{\cos,3}; A_{\sin,3}$	—	−21(3); −5.8(9)	—	
Wave 4: $A_{\cos,4}; A_{\sin,4}$	—	3.7(5); −2.9(5)	—	
Wave 5: $A_{\cos,5}; A_{\sin,5}$	—	0.04(3); 0.68(9)	—	

(and therefore a spectral slope  $\beta = -13/3$  in the timing residual spectrum, as follows from equations 2 and 3) would have a positive slope of  $2/3$  in  $\sigma_z$  (dashed lines).

Comparison of such theoretical slopes to the data is, however, non-trivial since the data are strongly affected by effects from sam-

pling and fitting. As an illustration of such effects, the top-left plot of Fig. 2 shows two  $\sigma_z$  curves derived from simulations. The first one is the full line that approximates the  $\sigma_z$  curve for PSR J1713+0747. In this case, the  $\sigma_z$  values of 1000 simulations of white noise with the timing rms and sampling of the PSR J1713+0747 data set were

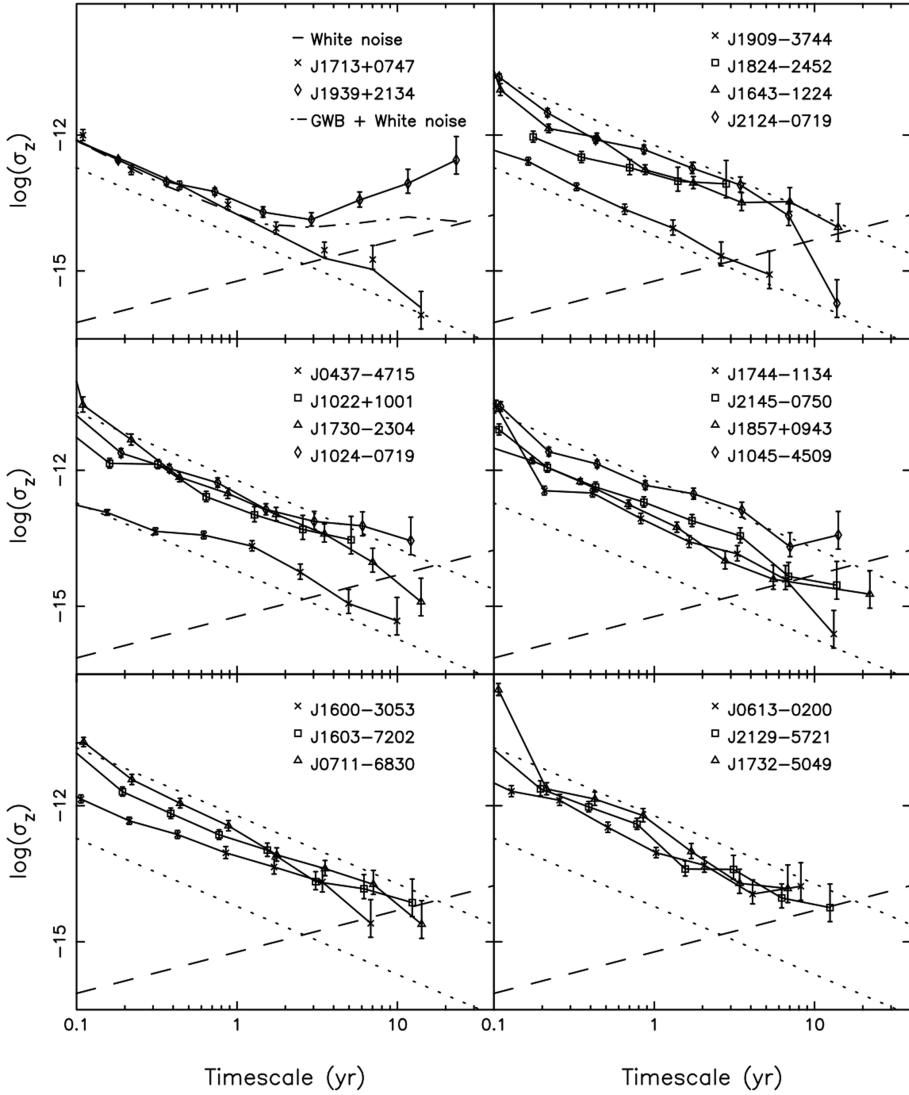
**Table 4.** Timing parameters for binary PSRs J0222+1001, J1600–3053, J1713+0747, J1857+0943, J1909–3744 and J2145–0750 (see caption of Table 3 for more information).

Fit and data set parameters						
Pulsar name	J1022+1001	J1600-3053	J1713+0747	J1857+0943	J1909-3744	J2145-0750
MJD range	52649.7–54528.5	52055.7–54544.6	49421.9–54546.8	46436.7–54526.9	52618.4–54528.8	49517.8–54547.1
Number of TOAs	260	477	392	376	893	377
rms timing residual ( $\mu\text{s}$ )	1.63	1.12	0.198	1.14	0.166	1.88
Reference epoch for $P, \alpha, \delta$ and DM determination	53 589	53 283	54 312	50 481	53 631	53 040
Measured quantities						
Right ascension, $\alpha$ (J2000.0)	10:22:58.003(3)	16:00:51.903798(11)	17:13:49.532628(2)	18:57:36.392909(13)	19:09:47.436612(8)	21:45:50.46412(3)
Declination, $\delta$ (J2000.0)	+10:01:52.76(13)	-30:53:49.3407(5)	+07:47:37.50165(6)	+09:43:17.2754(3)	-37:44:14.38013(3)	-07:50:18.4399(14)
Proper motion in $\alpha, \mu_\alpha$ $\cos \delta$ (mas $\text{yr}^{-1}$ )	-17.02(14)	-1.06(9)	4.92(10)	-2.64(3)	-9.510(7)	-9.66(15)
Proper motion in $\delta, \mu_\delta$ (mas $\text{yr}^{-1}$ )	-	-7.1(3)	-3.85(2)	-5.46(4)	-35.859(19)	-8.9(4)
Annual parallax, $\pi$ (mas)	1.8(6)	0.2(3)	0.94(10)	1.1(4)	0.79(4)	1.6(5)
Annual parallax, $\pi$ (mas)	10.261(2)	52.3262(10)	15.9915(2)	13.3004(4)	10.3934(2)	8.9977(14)
DM ( $\text{cm}^{-3}$ pc)	60.7794479762157(4)	277.9377070984926(17)	218.8118404414362(3)	186.494078620232(2)	339.31568740949071(10)	62.2958878569665(6)
Pulse frequency, $\nu$ (Hz)	-1.6012(2)	-7.3390(5)	-4.08379(3)	-6.20495(6)	-16.14819(5)	-1.15588(3)
Pulse frequency derivative, $\dot{\nu}$ ( $10^{-16} \text{ s}^{-2}$ )	-	-	-	-	-	-
Orbital period, $P_b$ (d)	7.8051302826(4)	14.3484577709(13)	67.825130963(17)	12.32719(4)	1.533449474590(6)	6.83893(2)
Orbital period derivative, $\dot{P}_b$ ( $10^{-13}$ )	-	-	41(20)	3(3)	5.5(3)	4(3)
Epoch of periastron passage, $T_0$ (MJD)	53587.3140(6)	53281.191(4)	54303.6328(7)	50476.095(8)	-	53042.431(3)
Projected semimajor axis, $x = a \sin i$ (s)	16.7654074(4)	8.801652(10)	32.3424236(3)	9.230780(5)	1.89799106(7)	10.1641080(3)
$\dot{x}$ ( $10^{-14}$ )	1.5(10)	-0.4(4)	-	-	-0.05(4)	-0.3(3)
Longitude of periastron, $\omega_0$ ( $^\circ$ )	97.75(3)	181.85(10)	276.5(2)	2170(6)	-	200.6318(1)
Orbital eccentricity, $e$ ( $10^{-5}$ )	9.700(4)	17.369(4)	-	-	-0.4(4)	-13(2)
$\kappa = e \sin \omega_0$ ( $10^{-8}$ )	-	-	-	-	-	-
$\eta = e \cos \omega_0$ ( $10^{-8}$ )	-	-	-	-	-	-
Ascending node passage, $T_{\text{asc}}$ (MJD)	-	-	-	-	5630.723214894(4)	-
Periastron advance, $\omega$ (deg $\text{yr}^{-1}$ )	-	-	-	-	-	0.06(6)
Sine of inclination angle, $\sin i$	0.73 <sup>a</sup>	0.8(4)	-	-	0.9990(7)	-
Inclination angle, $i$ ( $^\circ$ )	47 <sup>a</sup>	-	-	-	78.6(17)	-
Companion mass, $M_c$ ( $\text{M}_\odot$ )	1.05 <sup>a</sup>	0.6(15)	0.20(2)	0.27(3)	0.212(4)	-
Longitude of ascending node, $\Omega$ ( $^\circ$ )	-	-	-	-	67(17)	-

<sup>a</sup>From Hotan et al. (2006).

**Table 5.** Timing parameters for binary PSRs J0613–0200, J1045–4509, J1603–7202, J1643–1224, J1732–5049 and J2129–5721 (see caption of Table 3 for more information).

	Fit and data set parameters		
Pulsar name	J0613–0200	J1045–4509	J1603–7202
MJD range	51526.6–54527.3	49405.5–54544.5	50026.1–54544.7
Number of TOAs	190	401	212
rms timing residual (μs)	1.52	6.70	1.98
Reference epoch for P, $\alpha$ , $\delta$ and DM determination	53 114	53 050	53 024
	Measured quantities		
Right ascension, $\alpha$ (J2000.0)	06:13:43.975142(11)	10:45:50.189516(5)	16:03:35.67980(4)
Declination, $\delta$ (J2000.0)	−02:00:47.1737(4)	−45:09:54.1427(5)	−72:02:32.6985(3)
Proper motion in $\alpha$ , $\mu_\alpha$ cos $\delta$ (mas yr <sup>−1</sup> )	1.84(8)	−6.0(2)	−2.52(6)
Proper motion in $\delta$ , $\mu_\delta$ (mas yr <sup>−1</sup> )	−10.6(2)	5.3(2)	−7.42(9)
Annual parallax, $\pi$ (mas)	0.8(7)	3.3(38)	—
DM (cm <sup>−3</sup> pc)	38.782(4)	58.137(6)	38.060(2)
Pulse frequency, $\nu$ (Hz)	326.600 562 190 182(4)	133.793 149 594 456(2)	67.376 581 140 8911(5)
Pulse frequency derivative, $\dot{\nu}$ (10 <sup>−16</sup> s <sup>−2</sup> )	−10.2308(8)	−3.1613(3)	−0.70952(5)
Orbital period, $P_b$ (d)	1.198 512 5753(1)	4.083 529 2547(9)	6.308 629 6703(7)
Epoch of periastron passage, $T_0$ (MJD)	53113.98(2)	53048.98(2)	—
Projected semimajor axis, $x = a \sin i$ (s)	3.015 1325(10)	6.880 6610(4)	49577.9689(13)
$\dot{x}$ (10 <sup>−14</sup> )	−	1.8(5)	25.072 614(2)
Longitude of periastron, $\omega_0$ (°)	54(6)	242.7(16)	−4.9(5)
Orbital eccentricity, $e$ (10 <sup>−5</sup> )	0.55(6)	2.37(7)	321.850(3)
$\kappa = e \sin \omega_0$ (10 <sup>−6</sup> )	—	—	50.578(4)
$\eta = e \cos \omega_0$ (10 <sup>−6</sup> )	—	—	1.61(14)
Ascending node passage, $T_{\text{asc}}$ (MJD)	—	—	−9.41(13)
	—		
	53309.3307830(1)		
	—		



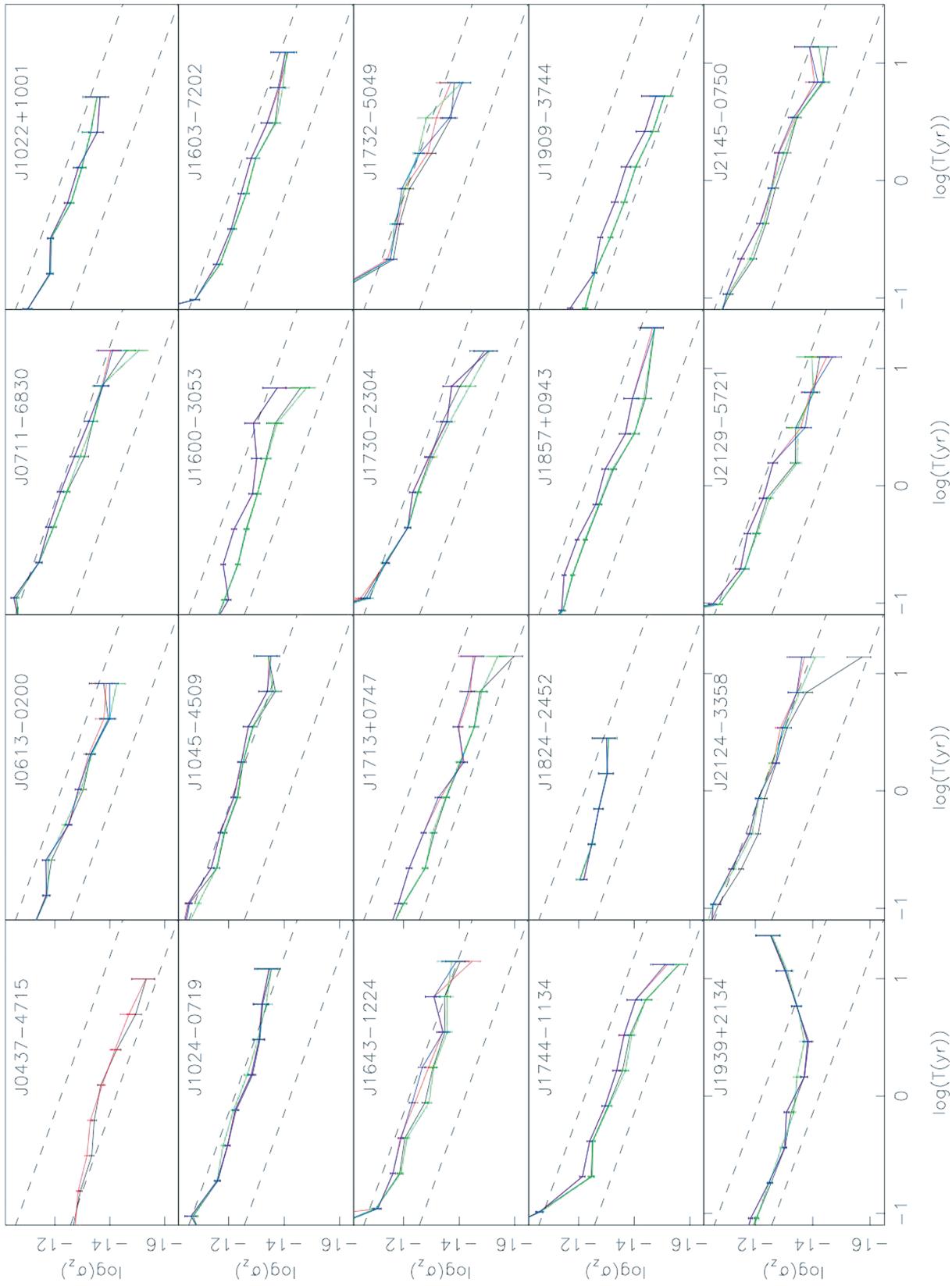
**Figure 2.**  $\sigma_z$  stability parameter for the 20 pulsars in our sample. The dotted slanted lines represent white noise levels of 100 ns (bottom line) and 10  $\mu$ s (top line); the dashed slanted line shows the steepness introduced to pre-fit residuals by a hypothetical GWB (see Section 1); pulsars whose curve is steeper than this line (like PSR J1939+2134) can therefore be expected to be of little use to PTA efforts on long time-scales. The top-left figure further shows the average  $\sigma_z$  values resulting from 1000 simulations of white noise residuals sampled at the times of the PSR J1713+0747 data set and fitted for the PSR J1713+0747 timing model parameters (full line). This demonstrates that the PSR J1713+0747 data do not – within the sensitivity provided by the  $\sigma_z$  statistic – contain a significant, steep red-noise process. The dash-dotted line in the top-left figure shows the average of 2000 simulations for white noise combined with a GWB, sampled at the times of the PSR J1939+2134 data set and fitted for pulse period and period derivative. These simulated results provide an example of the combined effect sampling and model fitting can have on the  $\sigma_z$  statistic, even in the case of white noise.

averaged. This curve is not perfectly parallel to the theoretical curve with slope  $-3/2$  due to sampling, varying TOA uncertainties and model fitting. Comparison of the white noise simulations with the actual PSR J1713+0747 data indicates that there is not a significant, steep red-noise process affecting the timing residuals for this pulsar. The second simulation in the top-left plot of Fig. 2 is the dot-dashed line, which is the average  $\sigma_z$  graph of 2000 simulations of white noise with an artificial GWB and the sampling of the PSR J1939+2134 data set, fitted for pulse period and spin-down. This simulated curve does not reach the theoretical slope of  $2/3$  because of the flattening off at low frequencies caused by sampling, fitting and leakage resulting from these. This simulation also demonstrates that the PSR J1939+2134 curve is significantly steeper than a simulated GWB, implying that this pulsar will most likely not be very useful for long-term PTA projects, although its low rms residual on

short time-spans might make it useful for detection of burst-type sources.

The  $\sigma_z$  graphs of our data are shown in Fig. 2. A comparison of these curves to those obtained from an unweighted analysis or from an analysis that does not contain the EFAC values described in Section 2.4 is presented in Fig. 3, a colour version of which is available online. This graph demonstrates that use of weighting or EFACs does not affect the data in any statistically significant manner, other than to decrease the effect of the white noise component in case of scintillating pulsars.

Comparison of the PSR J1713+0747 and PSR J1939+2134 data with the simulated curves shown, along with the invariability of the shape of  $\sigma_z$  plots to weighting or application of error factors, shows that the  $\sigma_z$  parameter provides a good first-order discrimination between pulsars that do not exhibit significant, steep red noise



**Figure 3.**  $\sigma_z$  graphs for all pulsars in our sample. The curves are for a weighted analysis with EFACs (solid line; black in the online version of this article), unweighted analysis with EFACs (dashed line; red online), weighted analysis without EFACs (dot-dashed line; green online) and unweighted analysis without EFACs (dotted line; blue online). The straight solid lines without error bars represent theoretical white noise at levels of 100 ns (bottom) and 10  $\mu\text{s}$  (top). The EFAC values of the PSR J0437-4715 data were lost in processing, so the dotted and dot-dashed lines are missing for that particular pulsar. For all other pulsars, all four curves are present, though they do frequently overlap.

(like PSR J1713+0747) and those that do have timing instabilities which could mask a GWB (like PSR J1939+2134).

### 3.2 Timing stability conclusions

Fig. 2 shows that PSR J1939+2134 has red noise with a level and steepness that will limit its use in GWB-detection efforts that last more than about 2 years. Four other pulsars (PSRs J0613–0200, J1024–0719, J1045–4509 and J1824–2452) show some indication of similar red noise, but longer timing and lower white noise levels are needed to determine this with statistical significance. For all the other pulsars, we have no evidence that the red noise that may be present in the timing residuals below the white noise level has a spectral index that prevents GWB detection on time-scales of 5 years to a decade.

We have been unable to detect timing instabilities with an amplitude and spectral slope that could mask a GWB in the timing data of PSRs J1713+0747 and J1744–1134, notwithstanding their long data spans and low timing rms which should make them highly sensitive to any low-frequency noise. Using equation (4), it can be shown that the data sets of PSRs J1713+0747 and J0437–4715 already meet the requirements for a 10-year long PTA experiment, proving that at least for some pulsars the timing stability and rms residual required to detect a GWB on time-scales of 10 years or more are achievable. The challenge for such long-term projects will therefore be to find more pulsars like these, or to replicate these results for other existing pulsars, by increasing the sensitivity of observing systems.

It must be noted that the study of irregularities in pulsar timing data (often referred to as ‘timing noise’) can be much more extensive than presented here. Given our main aim of assessing the impact on PTA science and the absence of clear timing noise in most of our data sets, precise modelling or bounding of timing irregularities as well as a thorough discussion of the potential sources of any observed timing instabilities has not been included in this analysis.

## 4 ANALYSIS OF RESIDUAL RMS

As an alternative to the long-term PTA detection efforts discussed in the previous section, a shorter term detection is possible if timing at lower residual rms is achievable. The standard scenario for a (relatively) short-term GWB detection by a PTA requires 5 years of weekly observations with a timing rms of 100 ns for 20 MSPs (Jenet *et al.* 2005). Since a residual rms of 100 ns has never been maintained over 5 years, the possibility that some intrinsic property of MSPs induces instabilities at that level remains open. In this section, we will address that issue by evaluating how much the timing rms of some of our most precisely timed pulsars may be reduced.

We separate the following three different categories of contributions to the pulsar timing residuals.

**Radiometer noise  $\sigma_{\text{Rad}}$ :** the Gaussian noise component that scales with the radiometer equation and which is mainly determined by the shape and S/N of the observed pulse profiles.

**Frequency-systematic effects  $\sigma_v$ :** this category of noise contributions contains most effects that produce timing residuals dependent on the observing frequency. This includes interstellar effects such as interstellar scintillation and DM variations.

**Temporal-systematic effects  $\sigma_t$ :** this category contains all time-dependent effects, such as calibration errors, instabilities in the observing systems, clock errors, errors in the Solar system ephemerides, GWs and intrinsic pulsar timing noise.

As it is impossible to get direct measures of the three contributions listed above, we base our analysis on the following three measurements.

**Total timing rms  $\sigma$ :** this is simply the timing residual rms of the data considered. It contains all three effects:

$$\sigma^2 = \sigma_{\text{Rad}}^2 + \sigma_v^2 + \sigma_t^2. \quad (6)$$

**Sub-band rms  $\sigma_{\text{sb}}$ :** in Section 4.2, we will introduce this new measure which is  $1/\sqrt{2}$  times the weighted rms of the offset between the residuals of two simultaneous observing bands with different centre frequencies. Since the observations in the two observing bands are simultaneous, their offset is determined by the radiometer noise and by frequency-systematic effects (as the observing bands are centred at slightly different frequencies). We can therefore write

$$\sigma_{\text{sb}}^2 = \sigma_{\text{Rad}}^2 + \sigma_v^2. \quad (7)$$

**Theoretical radiometer noise  $\sigma_{\text{Rad}}$ :** in Section 4.1, we will calculate  $\sigma_{\text{Rad}}$  directly from the pulse profiles used in our timing.

Using these three measures and equations (6) and (7), the three contributions to the timing residuals can be isolated, the results of which are described in Section 4.3.

Our analysis will be based on the CPSR2 data of PSRs J1909–3744, J1713+0747 and J1939+2134. We restrict this analysis to the CPSR2 data, because it is of superior quality to the data of older backend systems (see Section 2.2) and because it consists of the five most densely sampled years of observations. We focus on three of the most precisely timed pulsars in order to obtain the best limits on achievable residual rms. In doing so, we omit PSR J0437–4715 because the advanced calibration schemes used in its analysis (see van Straten 2004, 2006; Verbiest *et al.* 2008) complicates our efforts and because reported non-Gaussian noise in the timing data of this pulsar (Verbiest *et al.* 2008) may imply an inferior limit to that derived from PSRs J1909–3744 and J1713+0747. Note that the purpose of this analysis is to uncover the *potential* limit for high-precision timing: it is already known (see e.g. Section 3) that MSPs have different amounts of time-dependent noise, so the limit we will derive from PSRs J1909–3744 and J1713+0747 does not have to hold for all MSPs. However, it does suggest that other pulsars may achieve similar rms residual and that a PTA-size sample of 20 MSPs at such rms residual may mainly depend on increased sensitivity of current observing systems and new discoveries in ongoing and future surveys.

### 4.1 Theoretical estimation of radiometer noise, $\sigma_{\text{Rad}}$

The level at which the radiometer noise adds to the timing residuals can be determined based on the pulsar’s observed pulse profile shape and brightness, as described by van Straten (2006). Equation (13) of that publication provides the following measure (note we only consider the total intensity,  $S_0$ , to allow direct comparison with our timing results):

$$\sigma_{\text{Rad}} = P \times \sqrt{V} = P \times \left( 4\pi^2 \sum_{m=1}^{N_{\text{max}} \leq N/2} v_m^2 \frac{S_{0,m}^2}{\varsigma_0^2} \right)^{-0.5}, \quad (8)$$

where  $v_m$  is the  $m$ th frequency of the Fourier transform of the pulse profile,  $S_{0,m}^2$  is the total power at that frequency,  $\varsigma_0$  is the white noise variance of the profile under consideration,  $N$  is the total number of time bins across the profile and  $N_{\text{max}}$  is the frequency bin where the Fourier transform of the pulse profile reaches the white noise level,  $\varsigma_0$ .  $V$  is the expected variance in the phase offset or residual,  $P$  is

**Table 6.** Breakdown of weighted timing residuals for three selected pulsars. Given are the total timing rms of the  $\sim 5$  yr of CPSR2 data ( $\sigma$ ), the sub-band timing rms ( $\sigma_{\text{sb}}$ ), the radiometer noise ( $\sigma_{\text{Rad}}$ ), the temporal-systematic ( $\sigma_{\tau}$ ) and the frequency-systematic ( $\sigma_v$ ) contributions to the timing rms. All values are in ns and apply to 64 min integrations (see Section 4 for more information).

Pulsar name (1)	$\sigma$ (2)	$\sigma_{\text{sb}}$ (3)	$\sigma_{\text{Rad}}$ (4)	$\sigma_{\tau}$ (5)	$\sigma_v$ (6)
J1909–3744	166	144	131	83	60
J1713+0747	170	149	105	82	106
J1939+2134	283	124	64	254	106

the pulse period and  $\sigma_{\text{Rad}}$  is the residual rms predicted for the pulse profile considered.

In order to use equation (8) on our data, we first integrated all our pulse profiles together, weighted by S/N, after which equation (8) was applied to the final profile. Subsequently,  $\sigma_{\text{Rad}}$  was renormalized to 64 min integrations through use of the radiometer equation. In order to check this result, we also applied the equation to all individual pulse profiles contained in this analysis and averaged the results in a weighted way – resulting in the same answer, which is given in column four of Table 6. The value for PSR J1909–3744 shows that even at this low residual rms, radiometer noise dominates the timing rms. Applying this method to the other MSPs in our sample, we found that almost all our timing residuals are dominated by radiometer noise. For more than half of our sample of 20 MSPs,  $\sigma_{\text{Rad}}$  is of the order of a microsecond or more. This demonstrates the need for longer integration times, larger bandwidth and/or larger collecting area.

#### 4.2 Estimating frequency-dependent effects

As described in Section 2.2, the CPSR2 pulsar backend records two adjacent, 64 MHz-wide frequency bands simultaneously. This allows determination of a unique measure of a subset of timing irregularities, which we will call the ‘sub-band rms’,  $\sigma_{\text{sb}}$ :

$$\sigma_{\text{sb}} = \frac{1}{\sqrt{2}} \sqrt{\frac{\sum_i \frac{(r_{i,m} - r_{i,n})^2}{e_{i,mn}^2}}{\sum_i 1/e_{i,mn}^2}}, \quad (9)$$

where the sums run over all observing epochs  $i$ ,  $r_{i,m}$  and  $r_{i,n}$  are the residuals of either observing band (named m and n, respectively) at epoch  $i$  and  $e_{i,mn} = \sqrt{e_{i,m}^2 + e_{i,n}^2}$  is the average TOA error at epoch  $i$ . Effectively, the sub-band rms is  $1/\sqrt{2}$  times the weighted rms of the offset between the residuals of the two bands. This implies that it contains all contributions to the total rms that are not time-dependent but either statistically white or dependent on the observing frequency, as described earlier. Note, however, that many of these effects have both a temporal and frequency component. Given our sampling, it should therefore be understood that (specifically in the case of DM variations) only part of these effects is contained in  $\sigma_v$ , while the remaining contributions are contained in  $\sigma_{\tau}$ .

The sub-band rms for the three selected pulsars is presented in Column 3 of Table 6.

#### 4.3 Discussion

Based on equations (6) and (7) and the three measures  $\sigma$ ,  $\sigma_{\text{sb}}$  and  $\sigma_{\text{Rad}}$  determined in the preceding paragraphs, the three contributions to the rms ( $\sigma_{\text{Rad}}$ ,  $\sigma_v$  and  $\sigma_{\tau}$ ) can now be estimated. Their values

are presented in Columns 4, 5 and 6 of Table 6. In order to assess the potential for 100 ns timing of these sources over a 5-year time-scale, we will now discuss the possible means of reducing these three contributions.

The radiometer noise  $\sigma_{\text{Rad}}$  scales for different telescopes or observing systems according to the radiometer equation:

$$\sigma_{\text{Rad}} \propto \frac{T_{\text{sys}}}{A_{\text{eff}} \sqrt{Bt}}, \quad (10)$$

where  $B$  is the bandwidth of the observing system used,  $t$  is the integration time,  $A_{\text{eff}} = \eta \frac{\pi D^2}{4}$  is the effective collecting area of the telescope (with  $\eta$  the aperture efficiency and  $D$  the telescope diameter) and  $T_{\text{sys}}$  is the system temperature of the receiver.

The frequency-systematic contributions are not as easily scaled for different observing systems, but they can be decreased and research on this front is progressing (You et al. 2007; Hemberger & Stinebring 2008; Walker et al. 2008). Also, by reducing the radiometer noise, any measurements of DM variations will become more precise, which will enhance corrections for these effects and therefore decrease the contribution of  $\sigma_v$ . We also note that since these effects are frequency-dependent, the employment of very large bandwidth receivers or coaxial receiver systems, such as the 10/50 cm receiver at the Parkes observatory, may lead to highly precise determination and correction of these effects. Furthermore, increased collecting area and bandwidth may enable future timing observations at higher observational frequencies, which would limit the size of these effects. We therefore suggest that  $\sigma_v$  does not ultimately limit the achievable rms residual, but may largely be corrected for if current research and technological development progress.

The wide variety of sources that add to the temporal systematic make predictions about its future evolution hard. Sources such as intrinsic pulsar timing noise are (as yet) impossible to mitigate. Errors in the terrestrial clocks or in the Solar system ephemerides are expected to decrease as better models become available or as timing arrays provide their own improved solutions for these models. Instabilities in the observing system may to some degree be mitigated by improved calibration methods (van Straten 2004, 2006). Simultaneous observations of a single source at multiple observatories may also lead to detection and correction of instrumental instabilities and the time-dependent effect of DM variations may also be mitigated, as explained above.

Following from the above, we stress the fact that all contributions to  $\sigma_v$  and  $\sigma_{\text{Rad}}$  may be mitigated, but that certain contributions to  $\sigma_{\tau}$  cannot be corrected. This implies that this last class of effects will ultimately limit the residual rms that can be reached. We will therefore use the temporal-systematic contribution to the rms ( $\sigma_{\tau}$ ; Column 5 in Table 6) as an upper limit on the potential rms residual of the MSPs under investigation. Note that this is a conservative upper limit since significant portions of  $\sigma_{\tau}$  may be expected to be mitigated. However, without relative quantification of the various contributions to  $\sigma_{\tau}$ , this limit cannot be reliably decreased.

Given the discussion above, we note that the potential timing residual rms of PSRs J1909–3744 and J1713+0747 is predicted to be below 100 ns on a 5-year time-scale. This implies that the standard scenario of 100 ns timing over 5 years is possible provided techniques currently being developed for mitigation of frequency-dependent effects are successful, more sensitive observing systems are used and more bright, stable MSPs like PSRs J1909–3744 and J1713+0747 are found.

**Table 7.** Assumed parameters for future and ongoing PTA efforts.

PTA name	$N_{\text{Tel}}$	$B$ (MHz)	$D$ (m)	$\eta$	$T_{\text{sys}}$ (K)	Observing regularity	$T$ (yr)
Current	1	64	64	0.6	25	Weekly	5
Predicted PPTA	1	256	64	0.6	25	Weekly	10
NANOGrav	2	256	305; 100	0.5; 0.7	30; 20	Monthly	5
EPTA	5	128	100	0.7	30	Monthly	5
EPTA–LEAP	1 <sup>a</sup>	128	224	0.7	30	Monthly	5
Arecibo	1	512	305	0.5	30	Two-weekly	5
FAST <sup>b</sup>	1	400	500	0.36	20	Two-weekly	5
ASKAP <sup>c</sup>	40	256	12	0.8	50	Weekly	5
MeerKAT <sup>d</sup>	80	512	12	0.7	30	Weekly	5

*Note.* Besides the names of the different PTAs, the columns contain the number of telescopes  $N_{\text{Tel}}$ , the observing bandwidth  $B$ , the telescope diameter  $D$ , aperture efficiency  $\eta$ , system temperature  $T_{\text{sys}}$ , observing regularity and the duration of the project,  $T$ .

<sup>a</sup>Under the LEAP initiative, five 100-m class telescopes will be combined into an effective 224-m single telescope.

<sup>b</sup>Nan (2006), Jin, Nan & Gan (2008).

<sup>c</sup><http://www.atnf.csiro.au/projects/askap/specs.html>.

<sup>d</sup>Jonas (2007).

## 5 PROSPECTS FOR GRAVITATIONAL-WAVE DETECTION

Jenet *et al.* (2005) derived the expected sensitivity of a PTA to a GWB with given amplitude,  $A$ , for both homogeneous arrays (where all pulsars have comparable timing residuals) and inhomogeneous arrays. They also pointed out the importance of pre-whitening<sup>1</sup> the residuals to increase sensitivity at larger GWB amplitudes. In this section, we will build upon their analysis to provide more realistic predictions for ongoing and future timing arrays. We extend their analysis in three fundamental ways.

First, we use the rms timing residuals presented in Table 2. These results provide an inhomogeneous set of rms with a realistic spread. We assume that the residuals are statistically white and will therefore not change with the time-scale of the timing array project. Our analysis in Section 3 shows that for most pulsars this assumption is reasonable, especially on time-scales of the order of 5 years.

Secondly, we do not apply exactly the same algorithm as Jenet *et al.* (2005). In Appendix A, we present a derivation of PTA sensitivity to a GWB in a manner that provides some guidance on analysing the data. We assume that the pre-whitening and correlation are handled together by computing cross-spectra and estimate the amplitude of the GWB directly rather than using the normalized cross correlation function. We assume that the non-GWB noise is white, but can be different for each pulsar. Our results are very close to those of Jenet *et al.* (2005) and using our method we successfully reproduced the scaling law (equation 4). The analysis could be easily extended to include non-white noise if a model for the noise were available.

Finally, in order to generalize the results from our Parkes data to telescopes in other parts of the world, we scale the residuals based on realistic parameters for various PTA efforts listed in Table 7. In doing so, we scale  $\sigma_{\text{Rad}}$  (see Section 4) according to equation (10). As discussed in Section 4.3, some improvements in  $\sigma_v$  and  $\sigma_\tau$  can be expected in coming years, especially as the radiometer noise is

decreased. While quantification of any such improvement is practically impossible, we will apply the same radiometer scaling to  $\sigma_v$  as we apply to  $\sigma_{\text{Rad}}$  and assume  $\sigma_\tau$  to be constant at 80 ns for all pulsars at all telescopes. This may provide a slight disadvantage for larger telescopes, but overall we consider this a reasonable yet conservative approach.

### 5.1 Ongoing PTA projects

We consider the following five ongoing PTA efforts.

*Current* refers to the data presented in this paper, using the longest overlapping time-span of the sample: 5 years. This ignores the fact that the PSR J1824–2452 data set is shorter, but this globular cluster pulsar may not prove useful in a PTA project lasting longer than a few years anyway. We therefore assume that a replacement is found and has identical timing rms over a time-span of 5 years.

*Predicted PPTA* assumes the usage of 256 MHz of bandwidth at the Parkes telescope, which implies a four-fold bandwidth increase and therefore a two-fold decrease in timing rms. The PPTA is the only one to be considered for more than 5 years, mainly in order to demonstrate the large impact a doubling of campaign length can have, but also because several years of high-precision timing data with that bandwidth do already exist (Manchester 2008) for all 20 MSPs.

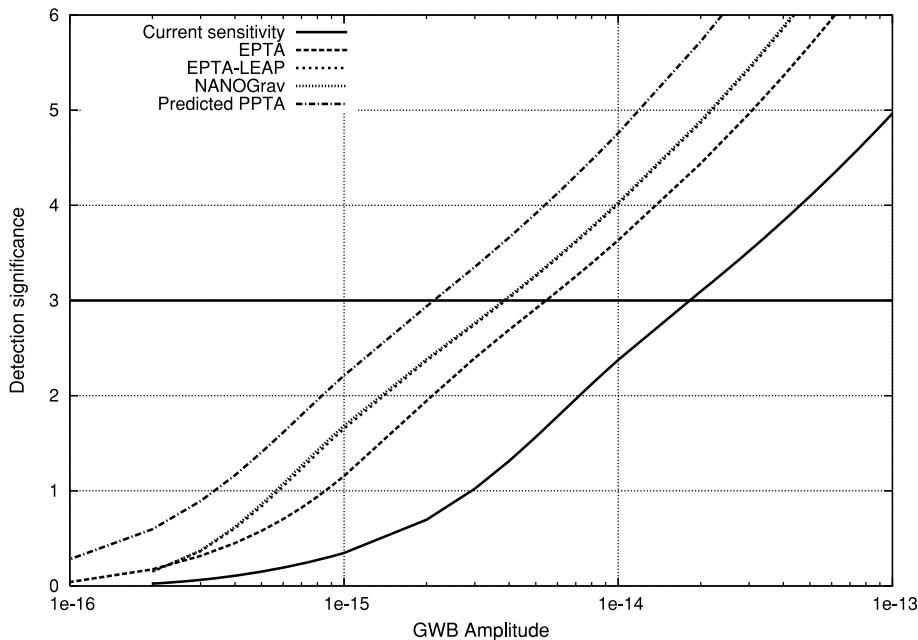
The North American Nanohertz Observatory for Gravitational Waves (NANOGrav) assumes Arecibo gain for the 10 least well-timed pulsars and Green Bank Telescope (GBT) gain for the 10 best-timed pulsars, in order to get a fairly equal rms for all 20 MSPs. (Since we consider  $\sigma_\tau$  as an upper limit on the rms residual, the advantage of Arecibo over the GBT is limited for the brightest sources.)

European Pulsar Timing Array (EPTA) assumes monthly observations with five 100-m class telescopes (Janssen *et al.* 2008).

*EPTA–LEAP*<sup>2</sup> interferometrically combines the five telescopes of the EPTA to form a single, larger one. This decreases the number of observations, but increases the gain.

<sup>1</sup> In this context, pre-whitening refers to a technique that flattens the power spectrum of a time series by means of weighting. This flattening optimizes the sensitivity of a PTA to steep red spectra such as those introduced by a GWB.

<sup>2</sup> Large European Array for Pulsars.



**Figure 4.** Sensitivity curves for different PTA efforts. Note the ‘NANOGrav’ and ‘EPTA–LEAP’ curves are practically coincident. Gravitational waves are predicted to exist in the range  $10^{-15}$ – $10^{-14}$  (see text and Table 7 for more information).

An important caveat to this analysis is that several of the pulsars under consideration cannot be observed with most northern telescopes, because of the telescope declination limits. We therefore assume stable MSPs to be discovered in the Northern hemisphere. As mentioned before, we also assume that progress will be made in the mitigation of frequency-dependent ISM and calibration effects. Finally, this analysis is based on the Parkes data presented in this paper and therefore assumes systematic effects to be at most at the level of the Parkes observing system used.

The sensitivity curves presented in Fig. 4 seem to justify cautious optimism for GWB detection through PTA experiments on time-scales of 5 to 10 years, provided current models of GWBs are correct. While none of the curves in Fig. 4 reaches the minimum predicted GWB amplitude of  $10^{-15}$  at a detection-significance level of three, their sensitivity can be expected to increase up to an order of magnitude through extension of the campaigns to a decade-long time-scale, as illustrated by the difference between the ‘predicted PPTA’ and ‘current sensitivity’ curves. The GWB predictions may, however, change if other effects such as black hole binary stalling occur. The models do, furthermore, rely on a substantial number of poorly determined input parameters, such as what fraction of galaxy growth happens by merging (Sesana, Vecchio & Volonteri 2009). Since only the merging of galaxies results in binary black holes and hence contributes to the GWB, this mass fraction is crucial for any reliable prediction of GWB strength.

As explained in Section 4.3, the temporal-systematic contribution to the rms,  $\sigma_\tau$ , is a conservative upper limit to the ultimate residual rms. In this analysis of PTA efforts, however, we have used the value of 80 ns as a hard lower limit on the timing rms,  $\sigma$ . This limits the potential for reduction of the rms and explains the equivalence of the NANOGrav and EPTA–LEAP efforts. Finally, the strong dependence on the time-scale,  $T$ , of the project underscores the importance of timing stability analysis over much longer time-spans and continued observing. While our  $\sigma_z$  analysis on PSR J1713+0747 provides the first evidence for high timing stability

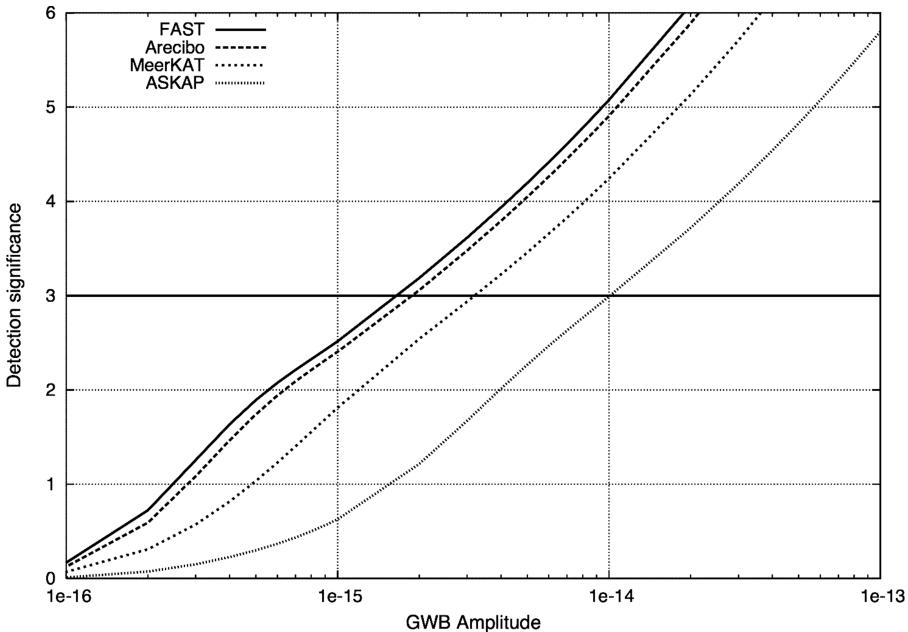
over time-scales beyond 10 years, such timing stability must still be demonstrated for many more MSPs.

## 5.2 Future PTA projects

With the completion of the Square Kilometre Array (SKA) pathfinders expected in 3 years time, we consider the potential of the Australian SKA Pathfinder (ASKAP), the South African Karoo Array Telescope (MeerKAT) and the Chinese Five hundred meter Aperture Spherical Telescope (FAST) for PTA programmes. ASKAP is primarily designed for H $\alpha$  surveys and therefore sacrifices point-source sensitivity for a wide field of view, whereas MeerKAT’s design is better suited for point-source sensitivity over a more limited field of view. FAST is an Arecibo-type single dish with a total diameter of 500 m of which 300 m is illuminated, resulting in a substantially larger sky coverage than is possible with Arecibo. The expected architecture for these telescopes is listed in Table 7 – note we assume phase-coherent combination of the signals of all ASKAP and MeerKAT dishes, effectively resulting in a single telescope of diameter 107 m for MeerKAT and 76 m for ASKAP.

The resulting sensitivity curves are drawn in Fig. 5, along with a hypothetical curve for the most sensitive telescope currently operational, the Arecibo radio telescope. This figure clearly shows the advantage MeerKAT holds over ASKAP for PTA work, in number of dishes, bandwidth and system temperature. The sensitivity of Arecibo is much higher than that of either interferometric prototype and is just inferior to FAST. As for the NANOGrav and EPTA–LEAP projects analysed earlier, the advantage of FAST over Arecibo is strongly limited by the bound of 80 ns we imposed on the achievable rms residual.

Note that the usefulness of Arecibo for PTA work is limited by the restricted sky coverage and hence available pulsars. While both MeerKAT and ASKAP can see large parts of the southern sky, the sky coverage of Arecibo as well as the short transit time makes an exclusively Arecibo-based PTA practically impossible; however, its



**Figure 5.** Sensitivity curves for the two main SKA pathfinders, Arecibo and FAST. Gravitational waves are predicted to exist in the range  $10^{-15}$ – $10^{-14}$  (see discussion in Section 5 and Table 7 for more information).

potential as part of a combined effort (Fig. 4) or in a global PTA is undeniable if the level of systematic errors is small compared to the radiometer noise. As for any Northern telescope, the usefulness of FAST will mostly depend on the discovery of good timing MSPs at positive declinations.

## 6 CONCLUSIONS

We have presented the first long-term timing results for the 20 MSPs constituting the PPTA. We have shown that only PSR J1939+2134 has timing instabilities that limit its use for long-term GWB efforts, while the PSR J1713+0747 data already demonstrate that the requirements for GWB detection on a time-scale of 10 to 15 years are achievable. Overall, the timing stability of the investigated MSPs was found to be encouraging even though potential timing instabilities were detected in four pulsars (in addition to PSR J1939+2134).

It was demonstrated that even on our most precisely timed MSPs white noise is a dominant contribution, suggesting that our residual rms will be much improved with current wide-bandwidth systems. We placed a conservative upper limit of  $\sim 80$  ns on intrinsic timing instabilities that will ultimately limit the residual rms. We interpreted this result in the context of ongoing and future PTA projects, demonstrating the realistic potential for GWB detection through pulsar timing within 5 years to a decade, provided technical and data reduction developments evolve as expected. For PTA efforts in the Northern hemisphere, the discovery of bright and stable MSPs in the northern sky will be crucial. Given the location of currently known MSPs, the prospects of the MeerKAT SKA pathfinder as a gravitational-wave detector are found to be particularly good.

## ACKNOWLEDGMENTS

The Parkes Observatory is part of the Australia Telescope which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. We thank the staff at the Parkes

Observatory for technical assistance and dedicated help over many years. We also acknowledge the support of the large number of collaborators and students who have assisted in the acquisition of the data presented in this paper over the last 14 years. JPWV acknowledges financial support provided by the Astronomical Society of Australia. XPY is supported by the National Natural Science Foundation of China (10803004) and by Natural Science Foundation Project CQ CSTC (2008BB0265).

## REFERENCES

- Backer D. C., Kulkarni S. R., Heiles C., Davis M. M., Goss W. M., 1982, *Nat*, 300, 615
- Bailes M. et al., 1994, *ApJ*, 425, L41
- Bailes M. et al., 1997, *ApJ*, 481, 386
- Bell J. F., Bailes M., Manchester R. N., Lyne A. G., Camilo F., Sandhu J. S., 1997, *MNRAS*, 286, 463
- Bhat N. D. R., Bailes M., Verbiest J. P. W., 2008, *Phys. Rev. D*, 77, 124017
- Boyle L. A., Buonanno A., 2008, *Phys. Rev. D*, 78, 043531
- Caldwell R. R., Battye R. A., Shellard E. P. S., 1996, *Phys. Rev. D*, 54, 7146
- Camilo F., Nice D. J., Shrauner J. A., Taylor J. H., 1996, *ApJ*, 469, 819
- Cognard I., Backer D. C., 2004, *ApJ*, 612, L125
- Damour T., Vilenkin A., 2005, *Phys. Rev. D*, 71, 063510
- Edwards R. T., Bailes M., 2001, *ApJ*, 553, 801
- Edwards R. T., Bailes M., van Straten W., Britton M. C., 2001, *MNRAS*, 326, 358
- Edwards R. T., Hobbs G. B., Manchester R. N., 2006, *MNRAS*, 372, 1549
- Foster R. S., Backer D. C., 1990, *ApJ*, 361, 300
- Foster R. S., Wolszczan A., Camilo F., 1993, *ApJ*, 410, L91
- Grishchuk L. P., 2005, *Phys. Uspekhi*, 48, 1235
- Hellings R. W., Downs G. S., 1983, *ApJ*, 265, L39
- Hemberger D. A., Stinebring D. R., 2008, *ApJ*, 674, L37
- Hobbs G., Lyne A. G., Kramer M., Martin C. E., Jordan C., 2004, *MNRAS*, 353, 1311
- Hobbs G. B., Edwards R. T., Manchester R. N., 2006, *MNRAS*, 369, 655
- Hobbs G. B. et al., 2009, *PASA*, 26, 103
- Hotan A. W., van Straten W., Manchester R. N., 2004, *Publ. Astron. Soc. Australia*, 21, 302

- Hotan A. W., Bailes M., Ord S. M., 2005, MNRAS, 362, 1267  
 Hotan A. W., Bailes M., Ord S. M., 2006, MNRAS, 369, 1502  
 Jacoby B. A., Bailes M., van Kerkwijk M. H., Ord S., Hotan A., Kulkarni S. R., Anderson S. B., 2003, ApJ, 599, L99  
 Jacoby B. A., Hotan A., Bailes M., Ord S., Kulkarni S. R., 2005, ApJ, 629, L113  
 Jaffe A. H., Backer D. C., 2003, ApJ, 583, 616  
 Janssen G. H., Stappers B. W., Kramer M., Purver M., Jessner A., Cognard I., 2008, in Bassa C., Wang Z., Cumming A., Kaspi V. M., eds, AIP Conf. Ser. 983, 40 Years of Pulsars: Millisecond Pulsars, Magnetars and More. Am. Inst. Phys., New York, p. 633  
 Jenet F. A., Hobbs G. B., Lee K. J., Manchester R. N., 2005, ApJ, 625, L123  
 Jenet F. A. et al., 2006, ApJ, 653, 1571  
 Jin C. J., Nan R. D., Gan H. Q., 2008, in Jin W. J., Platais I., Perryman M. A. C., eds, Proc. IAU Symp. Vol. 248, A Giant Step: From Milli- to Micro-Arcsecond Astrometry. Cambridge Univ. Press, Cambridge, p. 178  
 Johnston S. et al., 1993, Nat, 361, 613  
 Jonas J., 2007, Proc. of Science, From Planets to Dark Energy: the Modern Radio Universe. SISSA, Trieste, PoS(MRU)007  
 Kaspi V. M., Taylor J. H., Ryba M., 1994, ApJ, 428, 713  
 Kramer M. et al., 2006, Sci, 314, 97  
 Lee K. J., Jenet F. A., Price R. H., 2008, ApJ, 685, 1304  
 Löhmer O., Kramer M., Driebe T., Jessner A., Mitra D., Lyne A. G., 2004, A&A, 426, 631  
 Lorimer D. R. et al., 1995, ApJ, 439, 933  
 Lorimer D. R., Lyne A. G., Bailes M., Manchester R. N., D'Amico N., Stappers B. W., Johnston S., Camilo F., 1996, MNRAS, 283, 1383  
 Lyne A. G., Brinklow A., Middleditch J., Kulkarni S. R., Backer D. C., Clifton T. R., 1987, Nat, 328, 399  
 Maggiore M., 2000, Phys. Rep., 331, 283  
 Manchester R. N., 2008, in Bassa C., Wang Z., Cumming A., Kaspi V. M., eds, AIP Conf. Ser. 983, 40 Years of Pulsars: Millisecond Pulsars, Magnetars and More. Am. Inst. Phys., New York, p. 584  
 Martin C. E., 2001, PhD thesis, Univ. Manchester  
 Matsakis D. N., Taylor J. H., Eubanks T. M., 1997, A&A, 326, 924  
 Nan R., 2006, Sci. China G: Phys. Astron., 49, 129  
 Nice D. J., 2006, Adv. Space Res., 38, 2721  
 Ord S. M., Jacoby B. A., Hotan A. W., Bailes M., 2006, MNRAS, 371, 337  
 Romani R. W., 1989, in Ögelman H., van den Heuvel E. P. J., eds, Timing Neutron Stars, Timing a Millisecond Pulsar Array. Kluwer, Dordrecht, p. 113  
 Sandhu J. S., 2001, PhD thesis, California Institute of Technology  
 Sandhu J. S., Bailes M., Manchester R. N., Navarro J., Kulkarni S. R., Anderson S. B., 1997, ApJ, 478, L95  
 Sazhin M. V., 1978, Sov. Astron., 22, 36  
 Segelstein D. J., Rawley L. A., Stinebring D. R., Fruchter A. S., Taylor J. H., 1986, Nat, 322, 714  
 Sesana A., Vecchio A., Colacino C. N., 2008, MNRAS, 390, 192  
 Sesana A., Vecchio A., Volonteri M., 2009, MNRAS, 394, 2255  
 Splaver E. M., Nice D. J., Stairs I. H., Lommen A. N., Backer D. C., 2005, ApJ, 620, 405  
 Standish E. M., 2004, A&A, 417, 1165  
 Staveley-Smith L. et al., 1996, PASA, 13, 243  
 Taylor J. H., Weisberg J. M., 1982, ApJ, 253, 908  
 Toscano M., Sandhu J. S., Bailes M., Manchester R. N., Britton M. C., Kulkarni S. R., Anderson S. B., Stappers B. W., 1999, MNRAS, 307, 925  
 van Straten W., 2004, ApJ, 152, 129  
 van Straten W., 2006, ApJ, 642, 1004  
 van Straten W., Bailes M., Britton M., Kulkarni S. R., Anderson S. B., Manchester R. N., Sarkissian J., 2001, Nat, 412, 158  
 Verbiest J. P. W. et al., 2008, ApJ, 679, 675  
 Walker M. A., Koopmans L. V. E., Stinebring D. R., van Straten W., 2008, MNRAS, 388, 1214  
 Wyithe J. S. B., Loeb A., 2003, ApJ, 590, 691  
 You X.-P. et al., 2007, MNRAS, 378, 493

## APPENDIX A: PTA SENSITIVITY

In this Appendix, we derive a simplified formalism for estimating the sensitivity of a PTA to a stochastic and isotropic GWB of given amplitude,  $A$ . This derivation produces results equivalent to those resulting from equation (14) of Jenet et al. (2005), but is more readily implemented and inherently treats optimal weighting (or pre-whitening) of the pulsar power spectra.

The detection statistic is the sample cross-covariance of the residuals of two pulsars  $i$  and  $j$ , separated by an angle  $\theta_{i,j}$ :

$$R(\theta_{i,j}) = \frac{1}{N_s} \sum_{t=0}^T r_i(t) \times r_j(t), \quad (\text{A1})$$

where  $r_i(t)$  is the residual of pulsar  $i$  at time  $t$ ,  $N_s$  is the number of samples in the cross covariance and  $T$  is the data span. The expected value of  $R(\theta_{i,j})$  is the covariance of the clock error, which is 100 per cent correlated, plus the cross covariance of the GWB,  $\sigma_{\text{GW}}^2 \zeta(\theta_{i,j})$ . The clock error can be included in the fit, but one must also include its variance in the variance of the detection statistic. It is better to estimate the clock error and remove it, which also removes its ‘self noise’. So in subsequent analysis we neglect clock noise and effects of possible Solar system ephemeris errors. We model the pulsar timing residuals as a GWB term and a noise term:  $r(t) = r_{\text{GW}}(t) + r_N(t)$ , with variances  $\sigma_G^2$  and  $\sigma_N^2$ .  $\zeta(\theta_{i,j})$  is the cross-correlation curve predicted by Hellings & Downs (1983), as a function of the angle between the pulsars,  $\theta_{i,j}$ :

$$\zeta(\theta_{i,j}) = \frac{3}{2}x \log x - \frac{x}{4} + \frac{1}{2}$$

in which  $x = (1 - \cos \theta_{i,j})/2$ .

Since the detection significance will be limited by the variance in the sample cross covariance, we consider

$$\begin{aligned} \text{Var}[R(\theta_{i,j})] &= \text{Var} \left\{ \sum [(r_{\text{GW},i} + r_{N,i})(r_{\text{GW},j} + r_{N,j})/N_s] \right\} \\ &= \sigma_{G,i}^2 \sigma_{G,j}^2 \frac{[1 + \zeta(\theta_{i,j})^2]}{N_s} + \frac{\sigma_{N,i}^2 \sigma_{G,j}^2 + \sigma_{G,i}^2 \sigma_{N,j}^2}{N_s} \\ &\quad + \frac{\sigma_{N,i}^2 \sigma_{N,j}^2}{N_s}. \end{aligned} \quad (\text{A2})$$

After pre-whitening, this becomes (note our notation  $\sigma_{\text{PW}} = \varrho$ )

$$\text{Var}[R_{\text{PW}}(\theta_{i,j})] = \varrho_G^4 \frac{[1 + \zeta(\theta_{i,j})^2]}{N_s} + \varrho_G^2 \frac{(\varrho_{N,i}^2 + \varrho_{N,j}^2)}{N_s} + \frac{\varrho_{N,i}^2 \varrho_{N,j}^2}{N_s}, \quad (\text{A3})$$

in which we have used  $\varrho_{G,i}^2 = \varrho_{G,j}^2 = \varrho_G^2$ , which will be proved shortly.

We derive the gravitational-wave power from equations (3) and (2), for a GWB with spectral index  $\alpha = -2/3$ :

$$P_{\text{GWB}}(f) = K(f/f_{\text{ref}})^{-13/3}, \quad (\text{A4})$$

with  $K$  being a constant proportional to the amplitude of the GWB and  $f_{\text{ref}} = 1 \text{ yr}^{-1}$ .

Defining the corner frequency,  $f_c$ , as the frequency at which the gravitational-wave power equals the noise power enables us to use equation (A4) to determine the noise power:  $P_{\text{Noise}} = K(f_c/f_{\text{ref}})^{-13/3}$ .

As illustrated by Jenet et al. (2005), the steep spectral index of GWB-induced residuals implies that large gains in sensitivity can be achieved through optimal pre-whitening of the data. Assessment of the variance of both the GWB and noise components of the residuals after pre-whitening can most easily be done through integration

of the spectral powers, multiplied by the whitening filter,  $W(f)$ , which is a type of Wiener filter, designed to minimize the error in the estimation of  $\sigma_G$  and is of the form:  $W(f) = P_{\text{GWB}}/(P_{\text{GWB}} + P_{\text{Noise}})^2$ . Rescaling the weighting function thus defined, we get

$$W(f) = C \frac{(f/f_{\text{ref}})^{-13/3}}{\left[1 + (f/f_c)^{-13/3}\right]^2}, \quad (\text{A5})$$

with  $C$  being a normalization constant chosen for convenience to be

$$C = \left\{ \sum_f \frac{(f/f_{\text{ref}})^{-26/3}}{\left[1 + (f/f_c)^{-13/3}\right]^2} \right\}^{-1}. \quad (\text{A6})$$

The pre-whitened variances then become

$$\begin{aligned} \varrho_G^2 &= \sum_f K(f/f_{\text{ref}})^{-13/3} C \frac{(f/f_{\text{ref}})^{-13/3}}{\left[1 + (f/f_c)^{-13/3}\right]^2} \\ &= K, \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} \varrho_N^2 &= \sum_f K(f_c/f_{\text{ref}})^{-13/3} C \frac{(f/f_{\text{ref}})^{-13/3}}{\left(1 + (f/f_c)^{-13/3}\right)^2} \\ &= KC \sum_f \frac{(f_c f/f_{\text{ref}}^2)^{-13/3}}{\left[1 + (f/f_c)^{-13/3}\right]}, \end{aligned} \quad (\text{A8})$$

which justifies our choice for  $C$  and shows that, based on our weighting scheme,  $\varrho_{G,i}^2 = \varrho_{G,j}^2 = K$ , as used earlier.

Since the spectra are effectively bandlimited to  $f_c$  after pre-whitening, both the GWB and noise will have the same number of degrees of freedom, namely  $N_{\text{d.o.f.}} = 2T_{\text{obs}}f_c - 1$ , where  $T_{\text{obs}}$  is the length of the data span and therefore the inverse of the lowest frequency, implying there are  $T_{\text{obs}}f_c$  independent frequencies

measured below  $f_c$ . Since each frequency adds a real and imaginary part, there are twice as many degrees of freedom as there are independent frequency samples; quadratic fitting removes a single degree of freedom from the total. Note that  $\sqrt{N_{\text{d.o.f.},i} N_{\text{d.o.f.},j}}$  is the number of independent samples in the cross-covariance spectrum and therefore replaces  $N_s$  in equations (A1) and (A3).

The optimal least-squares estimator for  $K$  (and hence for the amplitude of the GWB), based on a given set  $R_{\text{PW}}(\theta_{i,j})$  with unequal errors, is (from equations A1 and A7)

$$\tilde{K} = \frac{\sum R_{\text{PW}}(\theta_{i,j})\zeta(\theta_{i,j})/\text{Var}(R_{\text{PW},i,j})}{\sum \zeta(\theta_{i,j})^2/\text{Var}(R_{\text{PW},i,j})}. \quad (\text{A9})$$

The variance of this estimator is

$$\text{Var}(\tilde{K}) = \frac{1}{\sum \zeta(\theta_{i,j})^2/\text{Var}(R_{\text{PW},i,j})}. \quad (\text{A10})$$

We can now write the expected S/N of a given timing array as the square root of the sum over all pulsar pairs of equation (A7) divided by the square root of equation (A10)

$$S = \sqrt{\sum_{i=1}^{N_{\text{psr}}-1} \sum_{j=i+1}^{N_{\text{psr}}} \frac{\varrho_G^4 \zeta^2 \sqrt{N_{\text{d.o.f.},i} N_{\text{d.o.f.},j}}}{\varrho_G^4 (1 + \zeta^2) + \varrho_G^2 (\varrho_{N,i}^2 + \varrho_{N,j}^2) + \varrho_{N,i}^2 \varrho_{N,j}^2}}. \quad (\text{A11})$$

Rewriting leads to

$$S = \sqrt{\sum_{i=1}^{N_{\text{psr}}-1} \sum_{j=i+1}^{N_{\text{psr}}} \frac{\zeta^2 \sqrt{N_{\text{d.o.f.},i} N_{\text{d.o.f.},j}}}{1 + \zeta^2 + (\varrho_i')^2 + (\varrho_j')^2 + (\varrho_i' \varrho_j')^2}}, \quad (\text{A12})$$

where  $\varrho_i' = \varrho_{N,i}/\varrho_G$ .

This paper has been typeset from a  $\text{\TeX}$ / $\text{\LaTeX}$  file prepared by the author.