RAA 2020 Vol. **20** No. **11**, 184(10pp) doi: 10.1088/1674-4527/20/11/184

http://www.raa-journal.org *http://iopscience.iop.org/raa*

Research in Astronomy and Astrophysics

Gravitational wave sensitivity curve of pulsar timing arrays affected by correlated noises

Shu Liu[1,2,3] and Ming-Lei Tong[1,2]

[1] National Time Center, Chinese Academy of Sciences, Xi'an 710600, China; *mltong@ntsc.ac.cn*
[2] Key Laboratory of Time and Frequency Primary Standard, Chinese Academy of Sciences, Xi'an 710600, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China

Received 2019 December 4; accepted 2020 March 26

Abstract The detection of gravitational waves (GWs) by pulsar timing arrays (PTAs) is not only a very important supplementation of the verification of general relativity but also a new window to study the evolution of supermassive binary black holes and the early universe. However, so far the detection sensitivity of PTAs is not good enough to catch signals of GWs due to the disturbance of various noises. In this paper we explore the influences of the correlated noises caused by the reference clock errors and solar system ephemeris errors in pulsar timing on the detection of stochastic gravitational waves background (GWB). We demonstrate the power-law integrated sensitivity curves of GWB detection under the impacts from these correlated noises. From the simulated data, we find that the influence of different reference time-scale is non-negligible, and the influence is even quite huge if one uses a very old version of solar system ephemeris. However, the impact from these correlated noises on the sensitivity curve is very limited for the real observational data released by international pulsar timing arrays (IPTA). We also calculate the signal-to-noise ratios based on the theoretical GW amplitude permitted by observations. Moreover, we study how the detection sensitivity increases with more pulsar number and longer observation.

Key words: gravitational waves — pulsars: general — time — ephemerides

1 INTRODUCTION

Multiple millisecond pulsars monitored by radio telescopes regularly form Pulsar Timing Arrays (PTAs). Gravitational wave (GW) detection is one of the main scientific goals of PTAs, because they can be regarded as galactic-scale GW detectors (Ferdman et al. 2010; Hobbs et al. 2010; Perera et al. 2019; Verbiest et al. 2010). Gravitational radiation interferes with the propagation path of the radio pulse between the pulsar and the Earth Observatory. Pulsar timing analyses give the differences, commonly known as timing residuals, between observed pulse times of arrival (TOAs), normally referred to as the barycentre of the solar system, and the predictions of a model for the pulsar properties. The timing residuals contains information about the GWs (Detweiler & Szedenits 1979; Estabrook & Wahlquist 1975; Sazhin 1978), and GW signals can be extracted by correlating the timing residuals of each pulsar pair (Detweiler & Szedenits 1979; Jenet et al. 2005). For the two polarization modes of plus and cross in the frame of general relativity, the correlation of timing residuals from different pulsar pairs manifests like a Hellings-Downs curve (Hellings & Downs 1983).

PTA is most sensitive to GWs in the 10^{-9} – 10^{-7} Hz frequency band, in which some typical sources are supermassive black-hole binaries (SMBHBs) with masses in the range of $\sim 10^7$–$10^{10}\,M_\odot$ during their slow, adiabatic inspiral phase (Jaffe & Backer 2003; Rajagopal & Romani 1995; Sesana et al. 2009). Theoretically, PTAs can also observe other GWs sources, such as cosmic strings (Kuroyanagi et al. 2013) and relic gravitational waves, a background originating during inflation in the early universe (Grishchuk 2001; Zhang et al. 2005; Zhao et al. 2013; Tong et al. 2016). The majority of SMBHBs are individually unresolvable, but the incoherent superposition of the very weak radiation from the many binaries in the population produce a stochastic GW background (GWB) which can be detected by PTAs (Sesana et al. 2008). Since the first GW event GW150914 was detected directly by Advance LIGO (Abbott et al. 2016), many coalescences of binary black hole were observed through GWs. Moreover, GWs generated by a binary neutron

star system GW170817 were detected by the network of Advanced LIGO (LIGO Scientific Collaboration et al. 2015) and Advanced Virgo (Acernese et al. 2015) on 2017 August 17. It is reliable that GWs could be detected by PTAs in the near future. If not, more stringent constraints will be placed on aspects of the assembly history of SMBHBs (Volonteri et al. 2003).

For the detection of GWs by PTAs, there are many noises which lower the detectability, such as spatially uncorrelated timing noise, stochastic monopolar clock-like signal, stochastic dipolar ephemeris-like signal, solar wind and so on. Numerous studies have already been carried out about these noises, including construction of a new timescale TT (IPTA16) (Hobbs et al. 2020), constraining the masses of the planet-moons systems tests (Caballero et al. 2018). In this paper, we will focus on the two correlated cases that the referenced clock errors and solar system ephemeris errors. As pointed by Tiburzi et al. (2016), the existence of the two correlated noises can affect the detection of GWB, even though GWB causes quadrupolar signal among different pulsar timing residuals, a different property from the above two correlated cases. We will analyze the effects of referenced clock errors and solar system ephemeris errors on the sensitivity curve of GWB detection by PTAs, respectively. Moreover, we will discuss the integrated sensitivity curve for stochastic GWB proposed by Thrane & Romano (2013), which reflects a more credible detection of GWB.

This article is organized as follows. In Section 2, we present the mathematical formulation of integrated sensitivity curves. In Section 3, the GW sensitivity curve of different clock- and ephemeris- combinations to isotropic stochastic signals was simulated. In Section 4, we then turn our attention to searches for GWB by arrays of pulsars with different numbers and observation time span. We conclude with a brief summary in Section 5.

2 INTEGRATED SENSITIVITY CURVE OF PULSAR TIMING ARRAYS

In this section we introduce the integrated sensitivity curve for stochastic GWB searches. The characteristic strain that the PTA is sensitive to scales linearly with f in the bandwidth range, The sensitivity curves are usually constructed by taking the ratio of the detector's noise power spectral density to its sky- and polarization-averaged response to a GW (Thrane & Romano 2013). But for stochastic GWs which has a power-law frequency dependence in the sensitivity band of the detectors, to illustrate the improvement in sensitivity that comes from integrating over frequency, we will show how the integrated sensitivity is computed for GWB. We follow the analysis of the paper (Thrane & Romano 2013).

The plane wave for the metric perturbation $h_{ab}(t,\boldsymbol{x})$ produced by a stochastic GWB can be expanded as (Allen & Romano 1999):

$$h_{ab}(t,\boldsymbol{x}) = \int_{-\infty}^{\infty} df \int_{S^2} d^2\hat{\Omega} \sum_A h_A(f,\hat{\Omega}) e^A_{ab}(\hat{\Omega}) e^{i2\pi f(t-\hat{\Omega}\cdot\boldsymbol{x}/c)}, \tag{1}$$

where f is the frequency of the GWs, $e^A_{ab}(\hat{\Omega})$ are the polarization tensors, and the index $A = +, \times$ denote the two independent polarizations. The unit vector $\hat{\Omega}$ identifies the propagation direction of a single GW plane.

The simplest GWB that could be considered is an isotropic, uncorrelated, unpolarized and stationary background. For this background, the quadratic expectation values have the form

$$\left\langle h_A(f,\hat{k}) h^*_{A'}\left(f',\hat{k}'\right)\right\rangle = \frac{1}{16\pi}\delta\left(f-f'\right)\delta_{AA'}\delta^2\left(\hat{k},\hat{k}'\right) S_h(f), \tag{2}$$

where

$$S_h(f) = \frac{3H_0^2}{2\pi^2}\frac{\Omega_{\rm gw}(f)}{f^3}, \tag{3}$$

is the GW power spectral density, and GWB is often described by power-law spectra. Hence, the fractional contribution of the energy density in GW to the total energy density required to close the universe can be written as

$$\Omega_{\rm gw}(f) = \Omega_\beta \left(\frac{f}{f_{\rm ref}}\right)^\beta, \tag{4}$$

where β is the spectral index and $f_{\rm ref}$ is a reference frequency typically $1\ \mathrm{yr}^{-1}$. The strain power spectral density of GWB is usually defined by characteristic strain amplitude as (Moore et al. 2015):

$$h_c(f) \equiv \sqrt{f S_h(f)}, \tag{5}$$

which also follows a power-law form:

$$h_c(f) = A_\alpha \left(\frac{f}{f_{\rm ref}}\right)^\alpha, \tag{6}$$

where A_α means the amplitude of the GWB at a pivot frequency $f_{\rm ref}$, and the index α is related to β as: $\beta = 2\alpha + 2$. Different GWB models have different values of α (Jenet et al. 2006). The index equal to $-2/3$ represents the background caused by SMBHB coalescences. The best current limit on the amplitude A_α of the characteristic strain of the stochastic isotropic GWB from SMBHBs is from Shannon et al. (2015), who report a value of $A_\alpha < 1\times10^{-15}$.

From the point of view of detection, the signal-to-noise ratio (SNR) is an interesting and important value. For a cross-correlation search for an unpolarized and

isotropic stochastic background in the network of PTAs, the expected SNR ratio is given by (Anholm et al. 2009):

$$\rho = \sqrt{2T}\left[\int_{f_{\min}}^{f_{\max}} df \sum_{I=1}^{M}\sum_{J>1}^{M} \frac{\Gamma_{IJ}^2(f) S_h^2(f)}{P_{nI}(f) P_{nJ}(f)}\right]^{1/2}, \quad (7)$$

where M is the number of pulsars, and the observation time T has been assumed to be the same coincident for each pulsar. The overlap reduction function $\Gamma_{IJ}(f)$ for the pulsar pair I and J, to a very good approximation, can be written as (Anholm et al. 2009):

$$\Gamma_{IJ}(f) = \frac{1}{(2\pi f)^2}\frac{1}{3}\zeta_{IJ}, \quad (8)$$

where

$$\begin{aligned}\zeta_{IJ} \equiv & \frac{3}{2}\left(\frac{1-\cos\psi_{IJ}}{2}\right)\log\left(\frac{1-\cos\psi_{IJ}}{2}\right) \\ & -\frac{1}{4}\left(\frac{1-\cos\psi_{IJ}}{2}\right)+\frac{1}{2}+\frac{1}{2}\delta IJ\end{aligned} \quad (9)$$

is the Hellings-Downs function (Hellings & Downs 1983), and ψ_{IJ} is the separated angle between the two pulsars. $P_{nI}(f)$ and $P_{nJ}(f)$ in Equation (7) are the auto power spectral densities in detectors I and J due to the noise via (Thrane & Romano 2013) :

$$P_n(f) = 2\Delta t\sigma^2, \quad (10)$$

where $1/\Delta t$ refers to the cadence of TOA measurements of the special pulsar, and σ is the corresponding root mean square of the timing noise. The limits of integration shown in Equation (7) are set as $1/T$ for the lower one and $1/(2\Delta t)$ for the upper one, respectively.

Similar to the definition of characteristic strain for GWs $h_c(f)$, the definition of the effective characteristic strain noise amplitude $h_{\rm eff}(f)$ is as follows (Thrane & Romano 2013)

$$h_{\rm eff}(f) \equiv \sqrt{f S_{\rm eff}(f)}, \quad (11)$$

and the effective strain noise power spectral density for the detector network is

$$S_{\rm eff} \equiv \left[\sum_{I=1}^{M}\sum_{J>1}^{M}\frac{\Gamma_{IJ}^2(f)}{P_{nI}(f)P_{nJ}(f)}\right]^{-1/2}. \quad (12)$$

The GW detection sensitivity curve is usually calculated based on the fractional contribution of the energy density in noises, which can be written as

$$\Omega_{\rm eff}(f) = \frac{2\pi^2 f^3 S_{\rm eff}(f)}{3H_0^2}. \quad (13)$$

However, for the GWB with power-law spectra, the power-law integrated sensitivity curve is more reasonable to evaluate GW detection. The power-law integrated sensitivity curve owns the improvement that comes from the broadband nature of the signal. Through the integration over frequency and considering the quantity relationships mentioned above, one has

$$\Omega_\beta = \frac{\rho}{\sqrt{2T}}\left[\int_{f_{\min}}^{f_{\max}} df \frac{(f/f_{\rm ref})^{2\beta}}{\Omega_{\rm eff}^2(f)}\right]^{-1/2}, \quad (14)$$

where $\Omega_{\rm eff}(f)$ can be converted by $S_{\rm eff}(f)$ using Equation (3). For a set of power-law indices and an arbitrary choice of the referenced frequency $f_{\rm ref}$, we can calculate the value of the amplitude Ω_β such that the integrated signal-to-noise ratio has some fixed value, e.g., $\rho = 1$. For each pair of values for β and Ω_β , we can plot $\Omega_{\rm gw}(f)$, whose envelope is the power-law integrated sensitivity curve (Thrane & Romano 2013):

$$\Omega_{\rm PI}(f) = \max_\beta\left[\Omega_\beta\left(\frac{f}{f_{\rm ref}}\right)^\beta\right]. \quad (15)$$

When there only exists white noise in the timing residuals, the detection rate is proportional to the weighted mean square value of the timing noise. However, in practice, the reference clock error and the solar system ephemeris error will also affect the detection of GWB. In the next section, we will analyze the effects of these two correlated noises on the power-law integrated sensitivity curve for GWB.

3 IMPACTS OF CORRELATED NOISES ON THE DETECTION OF GWB

The TOA measurements of pulsars are relied on the reference time, which is often served by the terrestrial time (TT) published yearly by the Bureau International des Poids et Mesures (BIPM) labelled by TT (BIPM). If a bad atomic clock serves as the reference time, then clock errors will be present in the pre-fit timing residuals of all pulsars in exactly the same way, and will affect the model parameter re-fittings for each pulsar. Similarly, the solar system ephemeris (SSE) errors will also affect both the timing residuals and the pulsar parameter fittings, since the TOAs measured at the observatory need to be converted to the solar system barycenter (SSB). However, the effects of the clock-like error and the ephemeris-like error on the timing residuals are different. In contrast from the reference clock errors whose effect is monopolar, the SSE errors will affect the timing residuals with a dipolar property. That is, for two pulsars located in opposite directions on the celestial sphere, the additional timing residuals of the two pulsars caused by SSE errors are exactly the opposite. Even though the effects of GWB on timing residuals are quadripolar, the clock errors and SSE errors are two factors disturbing the

detection of GWB from PTAs. The clock errors and SSE errors have been speculated on as sources of potential bias during the process of GW detection (Tiburzi et al. 2016), and the recent work to search for GWs is with clock- and ephemeris-free combination of timing data (Tinto & Hartwig 2018). However, below we will focus on the impacts of the reference clock errors and the SSE errors on the GWB sensitivity curve through the simulated pulsar timing data.

For the simulations, we use 20 millisecond pulsars from Parkes Pulsar Timing Array data released 1 (PPTA DR1) (Manchester et al. 2013) as our research object. The timing model parameters of each pulsar are used as the input quantities for simulations, but for all the pulsars the reference clock and SSE are set to be TT (BIPM2018) and DE436, respectively. During the simulating procedure, we set the cadence of the measurements for each pulsar to be 20 per year and the identical white timing noise to be 100 ns for all the pulsars. We find that the result shown in Thrane & Romano (2013) slightly overestimated the sensitivity. A detailed explanation can be seen in Appendix A.

Once the reference time scale or the SSE is changed, the timing model parameters should be re-fitted and the corresponding post-fit timing residuals are formed. It is worth to note that the change of the referenced clock or SSE will lead to additional timing noises for calculations in Equation (10). To study the effects of the reference clock errors, we use three different reference atomic time scales, TT (BIPM2018) and TT (TAI), respectively. Refer to the impacts from SSE errors, for example we choose JPL's DE200, DE405 and DE436, respectively. Moreover, in order to make a comparison with the case of the real observational data, we simulated all the 20 pulsars with a same time span of $T = 10$ yr. The post-fit timing residuals of all the pulsars with different time reference or ephemeris reference become the original information to start the analyses.

3.1 Influence of Referenced Clock Error on Detection of GWB

The clock time system referenced by the pulsar timing data has a good long-term stability, no frequency drift or periodic variation, and should be consistent with SI second. The reference clock time standard TT (BIPM2018) is currently recommended. The TAI system is a near real-time system weighted by key laboratories around the world. The TT (BIPM) scale has been revised every year by BIPM based on the data sets of TAI. It owns the long-term stability of TAI and the accuracy of the frequency standard, which make itself currently the most suitable time reference for pulsar timing.

Table 1 Pulsars of 15 Years from IPTA DR1. TT (BIPM2018) and DE436 were referenced.

Pulsar name	wrms timing residual (μs)
J0218+4232	6.664
J0437–4715	0.243
J0711–6830	2.024
J0751+1807	3.513
J1024–0719	3.134
J1045–0719	3.303
J1603–7202	1.936
J1643–1224	2.709
J1713+0747	0.310
J1730–2304	2.119
J1744–1134	0.877
J1857+0943	0.511
J2124–3358	2.932
J2129–5721	1.221
J2145–0750	1.228

To study the influence of different referenced clock on the sensitivity of detecting stochastic GWB, the solid black and dashed blue curves in Figure 1 are carried out using TT (BIPM2018), TT (TAI) separately. Here, TT (TAI) is referred to as TT realised by TAI, and differs from TAI by a constant: TT (TAI)=TAI+32.184s. Since the TT (BIPM) scale is revised on the basis of TT (TAI), TT (BIPM) is a time scale superior to TT (TAI). As shown in Hobbs et al. (2012), TT (TAI) fluctuates relative to TT (BIPM) although after a quadratic polynomial has been fitted and removed, which can be detected by ensemble pulsar time scale. Figure 1 shows that the sensitivity curve based the reference time of TT (TAI) is about 3 times worse than that based on TT (BIPM2018). This is consistent with the theoretical expectation that TT (BIPM) performs better than TT (TAI).

3.2 Influence of Planetary Ephemeris Error on Detection of GWB

The ideal celestial reference frame is based on the barycentric celestial reference system and the calculation of the origin depends on the mass of the planet and its orbit around it. Due to the estimation error of the planetary mass calculation, the origin position of the celestial reference frame is deviated from the ideal SSB. Therefore, if the position vector of the earth to the SSB provided by the planetary ephemeris is not completely accurate, it will directly affect the delay estimation of the pulse signal from the observation station to the SSB.

To study the influence of different versions of planetary ephemeris on the sensitivity of detecting stochastic GWB, the following curves in Figure 1 is carried out using different planetary ephemeris separately. It shows that GW detection ability with DE405 is improved roughly about

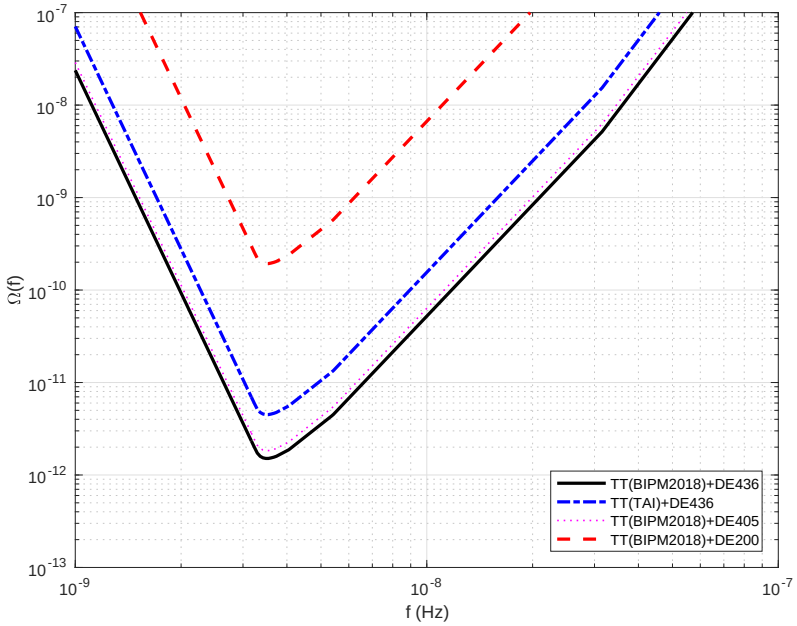

Fig. 1 Different power-law integrated sensitivity curves for different simulated pulsar data sets. The lowest sensitivity curve expressed by the solid black line is based on the standard clock TT (BIPM2018) and the ephemeris DE436. The dash-dotted blue line shows the sensitivity curve used TT (TAI) instead as the referenced time scale compared to the lowest one. The dashed red line represents the sensitivity curve based on DE200 and TT (BIPM2018), while the dotted pink line is based on DE405 and TT (BIPM2018). The thin solid black line shows the corresponding effective fractional energy density without frequency integration, that is, the usually unintegrated sensitivity curve based on TT (BIPM2018) and DE436.

two orders of magnitude on the basis of DE200. The big discrepancy is originated from the rough consideration of DE200 and the greatly improvement of the ephemeris starting from DE405. DE200 was created in 1981 and it includes nutations but not librations, while DE405 was released in 1998 and it added several years' extra data from telescopic, radar, spacecraft, and VLBI observations (of the Galileo spacecraft at Jupiter, in particular). The method of modeling the asteroids' perturbations was improved, although the same number of asteroids were modeled. The ephemeris was more accurately oriented onto the ICRF. Since DE405 is improved on the basis of DE200 for 14 years, that the accuracy of observation data is improved and the theoretical model is updated makes DE405 more accurate than DE200. Compared to the gap between DE200 and DE405, DE405 has been very close to DE436. So, the sensitivity curve calculated with DE405 is only slightly worse than that with DE436.

4 DETECTION ANALYSIS CONFRONTING WITH OBSERVED DATA

The power law integrated sensitivity curves analyzed above were based on the simulated TOA data that only contain white noise. However, in real data there always exist red noises more or less. In this section, we analyze the sensitivity curves and the observation strategy on them based real observational data.

After enough time has elapsed, pulsar timing arrays enter a new regime where the signal to noise only scales as $\sqrt{T}$. In addition, in this regime the quality of the pulsar timing data and the cadence become relatively unimportant, and the best strategy to increase the detectability of GWB in this regime is to increase the number of pulsars in the array (Siemens et al. 2013). In this section, we discuss the detectability of GWB with the two subsets of pulsars which are selected from the International Pulsar Timing Array data release 1 (IPTA DR1) [1]. To reduce the effects from red noises, we choose combination 'B' of IPTA DR1 because the red noises of some pulsars were partially suppressed. Both subsets include the pulsars of 10 yr span, but with different pulsar numbers, one include 15 pulsar while the other 25. To evaluate the SNR and the sensitivity curve easily, as can seen from Equations (7) and (14), we set all the pulsars has the same cadence equalling one every two weeks and the same data span. So, for all the chosen pulsars, we cut out the timing residual sequence of pulsars based on the time span standard of J0030+0451, which has a total observation time span of 10 yr. Only the timing residuals

[1] http://www.ipta4gw.org

Table 2 Pulsars of 10 Years from IPTA DR1

Pulsar name	wrms (μs) TT (BIPM2018) and DE436	wrms (μs) TT (TAI) and DE436	wrms (μs) TT (BIPM2018) and DE405
J0030+0451	1.497	1.495	1.500
J0034–0534	4.414	4.413	4.418
J0218+4232	6.578	6.573	6.580
J0437–4715	0.244	0.239	0.296
J0613–0200	1.070	1.071	1.072
J0621+1002	9.879	9.864	9.899
J0711–6830	1.989	1.987	1.991
J0751+1807	3.406	3.408	3.408
J1012+5307	1.660	1.661	1.662
J1024–0719	2.949	2.917	2.943
J1045–4509	3.055	3.053	3.061
J1600–3053	0.846	0.840	0.842
J1603–7202	1.813	1.813	1.820
J1640+2224	1.607	1.612	1.607
J1643–1224	2.652	2.668	2.642
J1713+0747	0.295	0.326	0.272
J1730–2304	2.101	2.111	2.109
J1744–1134	0.863	0.873	0.867
J1857+0943	0.514	0.514	0.514
J1909–3744	0.188	0.190	0.194
J1918–0642	1.553	1.553	1.551
J2033+1734	13.338	13.338	13.336
J2124–3358	2.794	2.792	2.793
J2129–5721	1.151	1.153	1.150
J2145–0750	1.223	1.222	1.219

during the common time span of 10 yr will be retained. The two subsets of the concrete pulsars' name are listed in Table 1 and Table 2, respectively, where initial information for calculation including the weighted root mean square timing residuals (wrms) and the positions of pulsars were also listed Table 2. Note that, since the post-fit timing residuals serve as the timing noise in Equation (10) now, the auto power spectral densities of the pulsars are different from each other.

Improving on earlier work (Kaspi et al. 1994), (Jenet et al. 2006) developed a frequentist technique in statistics, and had calculated an upper limit on $h_{\rm yr}$ for different values of α. Recently, Shannon et al. (2015) provided an upper limit of $h_{\rm yr} < 1 \times 10^{-15}$ at the 95% confidence level for $\alpha = -2/3$ using data from PPTA and available observations from the Arecibo Observatory. According to the ideal amplitude of GW which can be detected, we list the calculated SNR of two subsets of pulsars with different numbers in Table 3. According to Equations (6), (7) and (11), it can be found that the values of ρ are proportional to the square of characteristic strain GW amplitude h. However, it is worth to point out that, for an amplitude of $A_\alpha = 10^{15}$, these pulsar timing arrays operate at the *strong-signal limit* (Siemens et al. 2013), thus the corresponding values of SNR in Table 3 were overestimated.

According to Equation (15), Figure 2 plots the power-law integrated sensitivity curves for GWB using the two pulsar subsets which contains same time span but different

Table 3 ρ of Different Subsets

Pulsar Number	Time Span	A_α	ρ
25	10-year	10^{-15}	7.0797
		10^{-16}	0.0708
15	10-year	10^{-15}	3.8154
		10^{-16}	0.0382

pulsar numbers with $\rho = 1$ and $\rho = 5$, respectively. The minimum detectable energy density of a GWB will be reduced by about 40%, if one increases the number of pulsars from 15 to 25. That 90 % detection probability is typically reached when the average SNR is significantly larger than 3 (Siemens et al. 2013). Thus, in the following discussions, we typically fix SNR to be 5.

It is believed that the improvement would be even more significant if the pulsar number is about 25 or more. To study the effects of observational time span, we use the pulsar data set shown in Table 1 with 10-yr span and 15-yr span, respectively, to calculate the corresponding power law sensitivity curves. The results are shown in Figure 3, where one can find that data set with longer observational time span performs obviously better than that with shorter during all frequency band,which means that increasing the observational time span can enhance the detection ability.

We also want to analyze the real observed sensitivity curves under effects of clock errors and SSE errors. Here, we choose the 10-years observed data set as our researching object, and the SNR is set to be 5.

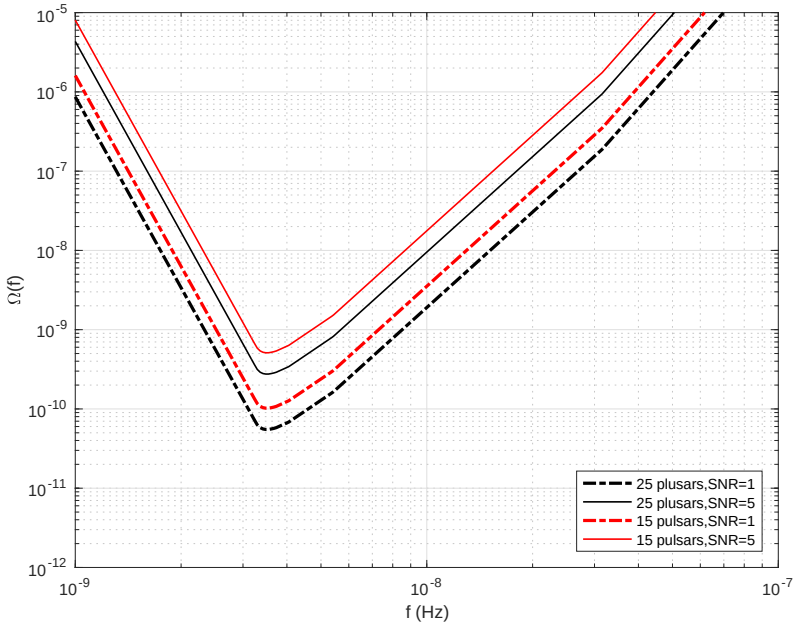

Fig. 2 Different power-law integrated sensitivity curves for different pulsar number data sets. The lowest sensitivity curve expressed by the dashed black line is based on the 10-yr time span data set in which 25 pulsars are included. The dash-dotted red line shows the sensitivity curve used 10-yr time span data set in which 15 pulsars are included. The solid lines show the same results but with $\rho = 5$.

Figure 4 plots the power-law integrated sensitivity curves for GWB using the 10-years subset, and the partially enlarged view of Figure 4 around 3×10^{-9} Hz. We have found that searches for GWB signals will benefit from the implementation with more accurate solar system ephemeris and referenced clock. Although from Figure 4, the curves drawn by changing the ephemeris and the referenced clock are almost coincident, from the partially enlarged view, the sensitivity enhancement of DE436 over that from pulsar experiments with DE405 can be observed in the lower-part of the integrated sensitivity curve. We can concluded that as the timing measurements' accuracy increases, the difference between the curves caused by changing the SSE and the reference clock will be more noticeable.

5 CONCLUSIONS

We discussed the impacts of two different correlated noises on the GWB detection. All the analyses are based on the power law sensitivity curve of GWB. First, we discussed the influences of two different reference atomic time scales in pulsar timing on the power law sensitivity curve of GWB. It was found that sensitivity would increase three times if one use TT (BIPM2018) as the reference time scale instead of TT (TAI). As TAI is realised by more than 500 atomic clocks at 76 laboratories for time keeping around the world, and is published monthly. Thus, TAI is in fact the most stable atomic time scale for a quasi real-time realization. Even so, if the reprocessed version TT (BIPM2018) serves as the reference time, the power law integrated sensitivity still has a significant improvement. Therefore, the influences of the reference clock error is nontrivial. Second, we analyzed the influences of different versions of DE ephemeris. Even though the sensitivity curve will be worse about two orders of magnitude based on DE200, the result based on DE405 is very close to the result based on DE421. Since the latest version of SSE is more and more accurate, the gaps between the two adjacent versions of SSE will get smaller and smaller. Therefore, in future the SSE error would be an unimportant factor on the detection sensitivity of GWB by PTAs.

For the real observation data sets, we selected millisecond pulsars from IPTA DR1 with time spans overlapping longer than 10 years , forming two pulsar subsets with different pulsar numbers separately. The timing residuals of all the participated pulsars are reserved according to the common duration, and the resulting wrms timing residuals serve as the timing noise to calculate the corresponding noise power spectral densities. First, we calculated the SNRs for the two pulsar subsets using the maximum amplitude of GWB permitted by observations and lower amplitude, respectively. It is proven that timing observations of more millisecond pulsars will indeed give an improvement in searching for stochastic GWB. Besides,

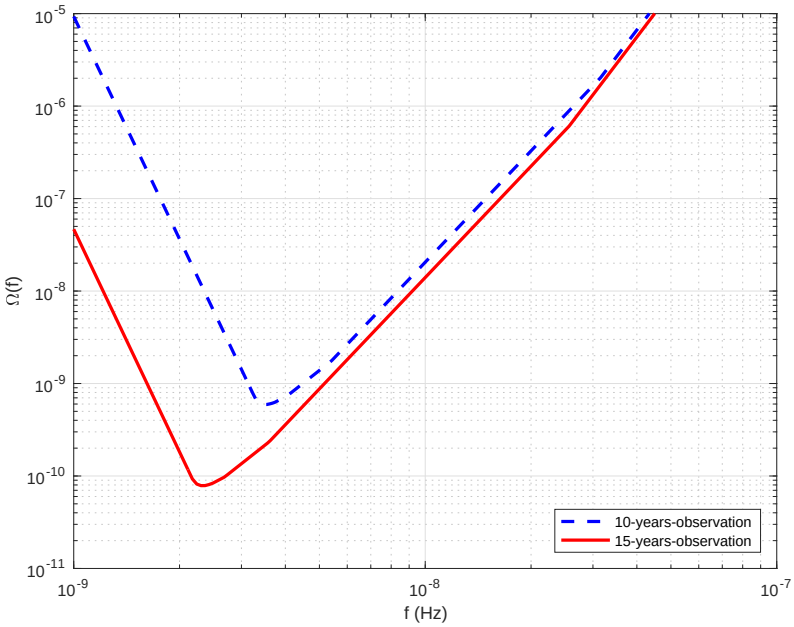

Fig. 3 Power law integrated sensitivity curves of the pulsars listed in Table 1 but with different observation durations. The blue dashed line represents the 10-years observation, and the red solid line represents the 15-years observation.

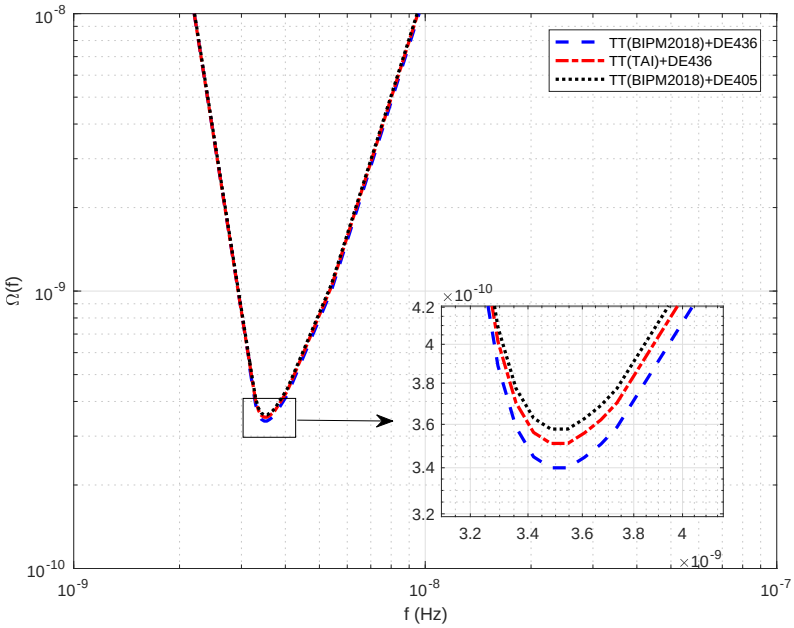

Fig. 4 Power law integrated sensitivity curve of 10-yr time span data set in which 25 pulsars are included, but calculated with different solar system ephemeris and referenced clock.

to estimate the benefit from the observation duration, we fixed 15 pulsars but with 10 years span and 15 years span, respectively. We found that the longer data set will gain an improvement in detection ability in all observed frequency bands. Therefore, the power law sensitivity curve will benefit from the observation strategy with arrays of more millisecond pulsars and longer observation time. Due to existences of unresolvable red noises, whether different reference times or different SSE models lead to little differences of the power-law sensitivity curves. However, this does not mean that the reference clock errors and SSE errors are not important for GWB detection.

In the analysis of NANOGrav 11-yr data set, the upper limit on a stochastic GWB was found to be sensitive to the SSE model assumed (Arzoumanian et al. 2018), which motivated new techniques such as BayesEphem (Vallisneri et al. 2020) to account for SSE uncertainties in the gravitational wave analysis. However, for the rapid improvement in pulsar timing precision, the more accurate reference time-scale and SSE seem to be increasingly important.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. U1831130 and U1531112), and the program of Youth Innovation Promotion Association CAS (2017450).

Appendix A: A COMPARISON OF CALCULATING THE INTEGRATED SENSITIVITY CURVES

In this appendix, we show a comparison between the results of the integrated sensitivity curves shown in Thrane & Romano (2013) and our results in this paper. We think that the result shown in Thrane & Romano (2013) (below we call it "Thrane's result" for short) overestimated the detection sensitivity of GWB. For comparison, we also set the cadence of the measurements for each pulsar to be 20 yr^{-1}, the root mean square of the white timing noise to be 100 ns identically and total observation time span $T = 5$, which are exactly the same as those employed in Thrane & Romano (2013). Therefore, the resulting power spectral density of the detector is exactly the same as each other, since it is only dependent on the cadence and the root mean square of the timing noise. However, the 20 pulsars we used are not completely the same as those employed in Thrane & Romano (2013). This difference will only lead to different *effective* number defined as follows,

$$\mathcal{N} \equiv \sum_{I=1}^{M} \sum_{J>1}^{M} \zeta_{IJ}^2, \tag{A.1}$$

where M is the total number of the participant pulsars, I and J denote the individual pulsars, and ζ_{IJ} is the Hellings-Downs function defined as Equation (9). The pulsar pairs of the 20 pulsars chosen from IPTA used in Thrane & Romano (2013) give an *effective* number $\mathcal{N} = 4.74$, while the pulsar pairs that we choose from PPTA give $\mathcal{N} = 4.73$. Due to Equation (14), this tiny difference will lead to negligible discrepancy of the integrated sensitivity curve.

For the calculations of the integrated sensitivity curve, we found that the results sensitively depended on the integration step during numerically integrating Equation (14). Thus, we tried many cases of the integration steps (denoted by df), and the corresponding results are shown in Figure A.1, where Thrane's result was also plotted for comparison. It is clear from Figure A.1 that different integration steps lead to different integrated sensitivity curves, especially for the most sensitive region. However, the integrated sensitivity curve will be convergent to a stable state once df is chosen to be small enough. On the contrary, if df is too large, the corresponding sensitivity curve will be overestimated. This can be explained as follows. For Equations(8)-(13), it is not hard to see that $\Omega_{\mathrm{eff}} \propto f^5$. Thus, the integration function will exhibit a red like spectrum for $\beta < 5$, i.e., the integration function $\propto f^a$ with $a < 0$. Therefore, for $\beta < 5$ and fixed integration limits, the integration implied in Equation (14) will be bigger than its true value if the integration step is set to be too large in the process of numerical calculation. Then, the sensitivity curve will be overestimated in turn. In Thrane's result, df was set to be $1/T$, which can be seen in their public code. So, we think the power-law integrated sensitivity curve in Thrane & Romano (2013) overestimated the sensitivity. There is another difference in that, the upper integration limit f_{max} hiding in Equation (14) was set to be $f_{\mathrm{max}} = 10^{-7}$ Hz, while $f_{\mathrm{max}} = 1/(2\Delta t)$ in our results. It is clear from Equation (14) that, a larger value of f_{max} will improve the detection sensitivity. This is the reason why Thrane's result is different from our result with the case of $df = 1/T$. In addition, the unintegrated sensitivity curve, $\Omega_{\mathrm{eff}}(f)$, was plotted in Figure A.1 for comparison. It can be seen that, even though the un-integrated sensitivity curve is lower than the integrated case at the lowest sensitivity region, the integrated sensitivity curve enhances the detectability at other regions. Moreover, it is interesting to find that, the most sensitive region is localized at f_{min} for the case of unintegrated sensitivity curve, while in the integrated case the most sensitive region will be localized around a frequency which is slightly larger than f_{min}.

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, Classical and Quantum Gravity, 32, 024001

Allen, B., & Romano, J. D. 1999, Phys. Rev. D, 59, 102001

Anholm, M., Ballmer, S., Creighton, J. D. E., Price, L. R., & Siemens, X. 2009, Phys. Rev. D, 79, 084030

Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2018, ApJ, 859, 47

Caballero, R. N., Guo, Y. J., Lee, K. J., et al. 2018, MNRAS, 481, 5501

Detweiler, S. L., & Szedenits, E., J. 1979, ApJ, 231, 211

Estabrook, F. B., & Wahlquist, H. D. 1975, General Relativity and Gravitation, 6, 439

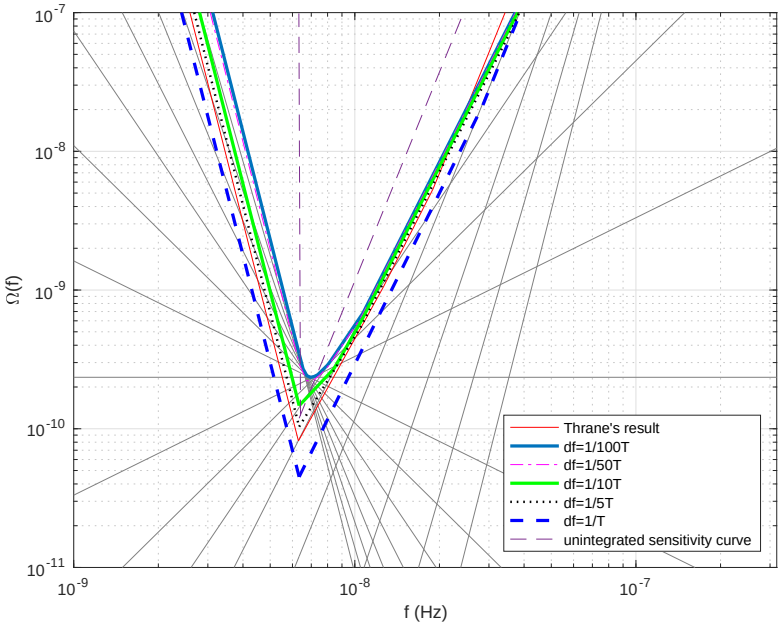

Fig. A.1 A comparison of the power-law integrated sensitivity curves between the results in Thrane & Romano (2013) and our results. The thin solid red line stands for Thrane's result. Our results of the power-law integrated sensitivity curves are shown with different line styles for the integration steps of $df = 1/T, 1/5T, 1/10T, 1/50T$, and $1/100T$, respectively. The dashed purple line shows the effective fractional energy density in noises $\Omega_{\rm eff}(f)$. $\rho = 5$, $T = 5$ yr and a cadence of 20 yr^{-1} are set throughout this figure.

Ferdman, R. D., van Haasteren, R., Bassa, C. G., et al. 2010, Classical and Quantum Gravity, 27, 084014

Grishchuk, L. P. 2001, Relic Gravitational Waves and Their Detection, eds. C. Lämmerzahl, C. W. F. Everitt, & F. W. Hehl, 562, (Springer-Verlag Berlin Heidelberg), 167

Hellings, R. W., & Downs, G. S. 1983, ApJL, 265, L39

Hobbs, G., Archibald, A., Arzoumanian, Z., et al. 2010, Classical and Quantum Gravity, 27, 084013

Hobbs, G., Coles, W., Manchester, R. N., et al. 2012, MNRAS, 427, 2780

Hobbs, G., Guo, L., Caballero, R. N., et al. 2020, MNRAS, 491, 5951

Jaffe, A. H., & Backer, D. C. 2003, ApJ, 583, 616

Jenet, F. A., Hobbs, G. B., Lee, K. J., & Manchester, R. N. 2005, ApJL, 625, L123

Jenet, F. A., Hobbs, G. B., van Straten, W., et al. 2006, ApJ, 653, 1571

Kaspi, V. M., Taylor, J. H., & Ryba, M. F. 1994, ApJ, 428, 713

Kuroyanagi, S., Miyamoto, K., Sekiguchi, T., Takahashi, K., & Silk, J. 2013, Phys. Rev. D, 87, 023522

LIGO Scientific Collaboration, Aasi, J., Abbott, B. P., et al. 2015, Classical and Quantum Gravity, 32, 074001

Manchester, R. N., Hobbs, G., Bailes, M., et al. 2013, PASA, 30, e017

Moore, C. J., Cole, R. H., & Berry, C. P. L. 2015, Classical and Quantum Gravity, 32, 015014

Perera, B. B. P., DeCesar, M. E., Demorest, P. B., et al. 2019, MNRAS, 490, 4666

Rajagopal, M., & Romani, R. W. 1995, ApJ, 446, 543

Sazhin, M. V. 1978, Soviet Ast., 22, 36

Sesana, A., Vecchio, A., & Colacino, C. N. 2008, MNRAS, 390, 192

Sesana, A., Vecchio, A., & Volonteri, M. 2009, MNRAS, 394, 2255

Shannon, R. M., Ravi, V., Lentati, L. T., et al. 2015, Science, 349, 1522

Siemens, X., Ellis, J., Jenet, F., & Romano, J. D. 2013, Classical and Quantum Gravity, 30, 224015

Thrane, E., & Romano, J. D. 2013, Phys. Rev. D, 88, 124032

Tiburzi, C., Hobbs, G., Kerr, M., et al. 2016, MNRAS, 455, 4339

Tinto, M., & Hartwig, O. 2018, Phys. Rev. D, 98, 042003

Tong, M.-L., Ding, Y.-H., Zhao, C.-S., et al. 2016, RAA (Research in Astronomy and Astrophysics), 16, 49

Vallisneri, M., Taylor, S. R., Simon, J., et al. 2020, ApJ, 893, 112

Verbiest, J. P. W., Bailes, M., Bhat, N. D. R., et al. 2010, Classical and Quantum Gravity, 27, 084015

Volonteri, M., Haardt, F., & Madau, P. 2003, ApJ, 582, 559

Zhang, Y., Yuan, Y., Zhao, W., & Chen, Y.-T. 2005, Classical and Quantum Gravity, 22, 1383

Zhao, W., Zhang, Y., You, X.-P., & Zhu, Z.-H. 2013, Phys. Rev. D, 87, 124012