

An independent test on the local position invariance of gravity with the triple pulsar PSR J0337+1715

Lijing Shao

Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany

E-mail: lijing.shao@aei.mpg.de

21 July 2017

Abstract. We design a direct test of the local position invariance (LPI) in the post-Newtonian gravity, using the timing observation of the triple pulsar, PSR J0337+1715. The test takes advantage of the large gravitational acceleration exerted by the outer white dwarf to the inner neutron star – white dwarf binary. Using machine-precision three-body simulations and dedicated Markov-chain Monte Carlo (MCMC) techniques with various sampling strategies and noise realizations, we estimate that the Whitehead’s parameter could have already been limited to $|\xi| \lesssim 0.4$ (95% CL), with the published timing data spanning from January 2012 to May 2013. The constraint is still orders of magnitude looser than the best limit, yet it is able to independently falsify Whitehead’s gravity theory where $\xi = 1$. In addition, the new test is immune to extra assumptions and involves full dynamics of a three-body system with a strongly self-gravitating neutron star.

PACS numbers: 04.80.Cc, 97.60.Gb, 04.25.-g

Keywords: gravitation, pulsar, strong equivalence principle, local position invariance

Submitted to: *Classical and Quantum Gravity*

1. Introduction

PSR J0337+1715, a millisecond pulsar (MSP) with a spin period $P \simeq 2.73$ ms, was discovered in a large-scale pulsar survey conducted with the Robert C. Byrd Green Bank Telescope (GBT) [1]. It is in a hierarchical triple system consisting of a neutron star (NS) with a gravitational mass $m_{\text{NS}} \simeq 1.44 M_{\odot}$, and two white dwarfs (WDs) with masses $m_{\text{WD,I}} \simeq 0.20 M_{\odot}$ and $m_{\text{WD,O}} \simeq 0.41 M_{\odot}$.[‡] The NS and the lighter WD are gravitationally bound as an *inner binary* with $P_{\text{b,I}} \simeq 1.63$ d that are, as a whole hierarchically bound to the outer WD with $P_{\text{b,O}} \simeq 327$ d. Two orbits are very circular with $e_{\text{I}} \simeq 6.9 \times 10^{-4}$ for the inner binary, and $e_{\text{O}} \simeq 3.5 \times 10^{-2}$ for the outer orbit. Two orbital planes are remarkably coplanar with an inclination $\lesssim 0.01^{\circ}$ due to the three-body dynamics in the formation of the system [2]. The apsides of two orbits are aligned with a difference $\lesssim 2^{\circ}$ due to the secular effects of the three-body interaction [1]. The 3-dimensional spatial trajectory of the inner binary for a time span slightly shorter than $P_{\text{b,O}}$ is illustrated in Figure 1.

The triple pulsar is identified immediately as a superb celestial laboratory to test the strong equivalence principle (SEP) by investigating the difference in the inertia and gravitational masses of the pulsar [1]. Later on such a test is extended to probe the equivalence between the passive and active gravitational masses as well, that will represent the first test of Newton’s third law with compact objects [3, 4]. Indeed, SEP is the founding principle of general relativity (GR) that deserves the strictest examination from every angle, to establish its precision as well as to look for new physics beyond GR. Will [5, 6] summarized the equivalence principles in gravity theories in a hierarchical way, from the weak equivalence principle (WEP) to the Einstein equivalence principle (EEP), and then to SEP, where the last one can be decomposed into three parts,

- the universality of free fall (UFF),
- the local Lorentz invariance (LLI),
- the local position invariance (LPI),

for non-self-gravitating bodies as well as for self-gravitating bodies. SEP describes the general rules for the outcome of gravitational experiments. It is indeed lying to the heart of GR, and actually, there are arguments that among the viable gravity theories, GR is the only one that respects SEP in its entirety [6]. Therefore, probing the building blocks of SEP probes the deepest foundational principle of GR [7].

The tests in the equivalence of different masses pertains to tests of UFF [1, 3]. In this paper we will study the possibility of using PSR J0337+1715 to test the LPI in gravity, which is another important ingredient of SEP. The motivation of the study is similar to the tests of UFF. The outer WD provides a substantial gravitational environment for the inner binary that is valuable to study UFF and LPI in gravity theories [8, 9, 10, 7].

[‡] In this paper, “PSR J0337+1715”, “the triple pulsar”, and “the neutron star” are used to refer to the pulsar, while “the triple system” and “the pulsar system” are used to refer to the three-body system.

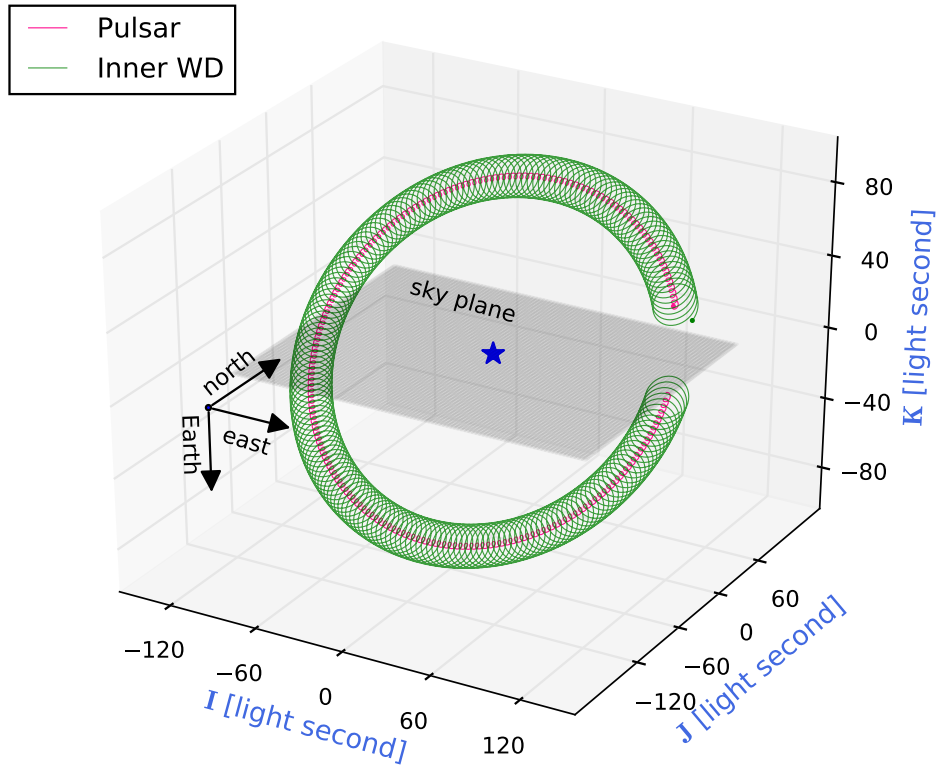


Figure 1. Illustration of the triple pulsar system in the $(\hat{\mathbf{I}}, \hat{\mathbf{J}}, \hat{\mathbf{K}})$ coordinate system where, the origin (marked as a blue star) is chosen to be the center of inertial masses of the triple system, $\hat{\mathbf{K}}$ points from the Earth to PSR J0337+1715, $\hat{\mathbf{I}}$ and $\hat{\mathbf{J}}$ respectively point to east and north in the sky plane. The longitude of ascending node for the outer orbit, which is generally not an observable in pulsar timing, is assumed to be $\Omega_{\text{O}} = 0^\circ$; for other nonzero values, a rotation around $\hat{\mathbf{K}}$ for an angle Ω_{O} is needed. The trajectories of the pulsar (in pink) and the inner WD (in green) start on MJD 55920.0 (December 25, 2011), and end on MJD 56233.9 (November 2, 2012) when the pulsar was ascending. The starting locations are indicated by small dots. The orbit of the outer WD is not shown here. These trajectories are integrated with the REBOUND package (see section 3).

In post-Newtonian gravity, the tests of UFF and LPI with self-gravitating bodies were done with the Sun-Earth-Moon system [11], and with the Milky Way – binary-pulsar systems [12, 13]. These studies would benefit greatly if the third body were exerting a larger gravitational effect on the other two bodies. The triple system precisely provides an ideal realization of this requirement in Nature. In addition, there is a strongly self-gravitating body involved (namely, the NS), some strong-field aspects could be studied, which is nearly impossible with weakly self-gravitating bodies alone (e.g. in the lunar laser ranging experiments [11]). Moreover, compared with tests using the Milky Way – binary-pulsar systems, the triple system probes a *dynamical* regime where the third body (namely the outer WD) does *react* to the gravitational dynamics. Although the limit on LPI violation that could have been provided by the triple pulsar is found to

be orders of magnitude looser than the best limit from solitary pulsars [19, 6], we still feel it worthy as an *independent* limit and it has various extra merits compared with previous studies (see discussions in section 4).

The paper is organized as follows. In the next section, we present some theoretical details for the orbital dynamics of PSR J0337+1715 in presence of LPI violation, in the parameterized post-Newtonian (PPN) gravity. Then in section 3 we simulate various mock time-of-arrival (TOA) data closely following the real observation in Ref. [1] (see Table 1) with a machine-precision N -body integrator. Dedicated parameter estimation using Markov-chain Monte Carlo (MCMC) techniques that take account of full correlations (of 17 parameters) is performed on these mock TOAs, and from MCMC chains we conservatively estimate the precision of the test. Section 4 discusses the relevance of the results, and makes some comparisons with limits obtained elsewhere.

2. Triple pulsar system with LPI violation

PPN formalism is the most popular framework in experimental gravity for testing alternative gravity theories. In PPN gravity theories are parametrized with ten generic PPN parameters that represent different physical properties of gravitation and take different values in different theories [14, 15, 5]. There also exists a generic framework called the standard-model extension (SME) [30, 31] which, using an effective field theory, casts the possible deviations from GR into new operators that can host anisotropic terms as well. Due to the large number of coefficients in SME, here we will focus on the LPI-violating PPN parameter, ξ . It is also called the Whitehead's parameter because of its first appearance in Whitehead's parameter-free gravity theory [16, 17, 18]. In GR, $\xi = 0$, and in Whitehead's gravity theory $\xi = 1$. The original Whitehead's gravity theory [16] was disproved by many experiments by now [17, 19, 6] (see Ref. [18] for an excellent review), but due to the importance of the SEP it does not hurt to have another *independent* test with distinct merits.

In the PPN formalism, one has an extra LPI-violating term in the N -body Lagrangian [5, 19],

$$L_\xi = -\frac{\xi}{2} \frac{G^2}{c^2} \sum_{i,j} \frac{m_i m_j}{r_{ij}^3} \mathbf{r}_{ij} \cdot \left[\sum_k m_k \left(\frac{\mathbf{r}_{jk}}{r_{ik}} - \frac{\mathbf{r}_{ik}}{r_{jk}} \right) \right], \quad (1)$$

where $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$, $r_{ij} \equiv |\mathbf{r}_{ij}|$, and the summation excludes terms that make any denominator vanish. In the Hamiltonian formalism, for a triple system like PSR J0337+1715 where finite-size effects are subdominant, L_ξ corresponds to an extra term in the interaction potential,

$$V_\xi = \xi \frac{G^2}{c^2} m_0 m_1 m_2 \left(\frac{\mathbf{r}_{01} \cdot \mathbf{r}_{12}}{r_{01}^3 r_{20}} + \frac{\mathbf{r}_{01} \cdot \mathbf{r}_{12}}{r_{12}^3 r_{20}} + \frac{\mathbf{r}_{01} \cdot \mathbf{r}_{20}}{r_{01}^3 r_{12}} + \frac{\mathbf{r}_{01} \cdot \mathbf{r}_{20}}{r_{20}^3 r_{12}} + \frac{\mathbf{r}_{12} \cdot \mathbf{r}_{20}}{r_{12}^3 r_{01}} + \frac{\mathbf{r}_{12} \cdot \mathbf{r}_{20}}{r_{20}^3 r_{01}} \right), \quad (2)$$

where in our case we use subscripts, 0, 1, 2, refer to the NS, the inner WD, and the outer WD, respectively. Due to the hierarchical structure for the triple system that we are considering, we have $r_{01} \ll r_{12} \sim r_{20}$. Therefore, the first and the third terms in the

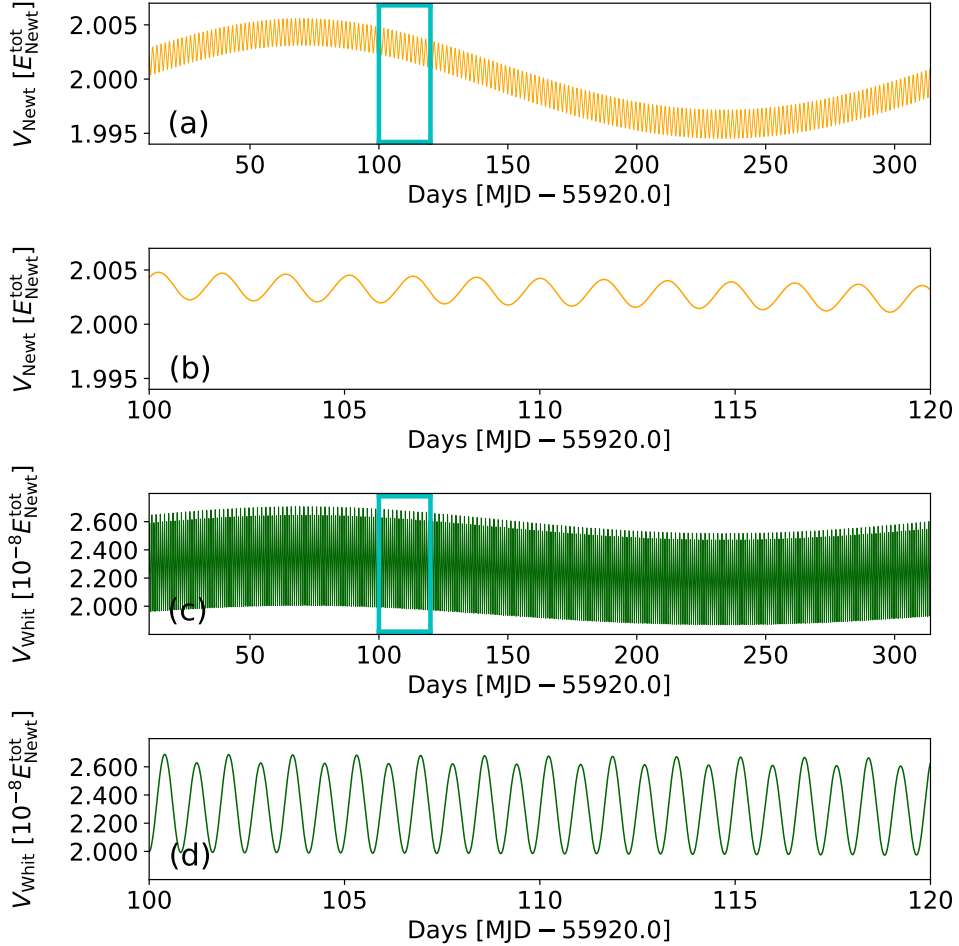


Figure 2. Panels (a) and (c) show respectively the Newtonian and the Whitehead potentials versus time for the triple pulsar system. The time span is the same as that of Figure 1, namely MJD 55920.0–56233.9. Panels (b) and (d) show a magnified view of the framed regions in panels (a) and (c) respectively.

parentheses contribute prominently to V_ξ . Nevertheless, we include all contributions in Eq. (2) in our numerical study below.

In Figure 2, we plot the Newtonian potential, $V_{\text{Newt}} \equiv \sum_{i \neq j} -\frac{Gm_i m_j}{2r_{ij}}$, and the Whitehead potential, $V_{\text{Whit}} \equiv V_\xi(\xi = 1)$, in the unit of total Newtonian orbital energy, $E_{\text{Newt}}^{\text{tot}} \equiv V_{\text{Newt}} + \sum_i \frac{1}{2} m_i v_i^2$ for the triple system. The total Newtonian orbital energy $E_{\text{Newt}}^{\text{tot}} \simeq -8.3 \times 10^{46}$ erg is conserved if the gravitational dissipation is ignored. According to the virial theorem, one has $\langle V_{\text{Newt}} \rangle = 2 \langle E_{\text{Newt}}^{\text{tot}} \rangle$ for Newtonian gravity, where $\langle \cdot \rangle$ denotes an average over time. As expected, V_{Newt} oscillates with a short timescale, $P_{\text{b,I}}$, and the oscillation is modulated with a large timescale, $P_{\text{b,O}}$. In contrast, V_{Whit} oscillates around its average value, $\langle V_{\text{Whit}} \rangle \simeq -2.0 \times 10^{39}$ erg, with a short timescale, $\frac{1}{2} P_{\text{b,I}}$. A faster oscillation for V_{Whit} can be understood based on the facts that, the dominant contribution to V_{Newt} comes from $-\frac{Gm_0 m_1}{r_{01}} \propto \frac{1}{r_{01}}$, while the dominant contributions to V_{Whit} come from the first and the third terms in the parentheses in Eq. (2), both

behaving as $\propto \frac{1}{r_{01}^2}$. There is less long-term *modulation* for V_{Whit} , compared with that of V_{Newt} , at the timescale of the outer orbit [see panels (a) and (c) in Figure 2]. The relative smallness in the modulation of V_{Whit} is understood that V_{Whit} is at the first post-Newtonian order which receives a smaller contribution from a wider orbit, by roughly a factor of $(\mathcal{V}_0^2/c^2) / (\mathcal{V}_1^2/c^2) = \mathcal{V}_0^2/\mathcal{V}_1^2 \simeq 0.03$ relative to the Newtonian order, where \mathcal{V}_1 and \mathcal{V}_0 are the characteristic *relative* velocities of the inner and the outer orbits.

Speaking of equations of motion for the triple system, for body $i \in \{0, 1, 2\}$ one has an extra acceleration term,

$$\delta \mathbf{a}_\xi^i = -\frac{1}{m_i} \nabla^i V_\xi(\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2), \quad (3)$$

in addition to the acceleration in the LPI-invariant gravity theory. In Eq. (3), ∇^i is the gradient operator for body i with its coordinate vector \mathbf{r}_i ; no summation is assumed for the *body index* i in Eq. (3). The explicit expressions of $\delta \mathbf{a}_\xi^i$ are tedious yet not inspiring, therefore, we do not give them here.

3. Mock simulation and parameter estimation

The timing data presented in Ref. [1] include radio observations with the GBT, the Arecibo telescope, and the Westerbork Synthesis Radio Telescope (WSRT), spanning from MJD 55930.9 to 56436.5 (~ 1.4 yr, or $\sim 4 \times 10^7$ s). Within such a duration, the inner and outer orbits revolve ~ 300 and ~ 1.5 cycles, respectively, while the pulsar spins ~ 16 billion cycles. The timing solution in Ref. [1] (see the second column in Table 1) is derived with $N_{\text{TOA}}^{(0)} = 26280$ TOAs and results in a weighted root mean squared residual $\sigma_{\text{TOA}}^{(0)} = 1.34 \mu\text{s}$.

Since the TOA data are not publicly available, we simulate mock TOAs closely following the observational characteristics, as was done in Ref. [3]. It is well known that there is no general analytical solution for the three-body problem in gravity given by algebraic expressions and integrals [20, 21]. To produce mock TOAs, we resort to a machine-precision N -body numerical integrator developed by Rein and his collaborators, which is implemented in the REBOUND package§ [22]. Specifically, we use a 15th-order integrator basing on the Gauß-Radau quadrature, IAS15, that uses adaptive time stepping, and keeps systematic errors well below machine precision over 10^9 orbits [23]. Such a precision meets the requirement posed by the accurate pulsar timing data [3].

For all simulations, we use the parameters of PSR J0337+1715 reported in Ref. [1] (see Table 1). Initial conditions for the system are worked out for MJD 55920.0 which is the reference epoch for all parameters. The LPI-violating modification in Eq. (3) is augmented to the IAS15 integrator [23]. We evolve the system in 3-dimensional space for a longer time than the observational span and cut data keeping the part that corresponds to the real observation. A spindown model for the pulsar is adopted with $f(t) = f_0 + \dot{f}t$ where t is the coordinate time. The Römer delay is obtained by

§ <https://github.com/hannorein/rebound>

projecting the 3-dimensional orbit of the pulsar onto the line of sight to the Earth [24]. Relativistic delays, e.g. the Einstein delay and the Shapiro delay, are ignored, due to the fact that they are not observable in PSR J0337+1715 yet [1]. An exception is the transverse Doppler delay due to the cross term of the velocities of two orbits, which is approximated for PSR J0337+1715 as $R(t) \simeq \frac{1}{c^2} \mathbf{x}_I(t) \cdot \mathbf{v}_O(t)$, where $\mathbf{x}_I(t)$ and $\mathbf{v}_O(t)$ are respectively the position vector of the inner orbit and the velocity vector of the outer orbit. It was shown in Ref. [3] that indeed it is a reasonable approximation. The end product of one integration is TOAs in the form of $N(t)$ with N the counting number of pulses.

As mentioned before, $N_{\text{TOA}}^{(0)} = 26280$ TOAs were collected with $\sigma_{\text{TOA}}^{(0)} = 1.34 \mu\text{s}$. Due to the computational cost in MCMC runs (see below), in our mock simulations we rescale it by a factor of 9 to $N_{\text{TOA}} = N_{\text{TOA}}^{(0)}/9 = 2920$ and $\sigma_{\text{TOA}} = \sigma_{\text{TOA}}^{(0)}/\sqrt{9} \simeq 0.447 \mu\text{s}$. Such a rescaling is reasonable because on average it still keeps about 10 TOAs per inner orbit.

Two strategies are used to sample TOAs [3],

- *uniform sampling*: $N_{\text{TOA}} = 2920$ TOAs are generated uniformly in time;
- *step sampling*: $N_{\text{TOA}} = 2920$ TOAs are generated with fake observing blocks once per week within which TOAs are separated by 10 seconds.

We consider the *step sampling* method more closely resembles the real observation, however, the sensitivities to LPI violation from two methods are extremely consistent (see below). In total, we simulate five noise realizations for each method, named as “TOA. k ” with $k \in \{0, 1, 2, 3, 4\}$ for *uniform sampling* and $k \in \{5, 6, 7, 8, 9\}$ for *step sampling*. Noises in TOAs are generated according to a Gaussian random number generator $\mathcal{N}(0, \sigma_{\text{TOA}}^2)$, and they are added to the “noiseless” TOAs to obtain mock TOAs. For different set of mock TOAs, the noise generation is independent to each other.

Mock TOAs are generated with the Whitehead’s parameter $\xi = 0$, with noises $\sigma_{\text{TOA}} \simeq 0.447 \mu\text{s}$ added homogeneously in the uncorrelated Gaussian form. We want to estimate the extent of these TOAs in constraining ξ . To achieve this task, following Refs. [1, 3], MCMC runs are set up to estimate 17 parameters, $\boldsymbol{\theta} \equiv \boldsymbol{\theta}_{\text{spin}} \cap \boldsymbol{\theta}_{\text{orbit}} \cap \boldsymbol{\theta}_{\xi}$, simultaneously in the model. Parameters in $\boldsymbol{\theta}$ include 2 spindown parameters in $\boldsymbol{\theta}_{\text{spin}}$, 14 orbital parameters in $\boldsymbol{\theta}_{\text{orbit}}$ (see Table 1 for definition of symbols), and 1 LPI-violating parameter in $\boldsymbol{\theta}_{\xi}$,

$$\boldsymbol{\theta}_{\text{spin}} \equiv \left\{ f_0; \dot{f} \right\}, \quad (4)$$

$$\boldsymbol{\theta}_{\text{orbit}} \equiv \left\{ (a \sin i)_I; P_{b,I}; \epsilon_{1,I}; \epsilon_{2,I}; T_{\text{asc},I}; (a \sin i)_O; P_{b,O}; \epsilon_{1,O}; \epsilon_{2,O}; T_{\text{asc},O}; (a \cos i)_I; (a \cos i)_O; q_I; \delta\Omega \right\}, \quad (5)$$

$$\boldsymbol{\theta}_{\xi} \equiv \{ \xi \}. \quad (6)$$

The PYTHON package of an affine-invariant MCMC ensemble sampler [25, 26], EMCEE||, is used to sample the 17-dimensional $\boldsymbol{\theta}$ -space. This algorithm has better performance

over traditional MCMC sampling methods (e.g., the traditional Metropolis-Hasting method), as measured by the smaller autocorrelation time and fewer hand-tuning input parameters. It transforms the sampling of the parameter space by an affine transformation such that the internal algorithm samples an isotropic density, and the efficiency is not limited by possibly large covariances among parameters [25, 26]. We use uniform priors for all parameters in $\boldsymbol{\theta}$, and choose good starting values to reduce computational cost. Convergence tests are performed in post-processing to ensure that the starting values do not influence our parameter estimation. The ranges of parameters are not limited, thus in principle they can take any values as long as they have support from the likelihood. For each mock TOA dataset, noiseless template TOAs are generated on the fly at every MCMC step according to $\boldsymbol{\theta}$ that is sampled by the kernel. These noiseless *template TOAs* are compared with the noisy *mock TOAs*, namely “TOA. k ” with $k \in \{0, 1, \dots, 9\}$. The kernel proceeds the sampling of $\boldsymbol{\theta}$ based on the difference between *template TOAs* and *mock TOAs*, characterized by $\chi^2(\boldsymbol{\theta})$. With this setting, the posterior is directly proportional to $e^{-\chi^2(\boldsymbol{\theta})/2}$.

In total, we have 10 MCMC runs for two sampling methods each with five different noise realizations. With EMCEE, 44 walkers are adopted for each run, and 44 chains for each set of mock TOAs are accumulated as the end product of MCMC runs for post-processing. There are about 300000 posterior samples for each mock dataset, and the first half of them are discarded as the BURN-IN phase [27]. The Gelman-Rubin statistic is used to assess the convergence of different chains [28], which tells that the runs have forgotten the starting values and are in equilibrium.

As was already demonstrated in Ref. [3], if the *template TOAs* use GR dynamics, the mock TOAs can reproduce the observational *uncertainties* of all orbital parameters in $\boldsymbol{\theta}_{\text{orbit}}$ within a factor of two for 13 parameters and a factor of three for the remaining one, while they underestimate the *uncertainties* of $\boldsymbol{\theta}_{\text{spin}}$. In the current case, the recovering template is a LPI-violating template with one extra degree of freedom, therefore, we expect to produce larger uncertainties, at least for some variables that are strongly correlated with ξ . This is indeed the case, as shown in Figure 3 for the dataset “TOA.3” as an example. We see that ξ strongly correlates with orbital parameters that pertain to the outer orbit with correlation coefficients $\rho = 0.98\text{--}0.99$, while it has relatively smaller correlations with the other parameters. The correlation with the parameters of the outer orbit makes the (marginalized) uncertainties of these parameters (see the third column of Table 1) larger than what was reported in Ref. [1] where the recovering template is LPI-invariant. The reason for large correlation, we suspect, is that the observational span only covers about 1.5 cycles for the outer orbit, which makes parameter estimation for these elements rather uncertain.¶ Worthy to mention

¶ Ideally it will be rather rewarding to study the parameter-estimation problem here with longer mock datasets, say, with $\gtrsim 6$ outer orbits, to investigate the reduction in the correlations between ξ and the outer orbital parameters, and the improvement in constraining ξ . However, currently we are limited by the speed of the three-body integration and the high dimensionality of the parameter space, thus a high computational cost in the MCMC runs. We hope the code can be speeded up in the future, and

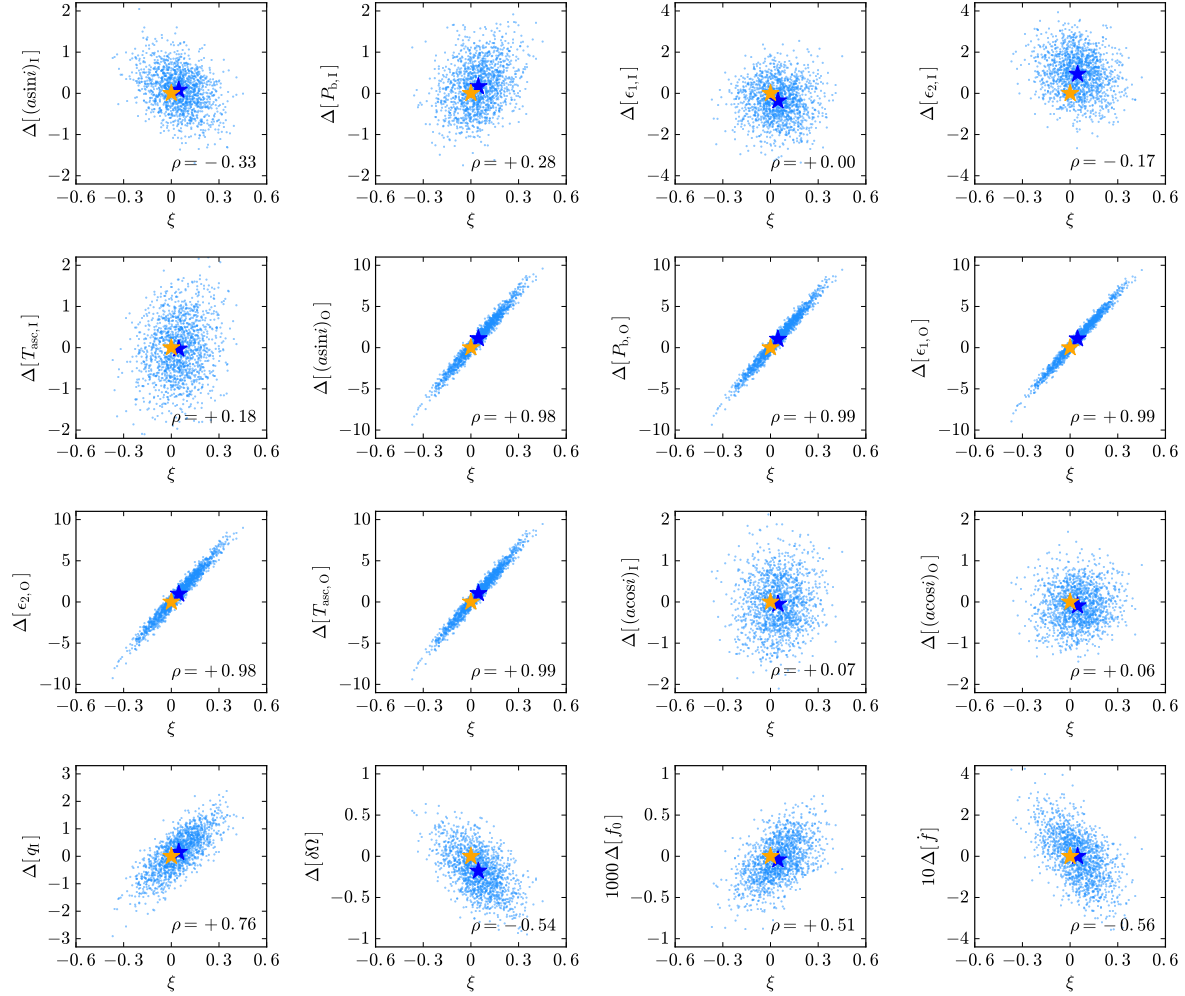


Figure 3. Correlations between ξ and the other 16 parameters from the parameter estimation of the mock dataset “TOA.3”. The quantity $\Delta[\mathcal{X}] \equiv (\mathcal{X} - \mathcal{X}^{(0)}) / \sigma_{\mathcal{X}}^{(0)}$, where $\mathcal{X}^{(0)}$ and $\sigma_{\mathcal{X}}^{(0)}$ are the values and uncertainties reported in Ref. [1] (see Table 1). Orange and blue stars are the *injected* and *recovered* values. The correlation coefficients are given at the bottom right corner in each panel. Only 1% of MCMC samples are shown for clarity.

that, in Figure 3 the parameter estimation recovers the *injected* parameters very well, as marked with stars. The parameter-estimation chains with other mock datasets, “TOA. k ” ($k \neq 3$), have similar results.

In Figure 4, the marginalized 1-dimensional posterior densities of ξ for all 10 runs are given by normalized histograms. We can see that though with different sampling strategies and different noise realizations, the posteriors on ξ are rather consistent, in terms of their means and variances. In real data, unlike the cases in our mock datasets where several noise realizations are simulated, only one noise realization is working and we only have one dataset. To make a conservative estimate on the expected constraint address this important question.

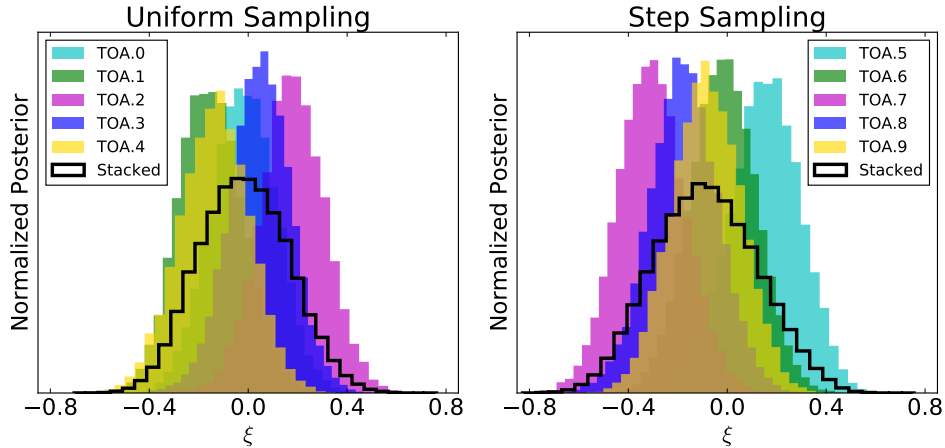


Figure 4. Marginalized and normalized 1-dimensional posterior densities for ξ from the parameter estimation of mock datasets with *uniform sampling* (left) and *step sampling* (right). The *stacked* posterior densities simply stack the posterior samples from different noise realizations with the same weight.

on ξ from PSR J0337+1715, we stack the posterior samples in each sampling method, and it is shown with black histograms in Figure 4. From the stacked posterior densities, we obtain a conservative sensitivity of PSR J0337+1715 in probing the LPI of gravity,

$$|\xi| \lesssim 0.4, \quad (95\% \text{ CL}), \quad (7)$$

for both sampling methods.

4. Discussion

Besides the generic value of ξ in the PPN framework [5, 6], and the specific example of Whitehead’s gravity theory ($\xi = 1$) [16], a class of theories called “quasilinear” theories of gravity could have a nonzero ξ [17]. In these theories, the PPN parameter β , that measures the nonlinearity in the superposition law for gravity, equals to ξ ; therefore the limit on ξ can be cast as a limit on β as well in these theories. Furthermore, as noted in Ref. [19], the constraint on ξ might also limit parameters in the *anisotropic* PPN framework of Ref. [29], and in the gravity sector of SME [30, 31, 32, 33, 34]. If the code can be speeded up significantly thus the computational cost can be reduced significantly in the future, similar analysis with multiple non-GR parameters will also be possible (for example, in the SME framework).

Compared with previous observational constraints on the Whitehead’s parameter [17, 35, 36, 37, 19], the expected constraint in Eq. (7) is worse than the current best limit from solitary pulsars [19] by orders of magnitude.⁺ Nevertheless, it has its own virtue. Firstly, it is a totally *independent* limit yet it is able to rule out the Whitehead’s

⁺ If the limit on ξ is converted to a limit on the anisotropy of the gravitational constant [5, 19], it is worse than the current best limit from solitary pulsars [19] by the same orders of magnitude.

gravity theory [16] alone, that adds to the “multiple deaths” [18] of that theory. Secondly, previous limits generally involve extra assumptions, for examples, the alignment of the Solar spin with the angular orbital momentum of the Solar system five billion years ago [36, 19] or the statistical assumptions of unknown angles in the cases of binary pulsars [37] and solitary pulsars [19]. The test proposed here is immune to extra assumptions. Thirdly, compared with the limits from gravimeters on Earth [17, 35] and the lunar laser ranging experiment [11], here we have a strongly self-gravitating body involved. For some gravity theories, for example the scalar-tensor gravity [38, 44], strong fields will amplify deviations from GR nonperturbatively. Even in the case of a perturbative expansion in the compactness \mathcal{C} , ξ might have a linear dependence on \mathcal{C} , as in the case of the Nordtvedt parameter [39, 5]. If strong fields are relevant, then the limit in Eq. (7) could have a large relative merit over weak-field ones due to the large compactness of the NS, $\mathcal{C}_{\text{NS}} \simeq 0.1$. Fourthly, compared with experiments that use the Milky Way – binary-pulsar systems, here it is a *dynamical* three-body system where the third body (the outer WD) reacts to the gravitational dynamics. Although it is not clear yet but we suspect that this might have a standing for some specific gravity theories.

The estimation in Eq. (7) is obtained with mock data simulated closely following the real observation for a time span about 1.4yr. In reality, new observations have accumulated more data with probably better qualities. Up to the time of writing, about 6 outer orbits are covered, compared with ~ 1.5 orbits used in the simulation. This will vastly break parameter degeneracy and reduce the strong correlations ($\rho = 0.98$ – 0.99) seen in Figure 3 with the elements of the outer orbit. Thus these new data are expected to give an even tighter limit than that by a naive rescaling. In future, new radio telescopes like the Five-hundred-meter Aperture Spherical Telescope (FAST) [40] and the Square Kilometre Array (SKA) [41, 42] will provide better sensitivities in obtaining TOAs for the triple system. Moreover, although triple pulsars are rare, there is a chance of hosting about 100 such systems in the Milky Way [1], and the SKA is going to discover almost all of them [43]. If an even tighter triple system is discovered, a better test of LPI in gravity could be conducted.

Lastly, we want to stress that, although our mock TOAs are able to reproduce major features of the observation in Ref. [1], they are nevertheless simplified compared with the complications in real data, e.g., the heteroscedasticity in TOAs from different telescopes (the GBT, the Arecibo telescope, and the WSRT), the removal of (probably time-dependent) interstellar dispersion, the irregular jumps between different observing sessions, and so on. Also, there will be correlations of parameters with the parallax and the proper motion of PSR J0337+1715, which could be resolved with the Very Long Baseline Array (VLBA) [1]. We are not expecting that simulated mock data to cover all these observational facts. Nevertheless, we believe that the simulations have include major features of the triple system. An analysis with the real data could settle the result firmly. We hope the analysis done here will simulate observers to analyze the real data to test LPI.

Acknowledgements

The author thanks Roland Haas and Hanno Rein for help with codes, and Vivien Raymond for discussion on statistics. The Markov-chain Monte Carlo runs were performed on the VULCAN cluster at the Albert Einstein Institute in Potsdam-Golm.

References

- [1] S. M. Ransom, I. H. Stairs, A. M. Archibald, et al. A millisecond pulsar in a stellar triple system. *Nature*, 505:520, 2014.
- [2] T. M. Tauris and E. P. J. van den Heuvel. Formation of the galactic millisecond pulsar triple system PSR J0337+1715 – a neutron star with two orbiting white dwarfs. *The Astrophysical Journal*, 781:L13, 2014.
- [3] L. Shao. Testing the strong equivalence principle with the triple pulsar PSR J0337+1715. *Physical Review D*, 93:084023, 2016.
- [4] H. Bondi. Negative mass in general relativity. *Reviews of Modern Physics*, 29:423, 1957.
- [5] C. M. Will. *Theory and Experiment in Gravitational Physics*, Cambridge: Cambridge University Press, 1993.
- [6] C. M. Will. The Confrontation between general relativity and experiment. *Living Reviews in Relativity*, 17:4, 2014.
- [7] L. Shao and N. Wex. Tests of gravitational symmetries with radio pulsars. *Science China – Physics, Mechanics & Astronomy*, 59:699501, 2016.
- [8] P. C. C. Freire, M. Kramer, and N. Wex. Tests of the universality of free fall for strongly self-gravitating bodies with radio pulsars. *Classical and Quantum Gravity*, 29:184007, 2012.
- [9] N. Wex. Testing relativistic gravity with radio pulsars. arXiv:1402.5594, 2014.
- [10] E. Berti, E. Barausse, V. Cardoso, et al. Testing general relativity with present and future astrophysical observations. *Classical and Quantum Gravity*, 32:243001, 2015.
- [11] J. G. Williams, S. G. Turyshev, and D. H. Boggs. Lunar laser ranging tests of the equivalence principle. *Classical and Quantum Gravity*, 29:184004, 2012.
- [12] I. H. Stairs, A. J. Faulkner, A. G. Lyne, et al. Discovery of three wide-orbit binary pulsars: Implications for binary evolution and equivalence principles. *The Astrophysical Journal*, 632:1060, 2005.
- [13] W. W. Zhu, I. H. Stairs, P. B. Demorest, et al. Testing theories of gravitation using 21-year timing of pulsar binary J1713+0747. *The Astrophysical Journal*, 809:41, 2015.
- [14] C. M. Will and K. Nordtvedt. Conservation laws and preferred frames in relativistic gravity. I. Preferred-frame theories and an extended PPN formalism. *The Astrophysical Journal*, 177:757, 1972.
- [15] K. Nordtvedt and C. M. Will. Conservation laws and preferred frames in relativistic gravity. II. Experimental evidence to rule out preferred-frame theories of gravity. *The Astrophysical Journal*, 177:775, 1972.
- [16] A. N. Whitehead. *The Principle of Relativity*, Cambridge: Cambridge University Press, 1922.
- [17] C. M. Will. Relativistic gravity in the Solar System. III. Experimental disproof of a class of linear theories of gravitation. *The Astrophysical Journal*, 185:31, 1973.
- [18] G. Gibbons and C. M. Will. On the multiple deaths of Whitehead’s theory of gravity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 39:41, 2008.
- [19] L. Shao and N. Wex. New limits on the violation of local position invariance of gravity. *Classical and Quantum Gravity*, 30:165020, 2013.
- [20] H. Bruns. *Acta Math.*, 11:25, 1887.
- [21] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*. Paris: Gauthier-Villars et fils, 1892.

- [22] H. Rein and S. Liu. REBOUND: An open-source multi-purpose N-body code for collisional dynamics. *Astronomy & Astrophysics*, 537:A128, 2012.
- [23] H. Rein and D. S. Spiegel. IAS15: A fast, adaptive, high-order integrator for gravitational dynamics, accurate to machine precision over a billion orbits. *Monthly Notices of the Royal Astronomical Society*, 446:1424, 2015.
- [24] J. H. Taylor. Pulsar timing and relativistic gravity. *Philosophical Transactions: Physical Sciences and Engineering*, 341:117, 1992.
- [25] J. Goodman and J. Weare. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.*, 5:65, 2010.
- [26] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. Emcee : The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125:306, 2013.
- [27] S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng. *Handbook of Markov Chain Monte Carlo*. London: Chapman and Hall/CRC, 2011.
- [28] A. Gelman and D. B. Rubin. Inference from Iterative simulation using multiple sequences. *Statistical Science*, 7:457, 1992.
- [29] K. Nordtvedt. Anisotropic parametrized post-Newtonian gravitational metric field. *Physical Review D*, 14:1511, 1976.
- [30] V. A. Kostelecký. Gravity, Lorentz violation, and the standard model. *Physical Review D*, 69:105009, 2004.
- [31] Q. G. Bailey and V. A. Kostelecký. Signals for Lorentz violation in post-Newtonian gravity. *Physical Review D*, 74:045001, 2006.
- [32] V. A. Kostelecký and J. D. Tasson. Matter-gravity couplings and Lorentz violation. *Physical Review D*, 83:016013, 2011.
- [33] L. Shao. New pulsar limit on local Lorentz invariance violation of gravity in the standard-model extension. *Physical Review D*, 90:122009, 2014.
- [34] L. Shao. Tests of local Lorentz invariance violation of gravity in the standard model extension with pulsars. *Physical Review Letters*, 112:111103, 2014.
- [35] R. J. Warburton and J. M. Goodkind. Search for evidence of a preferred reference frame. *The Astrophysical Journal*, 208:881, 1976.
- [36] K. Nordtvedt. Probing gravity to the second post-Newtonian order and to one part in 10^7 using the spin axis of the sun. *The Astrophysical Journal*, 320:871, 1987.
- [37] L. Shao, N. Wex, and M. Kramer. New tests of local Lorentz invariance and local position invariance of gravity with pulsars. *Proceedings of the Thirteenth Marcel Grossmann Meeting on General Relativity*, Singapore: World Scientific, p. 1704, 2015. arXiv:1211.6558
- [38] T. Damour and G. Esposito-Farèse. Nonperturbative strong-field effects in tensor-scalar theories of gravitation. *Physical Review Letters*, 70:2220, 1993.
- [39] K. Nordtvedt. Equivalence principle for massive bodies. II. Theory. *Physical Review*, 169:1017, 1968.
- [40] R. Nan, D. Li, C. Jin, et al. The five-hundred-meter aperture spherical radio telescope (FAST) project. *International Journal of Modern Physics D*, 20:989, 2011.
- [41] M. Kramer, D. Backer, J. Cordes, et al. Strong-field tests of gravity using pulsars and black holes. *New Astronomy Reviews*, 48:993, 2004.
- [42] L. Shao, I. Stairs, J. Antoniadis, et al. Testing gravity with pulsars in the SKA era. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 42, 2015. arXiv:1501.00058
- [43] E. Keane, B. Bhattacharyya, M. Kramer, et al. A cosmic census of radio pulsars with the SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, p. 40, 2015.
- [44] L. Shao, N. Sennett, A. Buonanno, M. Kramer, and N. Wex. Constraining nonperturbative strong-field effects in scalar-tensor gravity by combining pulsar timing and laser-interferometer gravitational-wave detectors. arXiv:1704.07561

Table 1. Parameters for the spindown of the pulsar and the orbits of the triple system. Values in the second column are from Ref. [1], while those in the third column are from the parameter estimation on the mock dataset “TOA.3” while allowing a nonzero whitehead parameter in the MCMC runs. Parenthesized numbers represent the $1\text{-}\sigma$ uncertainty in the last digits quoted.

Parameter	Observation	Simulation
SPINDOWN PARAMETERS		
Pulsar spin frequency, f_0	$365.953363096(11)$ Hz	$365.953363095999(3)$ Hz
Spin frequency derivative, \dot{f}	$-2.3658(12) \times 10^{-15}$ Hz s $^{-1}$	$-2.3658(1) \times 10^{-15}$ Hz s $^{-1}$
INNER KEPLERIAN PARAMETERS FOR PULSAR ORBIT		
Semimajor axis projected along line of sight, $(a \sin i)_I$	$1.21752844(4)$ ls	$1.21752844(2)$ ls
Orbital period, $P_{b,I}$	$1.629401788(5)$ d	$1.629401789(3)$ d
Eccentricity parameter, $\epsilon_{1,I} \equiv (e \sin \omega)_I$	$6.8567(2) \times 10^{-4}$	$6.8566(2) \times 10^{-4}$
Eccentricity parameter, $\epsilon_{2,I} \equiv (e \cos \omega)_I$	$-9.171(2) \times 10^{-5}$	$-9.169(2) \times 10^{-5}$
Time of ascending node, $T_{\text{asc},I}$	MJD 55920.407717436(17)	MJD 55920.40771744(1)
OUTER KEPLERIAN PARAMETERS FOR CENTER OF MASS OF INNER BINARY		
Semimajor axis projected along line of sight, $(a \sin i)_O$	$74.6727101(8)$ ls	$74.672711(2)$ ls
Orbital period, $P_{b,O}$	$327.257541(7)$ d	$327.25755(2)$ d
Eccentricity parameter, $\epsilon_{1,O} \equiv (e \sin \omega)_O$	$3.5186279(3) \times 10^{-2}$	$3.5186282(8) \times 10^{-2}$
Eccentricity parameter, $\epsilon_{2,O} \equiv (e \cos \omega)_O$	$-3.462131(11) \times 10^{-3}$	$-3.46212(3) \times 10^{-3}$
Time of ascending node, $T_{\text{asc},O}$	MJD 56233.935815(7)	MJD 56233.93582(2)
INTERACTION PARAMETERS		
Semimajor axis projected in plane of sky, $(a \cos i)_I$	$1.4900(5)$ ls	$1.4900(3)$ ls
Semimajor axis projected in plane of sky, $(a \cos i)_O$	$91.42(4)$ ls	$91.42(2)$ ls
Ratio of inner companion mass to pulsar mass, $q_I \equiv m_{\text{WD},I}/m_{\text{NS}}$	$0.13737(4)$	$0.13738(3)$
Difference in longitudes of ascending nodes, $\delta\Omega$	$2.7(6) \times 10^{-3}$ deg	$2.6(2) \times 10^{-3}$ deg