

UNIVERSITY OF BIRMINGHAM

Research at Birmingham

On tests of general relativity with binary radio pulsars

Del Pozzo, Walter; Vecchio, Alberto

DOI:

[10.1093/mnrasl/slw116](https://doi.org/10.1093/mnrasl/slw116)

License:

None: All rights reserved

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Del Pozzo, W & Vecchio, A 2016, 'On tests of general relativity with binary radio pulsars' Monthly Notices of the Royal Astronomical Society, vol. 462, no. 1, pp. L21-L25. <https://doi.org/10.1093/mnrasl/slw116>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This article has been accepted for publication in Monthly Notices of the Royal Astronomical Society ©: The Author(s) 2016, Published by Oxford University Press. All rights reserved.

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

On tests of general relativity with binary radio pulsars

W. Del Pozzo[★] and A. Vecchio

School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Accepted 2016 June 9. Received 2016 June 9; in original form 2016 April 29

ABSTRACT

The timing of radio pulsars in binary systems provides a superb testing ground of general relativity. Here we propose a Bayesian approach to carry out these tests, and a relevant efficient numerical implementation, that has several conceptual and practical advantages with respect to traditional methods based on least-squares fit that have been used so far: (i) it accounts for the actual structure of the likelihood function – and it is not predicated on the Laplace approximation which is implicitly built in least-squares fit that can potentially bias the inference – (ii) it provides the ratio of the evidences of any two models under consideration as the statistical quantity to compare different theories, and (iii) it allows us to put *joint* constraints from the monitoring of multiple systems, that can be expressed in terms of ratio of evidences or probability intervals of global (thus not system-dependent) parameters of the theory, if any exists. Our proposed approach optimally exploits the progress in timing of radio pulsars and the increase in the number of observed systems. We demonstrate the power of this framework using simulated data sets that are representative of current observations.

Key words: methods: data analysis – pulsars: general.

1 INTRODUCTION

Radio pulsars observations allow some of the most exquisite tests of two-body dynamics in general relativity (GR; Eardley 1975; Damour & Esposito-Farèse 1992, 1996; Kramer et al. 2006; Foster 2007); see also Stairs (2003) and Will (2014) for recent reviews. They rely on the monitoring of the time of arrival (TOA) of the radio pulses of these rotating neutron stars orbiting a companion compact object, a neutron star or a white dwarf (Bell, Camilo & Damour 1996; Bhat, Bailes & Verbiest 2008; Shao & Wex 2012; Freire et al. 2012; Shao 2014; Yagi et al. 2014). With the ever-increasing sensitivity, observational baseline and sky coverage of radio observations and with the advent of new major telescopes, such as the Five-hundred-metre Aperture Spherical Telescope (FAST) and the Square-Kilometre-Array (SKA), one can expect the discovery of new (possibly extreme) systems and the further increase of the accuracy in monitoring existing ones. As a consequence, progressively more stringent constraints will be placed on the two-body dynamics.

The TOA at the Earth of the radio pulses emitted by pulsars in binary systems depends on the Keplerian parameters of the binary as well as on a set of post-Keplerian parameters. The relations of the latter to the binary parameters depend on the specifics of the theory of gravity. Traditionally, the parameters values are obtained by means of a least-squares fit. Tests of the gravity are then carried out by checking consistency relations of overdetermined

parameters of the theory, and/or recovering probability intervals of a given parameter which is consistent with the theory (e.g. Burgay et al. (2003); Kramer et al. (2006); see however Zhu et al. (2015) for an hybrid Monte Carlo approach). This approach has proven very successful so far. None the less, it suffers from conceptual and practical limitations that hinder the exploitation of the full information content of these exquisite observations, present and future. Specifically, the fit implicitly assumes a quadratic approximation of the likelihood function, which in principle does not hold (see Vigeland & Vallisneri (2014) for further discussions). Moreover, all the tests hitherto performed are done on a system-by-system basis: they do not take advantage of the fact that much more stringent constraints can be put by combining information across multiple systems.

In this Letter we introduce an approach that addresses these shortcomings working in the framework of Bayesian inference: a test to compare different theories of gravity and to combine information from observations of several systems. We also provide a numerical implementation based on nested sampling that is both accurate and efficient. We illustrate the power of our approach by using a set of simulated data modelled around existing observations.

2 METHOD AND MODELS

The radio pulses' TOAs at the Solar system Barycentre, τ , are related to the emission time, t , at the location of the pulsar by $t = \tau + \Delta(\vec{\lambda})$, where we define $\Delta(\vec{\lambda})$ as the total contribution of the delays that depend on a set of unknown parameters $\vec{\lambda} = \{\theta, \vec{\lambda}_K, \vec{\lambda}_{PK}\}$. For convenience, we divide the parameters in two categories:

[★]E-mail: wdp@star.sr.bham.ac.uk

non-orbital parameters $\vec{\theta}$ – position, proper motion, propagation through the interstellar medium, white and red noise parameters, etc. – and orbital parameters $\{\vec{\lambda}_K, \vec{\lambda}_{pK}\}$. In any relativistic theory of gravity, the latter can be further divided into five Keplerian parameters, $\vec{\lambda}_K$, which are connected to the Keplerian description of the orbital dynamics, and a set of post-Keplerian parameters $\vec{\lambda}_{pK}$ that encode the higher order corrections to the Keplerian dynamics, and are specific to the theory of gravity under consideration. Observations of pulsars according to any relativistic theory of gravity are described by the same set of parameters $\vec{\theta}$ and $\vec{\lambda}_K$ but different post-Keplerian parameters $\vec{\lambda}_{pK}$.

In order to illustrate our proposed approach, in this Letter we consider three specific models for the orbital dynamics that are commonly considered in the analysis of pulsar timing data, based on the seminal work by Damour & Deruelle (1986) and Damour & Taylor (1992). Currently, only five post-Keplerian parameters are needed to describe the TOAs for the best systems – for the double pulsar, measurements of the relativistic spin-orbit precession rate come from radio eclipse models (Breton et al. 2008) – but our method is general and can include the appropriate additional parameters required to describe future data. We consider the Keplerian parameter vector $\vec{\lambda}_K = \{P_b, x_p, e, \omega_0, T_0\}$, where P_b is the orbital period, x_p the semi-major axis along the line of sight, e the orbital eccentricity, ω_0 the longitude of the periastron at a reference time T_0 , the epoch of the periastron. The post-Keplerian parameters $\vec{\lambda}_{pK}$ are: $\dot{\omega}$, the average rate of the periastron advance, γ_p , the amplitude of delays in arrival of pulses caused by the varying effects of the gravitational redshift and time dilation as the pulsar moves in its elliptical orbit at varying distances from its companion, \dot{P}_b , the rate of change of the orbital period, and r_p and s_p , the ‘range’ and ‘shape’, respectively, of the Shapiro time delay of the pulsar signal as it propagates through the curved space–time region near the companion. In any *specific* theory of gravity, $\vec{\lambda}_{pK}$ are (theory-dependent) functions of $\vec{\lambda}_K$ and the unknown binary component inertial masses – m_p , the mass of the pulsar and m_c the mass of the companion – the equation of state of the neutron stars (and in some cases the two angles that describe the pulsar spin axis).

In GR, the five post-Keplerian parameters are related to $\vec{\lambda}_K$ and the two masses by the following equations (Kramer & Wex 2009):

$$\dot{\omega} = 3T_\odot^{2/3} n_b^{5/3} \frac{1}{1 - e^2} (m_p + m_c)^{2/3}, \quad (1)$$

$$\gamma_p = T_\odot^{2/3} n_b^{-1/3} e \frac{m_c(m_p + 2m_c)}{(m_p + m_c)^{4/3}}, \quad (2)$$

$$r_p = T_\odot m_c, \quad (3)$$

$$s = \sin i = T_\odot^{-1/3} n_b^{2/3} x_p \frac{(m_p + m_c)^{2/3}}{m_c}, \quad (4)$$

$$\dot{P}_b = -\frac{192\pi}{5} T_\odot^{5/3} n_b^{5/3} f(e) \frac{m_p m_c}{(m_p + m_c)^{1/3}}. \quad (5)$$

Here we have defined $T_\odot = GM_\odot/c^3$, $n_b = 2\pi/P_b$ and $f(e)$ is given by $f(e) = [1 + (73/24)e^2 + (37/96)e^4]/(1 - e^2)^{7/2}$.

For generic conservative theories of gravity, the post-Keplerian parameters can be written in terms of *system independent parameters* as (Damour & Taylor 1992)

$$\dot{\omega} = n_b \left(\epsilon - \frac{1}{2}\xi + \frac{1}{2} \right) \frac{\beta^2}{1 - e^2}, \quad (6)$$

$$\gamma = \frac{e}{n_b} X_B (\tilde{G}_{OB} + \kappa + X_B) \beta^2, \quad (7)$$

$$r_p = \frac{1}{4n_b} X_B \tilde{G}_{OB} (\epsilon_{0B} + 1) \beta^3, \quad (8)$$

$$s_p = \frac{n_b X_A}{\beta X_B}, \quad (9)$$

where we have introduced \tilde{G}_{OB} which is the ratio between an ‘effective’ gravitational constant and the Newton’s constant G , $\epsilon = 3$ in GR, $\xi = 1$ in GR, $\kappa = 0$ in GR and finally $X_A = m_a/M_{\text{tot}}$, $X_B = m_b/M_{\text{tot}} = 1 - X_A$ and $\beta = (\mathcal{G}M_{\text{tot}}n_b)^{1/3}/c$ is a characteristic velocity for the relative orbital motion.

In summary, the three models that we consider are: (i) the *DD model* (Damour & Deruelle 1986), in which the Keplerian and post-Keplerian orbital parameters are all considered independent parameters, therefore $\{\vec{\lambda}_K, \vec{\lambda}_{pK} = \{\dot{\omega}, \gamma_p, r_p, s, \dot{P}_b\}\}$; (ii) the *DDGR model* in which we assume GR is the correct theory of gravity and therefore the parameters are $\{\vec{\lambda}_K, m_p, m_c\}$; and (iii) the *DDCG model* (Damour & Taylor 1992), in which the Keplerian orbital parameters are considered free while the post-Keplerian parameters are constrained to obey the relations imposed by a general metric theory, and therefore $\{\vec{\lambda}_K, \tilde{G}_{OB}, \epsilon, \xi, \kappa, \dot{P}_b\}$.

Given a set of observations from n pulsars $\vec{d} = \{\vec{d}_1, \dots, \vec{d}_n\}$, our purpose is to infer the parameters $\vec{\lambda}$ according to the predictions of some model \mathcal{H} for the dynamics of the system (and the background information \mathcal{I} at hand) and to evaluate the probability of the data under the specific model assumption, i.e. the *model evidence* or marginal likelihood. The posterior density function (PDF) for $\vec{\lambda}$ is obtained via Bayes’ theorem:

$$p(\vec{\lambda}|\vec{d}, \mathcal{H}, \mathcal{I}) = p(\vec{\lambda}|\mathcal{H}, \mathcal{I}) \frac{p(\vec{d}|\vec{\lambda}, \mathcal{H}, \mathcal{I})}{p(\vec{d}|\mathcal{H}, \mathcal{I})}. \quad (10)$$

Here $p(\vec{\lambda}|\mathcal{H}, \mathcal{I})$ is the prior probability density, $p(\vec{d}|\vec{\lambda}, \mathcal{H}, \mathcal{I})$ is the likelihood function and

$$p(\vec{d}|\mathcal{H}, \mathcal{I}) = \int d\vec{\lambda} p(\vec{\lambda}|\mathcal{H}, \mathcal{I}) p(\vec{d}|\vec{\lambda}, \mathcal{H}, \mathcal{I}) \quad (11)$$

is the *model evidence*. As we can assume that the observations of different pulsars are statistically independent, the likelihood factorizes into the product of each individual likelihood

$$p(\vec{d}|\vec{\lambda}, \mathcal{H}, \mathcal{I}) = \prod_{i=1}^n p(\vec{d}_i|\vec{\lambda}, \mathcal{H}, \mathcal{I}). \quad (12)$$

For observations of a single pulsar, say ‘1’, the likelihood is simply $p(\vec{d}_1|\vec{\lambda}, \mathcal{H}, \mathcal{I})$. Given several competing models $\mathcal{H}_{j=1, \dots, N}$, we can rank their effectiveness in explaining the observed data with the odds ratio:

$$\mathcal{O}_{i,j} \equiv \frac{p(\mathcal{H}_i|\mathcal{I}) p(\vec{d}|\mathcal{H}_i, \mathcal{I})}{p(\mathcal{H}_j|\mathcal{I}) p(\vec{d}|\mathcal{H}_j, \mathcal{I})}, \quad (13)$$

where $p(\mathcal{H}_i|\mathcal{I})/p(\mathcal{H}_j|\mathcal{I})$ is the prior odds and $p(\vec{d}|\mathcal{H}_i, \mathcal{I})/p(\vec{d}|\mathcal{H}_j, \mathcal{I})$ is the Bayes’ factor. Within a given theory (a model or hypothesis) \mathcal{H} , equation (10) provides all the information about the parameters of the theory. If one wishes to compare a given theory with alternatives, the odds ratio equation (13) is the statistically rigorous quantity to consider (Jaynes 2003).

The likelihood function (12) is constructed from the ‘timing residuals’ $\delta\vec{\tau}$, the difference between the predicted (for a given choice of the parameters of the model) TOAs and the observed TOAs. We take

the timing residuals to be distributed according to a multi-variate Gaussian, hence

$$p(\vec{d}_i | \vec{\lambda}, \mathcal{H}, \mathcal{I}) = (2\pi)^{-\frac{k_i}{2}} \|\mathbf{C}\|^{-1} \exp\left[-\frac{t(\delta\vec{t})\mathbf{C}^{-1}(\delta\vec{t})}{2}\right], \quad (14)$$

where t indicates the transpose operation. The expected statistical properties of the residuals are described by the covariance matrix \mathbf{C} . \mathbf{C} contains unknown parameters of the model, white and red-noise parameters which too need to be simultaneously estimated. The general form of our likelihood allows for a natural inclusion in the analysis of red noise as a Gaussian Process, e.g. van Haasteren & Vallisneri (2014), and of white noise processes as additional diagonal covariance matrices. For the results on simulated data presented in this Letter, we ignore these complications and assume that \mathbf{C} is diagonal with elements given by the known measurement uncertainties. Nevertheless, the inclusion of additional red or white noise parameters does not affect our results.

The marginalized posterior PDFs and the model’s evidence are obtained by numerical integrations over a large multi-dimensional parameter space, for which stochastic sampling techniques are particularly suitable. `TEMPO2` (Lentati et al. 2014) provides a specific sampling scheme that can be applied to this problem. We have developed an independent nested sampling algorithm tailored to the problem at hand, that is based on the scheme proposed by Veitch & Vecchio (2010), which is part of the analysis infrastructure to study compact binary systems with gravitational-wave laser interferometers (Veitch et al. 2015), and was used to estimate the parameters of the binary black hole merger associated with the transient GW150914 observed by Advanced LIGO and to perform tests of GR (Abbott et al. 2016a,b,c). The algorithm (written in `PYTHON`) provides a reliable estimate of the evidence integral as well as the marginalized PDFs on all model parameters. It implements a parallel scheme to generate new samples from the prior probability distribution. The timing model and thus the likelihood are computed through the timing software `TEMPO2` (Edwards et al. 2006; Hobbs et al. 2006) and the `PYTHON` wrapper `libstempo` (Vallisneri 2014). Any timing model supported by `TEMPO2` can therefore be immediately used in our code. The results presented in this Letter are obtained using 1024 ‘live points’ in the nested sampling – see for more technical details Veitch & Vecchio (2010) – using standard eight CPU cores in parallel. Each analysis (details are provided below) takes about 12 h to complete. This is therefore an efficient and generic approach that can be applied to a large range of observations.

Equation (14) is, in general, a sharp function of the parameters $\vec{\lambda}$, and a search over the whole prior volume is inefficient. We thus proceed as suggested in Lentati et al. (2014): from a first analysis of the data with `TEMPO2` we restrict the domain of integration for the evidence and PDFs calculation to a region of the parameter space which is centred around the `TEMPO2`’s best-fitting parameters with a width $\pm 5\sigma$, where σ is the nominal statistical error returned by `TEMPO2`. As long as the prior volume is large enough to capture the full likelihood function, this strategy is both numerically efficient and formally correct, since the likelihood outside this range does not contribute to the integral. Finally, we adopt uniform priors for all the parameters in the model.

3 RESULTS

In order to illustrate the power of our proposed approach, we present the results from analyses of simulated data of four hypothetical bi-

nary pulsar systems. We note that our method is not restricted to Double-Pulsar-like systems, but can operate on any class of binary systems (e.g. neutron star – white dwarf), thus probing different regions of the space of theories of gravity. In generating the parameters of the systems under consideration, we assume that GR is the correct description of gravity. The double pulsar J0737-3039 (Burgay et al. 2003) is one of the very best systems to provide tight constraints on theories of gravity, we thus consider a system and data set – System #1 in Table 1 – that qualitatively reproduce the observations reported in Kramer et al. (2006): a three year data set with a cadence of one per month and instrumental timing uncertainty of 20 μ s. The three additional systems and data sets have equal length and cadence as System #1 and the same (but statistically independent) timing noise. The parameters of the systems are given in Table 1. The TOAs are generated using the `fake` package of `TEMPO2` setting the noise realization to zero.

We illustrate our findings by discussing three key points. First, we show how constraints on GR can be rigorously quantified in terms of odds ratios between different models. Secondly, we show that the joint inference from the observations of many systems can yield more stringent evidence for – or against – a specific theory. Finally, constraints on system independent parameters that are characteristic of the theory improve when one combines the data from different systems. The main result of our analysis are summarized in Figs 1 and 2, and are discussed in detail below.

We analyse the data using the three models, DD, DDGR, and DDCG discussed in Section 2. The number of parameters that describe the models is 24, 18 and 24, respectively. All the models are described by the *same set* of non-orbital parameters $\vec{\theta} = \{v_A, \dot{v}_A, v_B, \dot{v}_B, \alpha, \delta, \mu_a, \mu_\delta, D\}$ and Keplerian parameters $\vec{\lambda}_K = \{P_b, x_a, x_b, e, \omega, T_0\}$. In addition for the DDGR model we have the parameters $\vec{\lambda}_{\text{pK}} = \{m_A, m_B\}$, for the DD model the parameters $\vec{\lambda}_{\text{pK}} = \{\dot{\omega}, \dot{P}_b, \gamma_A, \gamma_B, r_A, r_B\}$ and for the DDCG model the parameters $\vec{\lambda}_{\text{pK}} = \{\bar{G}_{\text{OB}}, \epsilon, \kappa, \xi\}$. We then proceeded with the calculation of the PDFs for all parameters as well as the evidence for the three models under consideration. With four independent systems the total number of unknown parameters in the analysis is 96, 72 and 96 for the DD, DDGR, and DDCG model, respectively. Fig. 1 shows one-dimensional marginalized PDFs for System #1 on $\vec{\lambda}_K, \vec{\lambda}_{\text{pK}}, m_1$ and m_2 under the three different model assumptions. The simulated values of the parameters are correctly recovered. Similar PDFs (not shown) are obtained for the remaining Systems. The widths and shapes of the PDFs vary considerably depending on the assumed orbital model. Our approach and its numerical implementation provides therefore a powerful tool to derive rigorous statistical information from pulsar timing.

Fig. 2 provides an example of the comparison of the models and the inference of global parameters of the theory when jointly analysing the data from multiple systems. The data were generated according to the DDGR model, and analysed in turn under the three models assumptions, DD, DDGR, and DDCG. Indeed, the results of the analysis shows that the DDGR model is vastly favoured, and as more systems are included in the analysis this conclusion becomes progressively stronger. The left-hand panel of Fig. 2 shows this result in a striking way. From a single system, the odds in favour of the DDGR model are $\mathcal{O}_{\text{DDGR,DD}} = 2.3 \times 10^3 : 1$ and $\mathcal{O}_{\text{DDGR,DDCG}} = 1.6 \times 10^3 : 1$ against the DD and DDCG model, respectively; the DDGR model is ~ 1000 times more probable than either alternatives. With the inclusion of more systems in the analysis the odds in favour of DDGR increase substantially, reaching values of $3 \times 10^{15} : 1$ and $10^{16} : 1$, respectively. These results

Table 1. The parameters of the simulated binary systems. Additional parameters not shown are the (arbitrary) values of the binary location in the sky – right ascension α and declination δ – the parallax and the TOAs standard deviation σ_{TOA} .

Parameter	System #1	System #2	System #3	System #4
Pulsar 1 spin frequency, ν_1 (s^{-1})	44.054 069 292 744	61.678 521 661 049	34.324 170 685 652	0.725 902 035 568
Pulsar 1 spin down rate, $\dot{\nu}_1$ (10^{-14}s^{-2})	-0.341 56	-4.923 60	-8.332 50	-4.809 14
Pulsar 2 spin frequency, ν_2 (s^{-1})	0.360 560 355 06	27.532 606 924 712	25.275 689 037 175	8.247 581 944 353
Pulsar 2 spin down rate, $\dot{\nu}_2$ (10^{-14}s^{-2})	-0.011 60	-6.595 30	-4.250 58	-3.833 87
Proper motion in α , μ_α (mas yr^{-1})	-3.3	-41.5	-5.5	-3.9
Proper motion in δ , μ_δ (mas yr^{-1})	2.6	-4.9	9.2	4.6
Dispersion measure, (pc cm^{-3})	48.920	11.1327	8.471	1.595
Orbital period, P_b (d)	0.102 251 562 48	0.708 530 130 25	0.095 368 400 65	0.044 953 054 03
Projected semi-major axis 1, a_1 (lt-s)	1.410 947 749 93	3.379 044 260 06	1.518 784 101 52	0.907 974 405 42
Projected semi-major axis 2, a_2 (lt-s)	1.511 721 662 41	4.327 006 302 38	1.207 194 088 62	0.809 195 604 00
Time of periastron passage, T_0 (MJD)	53 155.907 4280	53 252.218 8315	53 436.220 0040	53 892.267 6109
Angle of periastron, ω_1 (deg)	87.0331	155.3942	77.0200	20.6494
Orbital eccentricity, e	0.087 7775	0.875 4259	0.153 8017	0.012 0680
Mass of pulsar 1, m_1 (M_\odot)	1.3381	1.0504	1.1533	1.2986
Mass of pulsar 2, m_2 (M_\odot)	1.2489	1.3451	1.4510	1.4571

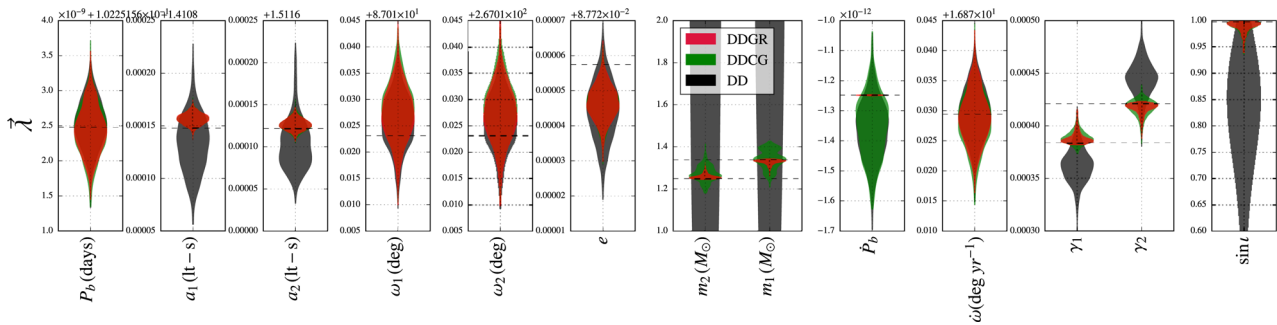


Figure 1. Violin plot summarizing the posterior distributions for the Keplerian and post-Keplerian parameters from the analysis of System #1 using the DD (black), DDGR (red) and DDCG (green) models. In each panel the thin dashed line indicates the value of the corresponding parameter used to generate the simulated data. We note the non-Gaussian nature of the PDFs for the mass parameters as well as for γ_1 and γ_2 , highlighting the failure of the Laplace quadratic approximation of the likelihood function.

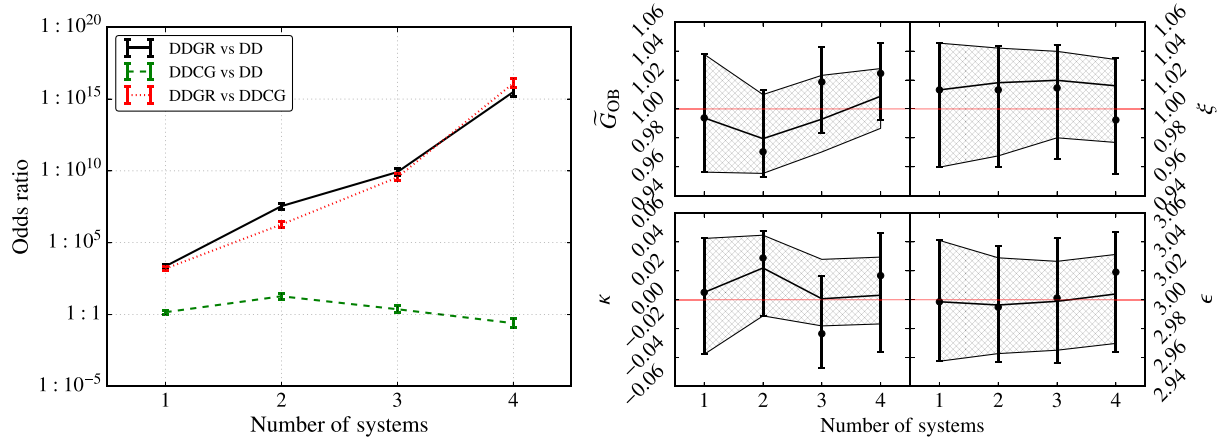


Figure 2. Results from the joint analysis of the four systems. Left-hand panel: cumulative odds ratios between the different models as a function of the number of systems included in the analysis. The error-bars indicate the theoretical uncertainty in the odds due to the discrete approximation of equation (11). The average uncertainty is 0.2 dex. Right-hand panel: 95 per cent probability interval computed from the joint PDF (hatched regions) for the global parameters \bar{G}_{OB} (top left), ϵ (top right), ξ (bottom left) and κ (bottom right) from the analysis on one (dashed), two (dotted), three (dot-dashed) and four (solid) systems. The solid horizontal line indicates the value predicted by general relativity, which has been used to generate the simulated observation data.

provide also a qualitative indication of how much more stringent the results from J0737-3039 could become as the time-span of the observations increases (assuming that the physics at work does follows the DDGR model). As for the DDCG model, the odds do not show any significant evidence in favour of DDCG against DD. From the analysis of one system DDCG shows odds against DD that are 1: 1.5 which from the analysis of four systems become 1: 1/4. These values are consistent with the numerical errors of the evidence evaluation. Correctly, the data do not provide evidence in favour of DDCG against DD.

The analysis of the data using the DDCG model, described by system independent parameters whose value in GR and other theories of gravity is known a priori (e.g. Kramer & Wex 2009) – \tilde{G}_{OB} , ϵ , ξ and κ – allows us to construct joint PDFs on these parameters, which in turns shows another advantage of our method. More strict limits on system-independent parameters can be obtained by a joint analysis of several systems. We stress that we perform the analysis by *simultaneously* fitting for all the parameters in the model for all the sources, and not by fixing the values of some (or all the other parameters) to a some arbitrary value (say the maximum likelihood or best-fitting value) and then estimating these global parameters. The right-hand panel of Fig. 2 shows the median of each of the parameters and the central 95 per cent probability interval coming from the *independent* analysis of each system, and the contraction of the 95 per cent probability interval produced in the *joint* analysis. The probability intervals both for the individual system analyses and the joint analysis correctly contain the GR values, which were used to generate the data sets. We find that as systems are added to the analysis the size of the probability interval decreases roughly as the square root of the number of systems considered. Furthermore, for the specific parameters of the simulated data that we consider here, one can reach an accuracy of a few percent at the 95 per cent confidence level on all the parameters. We note that, in reality, the measurement accuracy will also depend on the length of the observations and scale differently depending on the post-Keplerian parameter.

4 CONCLUSIONS

We have introduced a Bayesian approach to perform tests of GR using timing data of binary radio pulsars. It relies on the evaluation of the odds ratio between different theories on a system-by-system basis and in addition provides more stringent constrains by *combining* the results from the observation of multiple systems. If the specific theory of gravity that one wishes to test is described by global parameters (i.e. independent of the specific binary system), this approach also provides constraints on these parameters coming from the joint analysis of all the systems at hand. We have illustrated its power via an efficient numerical implementation (which

is appropriate for the large number of dimensions that characterize this problem) applied to a synthetic data set representative of plausible observations of several binary pulsars. This approach can and should be applied to existing and future data sets.

ACKNOWLEDGEMENTS

We thank Michael Kramer, Michele Vallisneri and John Veitch for useful comments and discussions. The Letter was funded in part by a Leverhulme Trust research project grant. The numerical simulations were performed on the Tsunami cluster of the University of Birmingham.

REFERENCES

- Abbott B. P. et al., 2016a, Phys. Rev. Lett., 116, 061102
 Abbott B. P. et al., 2016b, Phys. Rev. Lett., 116, 221101
 Abbott B. P. et al., 2016c, Phys. Rev. Lett., 116, 241102
 Bell J. F., Camilo F., Damour T., 1996, ApJ, 464, 857
 Bhat N. D. R., Bailes M., Verbiest J. P. W., 2008, Phys. Rev. D, 77, 124017
 Breton R. P. et al., 2008, Science, 321, 104
 Burgay M. et al., 2003, Nature, 426, 531
 Damour T., Deruelle N., 1986, Ann. Inst. Henri Poincaré Phys. Theor., 44, 263
 Damour T., Esposito-Farèse G., 1992, Phys. Rev. D, 46, 4128
 Damour T., Esposito-Farèse G., 1996, Phys. Rev. D, 54, 1474
 Damour T., Taylor J. H., 1992, Phys. Rev. D, 45, 1840
 Eardley D. M., 1975, ApJ, 196, L59
 Edwards R. T., Hobbs G. B., Manchester R. N., 2006, MNRAS, 372, 1549
 Foster B. Z., 2007, Phys. Rev. D, 76, 084033
 Freire P. C. C. et al., 2012, MNRAS, 423, 3328
 Hobbs G. B., Edwards R. T., Manchester R. N., 2006, MNRAS, 369, 655
 Jaynes E. T., 2003, Probability Theory: The Logic of Science. Cambridge Univ. Press, Cambridge
 Kramer M., Wex N., 2009, Class. Quantum Gravity, 26, 073001
 Kramer M. et al., 2006, Science, 314, 97
 Lentati L., Alexander P., Hobson M. P., Feroz F., van Haasteren R., Lee K. J., Shannon R. M., 2014, MNRAS, 437, 3004
 Shao L., 2014, Phys. Rev. D, 90, 122009
 Shao L., Wex N., 2012, Class. Quantum Gravity, 29, 215018
 Stairs I. H., 2003, Living Rev. Relativ., 6, 5
 Vallisneri M., 2014, libstempo, available at: <http://vallis.github.io/libstempo/>
 van Haasteren R., Vallisneri M., 2014, Phys. Rev. D, 90, 104012
 Veitch J., Vecchio A., 2010, Phys. Rev. D, 81, 062003
 Veitch J. et al., 2015, Phys. Rev. D, 91, 042003
 Vigeland S. J., Vallisneri M., 2014, MNRAS, 440, 1446
 Will C. M., 2014, Living Rev. Relativ., 17, 4
 Yagi K., Blas D., Yunes N., Barausse E., 2014, Phys. Rev. Lett., 112, 161101
 Zhu W. W. et al., 2015, ApJ, 809, 41

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