

A new pulsar timing model for scalar-tensor gravity with applications to PSR J2222–0137 and pulsar-black hole binaries

A. Batrakov¹, H. Hu¹, N. Wex¹, P. C. C. Freire¹, V. Venkatraman Krishnan¹, M. Kramer^{1,2},
Y. J. Guo¹, L. Guillemot^{3,4}, J. W. McKee^{5,6,7}, I. Cognard^{3,4}, and G. Theureau^{3,4}

¹ Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: abatrakov@mpi.fr-bonn.mpg.de

² Jodrell Bank Centre for Astrophysics, The University of Manchester M13 9PL, UK

³ Observatoire Radioastronomique de Nançay, Observatoire de Paris, Université PSL, Université d'Orléans, CNRS, 18330 Nançay, France

⁴ Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, Université d'Orléans/CNRS, 45071 Orléans Cedex 02, France

⁵ Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada

⁶ E.A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Kingston-upon-Hull HU6 7RX, UK

⁷ Centre of Excellence for Data Science, Artificial Intelligence and Modelling (DAIM), University of Hull, Cottingham Road, Kingston-upon-Hull HU6 7RX, UK

Received 18 October 2022 / Accepted 8 February 2023

ABSTRACT

Context. Scalar-tensor gravity (STG) theories are well-motivated alternatives to general relativity (GR). One class of STG theories, Damour–Esposito–Farèse (DEF) gravity, has a massless scalar field with two arbitrary coupling parameters. We are interested in this theory because, despite its simplicity, it predicts a wealth of different phenomena, such as dipolar gravitational wave emission and spontaneous scalarisation of neutron stars (NSs). These phenomena of DEF gravity can be tested by timing binary radio pulsars. In the methods used so far, intermediate phenomenological post-Keplerian (PK) parameters are measured by fitting the corresponding timing model to the timing data whose values are then compared to the predictions from the alternative theory being tested. However, this approach loses information between intermediate steps and does not account for possible correlations between PK parameters.

Aims. We aim to develop a new binary pulsar timing model ‘DDSTG’ (called after Damour, Deruelle and STG) to enable more precise tests of STG theories based on a minimal set of binary parameters. The expressions for PK parameters in DEF gravity are self-consistently incorporated into the model. PK parameters depend on two masses which are now directly fitted to the data without intermediate steps. The new technique takes into account all possible correlations between PK parameters naturally.

Methods. Grids of physical parameters of NSs were calculated in the framework of DEF gravity for a set of 11 equations of state. Automatic differentiation (AutoDiff) technique was employed, which aids in the calculation of gravitational form factors of NSs with a higher precision than in previous works. The pulsar timing program TEMPO was selected as a framework for the realisation of the DDSTG model. The implemented model is applicable to any type of pulsar companions. We also simulated realistic future radio-timing datasets for a number of large radio observatories for the binary pulsar PSR J2222–0137 and three generic pulsar-black hole (PSR-BH) systems.

Results. We applied the DDSTG model to the most recently published observational data for PSR J2222–0137. The obtained limits on DEF gravity parameters for this system confirm and improve previous results. New limits are also the most reliable because DEF gravity is directly fitted to the data. We argue that future observations of PSR J2222–0137 can significantly improve the limits and that PSR-BH systems have the potential to place the tightest limits in certain areas of the DEF gravity parameter space.

Key words. gravitation – binaries: close – gravitational waves – pulsars: general – stars: individual: J2222-0137

1. Introduction

General relativity (GR) proved to be the most successful theory of gravity for more than a century. Up to now, it has passed all experimental tests exceptionally well. The weak field regime has been verified by the Solar System experiments (Will 2014), whereas strong field effects have been tested especially well by the timing of binary pulsars (Wex & Kramer 2020). Among other things, pulsar tests have enabled precise tests of the radiative properties of gravity and the strong equivalence principle (SEP). Furthermore, the large-scale behaviour of gravity (low spacetime curvature and temporal variation) has been tested in cosmological observations (Clifton et al. 2012). Finally, in the last few years, ground-based gravitational wave (GW) detectors observed GWs from coalescing binary black holes (BHs) (Abbott et al.

2016, 2017c, 2019a, 2021b), binary neutron stars (NSs) (Abbott et al. 2017a,b, 2019b), and also BH-NS binaries (Abbott et al. 2021a), probing the hitherto unexplored highly dynamic, strong-field regime of gravity. Thus far, GR has accounted for all observed effects both in weak and strong gravitational fields, in the quasi-static and the highly dynamical regime, and on small as well as on large scales¹.

Despite such successes, there are still convincing reasons to investigate modified theories of gravity (Berti et al. 2015). GR describes a gravitational interaction by a single massless, spin-two tensor field $g_{\mu\nu}$, known as the metric of spacetime. Some of

¹ Within GR, the large scale requires the introduction of dark matter and dark energy (with the latter in the form of the cosmological constant Λ).

the best motivated alternative gravity theories, the scalar-tensor gravity (STG) theories, incorporate additional scalar degrees of freedom to mediate gravity. Such scalar fields arise naturally in higher dimensional theories in their low energy limits, for example in the old Kaluza-Klein theory (Kaluza 1921; Klein 1926) and string theories (Fujii & Maeda 2007). Scalar-tensor gravity theories are also motivated by cosmological questions of inflation, dark energy, and even the thus far undiscovered theory of quantum gravity (Clifton et al. 2012).

A scalar-tensor gravity theory which for many years was considered as the only natural competitor of GR is known as Jordan–Fierz–Brans–Dicke (JFBD) gravity (Jordan 1955, 1959; Fierz 1956; Brans & Dicke 1961)². JFBD gravity is a metric theory (matter and non-gravitational fields couple only to one specific spacetime metric) and it therefore fulfils the Einstein equivalence principle by design (Will 1993; Damour 2012). However, an additional scalar field is non-minimally coupled to matter via the choice of a special conformal coupling function; thus, SEP is violated (Will 2014). This coupling function depends on a single parameter ω_{BD} , which by now is tightly constrained, in particular with the pulsar in the stellar triple system (Voisin et al. 2020). Bergmann (1968) and Wagoner (1970) presented the most general STG theory with one scalar field which, while in action, is quadratic at most. The most general mono-scalar–tensor theory with second-order field equations is Horndeski gravity (Horndeski 1974). In 1992, Damour & Esposito-Farèse (1992b) presented a generic class of STG theories with an arbitrary number of scalar fields. In the scope of this paper, we mainly focus on Damour–Esposito-Farèse (DEF) gravity (Damour & Esposito-Farèse 1993). Damour–Esposito-Farèse gravity is a Bergman-Wagoner theory which has a massless scalar field and a quadratic coupling function with two arbitrary parameters.

Apart from being well-motivated, DEF gravity predicts interesting effects which are not present in weak fields of the Solar System but could be tested by employing pulsar astronomy. Neutron stars have strong gravitational fields and large gravitational binding energies due to their high compactness. Because of this, in STG theories they can have large gravitational form factors (also known as scalar charges), unlike weakly self-gravitating objects such as normal stars and white dwarfs (WDs). Consequently, while STG theories generally predict the emission of dipolar and monopolar GWs, which increase the rate of orbital period decay relative to the GR expectation, this is expected to be especially strong in asymmetric binary systems containing a NS and a WD, due to the significant differences in their compactness and scalar charges.

Furthermore, Damour & Esposito-Farèse (1993, 1996) found a specific phenomenon in NSs called ‘spontaneous scalarisation’. This fully non-perturbative effect excites the scalar field above the cosmological background value in NSs, allowing scalar charges of order unity, even if there is no deviation from GR in the weak-field regime. This could produce, for binary systems containing NSs with specific masses, a highly enhanced rate of dipolar GW emission.

Some of the tightest tests of gravity with strongly self-gravitating objects are provided by the high precision timing of binary pulsars (Stairs 2003; Wex 2014; Shao & Wex 2016; Wex & Kramer 2020; Kramer et al. 2021). These tests are performed in a quasi-stationary strong-field gravity regime, where

the gravitational field is strong near and inside NSs and produces high curvature, whereas the velocities are small compared to the speed of light $v/c \sim 10^{-3}$. The first timing model (BT) was introduced by Blandford & Teukolsky (1976) and it allowed for information from the timing of binary pulsars to be extracted. An extended model covering all relativistic effects in the dynamics of a binary system up to the first post-Newtonian level was proposed later by Damour & Deruelle (1985, 1986) (DD).

The DD model made it possible to perform self-consistent tests of gravity by means of the parametrised post-Keplerian (PPK) formalism (see, for instance, applications by Taylor & Weisberg 1989). In this generic framework, theory-independent Keplerian and post-Keplerian (PK) timing parameters, which quantify the relativistic motion of the pulsar and the propagation of its radio signals, are fitted to the observational timing data. Apart from the well-measured Keplerian parameters of the orbit, each PK parameter only depends on the masses of the pulsar and its companion. As soon as the expressions for the PK parameters in a particular gravity theory are known, these two masses can be derived from the PK parameters once at least two of them are known. If more than two PK parameters become available, then a test of the consistency of that gravity theory can be made. Later, Damour & Taylor (1992) expanded the DD model and introduced pulse structure parameters. They also showed that the DD model can be applied to a large set of fully conservative theories of gravity and presented phenomenological PK expressions in a framework of generic boost-invariant theories (the modified Einstein-Infeld-Hoffmann formalism).

However, such an independent measurement of PK parameters means that any correlations between them inadvertently reduce the sensitivity of their measurements. Taylor & Weisberg (1989) solved this problem for GR with the introduction of the DDGR model, where the GR expressions for PK parameters are incorporated in the model, and two masses are directly fitted to the observational timing data. This mitigates or breaks any correlations between the PK parameters.

In this work we present a new pulsar timing model – the ‘DDSTG’ model which is called after Damour, Deruelle and STG. The idea is very similar to that of the DDGR model: fitting directly for the masses of the components from pulsar timing data, but instead of GR using a particular member of the two-parametric DEF gravity. Apart from the theoretical predictions for all PK parameters within that theory, the model uses pre-calculated grids of physical parameters (gravitational form factors) of NSs in that particular gravity theory. The direct fit of two masses of the companions reduces the model to a minimal set of parameters and naturally solves all the issues of possible correlations between the observed parameters. This feature, achieved by the construction, makes the DDSTG model superior to the traditional methods, for example the ‘PK method’ based on the measurement of PK parameters.

As an application of the new DDSTG model, we use the most recently published timing data for PSR J2222–0137 from Cognard et al. (2017) and Guo et al. (2021b). This binary pulsar is interesting for several reasons: It is one of the closest binary pulsars known and has a good timing precision. A precise measurement of the variation of the orbital period (\dot{P}_b) and the high asymmetry in compactness between the pulsar and its companion allows for tight constraints on the emission of dipolar GWs. Moreover, the pulsar’s mass $m_p = 1.831(10) M_\odot$ (Guo et al. 2021b) lies in the range ($m_p \gtrsim 1.5 M_\odot$) subject to the spontaneous scalarisation effect; the non-detection of dipolar GW emission in this system has introduced strong constraints

² The first to formulate JFBD gravity was actually Willy Scherrer in the early 1940s (see Goenner 2012 for a historical review on the genesis of STG theories).

on the existence of that phenomenon, at least within the DEF framework (Zhao et al. 2022).

We organise the paper as follows: in Sect. 2, we introduce the basics of DEF gravity and present our machinery preparations necessary for the implementation of the new model. In Sect. 3, we discuss the standard timing models used thus far and introduce our new DDSTG model for testing STG theories. The subsequent sections are devoted to the applications of our new timing model. In Sect. 4, we provide a brief description of the Guo et al. (2021b) dataset on PSR J2222–0137 and apply the DDSTG model to it, discuss the limits on DEF gravity, and demonstrate that our new method does indeed provide tighter and more reliable constraints than previous methods. In Sect. 5, based on simulated datasets for the largest radio observatories, we discuss what limits on these theories might be achieved in the near future with continued timing of PSR J2222–0137. In Sect. 6, we show possible future constraints from reasonable pulsar-black hole (PSR-BH) binary systems. We also briefly discuss the possible origin of PSR-BH systems and how it affects the test if such a system is located in a globular cluster (GC). Section 7 is devoted to a discussion of other alternative gravity theories and possibilities of the DDSTG model extensions beyond DEF gravity. In Sect. 8 we discuss the effect of the time-varying gravitational constant as a perspective future extension of the DDSTG model. In Sect. 9 we summarise our results.

2. Damour–Esposito–Farèse gravity

This paper investigates STG theories, natural alternatives to GR. Specifically, we mainly focus on DEF gravity. In the following we summarise aspects of that class of gravity theories that are relevant for this paper.

2.1. Theoretical aspects

Damour–Esposito–Farèse gravity is a STG theory with one long range, massless scalar field φ non-minimally coupled to the curvature scalar. The theory is fully described by the action, which is most simply presented in the *Einstein frame* (Damour & Esposito-Farèse 1993, 1996)

$$S = \frac{c^4}{16\pi G_*} \int \frac{d^4x}{c} g_*^{1/2} (R_* - 2g_*^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi) + S_m[\psi_m; A^2(\varphi)g_{\mu\nu}^*], \quad (1)$$

where G_* is the bare gravitational coupling constant, g_* and R_* are the determinant and the Ricci scalar associated with the Einstein metric ($g_{\mu\nu}^*$). The last term in Eq. (1) describes an action associated with any matter fields (ψ_m). $A(\varphi)$ is the coupling function and takes a specific exponential form in DEF gravity: $A(\varphi) = \exp[\alpha_0(\varphi - \varphi_0) + \frac{1}{2}\beta_0(\varphi - \varphi_0)^2]$. Introduced quantities $\{\alpha_0, \beta_0\}$ are two arbitrary parameters of the theory and φ_0 is the scalar field at spatial infinity. The coupling function $A(\varphi)$ plays the role of a conformal factor connecting the ‘physical metric’ $\tilde{g}_{\mu\nu} = A^2(\varphi)g_{\mu\nu}^*$ with the Einstein metric. The metric $\tilde{g}_{\mu\nu}$ is the one measured by laboratory clocks and rods due to the universal coupling of matter to this metric. The bare gravitational constant is simply related to the one measured in a Cavendish experiment by $G_{\text{Cav}} = G_*(1 + \alpha_0^2)$. GR corresponds to vanishing coupling parameters, that is $\alpha_0 = \beta_0 = 0$, and JFBD gravity is recovered by setting $\beta_0 = 0$ while keeping $\alpha_0 \neq 0$.

The field equations are derived by variation of the action in Eq. (1) and are most simply formulated in terms of the pure-spin

field variables ($g_{\mu\nu}^*, \varphi$) in the Einstein frame:

$$R_{\mu\nu}^* = 2\partial_\mu \varphi \partial_\nu \varphi + \frac{8\pi G_*}{c^4} \left(T_{\mu\nu}^* - \frac{1}{2} T^* g_{\mu\nu}^* \right), \quad (2)$$

$$\square_{g^*} \varphi = -\frac{8\pi G_*}{c^4} \alpha(\varphi) T^*, \quad (3)$$

with a material stress-energy tensor $T_*^{\mu\nu} = 2cg_*^{-1/2} \delta S_m / \delta g_{\mu\nu}^*$ and $\alpha(\varphi) = \partial \ln A(\varphi) / \partial \varphi$, which measures the coupling strength between the scalar field and matter.

All weak-field deviations from GR may be described in terms of the asymptotic behaviour of $\alpha(\varphi)$ at spatial infinity. The theory parameters $\alpha_0 = \alpha(\varphi_0)$, $\beta_0 = \beta(\varphi_0)$ can be interpreted as asymptotic values of the function $\alpha(\varphi)$ and its derivative $\beta(\varphi) = \partial \alpha(\varphi) / \partial \varphi$.

In this work, without loss of generality, we use the formulation of DEF gravity with $\varphi_0 \equiv 0$, two parameters $\{\alpha_0, \beta_0\}$ appearing in the form of $A(\varphi)$ and $A(\varphi_0) = 1$. There is an equivalent formulation in terms of parameters $\{\beta_0, \varphi_0\}$, where $A(\varphi) = \exp(\beta_0 \varphi^2 / 2)$ and $\varphi_0 = \alpha_0 / \beta_0$. But the latter formulation is not well-determined for the JFBD limit $\beta_0 = 0$ (Damour 2007).

The parameterised post-Newtonian (PPN) framework allows for a wide range of metric theories of gravity to be described in their weak field approximation by ten independent PPN parameters. Damour–Esposito–Farèse gravity is a fully conservative theory and has only two PPN parameters that deviate from their GR values, which are also called Eddington parameters (Will 1993):

$$\gamma_{\text{Edd}} = 1 - \frac{2\alpha_0^2}{1 + \alpha_0^2}, \quad (4)$$

$$\beta_{\text{Edd}} = 1 + \frac{\beta_0 \alpha_0^2}{2(1 + \alpha_0^2)^2}. \quad (5)$$

The PPN parameter γ_{Edd} measures the spatial curvature induced by unit rest mass and β_{Edd} is a measure of the amount of non-linearity in the superposition law of gravity. Solar System experiments put limits on these parameters. The $|\gamma_{\text{Edd}} - 1| \lesssim 2.3 \times 10^{-5}$ limit comes from the Shapiro time delay observed by the Cassini spacecraft (Bertotti et al. 2003; Will 2014). The $|\beta_{\text{Edd}} - 1| \lesssim 8 \times 10^{-5}$ limit comes from observations of the perihelion shift of Mercury (Will 2014). The Cassini experiment yields a direct limit on $\alpha_0^2 \lesssim 1.15 \times 10^{-5}$, whereas β_0 remains unconstrained from weak-field Solar System experiments, since α_0 can be arbitrarily small.

2.2. Scalarisation of neutron stars in DEF gravity

To place firm limits on β_0 parameter, we need to explore the strong-field effects of DEF gravity. One can expect deviations from GR in the presence of highly compact massive objects, in particular for NSs, due to the scalar field sourced by the strong internal gravitational field of the star. When placing limits on β_0 it is important that for certain areas in the DEF gravity parameter space one has a significant growth of the scalar field in the NS interior, even if the deviations for weakly self-gravitating bodies are very small, or even zero (Damour & Esposito-Farèse 1992a, 1993, 1996). The origin of this spontaneous scalarisation lies in the tachyonic instability of the field equations happening when a NS reaches a sufficiently high compactness. The mechanism of this instability is similar to what happens in ferromagnets during spontaneous magnetisation (Damour & Esposito-Farèse 1996). It is a fully non-linear non-perturbative effect, thus it requires accurate techniques to be properly analysed.

For testing DEF gravity, it is crucial to introduce three gravitational form factors appearing in the expressions for PK parameters

$$\alpha_A = \left. \frac{\partial \ln m_A}{\partial \varphi_0} \right|_{\bar{m}_A}, \quad \beta_A = \left. \frac{\partial \alpha_A}{\partial \varphi_0} \right|_{\bar{m}_A}, \quad k_A = - \left. \frac{\partial \ln I_A}{\partial \varphi_0} \right|_{\bar{m}_A}, \quad (6)$$

where all the derivatives are taken for a fixed baryonic mass (\bar{m}_A). The important quantity α_A counts the effective couple strength between the ambient scalar field and a NS. The subscript A refers to the star A. α_A is a strong-field counterpart to the parameter α_0 , which is its weak field approximation. If a NS mass exceeds the critical mass of m_{crit} then α_{NS} can become suddenly of order of unity $\alpha_{\text{NS}} \sim O(1)$ due to the scalarisation effect, while its weak field counterpart stays close to zero $\alpha_0 \sim 0$. The value of the critical mass depends on β_0 and the chosen equation of state (EOS). The second parameter β_A reflects a change of the scalar charge when the asymptotic scalar field changes. The derivative β_A can also obtain very large values in the transition scalarisation region, where α_A changes fast with the change of a mass. The last quantity k_A describes the field dependence of the moment of inertia I_A and can become important in the eccentric systems.

2.3. Calculating grids of NSs

To test DEF gravity in precision timing experiments of binary pulsars, one must know the exact properties of NSs in that specific class of alternative gravity. In our work, we follow the procedure described in [Damour & Esposito-Farèse \(1996\)](#) for calculating NSs in DEF gravity. We assume the slowly rotating NS approximation with a stationary axisymmetric metric. Assuming this axial symmetry and neglecting terms of rotational velocity squared (or higher), the Einstein equations can be written as a system of eight first-order ordinary differential equations (ODEs) for a set of radial functions (see the representation by [Damour & Esposito-Farèse 1996](#)). The equations are complemented by appropriate boundary conditions placed at spatial infinity and the NS centre. Boundary conditions at spatial infinity and thus properties of NSs depend on the values of the theory parameters $\{\alpha_0, \beta_0\}$.

For solving NS structures, we developed a modular program in the JULIA language ([Bezanson et al. 2017](#)). The program enables us to calculate both single NSs for selected parameters and grids of NSs for many desired parameters. The calculations are performed accurately and efficiently using the methods of parallel programming. Currently, DEF gravity is implemented as the chosen theory of gravity, but our program allows one to use any coupling function $A(\varphi)$ to extend our tests beyond DEF gravity parametrisation.

The internal structure and NSs parameters significantly depend on the selected EOS. The strong dependence of the scalarisation phenomenon on the choice of a EOS was extensively analysed by [Shao et al. \(2017\)](#) and [Zhao et al. \(2022\)](#) in order to put constraints on this non-linear phenomenon from radio pulsars. In our work, we use a piece-wise polytropic approximation for the EOSs following the procedure in [Read et al. \(2009\)](#). We select 11 EOSs that allow us to cover a range from soft to stiff while still being in agreement with the maximum mass observed for a NS (for more information about selected EOSs we refer to [Appendix A.2](#)). For tests in the present paper, we select the MPA1 high-density EOS ([Müther et al. 1987](#)) in its piece-wise polytropic form ([Read et al. 2009](#)) with three polytropic pieces. MPA1 is a stiff EOS with a maximum mass of a NS equal to $2.46 M_{\odot}$. For low densities,

we use an analytic form of the SLy EOS proposed by [Douchin & Haensel \(2001\)](#) and approximated by four polytropic pieces ([Read et al. 2009](#)). In general stiff EOSs (e.g. MPA1) produce more conservative limits in most parts of the DEF gravity parameter space.

In our program, we utilise automatic differentiation (AutoDiff) technique (see e.g. [Margossian 2018](#)) for calculating derivatives of NSs quantities. AutoDiff is a powerful alternative to numerical differentiation deprived of numerical errors (for detailed explanation see [Appendix A.1](#)). AutoDiff algorithm utilises exact expressions of derivatives for elementary functions and a chain rule to calculate complex derivatives with the working machine-precision. Thus our program with applied AutoDiff technique calculates gravitational form factors $\{\alpha_A, \beta_A, k_A\}$, presented as derivatives in [Eq. \(6\)](#), with very high numerical precision.

Using the developed program, we calculate gravitational form factors and masses for an extensive grid of DEF parameters and central pressures $\{\alpha_0, \beta_0, p_c\}$ assuming a specific EOS. Our grids cover ranges of $\alpha_0 \in [-10^{-1}, -10^{-5}]$, $\beta_0 \in [-6, +10]$, $p_c \in [10^{34}, 10^{36}] \text{ dyn cm}^{-2}$ and are sampled linearly for $\log \alpha_0$, $\log p_c$ and β_0 . The size of each grid is $\{101, 351, 121\}$ points respectively. More technical details about the procedure of calculating NS properties, AutoDiff technique, and the structure of pre-calculated grids can be found in [Appendix A](#). We stick to pre-calculating grids because it is numerically costly to calculate NS structures for a particular mass on the fly. Further one can perform interpolation over these pre-calculated grids or methods beyond interpolation, e.g reduced-order models in [Guo et al. \(2021a\)](#). The achieved accuracy of grid values and further interpolation over the grid is sufficient for our purposes.

The accuracy of the calculations can be checked by comparison of the scalar charge (α_A) calculated as the derivative from [Eq. \(6\)](#) $\alpha_{A, \text{deriv}}$ and its direct value from the asymptotic behaviour of the external scalar field obtained after integration $\alpha_{A, \text{asympt}}$ ([Damour & Esposito-Farèse 1993](#)). The relative accuracy expressed by $\delta_{\text{rel}} = |1 - \alpha_{A, \text{deriv}}/\alpha_{A, \text{asympt}}|$ lies in the range $\delta_{\text{rel}} \sim 10^{-7} - 10^{-13}$ for our calculations. The calculated scalar charges are significantly more accurate compared to previous works, for instance, [Anderson & Yunes \(2019\)](#) with typical accuracy of $\delta_{\text{rel}} \sim 10^{-1} - 10^{-5}$. However, it is essential to mention that the accuracy of several percent is already sufficient for testing DEF gravity with binary pulsar timing.

Besides, we also find that the NS structure appears to be very sensitive to internal thermodynamic consistency. The piece-wise polytropic EOSs fulfil the thermodynamic consistency by their definition. However, often one uses tabulated EOSs instead, which must be interpolated. Unfortunately, a commonly used log-log interpolation of tabulated EOSs can lead to significant numerical errors in the scalar derivatives of [Eq. \(6\)](#) because of its thermodynamic inconsistency³. For this reason, we use piece-wise polytropic approximations instead of tabulated EOSs,

³ Doing an interpolation of a tabulated EOS in a way that correctly accounts for the first law of thermodynamics is not a trivial problem and requires a formulation in terms of Helmholtz free energy ([Swesty 1996](#)). If one uses a simple linear interpolation in $\{\log n, \log p\}$ and $\{\log n, \log \varepsilon\}$ planes for a tabulated EOS (n is the baryon number density and ε is the energy density), the accuracy check can fail by the amount of $\delta_{\text{rel}} \sim O(1) - O(0.1)$ in the region of strong scalarisation. The reason is that such interpolations do not obey the first law of thermodynamics $\frac{dn}{d\varepsilon} = \frac{n}{\varepsilon+p}$ which is assumed for structure equations and defines the baryonic mass \bar{m}_A ([Hartle 1967](#); [Damour & Esposito-Farèse 1996](#)). As a consequence, this not only affects NS masses but even more so the derivatives in [Eq. \(6\)](#), including the expression of α_A from [Eq. \(6\)](#) as the derivative of the mass which enters the accuracy check.

which is absolutely sufficient for conducting our tests and for exploring the EOS dependence of pulsar constraints on DEF gravity.

3. A new timing model for scalar-tensor gravity

To place constraints on a gravity theory from the observations of a binary pulsar, one must know its predictions for the binary system's motion and the propagation of the electromagnetic signal in the curved spacetime of the binary system. The timing formula is the tool to capture the relativistic effects predicted by the theory from a sequence of pulse arrival times on Earth. The timing formula relates the observed (topocentric) time of arrival (TOA) of the pulse and its time of emission.

3.1. Binary pulsar timing models

The first timing model 'BT' was introduced by [Blandford & Teukolsky \(1976\)](#) in order to describe the timing of the first binary pulsar, PSR B1913+16, which had been discovered earlier by [Hulse & Taylor \(1975\)](#). The BT model assumes that the pulsar and its companion follow a Keplerian motion with additional secular changes in the Keplerian parameters of the orbit. utilised Keplerian parameters are the orbital period (P_b), the epoch of periastron passage (T_0), the orbital eccentricity (e), the longitude of periastron (ω), and the projected semi-major axis of the pulsar's orbit (x). The model accounts for a combination of a special-relativistic time dilation and a gravitational redshift; this periodic effect is described by an extra parameter called the Einstein delay (γ). The BT model accounts for (linear-in-time) secular changes in P_b , e , ω and x , introducing PK parameters: rates of change of the orbital period (\dot{P}_b), eccentricity (\dot{e}), longitude of periastron ($\dot{\omega}$) and projected semi-major axis (\dot{x}). These secular changes can be caused by both relativistic and astrometric effects ([Damour & Taylor 1992](#); [Lorimer & Kramer 2004](#)).

Later, [Damour & Deruelle \(1986\)](#) proved that all of the independent $O(v^2/c^2)$ timing effects could be described in a simple mathematical way for a wide range of alternative gravity theories. They developed a phenomenological (i.e. theory independent) timing model based on the full first post-Newtonian description of the two-body problem, which we refer to as the 'DD' model that uses a quasi-Keplerian solution to the equations of motion of a two-body problem. This model allows working within the PPK approach. [Damour & Taylor \(1992\)](#) then showed that the DD model could be used to constrain a wide range of conservative theories of gravity obeying the modified Einstein–Infeld–Hoffmann (mEIH) framework (see also [Will 1993](#)).

For the binary system part, the timing formula of the DD model reads as:

$$t_b - t_0 = D^{-1} [T + \Delta_R(T) + \Delta_E(T) + \Delta_S(T) + \Delta_A(T)], \quad (7)$$

where t_b is the Solar System barycentric (infinite-frequency) arrival time, t_0 is a chosen reference epoch, T is the pulsar's proper time, and D is a Doppler factor accounting for the relative radial motion of the centre of mass of the binary system and the Solar System barycentre. The quantities Δ_i in Eq. (7) are different time delays introducing corrections due to internal binary effects.

Splitting the timing formula into a set of different contributions is to some extent a coordinate-dependent concept, however, within the approximations used it is convenient to work with

these individual expressions for timing delays. The term $\Delta_R(T)$ is called 'Roemer delay' and counts for the classical light travel time through the binary. $\Delta_E(T)$ is the 'Einstein delay' and relates the proper emission time to the coordinate time of emission. $\Delta_S(T)$ is the 'Shapiro delay' arising from the effect of the gravitational potential of the companion on the propagation of the pulsar signals. The 'aberration delay' $\Delta_A(T)$ places corrections due to the periodic changes in the direction of pulse emission (as seen in the frame of the rotating pulsar) while the pulsar follows its binary motion.

All the expressions depend on three sets of parameters. Keplerian parameters are

$$\{p^K\} = \{P_b, T_0, e_0, \omega_0, x_0\} \quad (8)$$

and remain the same as for the BT model (the subscript 0 means a value at a given epoch). Separately measurable PK parameters are

$$\{p^{PK}\} = \{k, \gamma, \dot{P}_b, r, s, \delta_\theta, \dot{e}, \dot{x}\}. \quad (9)$$

The DD model introduces a periastron-shift parameter $k = \omega P_b / (2\pi)$, 'Shapiro shape' ($s = \sin i$, where i is the inclination angle) and 'Shapiro range' (r) parameter for the signal propagation, and a relativistic deformation of the orbit (δ_θ). Not separately measurable PK parameters are

$$\{q^{PK}\} = \{\delta_r, A, B, D\}, \quad (10)$$

including a second parameter for the relativistic deformation (δ_r), two aberration parameters (A and B) and a Doppler factor (D).

The mentioned delays in the framework of the DD model are presented by expressions

$$\Delta_R(T) = x \sin \omega [\cos u - e(1 + \delta_r)] + x [1 - e^2(1 + \delta_\theta)^2]^{1/2} \cos \omega \sin u, \quad (11)$$

$$\Delta_E(T) = \gamma \sin u, \quad (12)$$

$$\Delta_S(T) = -2r \ln \left\{ 1 - e \cos u - s [\sin \omega (\cos u - e) + (1 - e^2)^{1/2} \cos \omega \sin u] \right\}, \quad (13)$$

$$\Delta_A(T) = A \{ \sin [\omega + A_e(u)] + e \sin \omega \} + B \{ \cos [\omega + A_e(u)] + e \cos \omega \}, \quad (14)$$

where the secular changes incorporated in the projected semi-major axis x and eccentricity e are given by

$$x = x_0 + \dot{x}(T - T_0), \quad (15)$$

$$e = e_0 + \dot{e}(T - T_0). \quad (16)$$

The true anomaly (A_e) and ω depend on the eccentric anomaly (u) via following the relations

$$A_e(u) = 2 \arctan \left[\left(\frac{1+e}{1-e} \right)^{1/2} \tan \frac{u}{2} \right], \quad (17)$$

$$\omega = \omega_0 + k A_e(u), \quad (18)$$

where u is itself a function of T and defined by solving Kepler's equation

$$u - e \sin u = 2\pi \left[\left(\frac{T - T_0}{P_b} \right) - \frac{1}{2} \dot{P}_b \left(\frac{T - T_0}{P_b} \right)^2 \right]. \quad (19)$$

The parameters in Eq. (10) cannot be measured separately from the parameters in Eqs. (9) and (8) as shown by Damour & Deruelle (1986). It means that all not separately measurable parameters can be fully absorbed into the change of other parameters. More details about measurable parameters and the connection between observed parameters and the intrinsic parameters can be found in Damour & Taylor (1992). The redefinition of orbital parameters can absorb the Doppler factor (D). Such a redefinition does not affect the tests because it is only a rescaling of physical units, thus D is set to one⁴. The aberration parameters A and B can be absorbed as well by redefinition of T_0 , x , e , δ_r , and δ_θ .

3.2. A timing model in DEF gravity

The DD model can be applied to a wide range of fully conservative theories of gravity. Damour–Esposito–Farèse gravity is a fully conservative theory (Will 1993), thus we can investigate it in the framework of the DD model.

In different theories of gravity, the functional dependence of the PK parameters, i.e.

$$p_i^{\text{PK}} = f_i^{\text{theory}}(m_p, m_c; P_b, e_0, x_0; \text{EOS}), \quad (20)$$

can differ substantially because of strong-field effects involving a highly compact NS and its companion. The PK parameters are functions of two masses, thus one can perform gravity tests if at least three PK parameters are measurable. In general, if there are n measured PK parameters, one can do $n - 2$ independent tests for a given gravity theory. A common procedure of making tests using expressions for PK parameters is discussed in detail in Sect. 4, devoted to applications of a new model and comparison between different techniques.

The core part of the new timing model is the predictions for PK parameters (20) in DEF gravity. We use expressions for PK parameters in DEF gravity provided in the literature (Damour & Esposito-Farèse 1992b, 1993, 1996). However, for parameters δ_θ , δ_r , A and B there are only more general expressions. A phenomenological theory-independent description of PK parameters from which we derive their expressions in DEF gravity is presented in Damour & Taylor (1992). The expressions of PK parameters and details are given in Appendix B.

The aberration parameters A and B can also be calculated in DEF gravity, provided we know the system’s geometry. The applications we use in the paper assume a special situation when the pulsar’s spin axis is aligned to the orbital angular momentum of the system. The alignment is a reasonable assumption for a recycled pulsar with a WD companion due to the preceding accretion process during the system’s evolution. In this special case, A is calculated according to Eq. (B.7) assuming angles $\eta = -\pi/2$ and $\lambda = i$ (Damour & Taylor 1992). The second aberration parameter $B = 0$ in this special situation. In general, an alignment may not be the case for a particular binary pulsar (e.g. double NS systems), but the new model can also handle this misalignment. Once the system’s geometry is assumed, real A and B values can be calculated for a given epoch without the necessity to redefine other parameters⁵.

⁴ We note, however, that a temporal change of D is of relevance for pulsar timing experiments, as we discuss later.

⁵ If the spin of the pulsar is misaligned, this can lead to a further complication of the analysis. In such a case geodetic precession leads to apparent changes in the binary parameters (Damour & Taylor 1992; Lorimer & Kramer 2004).

The PK parameters from Eq. (20) depend on the orbital parameters of the system and the physical properties of the pulsar. However, we also have to know properties of the companion. The PK expressions (B.1–B.13) depend on gravitational form factors of the companion $\{\alpha_c, \beta_c\}$. The quantity k_c does not appear in the expressions because they take into account only leading order terms, and is therefore of no interest. These quantities, in turn, depend on the type of companion. If the companion is a NS, they are calculated in the same way as for the primary pulsar using Eq. (6). If the companion is a BH, then the gravitational form factors all vanish:

$$\alpha_{\text{BH}} = 0, \quad \beta_{\text{BH}} = 0, \quad k_{\text{BH}} = 0. \quad (21)$$

This is a consequence of the ‘no-hair’ theorem; the BH is not scalarised in DEF gravity (Hawking 1972; Damour & Esposito-Farèse 1992a). Section 6 is dedicated to a more thorough discussion on PSR–BH systems and the results which we can obtain from timing of a pulsar in a relativistic orbit with a BH. For a WD companion, the gravitational form factors are approximated by their weak field expressions

$$\alpha_{\text{WD}} \simeq \alpha_0, \quad \beta_{\text{WD}} \simeq \beta_0, \quad k_{\text{WD}} \simeq 0. \quad (22)$$

Such an approximation is sufficient for our purposes because WDs are not compact enough to show the strong field effects present in NSs and BHs. The next order approximation for weakly self-gravitating objects is $\alpha_A \simeq \alpha_0(1 - 2s_A)$, where $s_A \simeq G_* m_A / (Rc^2)$ is the sensitivity (R is the object’s radius) and has a typical value of $s_{\text{WD}} \sim 10^{-5} - 10^{-3}$ for a WD (Damour & Esposito-Farèse 1992a, 1993; Will 2018b). Thus the usage of weak field counterparts $\{\alpha_0, \beta_0\}$ for a WD companion instead of precisely calculated values is justified.

To summarise, the timing model in DEF gravity is defined by two theory parameters $\{\alpha_0, \beta_0\}$, the chosen EOS for a NS, and the type of the companion among {WD, NS, BH}. It concludes the theoretical description of the DDSTG model.

3.3. DDSTG implementation into TEMPO

Once we obtain the theoretical part of the new DDSTG timing model, we need to apply it to the timing data. We implemented the DDSTG model into one of the commonly used pulsar timing software – the TEMPO⁶ program (Nice et al. 2015). The local implementation of the DDSTG model in TEMPO and precalculated grids of masses and gravitational form factors are supplied with the paper⁷. The authors intend to make the DDSTG model a part of the official TEMPO distribution to ensure forward compatibility.

There is a standard procedure of analysing radio pulsar timing data. The preprocessed TOAs are obtained after radio observations of the desired pulsar system. During the procedure, TEMPO reads TOAs, parameters of the binary model, and some coded instructions from supplied files. Then TEMPO fits the selected timing model accounting for the transformation to the Solar System barycentre, pulsar rotation, and its spin down for a chosen binary model.

Specifically for the DDSTG model, a user selects theory parameters $\{\alpha_0, \beta_0\}$, a EOS, and a type of a companion. During the initialisation TEMPO reads pre-calculated

⁶ <http://tempo.sourceforge.net>

⁷ https://github.com/AlexBatrakov/tempo-13.103_ddstg

three-dimensional (3D) grids of gravitational form factors and NS masses $\{\alpha_A, \beta_A, k_A, m_A\}$ (see Sect. 2.3) for the chosen EOS. Each 3D grid depends on $\{\alpha_0, \beta_0, p_c\}$ parameters, where p_c is the central pressure. Then gravitational form factors and masses are interpolated for the selected $\{\alpha_0, \beta_0\}$ values and saved into a smaller 1D grids (depending only on p_c) in the program memory. The final 1D grids remain unchanged and are used to interpolate gravitational form factors for a particular theory in the mass range during the fitting procedure of TEMPO.

TEMPO with the selected DDSTG model fits two masses: the total mass (m_{tot}) of the system and the companion's mass (m_c). Every time when these masses change, the model recalculates the gravitational form factors for the pulsar and the companion. Once the gravitational form factors (Eq. (6)) are known, the model calculates all PK parameters (Eqs. (9) and (10)) using equations given in Appendix B. These calculated PK parameters are used afterwards in the timing formula (7). The details about the DDSTG implementation and the description of model parameters can be found in Appendix C.

In the timing formula (7), the Solar System barycentric arrival time (t_b) is known. However, to obtain the proper pulsar time (T), one has to perform the inversion of the timing formula, i.e. get T as a function of t_b : $T = t_b - \bar{\Delta}(t_b)$. The original DD model utilises an approximate analytic inversion, which sometimes is not accurate enough, for example, for the double pulsar PSR 0737–3039A (Kramer et al. 2021). In contrast, our DDSTG model performs accurate numerical inversion. Equation (7) is solved iteratively for T for each TOA. The discrepancy due to an inaccurate inversion may influence the test for precise timing in the future or PSR-BH systems discussed further below. Nowadays, numerical inversion is considered a new standard and implemented in the DDS model of TEMPO (Kramer et al. 2021) and generally used in PINT (Luo et al. 2021).

The DDSTG model produces the best fit to the data for DEF gravity with a selected set of parameters and a given EOS. The output of TEMPO stays the same as for other binary models, for instance, the calculated χ^2 value and root mean square error of fit. The χ^2 statistics may be applied for further tests to compare the results for different theory parameters.

3.4. Advantages of DDSTG

The tests with pulsar data allow for constraints to be placed on the gravity theory parameters. The most simple and common way to achieve this is to compare the measured PK parameters with the theoretical predictions from the theory. PK parameters are obtained by fitting a phenomenological model to the observational data, for instance, the DD model. A phenomenological model estimates PK parameters and their uncertainties.

Within a particular gravity theory, each PK parameter depends on the masses of the pulsar and the companion and thus corresponds to a curve in a mass-mass space. Together with the measurement error, each PK parameter produces a strip in the $\{m_p, m_c\}$ space. If there are two measured parameters, one may find the intersection area and obtain the estimated mass values. If the number of measured parameters is more than two, one can perform a test of that particular gravity theory. The test is passed if all three curves intersect in the range of errors at one point. Generally, n measured PK parameters result in $n - 2$ independent tests.

If a gravity theory has arbitrary parameters, tests can be done for any fixed values of theory parameters. For DEF gravity, this

procedure results in an ‘allowed’ area in $\{\alpha_0, \beta_0\}$ space that pass gravity tests within the measurement uncertainties. This area is bounded within the selected confidence limit by a curve, which is usually plotted. To date, the regions allowed by different tests have always included GR.

If the companion is optically bright, one can get information about the system from optical observations. For example, for a binary pulsar system with a bright WD it is possible to obtain the mass ratio and the companion mass using high-resolution optical spectroscopy of the companion (e.g. Antoniadis et al. 2012, 2013). In such cases, we do not have to measure three independent PK parameters based on the radio data, but we can combine several multi-wavelength constraints.

Another issue are possible correlations between PK and other parameters. Observed correlations can come from the theoretical correlation of parameters within the binary model (e.g. $T_0 \leftrightarrow \omega$ and $P_b \leftrightarrow \dot{\omega}$ in a low eccentric case) and from a non-uniform data sampling (e.g. a non-linear correlation in a parametrisation of Shapiro delay $r \leftrightarrow s$). Often in the past, measured PK parameters were published without any information about observed correlations. This additional information from the timing data is lost in such cases. These days it became a common practice to publish observed correlations and even provide them explicitly. Anderson & Yunes (2019) and Anderson et al. (2019) accounted for possible observed and theoretical correlations between PK parameters within DEF and Mendes-Ortiz gravity (MO, Mendes & Ortiz 2016) via computationally highly demanding Markov chain Monte Carlo (MCMC) simulations utilising the correlation matrix from the TEMPO output. However, this elaborate approach would not work if the relativistic effect is not well-measured or the dependence between the parameters is highly non-linear. If the relation is non-linear, the correlation matrix gives information only about the linear contribution. To fully account for a non-linear relation, one has to obtain a full probability density function in the parameter space.

Contrary to the PK method, the DDSTG model accounts for all even non-linear correlations naturally and breaks them by a direct fit of the masses. This was already one of the main advantages of the DDGR model; and is one of the reasons why the DDSTG model uses the same superior approach while extending it to STG theories. The model accounts for all effects internally and extracts all the information from the timing data. The information is not lost even if the correlating parameter cannot be measured but influences other parameters. Such possible correlations are relevant for PSR J1141–6545 (Venkatraman Krishnan 2019; Venkatraman Krishnan et al. 2020), where there is a correlation between the time dilation parameter (γ) of the Einstein delay and the rate of change of the projected semi-major axis due to the spin-orbit coupling caused by the fast rotating WD companion (\dot{x}^{SO}). Furthermore, the Shapiro delay cannot be measured independently but is still essential (see e.g. Bhat et al. 2008). We expect the DDSTG model to be of particular advantage for PSR J1141–6545 and there will be a dedicated paper (Venkatraman Krishnan et al., in prep.) about applying the new model to this system.

Compared to conventional methods, the constraints obtained by the DDSTG model may, in general, become either tighter or weaker depending on the particular case. However, the resulting restrictions are more reliable by the construction. This weakening of the limits can happen because of unaccounted correlations.

4. Application to PSR J2222–0137

As a demonstration of the DDSTG model, we now apply it to published timing data of a binary pulsar. We select PSR J2222–0137 because of its unusual characteristics, which we now list in detail.

4.1. About the system

We are particularly interested in recycled pulsars because of their, in general, better timing precision essential for precise gravity tests. The radio pulsar PSR J2222–0137 was discovered in the Green Bank Telescope (GBT) 350 MHz drift scan pulsar survey (Boyles et al. 2013). It is a mildly recycled pulsar which has a spin period (P) of 32.8 ms. The pulsar is in a binary system with P_b of 2.44576 days and x of 10.848 light-seconds. We also expect the pulsar’s spin axis to be aligned to the orbital angular momentum of the system. The alignment happens as a consequence of the pulsar recycling process, when the pulsar accretes matter from the companion.

PSR J2222–0137 is already known as a unique laboratory for testing gravity theories because of its special characteristics; for more detailed information, we refer the reader to Cognard et al. (2017) and Guo et al. (2021b). It is one of the closest pulsars known, with the most precise distance measured with Very Large Baseline Interferometry (VLBI, see Deller et al. 2013) and excellent timing precision. The system has a highly significant detection of the Shapiro delay as well as the measured rate of advance of periastron ($\dot{\omega}$), which yield precise mass measurements and $\sim 1\%$ test of the GR predictions for the Shapiro delay (Guo et al. 2021b).

The most important characteristic in the scope of this paper is the precise measurement of \dot{P}_b . Given the precise masses, this can be compared with a precise prediction for the orbital decay due to GW damping, furthermore the kinematic contributions to \dot{P}_b (see Sect. 4.4) can be estimated precisely because of the good distance measurement. The \dot{P}_b measurement constrains dipolar GW emission in this system, which would arise in DEF gravity (Damour & Esposito-Farèse 1992a) because of the large difference in the compactness of the components (NS and WD). Finally, the mass of the pulsar ($m_p = 1.831(10) M_\odot$) (Guo et al. 2021b) lies in the range where spontaneous scalarisation happens, yielding strong limits on this highly non-linear phenomenon (Shao et al. 2017; Zhao et al. 2022).

4.2. Observational data

The PSR J2222–0137 timing data used in this work are the same as used by Guo et al. (2021b). This includes observations of this pulsar going back to its original follow-up, which started on 2009 June 23 using the Green Bank Telescope (GBT); however these ended on 2011 December 26. Regular observations with the three largest European radio telescopes (3ERT) started the following years: the Nançay Radio Telescope (NC) on 2012 October 4, The Lovell Telescope (LT) at Jodrell Bank on 2012 November 20 and the dense orbital campaigns with the Effelsberg 100-m radio telescope (EFF) on 2015 October 26. These observations continue to the present day; however Guo et al. (2021b) only analysed the data obtained until 2021 May 2; their processing and how the resulting ToAs were analysed is described in detail by Cognard et al. (2017) and Guo et al. (2021b).

Observations with the MeerKAT radio telescope array and the Five hundred meter Aperture Spherical Telescope (FAST) started in 2019 September 24 and 2020 October 5, but these

timing data were not used by Guo et al. (2021b). However, the latter authors did use FAST data for a detailed study of the emission properties of PSR J2222–0137. In this work, we also extend the existing data by simulated data assuming the same timing properties of all of these telescopes to simulate how the timing parameters for this system will improve in the near future (see Sect. 5).

4.3. Mass-mass diagrams in DEF gravity

The illustrative and straightforward way to explore if a given gravity theory agrees with binary pulsar observational data is to calculate mass-mass diagrams. As mentioned in Sect. 3.4, firstly, one measures values of PK parameters with their uncertainties by fitting the phenomenological DD binary model to binary pulsar data. Then the observed PK values can be compared with their theoretical predictions of a specific gravity theory. The PK parameters from Eq. (20) depend on the mass of the pulsar (m_p) and the mass of the companion (m_c) (see Appendix B). For PSR J2222–0137 we use the measured PK parameters from Guo et al. (2021b).

To illustrate how the tests are performed, we select two points in DEF gravity parameter space near the scalarisation region with large negative β_0 and calculate mass-mass diagrams for them. The test is sensitive to the predicted dipolar contribution to \dot{P}_b , which rises dramatically in the region of spontaneous scalarisation. The first point with $\alpha_0 = -10^{-4}$, $\beta_0 = -4.3$ does not predict a strong enough dipolar contribution and passes the test within a 1σ limit. The corresponding mass-mass diagram is presented in the left panel of Fig. 1. In contrast, the second point with $\alpha_0 = -10^{-4}$ and a bit smaller $\beta_0 = -4.35$ is already excluded because of a rather strong scalarisation of the pulsar leading to significant dipolar GW damping (see the right panel of Fig. 1). The \dot{P}_b curve nicely shows that a significant enhancement in the scalarisation happens for a particular interval in the mass range defined by the EOS.

4.4. Various contributions to \dot{P}_b

At this point we need to discuss the different contributions to the observed \dot{P}_b^{obs} and their influence on our results. The observed orbital decay of the system consists of many terms

$$\dot{P}_b^{\text{obs}} = \dot{P}_b^{\text{GW}} + \dot{P}_b^{\text{Gal}} + \dot{P}_b^{\text{Shk}} + \dot{P}_b^{\text{Tid}} + \dot{P}_b^{\dot{M}} + \dot{P}_b^{\dot{G}}, \quad (23)$$

where the first term \dot{P}_b^{GW} is due to GW damping and can include dipolar and monopolar terms in DEF gravity, besides the general quadrupolar prediction of GR. The full expressions of different GW contributions are presented in Appendix (B.9). The next two terms are of kinematic origin: the Shklovskii effect (\dot{P}_b^{Shk}) and the Galactic contribution (\dot{P}_b^{Gal}). They are the result of a time-varying Doppler factor D due to an (apparent) radial acceleration between the pulsar binary and the Solar System (Damour & Taylor 1991). The last three terms come from tidal effects, mass loss in the system, and a possible temporal variation of the gravitational constant G .

We are mainly interested in measuring the GW emission term (\dot{P}_b^{GW}) and comparing it with the prediction of DEF gravity. The most significant additional effects in PSR J2222–0137 come from kinematic contributions \dot{P}_b^{Gal} and \dot{P}_b^{Shk} . These two effects arise beyond the binary system and can be combined in the overall external contribution

$$\dot{P}_b^{\text{ext}} = \dot{P}_b^{\text{Gal}} + \dot{P}_b^{\text{Shk}}. \quad (24)$$

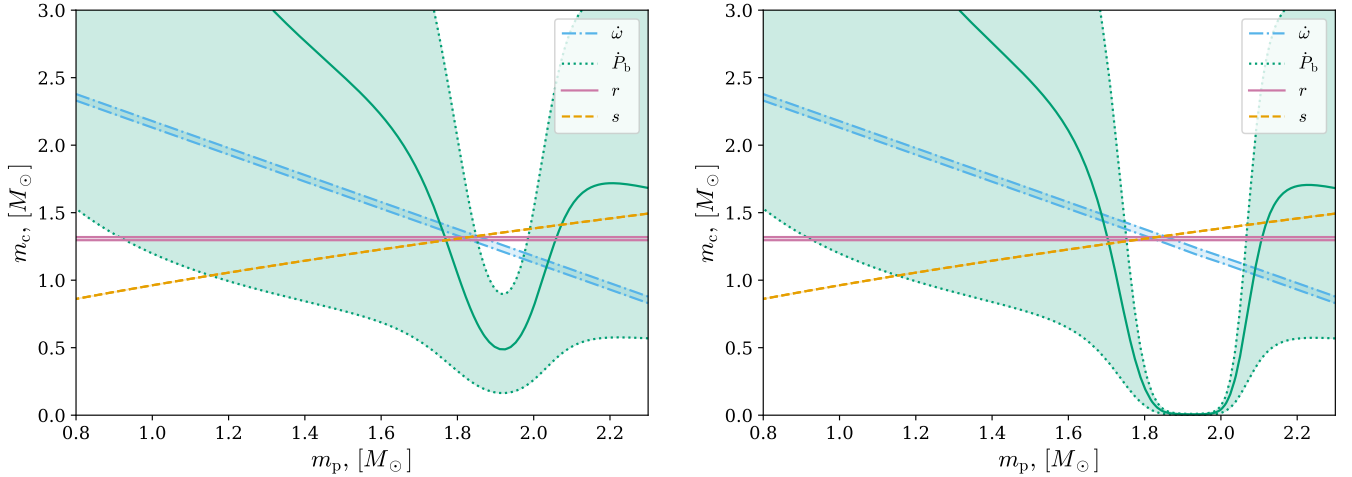


Fig. 1. Mass-mass diagrams for the PSR J2222–0137 system at the zone of high non-linearity where a fast transition to a strongly scalarised NS happens. PK parameters are obtained by fitting the DD model to the timing data from Guo et al. (2021b). Calculations are performed within DEF gravity for the MPA1 EOS allowing NSs with maximum mass in GR of $M_{\text{GR}} \simeq 2.46 M_{\odot}$. Left panel corresponds to the point in the DEF gravity parameter space with $\alpha_0 = -10^{-4}, \beta_0 = -4.3$, which is not excluded by the test. Right panel shows the same test but for $\alpha_0 = -10^{-4}, \beta_0 = -4.35$. The prediction of a strong dipolar contribution \dot{P}_b^D fails the test for the more negative $\beta_0 = -4.35$. The axes correspond to the masses of the pulsar m_p and the companion c . The shadowed area is the allowed region at 68% confidence level (CL) limit for a corresponding PK parameter. The solid green line corresponds to the observed value of \dot{P}_b^{GW} due to the GW emission.

The last three terms in Eq. (23) are parts of the internal contribution

$$\dot{P}_b^{\text{int}} = \dot{P}_b^{\text{GW}} + \dot{P}_b^{\text{Tid}} + \dot{P}_b^{\text{M}} + \dot{P}_b^{\text{G}}, \quad (25)$$

and can be neglected for PSR J2222–0137. Thus the internal contribution has only one significant term left, the GW term. For the discussion about possible account of \dot{P}_b^{G} we refer the reader to Sect. 8. Within the DDSTG model, the parameter ‘XPBDOT’ in TEMPO is used for treating both external and internal effects $\dot{P}_b^{\text{ext}} + \dot{P}_b^{\text{int}} - \dot{P}_b^{\text{GW}}$ simultaneously subtracting the GW emission term. In our case XPBDOT accounts for only external Shklovskii and Galactic effects.

There is an extensive analysis of the external terms for PSR J2222–0137 by Guo et al. (2021b) and their determined values are: $\dot{P}_b^{\text{Gal}} = -0.0142(13) \times 10^{-12} \text{ s s}^{-1}$ and $\dot{P}_b^{\text{Shk}} = 0.2794(12) \times 10^{-12} \text{ s s}^{-1}$. These values correspond to the external variation of the observed orbital period of

$$\dot{P}_b^{\text{ext}} = 0.2652(18) \times 10^{-12} \text{ s s}^{-1}, \quad (26)$$

which is consistent within 2σ with the total observed value

$$\dot{P}_b^{\text{obs}} = 0.2509(76) \times 10^{-12} \text{ s s}^{-1}, \quad (27)$$

leaving the internal contribution $\dot{P}_b^{\text{int}} = \dot{P}_b^{\text{obs}} - \dot{P}_b^{\text{ext}} = -0.0143(78) \times 10^{-12} \text{ s s}^{-1}$ consistent with the GR prediction for quadrupolar GW emission $\dot{P}_b^{\text{GR}} = -0.00809(5) \times 10^{-12} \text{ s s}^{-1}$ (Guo et al. 2021b).

According to Damour & Taylor (1991), Nice & Taylor (1995), and Lazaridis et al. (2009), the Galactic differential acceleration may be analytically approximated with the expression

$$\frac{\dot{P}_b^{\text{Gal}}}{P_b} = -\frac{K_z |\sin b|}{c} - \frac{\Theta_0^2}{cR_0} \left(\cos l + \frac{\beta}{\beta^2 + \sin^2 l} \right) \cos b, \quad (28)$$

where l is Galactic longitude, b the Galactic latitude and $\beta = (d/R_0) \cos b - \cos l$. The quantity K_z is the vertical component of the Galactic acceleration, which for Galactic heights

$z \equiv |d \sin b| \leq 1.5 \text{ kpc}$ can be approximated with sufficient accuracy by

$$K_z (10^{-9} \text{ cm s}^{-2}) \simeq 2.27 z_{\text{kpc}} + 3.68 (1 - e^{-4.31 z_{\text{kpc}}}), \quad (29)$$

where $z_{\text{kpc}} \equiv z(\text{kpc})$ (Holmberg & Flynn 2004; Lazaridis et al. 2009). From GRAVITY Collaboration (2021) we can take a value for the Sun’s Galactocentric distance $R_0 = 8275 \pm 9 \pm 33 \text{ pc}$. The Galactic circular velocity at the location of the Sun (Θ_0) is taken to be $240.5(41) \text{ km s}^{-1}$ (see Guo et al. 2021b, and references therein).

The Shklovskii contribution (Shklovskii 1970) can be calculated using

$$\dot{P}_b^{\text{Shk}} = \frac{(\mu_\alpha^2 + \mu_\delta^2) d}{c} P_b, \quad (30)$$

where μ_α and μ_δ are the proper motion in right ascension (RA) and declination (DEC), respectively, and d is the distance to the pulsar. The astrometric values and uncertainties for PSR J2222–0137 are taken from Guo et al. (2021b). For PSR J2222–0137, the uncertainty in \dot{P}_b^{ext} is small compared to the precision of the \dot{P}_b measurement. The uncertainty of \dot{P}_b^{ext} in our case can be neglected and therefore we can provide the fixed XPBDOT parameter in TEMPO.

4.5. Results of applying DDSTG

Our main goal is to obtain limits on DEF gravity parameters $\{\alpha_0, \beta_0\}$ by applying the DDSTG model. TEMPO allows one to calculate a χ^2 value for every particularly selected pair of values (α_0, β_0) . Therefore we obtain χ^2 values for a grid of parameters and compare them to each other. For straightforward comparison we subtract the minimum χ_{min}^2 over the $\{\alpha_0, \beta_0\}$ grid from calculated χ^2 . The location of the minimum χ_{min}^2 is statistically in agreement with GR ($\alpha_0 = 0, \beta_0 = 0$). The shifted quantity $\Delta\chi^2 = \chi^2 - \chi_{\text{min}}^2$ has a $\chi^2(\text{d.o.f.} = 2)$ distribution with two degrees of

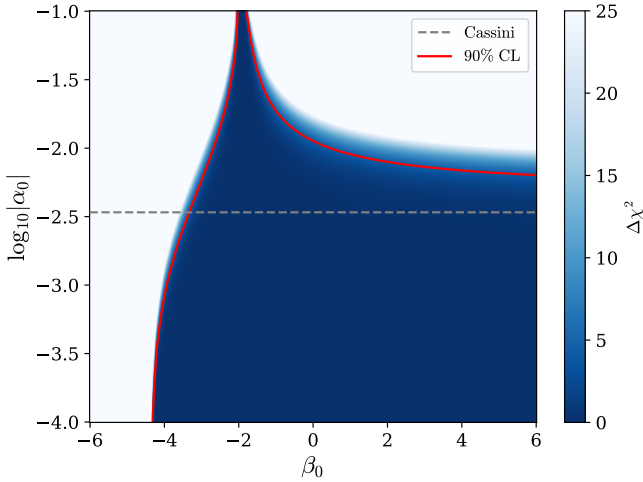


Fig. 2. $\Delta\chi^2$ map in DEF gravity parameter space for PSR J2222–0137. The test is performed by applying the DDSTG model to the timing data of Guo et al. (2021b) and assuming the rather stiff MPA1 EOS. The red line corresponds to 90% CL limit ($\Delta\chi^2 \simeq 4.6$), the area above the grey line is restricted by Cassini mission. GR with $\alpha_0 = 0, \beta_0 = 0$ lies beyond the plotted domain in the blue region at the bottom at the infinity.

freedom. Finally, we construct contours of a fixed $\Delta\chi^2$ value corresponding to desired confidence levels. In this paper, we adhere to 90% confidence level limits with $\Delta\chi^2 \simeq 4.6$.

During each run, TEMPO fits spin parameters, astrometric parameters, Keplerian parameters, two masses, and other parameters of interest, for example, \dot{x} . A value of the external contribution to the rate of the orbital period change (\dot{P}_b^{ext}) is fixed and selected, so that it fully accounts for Shklovskii effect and the Galactic \dot{P}_b contributions (Shklovskii 1970; Brumberg et al. 1975; Damour & Taylor 1992). The resulting $\Delta\chi^2$ map in $\{\alpha_0, \beta_0\}$ space for PSR J2222–0137 is presented in Fig. 2. The limit on DEF gravity parameters is placed by a contour of 90% confidence level limit of χ^2 (d.o.f. = 2) statistics.

Generally, the uncertainty of \dot{P}_b^{ext} can be important for other systems where it is not well-constrained. In this case, the DDSTG model allows a complete description with account for all the uncertainty due to external effects \dot{P}_b^{ext} . First we have to assume a prior distribution for the \dot{P}_b^{ext} value. Then we have to calculate the χ^2 value for 3D map with $\{\alpha_0, \beta_0, \dot{P}_b^{\text{ext}}\}$ as parameters of the three axes. Finally, we marginalise the probability density for the \dot{P}_b^{ext} axis using Bayes' theorem and obtain the corrected 2D map of χ^2_{corr} . The details of the procedure are presented in Appendix E.

Calculating an extensive grid in a 2D parameter space with a lot of points (e.g. ~ 500) for each axis to obtain a finely resolved contour of the desired $\Delta\chi^2$ value is too computationally demanding. In this work, we use an adaptive mesh refinement technique to trace the location of the desired contour. We start from the sparse grid with 9×9 points in the $\{\alpha_0, \beta_0\}$ space. Then, the algorithm resolves all the cells that can have the contour line inside them. The algorithm repeats the refinement of important cells until we obtain the desired curve fineness.

For a $N \times N$ grid, the naive approach utilises $O(N^2)$ iterations, whereas adaptive mesh refinement has only $O(N \log_2 N)$ complexity. For the present work we use the refinement level of six, resulting in the final grid with 513×513 points. We need such a large resolution in the parameter space because of the

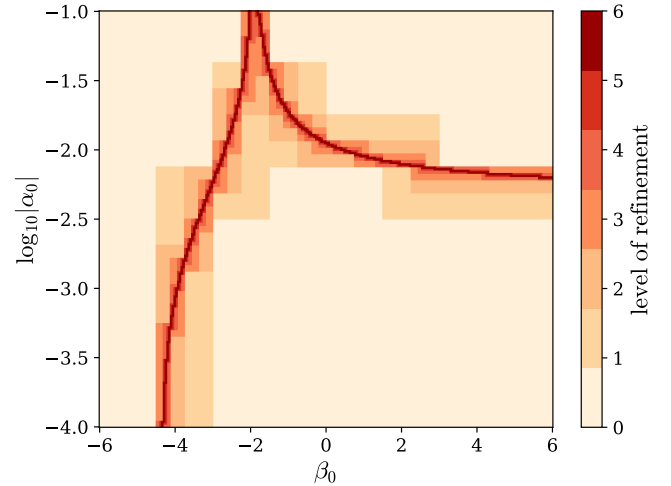


Fig. 3. Map of the refinement level in the DEF gravity parameter space. The test is the same as in Fig. 2. Areas in the parameter space with a higher refinement level have an exponentially growing resolution. The adaptive refinement procedure resolves the contour of 90% CL limit and saves computational time dramatically.

‘horn’ feature at large α_0 values near $\beta_0 \simeq -2$. A simple analysis shows that adaptive refinement requires one to calculate 56 times fewer points to resolve a single contour for a grid with 513×513 points. In Fig. 3, we present an adaptive refinement map corresponding to the search of a contour from Fig. 2. A high level of refinement is performed only along the contour of 90% CL limit. In Appendix D, one can also find the mass-mass diagram showing what happens in the region of the horn in terms of PK parameters.

4.6. Comparison between DDSTG and the PK method

In this section, we compare the constraining power of the newly developed approach of this paper with that of existing procedures. For this reason we perform the test with the traditional PK method based on the PK parameters from the timing data. The corresponding PK parameters are measured by means of fitting the DD model (Damour & Deruelle 1986) to the same timing data. Then for each specific choice of the theory parameters $\{\alpha_0, \beta_0\}$ we fit two masses m_p and m_c to minimise the χ^2 value. The χ^2 is calculated by comparison of the observed PK parameters p_i^{obs} and predicted values p_i^{theory} from the theory

$$\chi^2(p_i^{\text{obs}}; m_p, m_c) = \sum_i (\sigma_{p_i}^{\text{obs}})^{-2} (p_i^{\text{theory}}(m_p, m_c) - p_i^{\text{obs}})^2, \quad (31)$$

where $\sigma_{p_i}^{\text{obs}}$ is the standard deviation of the i -th measured PK parameter. Finally, we calculate the grid of χ^2 values over the desired $\{\alpha_0, \beta_0\}$ space with the same settings as for the DDSTG approach and shift χ^2 values by its minimum χ_{min}^2 . The detailed explanation of this common PK method can be found in Damour & Esposito-Farèse (1998). For both methods the minimum χ_{min}^2 is in statistical agreement with the GR value χ_{GR}^2 .

Figure 4 shows the comparison between the two methods. Each point on the plot presents a particular pair of $\{\alpha_0, \beta_0\}$ parameters. The analysis shows that, for PSR J2222–0137, the DDSTG model produces higher or equal values of $\Delta\chi^2$ compared to the traditional PK method. Despite the absence of any strong correlations between the PK parameters in PSR J2222–0137, the

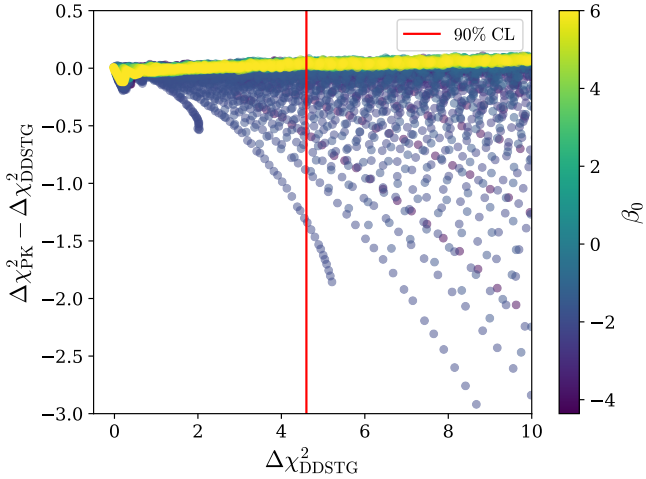


Fig. 4. Comparison in the constraining power between the DDSTG model and the method based on measured PK parameters with the DD model. Both methods calculate χ^2 values for a grid in $\{\alpha_0, \beta_0\}$ space, each point corresponds to a unique theory. β_0 values are presented by the colour, while α_0 values cover the $[-10^{-1}, -10^{-4}]$ range. Red vertical line corresponds to the 90% CL limit, all points to the left are allowed by the test.

new approach produces slightly more restrictive results for the area with negative β_0 . For certain areas in $\{\alpha_0, \beta_0\}$ space the difference in the shifted $\Delta\chi^2 = \chi^2 - \chi_{\min}^2$ values becomes statistically significant when we calculate contours of 90% confidence level limit ($\Delta\chi^2 \simeq 4.6$).

5. Predictions for PSR J2222–0137

Our next step is to estimate what enhancement in limits we can expect with future observations of PSR J2222–0137. For this purpose we simulate fake TOAs for a set of radio observatories, assuming realistic timing precision estimated from real timing data.

5.1. Simulated datasets for FAST, MeerKAT, and 3ERT

The PSR J2222–0137 timing data used up to this point are described in Sect. 4.2. We now describe our simulations, which show what we might be able to achieve in the near and foreseeable future with timing from this system.

We simulate TOAs spanning 10 yr from 2021 to 2030 based on the current precision of the TOAs obtained with the 3ERT telescopes, as well as the TOA precisions from ongoing observations from MeerKAT and FAST. These simulations are conservative since they assume that there will be no improvement in the existing capabilities at these telescopes.

For TOAs from FAST, the radiometer noise reduces significantly thanks to its large collecting area, while the jitter noise becomes the primary limitation of timing precision. We find, however, that increasing the integration time to 15 min can largely reduce the jitter noise and eliminate its contribution to the timing precision (as it scales with the number of averaged pulses (N_p) as $\sigma_J \propto 1/\sqrt{N_p}$ (Lorimer & Kramer 2004). Therefore, the median TOA uncertainty from 15-min TOAs are adopted in the simulation.

Table 1 lists the telescopes assumed in our simulation, with the information on their effective diameter (D_{eff}), observing bandwidth, and TOA uncertainties at L-band. All TOA

Table 1. Parameters of telescopes used in simulations.

Telescope	D_{eff} (m)	Bandwidth (MHz)	σ_{TOA} (μs)
Effelsberg	100	200/400	2.67
Nançay	94	512	2.02
Lovell	76	400	3.66
MeerKAT	108	856	0.84
FAST	300	500	0.12

Notes. Effective diameter of telescopes, observing bandwidth, and TOA uncertainties of PSR J2222–0137 used in the simulations. All information is based on the L-band data from real observations, and is scaled to 15-min integration over the full bandwidth.

uncertainties are scaled to an integration time of 15 min over the full bandwidth. For each telescope, we assume one full orbit observation (~ 60 h) per year, and split the observations into a monthly cadence, i.e. 5 h per month⁸, to allow a good estimation of timing parallax (which requires a good coverage of Earth’s orbit). This is important for the estimation of uncertainties in the external \dot{P}_b contributions shown in the next section.

The simulations are performed using the program developed in Hu et al. (2020). First, we simulate TOAs based on the above assumptions, and add the TOAs from Effelsberg, Nançay, and Lovell telescopes together to be compared with MeerKAT and FAST. For 3ERT, the simulated TOAs are combined with the existing TOAs in Guo et al. (2021b). We then adjust the TOAs to perfectly match the timing parameters measured in Guo et al. (2021b), and add a Gaussian white noise to each TOA based on its σ_{TOA} . Finally, we fit for timing parameters and obtain their uncertainties, among which \dot{P}_b and timing parallax are of most important here. The whole process is done with the TEMPO DDSTG model.

5.2. Contributions to the uncertainty of \dot{P}_b

The predicted uncertainties of \dot{P}_b are shown in Fig. 5, where the pink, orange, and blue lines show the improvement in time with the simulated data from 3ERT, MeerKAT, and FAST, respectively. As discussed in Sect. 4.4, we also need to account for uncertainties from external effects, \dot{P}_b^{ext} . They are also expected to become more precise in the future because of anticipated improvements in the model for the Galactic gravitational potential, distance, and proper motion from future observations. Correspondingly improved values for \dot{P}_b^{Gal} and \dot{P}_b^{Shk} may then be estimated using Eqs. (28) and (30). For the Shklovskii effect, we find that it is mostly limited by the uncertainty in the distance, which comes from VLBI parallax or timing parallax measurement, whichever is better. With the simulated FAST data, the measurements of timing parallax and proper motion improve quickly with time. In particular, the uncertainty of timing parallax will soon surpass the VLBI parallax and hence improve the distance measurement. The corresponding uncertainty in \dot{P}_b^{Shk} will decrease with time as shown by the brown line in Fig. 5. This line is below the predicted uncertainties from observed \dot{P}_b most of the time, except at the very end of the simulation when compared to $\Delta\dot{P}_b^{\text{obs}}$ with FAST (blue line). In addition, future VLBI

⁸ This may be unrealistic as telescopes may be oversubscribed. However, we found, e.g. for FAST, that with a more realistic assumption of 30 min per month, the uncertainty in the observed \dot{P}_b is only 5% worse than observing 5 h per month after 10-yr timing.

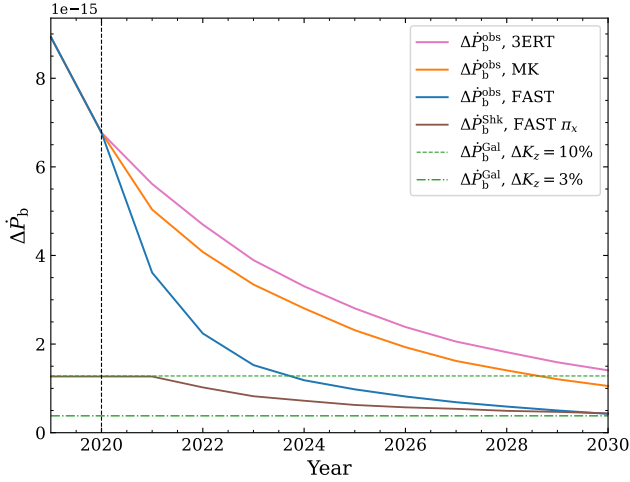


Fig. 5. Comparison of different contributions in the uncertainty of \dot{P}_b for simulated data from 2021 to 2030. The pink, orange and blue lines show the uncertainty of \dot{P}_b using the simulated data from 3ERT (EFF, NC, LT), MeerKAT (MK), and FAST, respectively. The brown line indicates the uncertainty of \dot{P}_b^{Shk} when using the timing parallax (π_x) measured from simulated FAST data. The green lines indicate the uncertainties of \dot{P}_b^{Gal} when assuming 10% (dashed) and 3% (dash-dotted) uncertainties in the vertical component of Galactic acceleration K_z .

parallax measurements will likely be improved so that Shklovskii effect will not be a limiting factor for \dot{P}_b .

As for the contribution from the Galactic acceleration, a typical uncertainty in its vertical component (ΔK_z) is about 10%, which contributes the most in $\Delta \dot{P}_b^{\text{Gal}}$, shown as the green dashed line in Fig. 5. With FAST, \dot{P}_b will then be limited by ΔK_z from 2024 onwards, if there is no improvement for this quantity. In fact, we find that ΔK_z needs to be improved to $\lesssim 3\%$ (see the green dash-dotted line) to not limit the precision of \dot{P}_b before 2030. The uncertainties in the Galactic potential do not limit the precision in this case. For the scope of further analysis, we assume that with future observations on pulsar timing, VLBI parallax, and Galactic acceleration, the precision of \dot{P}_b will not be limited by Shklovskii and Galactic effects. For Shklovskii contribution it is a reasonable assumption at least up to 2030. We also expect our knowledge about the Galactic potential to improve in the future especially in the proximity of the Solar System, thus we assume the improvement of ΔK_z for the selected very close pulsar to be 3%.

5.3. Potential future constraints on DEF gravity from PSR J2222–0137

We apply the same techniques as described in Sect. 4 including the adaptive mesh refinement for contours. The map of $\Delta\chi^2$ for a combined simulated data is presented in Fig. 6. By 2030, we can expect a significant improvement in the limits for large positive β_0 . This region is susceptible to dipolar gravitational emission. The significant improvement in \dot{P}_b measurement pushes the limit below the Cassini limit.

In Fig. 6, we also present the comparison of constraining power for different observatories. The tightest constraints may be obtained for the combination of all three observatories. However, the precision of FAST is high enough to be significantly constraining on its own.

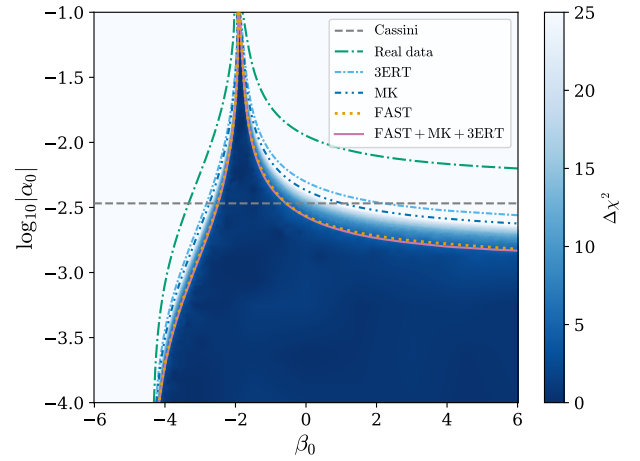


Fig. 6. $\Delta\chi^2$ map for PSR J2222–0137 from simulated data (2021–2030) combining different observatories and using the MPA1 EOS. Assumed one orbit per year for FAST, MeerKAT (MK), and 3ERT (EFF, NC, LT) based on their observation precision, appending to the existing dataset (Guo et al. 2021b). Each contour corresponds to 90% CL limits on DEF gravity parameters from different datasets, grey line depicts the limit from the Cassini mission.

6. Simulations for PSR-BH systems

As a further application of the new timing model, we investigate what limits on DEF gravity we can obtain from a PSR-BH system. As mentioned above, our DDSTG implementation already allows for a BH to be specified as a companion, even though such a system has so far not been found. We apply the same technique used in Sects. 4 and 5 to synthetic TOAs for three different hypothetical PSR-BH systems.

6.1. Simulated FAST datasets

To investigate possible limits on DEF gravity from PSR-BH systems, we select presumed realistic parameters for simulating fake TOAs. These systems are assumed to comprise a $1.4 M_\odot$ pulsar and a $10.0 M_\odot$ BH in a highly eccentric orbit ($e = 0.6$). As PSR-BH systems are more likely to reside in GCs (see the discussion in Sect. 6.4 and references therein), we assume that our hypothetical systems are located in the GC M15 and take the position (RA and Dec) and dispersion measure (DM) of PSR B2127+11C (M15C) for our simulations. To investigate how the limits on DEF gravity depend on the orbital period, we consider three cases with orbital periods of 3 days, 1 days, and 8 h, respectively. The selected parameters for PSR-BH systems are presented in Table 2.

We simulate datasets with two different TOA uncertainties (for 15-min integration time): moderate $10 \mu\text{s}$ (which is typical for pulsars in GCs), and precise $1 \mu\text{s}$. Two TOA uncertainties help us to explore the dependence of the DEF gravity test on the TOA precision. We effectively assume a recycled millisecond pulsar (MSP), because they in general produce a better timing precision. We assume 6 h of timing observation on these systems every two weeks, each session starts at a random orbital phase. All simulations cover a time span of 5 yr and have the same number of TOAs ($n_{\text{TOA}} = 3144$).

6.2. Tests of DEF gravity in case of a BH companion

Binary systems consisting of a pulsar and a white dwarf (PSR-WD) are particularly interesting for constraining STG theories

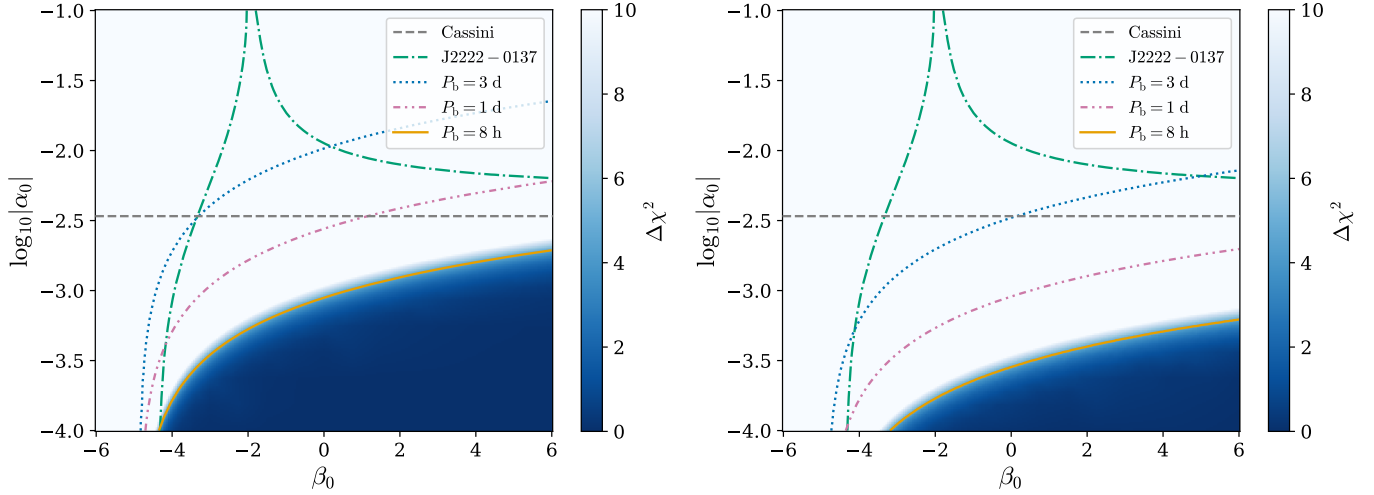


Fig. 7. $\Delta\chi^2$ map in the DEF gravity parameter space from the simulated 5-yr timing data for three different PSR-BH systems. The left panel corresponds to $10\ \mu\text{s}$ TOA uncertainty, the right panel to $1\ \mu\text{s}$. Solid lines show the 90% CL limits for PSR-BH systems with different orbital periods: 3 days, 1 day and 8 h. The grey dashed line depicts the limit from the Cassini mission and green dash-dotted line is the current limit from PSR J2222–0137 discussed in Sect. 4 (see Fig. 2).

Table 2. Properties of PSR-BH systems taken for investigation.

Parameters	Values
RA (α , J2000)	21:20:01.2
Dec (δ , J2000)	12:10:38.2
DM (pc cm^{-3})	67.1
Pulsar mass (m_p, M_\odot)	1.4
BH mass (m_{BH}, M_\odot)	10.0
Eccentricity (e)	0.6
Orbital period (P_b , d)	{1/3, 1, 3}
TOA uncertainty (μs)	{1, 10}
Number of TOAs	3144

due to their high asymmetry in compactness ($\alpha_{\text{WD}} \approx \alpha_0$). The theory predicts a higher rate of orbital energy loss due to dipolar radiation from such asymmetric systems (Will 1993; Damour & Esposito-Farèse 1996). However, generally, PSR-BH systems are expected to be even more asymmetric up to a certain positive β_0 , as was pointed out by Damour & Esposito-Farèse (1998). As a result of the no-scalar-hair theorem to BHs in DEF gravity one has (Hawking 1972; Damour & Esposito-Farèse 1992a)

$$\alpha_{\text{BH}} = 0, \quad \beta_{\text{BH}} = 0. \quad (32)$$

This approximation is only valid for stationary BHs with an asymptotically flat spacetime and an asymptotically constant scalar field (Berti et al. 2013), which is not true anymore in the presence of a compact scalarised companion. However, a justification that it is still an excellent approximation can be found in Liu et al. (2014). The pulsar orbiting in an eccentric orbit will eventually induce a time-dependent scalar field at the location of the BH. This scalar field results in an induced effective scalar charge ($\alpha_{\text{BH}}^{\text{induced}}$) which, however, is totally negligible ($\lesssim 8 \times 10^{-14} \alpha_p$; cf. Eq. (50) in Liu et al. 2014).

The absence of scalar charges for the BH results in simplified relations of the PK parameters. The main consequence of a BH presence is that all PK parameters except \dot{P}_b are identical to their GR expressions with an appropriate rescaling of the

Table 3. Orbital period properties of simulated PSR-BH systems.

P_b (days)	\dot{P}_b^{GR} ($10^{-12} \text{ s s}^{-1}$)	$\Delta\dot{P}_b, 10\ \mu\text{s}$	$\Delta\dot{P}_b, 1\ \mu\text{s}$
3	−0.22220	0.08031	0.00827
1	−1.38661	0.01723	0.00185
1/3	−8.65279	0.00532	0.00055

Notes. The rate of the orbital period change in GR (\dot{P}_b^{GR}) for selected PSR-BH systems and corresponding uncertainty in \dot{P}_b measurement for three different orbital periods and two TOA uncertainties. The values of $\dot{P}_{b,\text{GR}}$ and $\Delta\dot{P}_b$ are given in $10^{-12} \text{ s s}^{-1}$.

masses. In the GW damping sector, all the multipoles, for example, monopolar, dipolar and quadrupolar modes, are still present. Thus PSR-BH systems are susceptible but only to the test of GW damping via \dot{P}_b (Damour & Esposito-Farèse 1992a; Mirshekari & Will 2013; Liu et al. 2014).

6.3. Results

The limits obtained from the DDSTG model for the simulated PSR-BH systems are shown in Fig. 7. For the tests presented here, we do not include any potential uncertainty in the observed \dot{P}_b^{obs} value due to the external effects. The restrictive power of the test increases for more relativistic systems with shorter orbital periods. To place new limits on DEF gravity with a moderate TOA precision of $10\ \mu\text{s}$, the orbital period should be a fraction of a day.

Another way to improve the restrictive power is to increase the precision of the TOAs (compare left and right panels of Fig. 7). The obtained limits strongly depend on the accuracy of \dot{P}_b measurement as most of the deviations from GR come from the predicted dipolar GW emission. Table 3 shows the predicted GR values of \dot{P}_b and compares them to the corresponding $\Delta\dot{P}_b$ uncertainties. The lower uncertainty and the higher the predicted value the better the test.

We expect a dramatic improvement of limits on DEF gravity from a relativistic PSR-BH system. Due to higher asymmetry, the

test is extremely sensitive to the precision of \dot{P}_b measurement. To answer the possibility of obtaining accurate enough TOAs, we have to argue where we can find such a system.

6.4. Origin of PSR-BH systems

A binary pulsar system with a stellar-mass BH may originate from several different evolutionary scenarios. The first way is a standard evolution of a massive binary system (Voss & Tauris 2003). It results in a PSR-BH system with a young slow spin-period (~ 0.1 – 1 s) pulsar and a wide, eccentric orbit. The second path is a reversal mechanism (Sipior et al. 2004) taking place under a specific set of circumstances. The pulsar is formed first and later spun-up by accretion during the red-giant phase of the companion (Pfahl et al. 2005), later becomes a BH. This is similar to the origin of PSR J1141–6545, a system where a massive WD star formed before a more massive NS (Tauris & Sennels 2000). The reversal mechanism may give a more desirable result of a recycled pulsar in a system with a BH, because recycled pulsars have generally more precise timing and a more stable rotation than the slow ‘normal’ pulsars (Verbiest et al. 2009).

Moreover, the third possible way to form a PSR-BH system is a multiple body encounter, happening in regions of high stellar density, for example, GCs and the Galactic Centre region (Verbiest et al. 2009; Clausen et al. 2014). Such encounters are the reason why there are $\sim 10^3$ times more low-mass X-ray binaries (LMXBs) per unit stellar mass in GCs than in the Galactic disk (Clark 1975), which results in a similarly enhanced proportion of MSPs. In these encounters, an old, inactive NS approaches a main sequence (MS) binary so closely that a chaotic interaction ensues. The most likely result of such an interaction is that the two most massive objects (in this case the old, recycled NS and the more massive MS star) will form a more compact binary system, with the lighter MS star component being ejected at high velocity (see review by Phinney 1992). That MS will evolve, fill its Roche lobe, and start transferring matter onto the NS, which is spun up in the process, this is the aforementioned LMXB stage.

If left undisturbed, many of these then evolve into binary MSP systems. A consequence of this is that GCs not only have many MSPs, but that the number of MSPs in each cluster appears to be roughly proportional to its rate of stellar encounters (Γ) (Verbunt & Hut 1987).

However, in some GCs – especially those with collapsed cores – stellar densities are so high that only few binaries evolve without being disturbed; these GCs have a large interaction rate per binary (γ_{GC}) (Verbunt & Freire 2014). This means that, even after being recycled, an MSP has a high probability of undergoing further (‘secondary’) exchanges.

In such exchanges, an incoming massive star interacts chaotically with the components of either a LMXB or a MSP binary. Again, the least massive object is the most likely to be ejected, in this case that will be the low-mass star that recycled (or was still recycling) the NS. A new binary will form, consisting of the NS and the intruding massive star. If the latter is degenerate, then the system will not undergo accretion and the orbital circularisation that comes with it, but will keep the orbital eccentricity it acquired after formation.

Several such obvious products of secondary exchange encounters have been found in GCs, invariably with a large γ_{GC} : PSR B2127+11C, in the core-collapsed M15 GC (Prince et al. 1991; Jacoby et al. 2006), PSR J0514–4002A, in NGC 1851 (Freire et al. 2004; Ridolfi et al. 2019; recently two additional such systems, D and E, were found in the same GC, see

Ridolfi et al. 2022), PSR J1807–2500B, in the core-collapsed NGC 6544 (Lynch et al. 2012), PSR J835–3259A, in the core-collapsed NGC 6652 (DeCesar et al. 2015) and PSR J1823–3021G, in the core-collapsed NGC 6624 (Ridolfi et al. 2021). These discoveries suggest (but by no means assure us) of the possibility that similar secondary exchange encounters might produce MSP binaries with even more massive companions, such as BHs. As for the massive MSP binaries above, such MSP-BH systems would be preferentially produced in the GCs with large γ_{GC} ; this is one of the reasons why they are targeted by the MeerKAT TRAPUM survey (Ridolfi et al. 2021).

These secondary exchange products are that their orbital periods vary between 8 h in the case of B2127+11C and 18.8 d in the case of PSR J0514–4002A, their orbital eccentricities are between 0.38 and 0.90. Thus the simulated PSR-BH systems listed in Table 2 have realistic orbital periods and eccentricities.

6.5. Influence of a globular cluster origin

For the pulsars observed in GCs, the derivatives of the spin period, P , and in the case of binary pulsars, the derivative of the orbital period, \dot{P}_b are contaminated (and in most cases dominated) by the line-of-sight component of the acceleration of the pulsar (or binary) in the gravitational potential of the GC (a_{GC}); this enters Eq. (23) as an additional term that is similar to \dot{P}_b^{Gal} :

$$\dot{P}_b^{GC} = \frac{a_{GC}}{c} P_b. \quad (33)$$

For pulsars in GCs, the only radiative test done to date was with PSR B2127+11C (Jacoby et al. 2006). The orbital decay observed in this system is $-3.95 \pm 0.13 \times 10^{-12} \text{ s s}^{-1}$, which is within $\sim 3\%$ of the predicted value. That test, however, is limited by the fact that the maximum value for $|a_{GC}/c| \sim 6 \times 10^{-18} \text{ s}^{-1}$ (Phinney 1993), i.e., in this case, the maximum value for $|\dot{P}_b^{GC}| \sim 0.17 \times 10^{-12} \text{ s s}^{-1}$. This is of the same order as their measurement precision, which means that this radiative test cannot be improved, unless one could measure the acceleration of the pulsar independently. The situation would be much worse if PSR B2127+11C were closer to the cluster centre, where accelerations are much larger. For instance, PSR B2127+11A, a couple of arcseconds from the centre, has $P = 0.1106 \text{ s}$ and $\dot{P} = -2.107 \times 10^{-17}$; this implies that $|a_{GC}/c| > 1.9 \times 10^{-16} \text{ s}^{-1}$ and thus $|\dot{P}_b^{GC}| > 5.5 \times 10^{-12} \text{ s s}^{-1}$, a value larger than the orbital decay predicted by GR.

For the three PSR-BH systems in Table 3, the situation would be similar. If they were at the locations of PSR B2127+11C, their values of P_b would imply that the maximum values for \dot{P}_b^{GC} would be 9, 3 and 1 times larger respectively, i.e., 1.53, 0.51 and $0.17 \times 10^{-12} \text{ s s}^{-1}$. This corresponds to, respectively, ~ 6.9 , ~ 0.37 and 0.02 times the values of \dot{P}_b^{GR} in Table 3. That means that, for the latter system, a ~ 50 - σ test of the radiative properties of a PSR-BH system would be possible. These numbers illustrate the immense advantage of an increasingly shorter P_b , either for GW tests in GCs or in the Galaxy: on one hand, the \dot{P}_b^{GR} increases with $P_b^{-5/3}$, the ‘polluting’ part \dot{P}_b^{GC} decreases as P_b . Thus, the significance of a particular GW test grows, everything else being identical, with $P_b^{-8/3}$.

However, if the 8-h PSR-BH system is placed at the location of PSR B2127+11A, the test would lose its significance almost entirely. Fortunately, even in this case, we can get a firm upper limit for $|\dot{P}_b^{GR}|$; the reason is that we also measure the spin period derivative. That is also affected by the acceleration in the cluster

and by other terms, such as the Shklovskii effect. Adding the equations for \dot{P} and \dot{P}_b , we obtain:

$$\dot{P}_b^{\text{GR}} = \dot{P}_b^{\text{obs}} - P_b \left(\frac{\dot{P}}{P} \right)^{\text{obs}} + \frac{1}{2\tau_c} P_b. \quad (34)$$

We note that all the terms on the right can be measured precisely, except for the characteristic age of the pulsar, τ_c , which cannot be measured independently. Therefore, the last term then quantifies the uncertainty of the test, which is, again, proportional to P_b , which implies that the significance of the test is again proportional to $P_b^{-8/3}$.

However, that last term is necessarily positive. If we assume the pulsar is extremely old, then this term will be very small and we obtain, from the other two terms, a hard lower limit for \dot{P}_b^{GR} . Since the latter is negative, this represents a hard upper limit of its magnitude. This means that such a test could still, in principle, falsify GR. An upper limit on the last term can be obtained from the lowest likely value for τ_c ; for most MSPs $\tau_c < 1$ Gyr. This means that, for the 8-h PSR-BH system, this unknown term would be at most $0.46 \times 10^{-12} \text{ s s}^{-1}$, implying a $\sim 20\text{-}\sigma$ test of GR. Thus, this test will be the more significant the larger τ_c is compared to the orbital decay timescale for the binary.

Despite such possible mitigation, it is clear that the location in a GC always degrades the quality of radiative tests. For the 8-h PSR-BH system discussed above, significances of 50 and 20 σ in the measurement of \dot{P}_b represent a significant degradation relative to the tests listed in Table 3, where for a 8-h PSR-BH system timed with 10 μs the significance of the \dot{P}_b measurement is larger than 1600.

7. DDSTG and other gravity theories

The approach developed in this work is not restricted to DEF gravity. It can be extended straightforwardly to investigate a larger set of alternative gravity theories. Broadly speaking, every theory that maps to the phenomenological DD model (Damour & Deruelle 1986; Damour & Taylor 1992) can be put into the DDSTG framework with a new implementation of appropriate PK formulae. The DD model is a quasi-Keplerian solution in the first post-Newtonian approximation to the dynamics of a two-body system within the modified Einstein-Infeld-Hoffmann (mEIH) framework (Damour & Taylor 1992; Will 1993, 2018b). The mEIH formalism covers a large set of fully conservative gravity theories without the ‘Whitehead term’ in the post-Newtonian limit. Later it was extended by Will (2018a) from fully conservative to semi-conservative theories of gravity. However, this extension requires additional terms in the mEIH Lagrangian, which are not accounted for in the DD model.

The DDSTG model can be directly applied to a certain class of gravity theories without a modification of the PK parameter equations. For example, DEF gravity is a specific case of a massless mono-scalar tensor gravity theory with a particular expression of the conformal coupling $A(\varphi)$. The DDSTG model covers STG theories with any conformal coupling function depending on two arbitrary parameters $\{\alpha_0, \beta_0\}$ and a single massless scalar field. To work, the model only requires pre-calculated gravitational form factors for the new coupling function. In case of a more complex coupling function depending on a higher number of parameters the TEMPO code needs to be adapted.

Recently, Mendes and Ortiz (MO, Mendes & Ortiz 2016) introduced an extension of DEF gravity. MO gravity is an example of a theory that can be easily incorporated into the DDSTG

model. Its difference from DEF gravity is in the form of the conformal coupling

$$A(\varphi) = \left[\cosh \left(\sqrt{3}\beta_0\varphi \right) \right]^{1/(3\beta_0)}, \quad (35)$$

where β_0 is a free parameter. The second parameter α_0 is hidden in MO gravity to the scalar field at infinity. MO theory received attention in recent years and was originally introduced as an analytical approximation to a more fundamental theory, where the action includes quadratic terms of the scalar field coupled to curvature. The developed framework can therefore be extended straightforwardly to MO gravity without the change of the TEMPO implementation.

8. DDSTG and a time-varying gravitational constant

Another interesting extension of the DDSTG model – planned for a future release – is the inclusion of the effects of a temporal variation of the gravitational constant (G). A time-evolving asymptotic scalar field (φ_0) of a gravitating system, generally, leads to a temporal variation of the local gravitational constant. In Damour–Esposito–Farèse gravity, such a change of the Newtonian gravitational constant as measured in the Solar System reads

$$\frac{\dot{G}_{\text{Cav}}}{G_{\text{Cav}}} = 2 \left(1 + \frac{\beta_0}{1 + \alpha_0^2} \right) \alpha_0 \dot{\varphi}_0 \quad (36)$$

(see e.g. Uzan 2011). Generally, one expects a temporal change in φ_0 to arise from the expansion of the universe and $\dot{\varphi}_0$ to result from a cosmological model based on DEF gravity (see e.g. Damour & Nordvedt 1993). However, as part of a more agnostic approach, $\dot{\varphi}_0$ can be treated as an additional, independent timing parameter.

The main impact on the orbital motion of a binary system that arises from a time-varying gravitational constant is a secular change in the orbital period (Damour et al. 1988). For two weakly self-gravitating masses, one simply has $\dot{P}_b^G/P_b = -2\dot{G}/G$. However, for binary pulsar systems, this simple expression needs to be extended by body-dependent contributions (Nordvedt 1990, 1993). One then has

$$\frac{\dot{P}_b^G}{P_b} = -2 \left(\frac{\dot{G}_{\text{Cav}}}{G_{\text{Cav}}} \right) \mathcal{F}_{AB}, \quad (37)$$

where the factor \mathcal{F}_{AB} accounts for all the corrections related to the strong gravitational fields of the pulsar and its companion, if the latter is also a NS. More specifically, following Nordvedt (1993), \mathcal{F}_{AB} accounts for a change in the body-dependent part in the effective gravitational constant G_{AB} as well as a change in the masses resulting directly from $\dot{\varphi}_0$.

It has been demonstrated by Wex (2014) that, depending on a parameter space, a pulsar mass, and a EOS, strong-field effects can considerably enhance the effect of a time-varying gravitational constant, i.e. $\mathcal{F}_{AB} \gg 1$. Consequently, accounting for \dot{P}_b^G in binary pulsar tests not only provides an independent test for a varying gravitational constant but also probes strong-field aspects related to a time-varying gravitational constant.

⁹ Detailed expressions for \mathcal{F}_{AB} will be given in a future publication. See also Wex (2014).

9. Discussion and conclusions

In this work, we developed a new, improved approach for testing STG theories. We examined a specific class of STG theories known as DEF gravity. This approach is based on a new timing model, called the DDSTG model, which is an extension of the DD model to work within DEF gravity. Analysis of pulsar timing data with this model overcomes some of the problems of conventional methods, which we have discussed in detail. The DDSTG timing model uses theoretical predictions for PK parameters in DEF gravity and therefore uses a minimal set of binary parameters. For that reason, it accounts for all the possible correlations between these parameters. All the information from the observational data is used to provide the most reliable tests of an alternative theory directly, without intermediate steps with phenomenological parameters of the DD model.

As a demonstration of the DDSTG model, we applied it to the most recently published timing data of the binary pulsar, PSR J2222–0137, described and used by Guo et al. (2021b). This system is of great importance for testing alternative gravity theories because it is very close to us (resulting in the most precise VLBI distance for any pulsar), it has precise timing, and it shows a set of well-measured relativistic effects. The system has a high asymmetry in the compactness between the components: it comprises a massive NS ($m_{\text{NS}} \sim 1.82 M_{\odot}$) and a massive WD ($m_{\text{WD}} \sim 1.31 M_{\odot}$). This high asymmetry results, for some areas in the DEF gravity parameter space, in the prediction of a very strong dipolar GW contribution to the rate of orbital decay (\dot{P}_b^D). The non-detection of dipolar GWs in this system is used to constrain the DEF gravity parameter space. Moreover, the mass of the pulsar lies in the ‘scalarisation gap’ ($m_{\text{NS}} \gtrsim 1.5 M_{\odot}$); this means that strict limits can be placed on the occurrence of spontaneous scalarisation, a highly non-linear phenomenon (Zhao et al. 2022).

The results from the new method confirmed and improved the existing limits on DEF gravity parameters from this system. The DDSTG model appeared to be more constraining in the area near spontaneous scalarisation ($\beta_0 \lesssim -4.0$) when compared to the commonly used PK method. This suggests that the combination of the DDSTG model with a EOS agnostic approach can improve limits placed on the spontaneous scalarisation.

Moreover, we applied the DDSTG timing model to the simulated TOA for PSR J2222–0137 covering the period of 2021–2030 to see what improvement we can expect from that system in the future. The mock timing datasets were simulated for several large observatories: FAST, 3ERT (EFF, NC, LT), and MeerKAT with TOA uncertainties based on the real timing data. We have discussed the future importance of kinematic contributions (Shklovskii and Galactic) to \dot{P}_b , and consequently to the precision of these tests. The main limiting factor to \dot{P}_b comes from the uncertainty in the Galactic contribution (\dot{P}_b^{Gal}). In particular, the limit comes from the current uncertainty in the vertical component of the Galactic acceleration (ΔK_z). Our analysis predicts that this limitation will disappear if the current uncertainty of $\Delta K_z \sim 10\%$ can be improved to $\Delta K_z \lesssim 3\%$, which is likely to be the case in the future with improvement of models for the gravitational potential of our Galaxy. Future observations are expected to significantly improve the limits on DEF gravity, especially with the use of FAST data.

One of the most promising systems for testing gravity, which we hope to have in the near future, are binary PSR-BH systems. We simulated artificial timing datasets for three eccentric PSR-BH systems with reasonable orbital parameters for three different orbital periods. The results of applying the DDSTG

model to the simulated PSR-BH data strongly depended on the precision of the \dot{P}_b measurement. We have briefly discussed possible evolution scenarios leading to the formation of the PSR-BH system, such as GC origin and a reversal mechanism. Depending on the place of origin, there might be issues in obtaining a precise intrinsic \dot{P}_b , that is to say in accounting for the contamination of \dot{P}_b from a kinematic contribution due to the acceleration of the system in the gravitational field of the GC. Depending on the timing precision and orbital properties, PSR-BH systems can place stringent limits on DEF gravity.

In the future, the DDSTG model can be applied to a range of different binary pulsar systems to improve the limits on DEF gravity. We expect especially interesting results from PSR J1141–6545 where the DDSTG model is expected to be superior to standard approaches (Venkatraman Krishnan 2019). PSR J1141–6545 is an asymmetric PSR-WD system in a 4.7 h orbit with significant spin-orbit coupling due to the fast rotating WD. Due to the spin-orbit coupling, the system shows a change in the projected semi-major axis \dot{x}^{SO} , which has a strong correlation with the time dilation parameter γ . The latter parameter is caused by the precession which cannot be calculated because we do not know the exact spin properties of the WD. Moreover, PSR J1141–6545 shows a weak Shapiro delay in the timing data which the DDSTG model can fully exploit, resulting in a significant improvement to the test. The high asymmetry in the compactness between the components ($m_p \sim 1.26 M_{\odot}$ NS and $m_c \sim 1.02 M_{\odot}$ WD) makes this system a perfect tool for radiative tests of gravity. We expect the DDSTG model to be particularly advantageous because it accounts for both possible correlations and weak relativistic effects (Venkatraman Krishnan et al., in prep.).

Another perspective system to perform DEF gravity tests with the DDSTG model is the double pulsar, PSR 0737–3039A which shows the largest number of PK parameters in the timing data (Kramer et al. 2021). The system consists of two radio pulsars with masses of $m_p \simeq 1.34 M_{\odot}$ and $m_c \simeq 1.25 M_{\odot}$. Properties of NSs in DEF gravity (gravitational form factors) strongly depend on the choice of a EOS, which in turn affects the test. Thus to put reliable limits on DEF gravity from PSR 0737–3039A independent of a choice of a EOS, one must perform a EOS-agnostic analysis. An EOS-agnostic test means that it is performed with a set of EOSs which are diverse in their properties (see Voisin et al. 2020, who used such a EOS-agnostic approach to constrain DEF gravity with a pulsar in a stellar triple system). With this paper, we have included a set of 11 EOSs varying from soft to stiff (see Appendix A.2) which can be used for such agnostic tests in the future.

Acknowledgements. We thank Lijing Shao for carefully reading the manuscript and giving important comments. We further thank Thomas Tauris and Selma de Mink for useful discussions and suggestions. AB and HH are members of the International Max Planck Research School for Astronomy and Astrophysics at the Universities of Bonn and Cologne. HH acknowledges the support by the Max-Planck Society as part of the ‘LEGACY’ collaboration with the Chinese Academy of Sciences on low-frequency gravitational wave astronomy. This study is partly based on observations with the 100-m telescope of the MPIfR (Max-Planck-Institut für Radioastronomie) at Effelsberg. The Nançay Radio Observatory is operated by the Paris Observatory, associated with the French Centre National de la Recherche Scientifique (CNRS). LG, IC, and GT acknowledge financial support from the ‘Programme National Gravitation, Références, Astronomie, Métrologie (PNGRAM)’ of CNRS/INSU, France. J.W. McKee gratefully acknowledges support by the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference #CITA 490888-16]. This publication made use of open source libraries including DIFFERENTIALEQUATIONS.JL (<https://diffeq.sciml.ai/>) (Rackauckas & Nie 2017), OPTIM.JL (Mogensen & Riseth 2018), FORWARDDIFF.JL (Revels et al. 2016), MATPLOTLIB (Hunter 2007), along with a pulsar analysis package TEMPO (<http://tempo.sourceforge.net/>) (Nice et al. 2015).

References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *Phys. Rev. Lett.*, **116**, 221101
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, *Phys. Rev. Lett.*, **119**, 161101
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, *ApJ*, **848**, L12
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, *Phys. Rev. Lett.*, **118**, 221101
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019a, *Phys. Rev. D*, **100**, 104036
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019b, *Phys. Rev. Lett.*, **123**, 011102
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2021a, *ApJ*, **915**, L5
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2021b, *Phys. Rev. D* **103**, 122002
- Anderson, D., & Yunes, N. 2019, *Class. Quant. Grav.*, **36**, 165003
- Anderson, D., Freire, P., & Yunes, N. 2019, *Class. Quant. Grav.*, **36**, 225009
- Antoniadis, J., van Kerkwijk, M. H., Koester, D., et al. 2012, *MNRAS*, **423**, 3316
- Antoniadis, J., Freire, P. C. C., Wex, N., et al. 2013, *Science*, **340**, 448
- Bergmann, P. G. 1968, *Int. J. Theor. Phys.*, **1**, 25
- Berti, E., Cardoso, V., Gualtieri, L., Horbatsch, M., & Sperhake, U. 2013, *Phys. Rev. D*, **87**, 124020
- Berti, E., Barausse, E., Cardoso, V., et al. 2015, *Class. Quant. Grav.*, **32**, 243001
- Bertotti, B., Iess, L., & Tortora, P. 2003, *Nature*, **425**, 374
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. 2017, *SIAM Rev.*, **59**, 65
- Bhat, N. D. R., Bailes, M., & Verbiest, J. P. W. 2008, *Phys. Rev. D*, **77**, 124017
- Blandford, R., & Teukolsky, S. A. 1976, *ApJ*, **205**, 580
- Boyles, J., Lynch, R. S., Ransom, S. M., et al. 2013, *ApJ*, **763**, 80
- Brans, C., & Dicke, R. H. 1961, *Phys. Rev.*, **124**, 925
- Brumberg, V. A., Zeldovich, I. B., Novikov, I. D., & Shakura, N. I. 1975, *Sov. Astron. Lett.*, **1**, 2
- Clark, G. W. 1975, *ApJ*, **199**, L143
- Clausen, D., Sigurdsson, S., & Chernoff, D. F. 2014, *MNRAS*, **442**, 207
- Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, *Phys. Rep.*, **513**, 1
- Cognard, I., Freire, P. C. C., Guillemot, L., et al. 2017, *ApJ*, **844**, 128
- Damour, T. 2007, arXiv e-prints [arXiv:0704.0749]
- Damour, T. 2012, *Class. Quant. Grav.*, **29**, 184001
- Damour, T., & Deruelle, N. 1985, *Ann. Inst. H. Poincaré (Physique Théorique)*, **43**, 107
- Damour, T., & Deruelle, N. 1986, *Ann. Inst. H. Poincaré (Physique Théorique)*, **44**, 263
- Damour, T., & Esposito-Farèse, G. 1992a, *Class. Quant. Grav.*, **9**, 2093
- Damour, T., & Esposito-Farèse, G. 1992b, *Phys. Rev. D*, **46**, 4128
- Damour, T., & Esposito-Farèse, G. 1993, *Phys. Rev. Lett.*, **70**, 2220
- Damour, T., & Esposito-Farèse, G. 1996, *Phys. Rev. D*, **54**, 1474
- Damour, T., & Esposito-Farèse, G. 1998, *Phys. Rev. D*, **58**, 1
- Damour, T., & Nordtvedt, K. 1993, *Phys. Rev. D*, **48**, 3436
- Damour, T., & Taylor, J. H. 1991, *ApJ*, **366**, 501
- Damour, T., & Taylor, J. H. 1992, *Phys. Rev. D*, **45**, 1840
- Damour, T., Gibbons, G. W., & Taylor, J. H. 1988, *Phys. Rev. Lett.*, **61**, 1151
- DeCesar, M. E., Ransom, S. M., Kaplan, D. L., Ray, P. S., & Geller, A. M. 2015, *ApJ*, **807**, L23
- Deller, A. T., Boyles, J., Lorimer, D. R., et al. 2013, *ApJ*, **770**, 145
- Douchin, F., & Haensel, P. 2001, *A&A*, **380**, 151
- Fierz, M. 1956, *Helv. Phys. Acta*, **29**, 128
- Fonseca, E., Cromartie, H. T., Pennucci, T. T., et al. 2021, *ApJ*, **915**, L12
- Freire, P. C., Gupta, Y., Ransom, S. M., & Ishwara-Chandra, C. H. 2004, *ApJ*, **606**, L53
- Fujii, Y., & Maeda, K.-i. 2007, *The Scalar-Tensor Theory of Gravitation* (Cambridge University Press)
- Goenner, H. 2012, *Gen. Relativity Grav.*, **44**, 2077
- GRAVITY Collaboration (Abuter, R., et al.) 2021, *A&A*, **647**, A59
- Guo, M., Zhao, J., & Shao, L. 2021a, *Phys. Rev. D*, **104**, 104065
- Guo, Y. J., Freire, P. C. C., Guillemot, L., et al. 2021b, *A&A*, **654**, A16
- Hartle, J. 1967, *ApJ*, **150**, 1005
- Hawking, S. W. 1972, *Commun. Math. Phys.*, **25**, 167
- Holmberg, J., & Flynn, C. 2004, *MNRAS*, **352**, 440
- Horndeski, G. W. 1974, *Int. J. Theor. Phys.*, **10**, 363
- Hu, H., Kramer, M., Wex, N., Champion, D. J., & Kehl, M. S. 2020, *MNRAS*, **497**, 3118
- Hulse, R. A., & Taylor, J. H. 1975, *ApJ*, **195**, L51
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Jacoby, B. A., Cameron, P. B., Jenet, F. A., et al. 2006, *ApJ*, **644**, L113
- Jordan, P. 1955, *Schwerkraft und Weltall, Die Wissenschaft* (Vieweg)
- Jordan, P. 1959, *Z. Phys.*, **157**, 112
- Kaluza, T. 1921, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (Berlin), 966
- Klein, O. 1926, *Zeitsch. Phys.*, **37**, 895
- Kramer, M., Stairs, I. H., Manchester, R. N., et al. 2021, *Phys. Rev. X*, **11**, 041050
- Lazaridis, K., Wex, N., Jessner, A., et al. 2009, *MNRAS*, **400**, 805
- Liu, K., Eatough, R. P., Wex, N., & Kramer, M. 2014, *MNRAS*, **445**, 3115
- Lorimer, D. R., & Kramer, M. 2004, *Handbook of Pulsar Astronomy*, 4 (Cambridge University Press)
- Luo, J., Ransom, S., Demorest, P., et al. 2021, *ApJ*, **911**, 45
- Lynch, R. S., Freire, P. C. C., Ransom, S. M., & Jacoby, B. A. 2012, *ApJ*, **745**, 109
- Margossian, C. C. 2018, *WIREs*, **9**, e1305
- Mendes, R. F. P., & Ortiz, N. 2016, *Phys. Rev. D*, **93**, 124035
- Mirshakari, S., & Will, C. M. 2013, *Phys. Rev. D*, **87**, 084070
- Mogensen, P. K., & Riset, A. N. 2018, *J. Open Source Softw.*, **3**, 615
- Müther, H., Prakash, M., & Ainsworth, T. L. 1987, *Phys. Lett. B*, **199**, 469
- Nice, D. J., & Taylor, J. H. 1995, *ApJ*, **441**, 429
- Nice, D., Demorest, P., Stairs, I., et al. 2015, *Tempo: Pulsar timing data analysis Astrophysics Source Code Library [record ascl:1509.002]*
- Nordtvedt, K. 1990, *Phys. Rev. Lett.*, **65**, 953
- Nordtvedt, K. 1993, *ApJ*, **407**, 5
- Pfahl, E., Podsiadlowski, P., & Rappaport, S. 2005, *ApJ*, **628**, 343
- Phinney, E. S. 1992, *Philos. Trans. Roy. Soc. London A*, **341**, 39
- Phinney, E. S. 1993, *ASP Conf. Ser.*, **50**, 141
- Prince, T. A., Anderson, S. B., Kulkarni, S. R., & Wolszczan, W. 1991, *ApJ*, **374**, L41
- Rackauckas, C., & Nie, Q. 2017, *J. Open Res. Softw.*, **5**
- Read, J. S., Lackey, B. D., Owen, B. J., & Friedman, J. L. 2009, *Phys. Rev. D*, **79**, 124032
- Revels, J., Lubin, M., & Papamarkou, T. 2016, arXiv e-prints [arXiv:1607.07892]
- Ridolfi, A., Freire, P. C. C., Gupta, Y., & Ransom, S. M. 2019, *MNRAS*, **490**, 3860
- Ridolfi, A., Gautam, T., Freire, P. C. C., et al. 2021, *MNRAS*, **504**, 1407
- Ridolfi, A., Freire, P. C. C., Gautam, T., et al. 2022, *A&A*, **664**, A27
- Shao, L., & Wex, N. 2016, *Sci. China Phys. Mech. Astron.*, **59**, 699501
- Shao, L., Sennett, N., Buonanno, A., Kramer, M., & Wex, N. 2017, *Phys. Rev. X*, **7**, 041025
- Shklovskii, I. S. 1970, *Sov. Astron.*, **13**, 562
- Splaver, E. M., Nice, D. J., Arzoumanian, Z., et al. 2002, *ApJ*, **581**, 509
- Sipior, M. S., Portegies Zwart, S., & Nelemans, G. 2004, *MNRAS*, **354**, L49
- Stairs, I. H. 2003, *Living Rev. Relativity*, **6**, 5
- Swesty, F. D. 1996, *J. Comput. Phys.*, **127**, 118
- Tauris, T. M., & Sennels, T. 2000, *A&A*, **355**, 236
- Taylor, J. H., & Weisberg, J. M. 1989, *ApJ*, **345**, 434
- Uzan, J.-P. 2011, *Living Rev. Relativity*, **14**, 2
- Venkatraman Krishnan, V. 2019, PhD thesis, Swinburne University of Technology, Australia
- Venkatraman Krishnan, V., Bailes, M., van Straten, W., et al. 2020, *Science*, **367**, 577
- Verbiest, J. P. W., Bailes, M., Coles, W. A., et al. 2009, *MNRAS*, **400**, 951
- Verbunt, F., & Freire, P. C. C. 2014, *A&A*, **561**, A11
- Verbunt, F., & Hut, P. 1987, *IAU Symposium*, **125**, 187
- Voisin, G., Cognard, I., Freire, P. C. C., et al. 2020, *A&A*, **638**, A24
- Voss, R., & Tauris, T. M. 2003, *MNRAS*, **342**, 1169
- Wagoner, R. V. 1970, *Phys. Rev. D*, **1**, 3209
- Wex, N. 2014, arXiv e-prints [arXiv:1402.5594]
- Wex, N., & Kramer, M. 2020, *Universe*, **6**, 156
- Will, C. M. 1993, *Theory and Experiment in Gravitational Physics* (Cambridge: Cambridge University Press)
- Will, C. M. 2014, *Living Rev. Relativity*, **17**, 4
- Will, C. M. 2018a, *Class. Quant. Grav.*, **35**, 085001
- Will, C. M. 2018b, *Theory and Experiment in Gravitational Physics*, 2nd edn. (Cambridge University Press)
- Zhao, J., Freire, P. C. C., Kramer, M., Shao, L., & Wex, N. 2022, *Class. Quant. Grav.*, **39**, 11LT01

Appendix A: Details on calculating grids of NSs

We select slowly rotating axisymmetric approximation of NSs following Damour & Esposito-Farèse (1996). The internal structure of such a NS in DEF gravity is described by a set of 8 ordinary differential equations depending on the radial variable (r). To calculate a single NS the program solves the system of ODEs by applying appropriate boundary conditions at the centre of a NS and spatial infinity. We use a shooting technique to match initial internal parameters with the external structure of the spacetime. The differential equations are solved by means of the Julia library DIFFERENTIALEQUATIONS.JL¹⁰ (Rackauckas & Nie 2017).

A particular solution is determined by the choice of two theory parameters $\{\alpha_0, \beta_0\}$, a EOS and a central pressure (p_c). The central value of the scalar field (φ_c) is determined by the shooting procedure to correspond to the scalar field value at spatial infinity $\varphi_0 = \varphi(r = \infty)$, which in our case is, without loss of generality, set to zero. Thus φ_c is not an arbitrary parameter. The central pressure uniquely corresponds to the gravitational mass m_A for most masses of interest and fixed theory parameters $\{\alpha_0, \beta_0\}$. In general, this is not the case – there may be several stable solutions for a fixed mass in the area of strong scalarisation. However, multiple solutions happen for large NS masses and very negative $\beta_0 < -4.5$ which is already ruled out and thus of no interest.

In the recent papers devoted to the calculation of NSs in STG theories (Anderson & Yunes 2019), the gravitational form factors from Eq. (6) (they are also often called ‘scalar charges’) are calculated using numerical differentiation. Numerical differentiation is performed on a grid of dependent variables. The precision of this approach is determined by two factors: a) the error in the calculation of the desired quantity and b) the error due to the numerical differentiation formula. The first error depends on the precision of the numerical integration of the structure equations. The second error strongly depends on the fineness of the grid on which the derivatives are calculated. If the grid is too sparse the formula such as the central difference formula $f'(x) \approx [f(x + \delta x) - f(x - \delta x)] / [2\delta x]$ is not accurate enough. On the other hand, if the grid is too fine, the error arises because of the subtraction of two close numbers in the computer memory. The total relative errors for calculating gravitational form factors usually lie in the range of $\delta_{\text{rel}} \sim 10^0 - 10^{-5}$. In our work, we propose a more precise way to calculate all the quantities.

A.1. Automatic differentiation

In our program, we utilise automatic differentiation (AutoDiff) technique. It is a special technique to calculate derivatives, when the algorithm knows the exact expressions for derivatives of all elementary functions and uses the chain rule to unfold complex derivatives. A review on AutoDiff can be found, for instance, in Margossian (2018). We use an implementation of AutoDiff technique in Julia from ForwardDiff.jl library, which is a part of a broad JuliaDiff framework (Revels et al. 2016). AutoDiff calculates complex derivatives of quantities with respect to their arguments with the machine-precision (for Double Float numbers the corresponding precision is $\delta \sim 2 \times 10^{-16}$ on 64-bit systems). It is done by internal implementation of sophisticated arithmetic on a special type of numbers (dual numbers) in the programming language. A derivative is calculated exactly in the desired point and does not require points in the vicinity. The

precision of the method is constant and thus does not depend anymore on the fineness of the grid.

Automatic differentiation enables us to simultaneously calculate both the function $F = F(x, y)$ and its derivatives with respect to arguments $\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}$. It takes twice as long as calculating a mere function itself and both are done with the machine-precision. To calculate a complex derivative, then the arguments are also functions $B = B(x, y)$, $C = C(x, y)$ we apply the chain rule:

$$\frac{\partial A(B, C)}{\partial B} \Big|_C = \left(\frac{\partial A}{\partial x} \frac{\partial C}{\partial y} - \frac{\partial A}{\partial y} \frac{\partial C}{\partial x} \right) / \left(\frac{\partial B}{\partial x} \frac{\partial C}{\partial y} - \frac{\partial B}{\partial y} \frac{\partial C}{\partial x} \right). \quad (\text{A.1})$$

In our program, we calculate the gravitational form factors from Eq. (6), which are complex derivatives of quantities (e.g. a mass (m_A) and a moment of inertia (I_A) of a NS) taken with respect to external parameters (the value of the scalar field at spatial infinity (φ_0)). The properties of NSs are obtained through numerical integration of the structure equations and depend on the central scalar field (φ_c) and central pressure (p_c). For example for α_A using Eq. (A.1) we simply obtain:

$$\alpha_A = \frac{\partial \ln m_A}{\partial \varphi_0} \Big|_{\bar{m}_A} = \left(\frac{\partial \ln m_A}{\partial \varphi_c} \frac{\partial \bar{m}_A}{\partial p_c} - \frac{\partial \ln m_A}{\partial p_c} \frac{\partial \bar{m}_A}{\partial \varphi_c} \right) / \left(\frac{\partial \varphi_0}{\partial \varphi_c} \frac{\partial \bar{m}_A}{\partial p_c} - \frac{\partial \varphi_0}{\partial p_c} \frac{\partial \bar{m}_A}{\partial \varphi_c} \right), \quad (\text{A.2})$$

where every simple derivative is calculated with the machine-precision due to AutoDiff.

As a result, the total precision of calculated gravitational form factors is determined only by the accuracy of the differential equation solver and may be easily enhanced. We use the Julia library DIFFERENTIALEQUATIONS.JL which allows us to use AutoDiff on the results of the numerical integration. The numerical integration can take advantage of AutoDiff itself because the Jacobian of the system can be calculated more accurately and used to integrate the ODE system. We obtain a higher precision if we select a finer adaptive step in the radial variable r . For our grids we set the integrator to the relative error of $\delta_{\text{Int,rel}} \sim 10^{-11}$. The obtained relative precision for gravitational form factors typically lies in $\delta_{\text{rel}} \sim 10^{-7} - 10^{-13}$ range, which is more precise than previous calculations.

A.2. Pre-calculated grids of neutron stars

To place limits on the DEF gravity parameters $\{\alpha_0, \beta_0\}$ we need to know the gravitational form factors $\{\alpha_A, \beta_A, k_A\}$ in Eq. (6) as functions of the mass m_A . To approximate these relations we have to calculate grids of NSs. We calculate grids for several EOSs, including MPA1 used in this paper. Each grid contains 4 calculated values: masses m_A and gravitational form factors $\{\alpha_A, \beta_A, k_A\}$ for a range of values of $\{\alpha_0, \beta_0, p_c\}$ with number of points in each axis respectively $\{101, 351, 121\}$. The grid properties are shown in the Table A.1. α_0 and p_c are selected to be logarithmically distributed. Whereas, the parameter β_0 is piece-wise linearly distributed. The grid for p_c is more dense near the common masses of NSs $1.4 - 2.0 M_\odot$. We also include more points in the region of spontaneous scalarisation with $\beta_0 \in [-5.0, -4.0]$. The resulting calculated gravitational form factors are saved in files which then can be used in DEF gravity tests, the DDSTG model in particular.

With the developed DDSTG model we supply grids for 11 different EOSs ranging from soft to stiff. Each EOS satisfies the

¹⁰ <https://diffeq.sciml.ai>

Table A.1. Properties of the pre-calculated grids of NSs.

Axis	Type	# of points	Range
α_0	Log	101	$[-10^{-1}, -10^{-5}]$
β_0	Lin	21	$[-6.0, -5.0]$
		51	$[-5.0, -4.0]$
		281	$[-4.0, +10.0]$
p_c (dyn/cm ²)	Log	21	$[10^{34}, 5 \times 10^{34}]$
		81	$[5 \times 10^{34}, 5 \times 10^{35}]$
		21	$[5 \times 10^{35}, 10^{36}]$

Notes. The grids are 3D and depend on $\{\alpha_0, \beta_0, p_c\}$. Each axis of a grid can be divided into several intervals and be either linearly or logarithmically separated.

requirement on a maximal mass in GR: $M_{\text{NS,max}}^{\text{GR}} > 2M_{\odot}$ (Antoniadis et al. 2013; Fonseca et al. 2021). Table B.1 presents the selected EOSs with their maximum masses in GR. In our work we use the piece-wise polytropic approximation described in Read et al. (2009).

Appendix B: PK parameters in DEF gravity

In this Appendix we present the expressions for the PK parameters in DEF gravity as functions of the masses of pulsar (m_p) and companion (m_c), and the three Keplerian timing parameters P_b , e , and x . The formulae for the PK parameters $\{\gamma, k, r, s\}$ can be found in Damour & Esposito-Farèse (1992a) and Damour & Esposito-Farèse (1996). For the PK parameters $\{\delta_r, \delta_\theta, A, B\}$ there are only more general theory-independent presentations in Damour & Taylor (1992), from which the expressions in DEF gravity, however, can be derived straightforwardly. The total set of PK parameters in DEF gravity reads:

$$\gamma = \frac{e}{n} \frac{X_c \beta_O^2}{1 + \alpha_p \alpha_c} \left[X_c (1 + \alpha_p \alpha_c) + 1 + k_p \alpha_c \right], \quad (\text{B.1})$$

$$k = \frac{3\beta_O^2}{1 - e^2} \left[\frac{1 - \frac{1}{3}\alpha_p \alpha_c}{1 + \alpha_p \alpha_c} - \frac{X_p \beta_c \alpha_p^2 + X_c \beta_p \alpha_c^2}{6(1 + \alpha_p \alpha_c)^2} \right], \quad (\text{B.2})$$

$$r = \frac{G_* m_c}{c^3}, \quad (\text{B.3})$$

$$s = \frac{nx}{X_c \beta_O} \left[1 + \frac{1}{3} \left(\frac{9 - \alpha_p \alpha_c}{1 + \alpha_p \alpha_c} - X_p X_c \right) \beta_O^2 \right], \quad (\text{B.4})$$

$$\delta_r = \frac{\beta_O^2}{m_{\text{tot}}^2 (1 + \alpha_p \alpha_c)} \left[(3 - \alpha_p \alpha_c) m_p^2 + (6 - \alpha_p \alpha_c - k_p \alpha_c) m_p m_c + (2 - \alpha_p \alpha_c - k_p \alpha_c) m_c^2 \right], \quad (\text{B.5})$$

$$\delta_\theta = \frac{\beta_O^2}{m_{\text{tot}}^2 (1 + \alpha_p \alpha_c)} \left[\left(\frac{7}{2} - \frac{1}{2} \alpha_p \alpha_c \right) m_p^2 + (6 - \alpha_p \alpha_c - k_p \alpha_c) m_p m_c + (2 - \alpha_p \alpha_c - k_p \alpha_c) m_c^2 \right], \quad (\text{B.6})$$

$$A = -\frac{\beta_O X_c}{2\pi v_p (1 - e^2)^{1/2}} \frac{\sin \eta}{\sin \lambda}, \quad (\text{B.7})$$

$$B = -\frac{\beta_O X_c}{2\pi v_p (1 - e^2)^{1/2}} \frac{\cos i \cos \eta}{\sin \lambda}, \quad (\text{B.8})$$

$$\dot{P}_b = \dot{P}_b^{\varphi, \text{mon}} + \dot{P}_b^{\varphi, \text{dip}} + \dot{P}_b^{\varphi, \text{quad}} + \dot{P}_b^{g, \text{quad}}, \quad (\text{B.9})$$

Table B.1. EOSs selected for the calculation of grids of NSs.

Name	M_{max} (M_{\odot})	$R_{1.4}$ (km)	Name	M_{max} (M_{\odot})	$R_{1.4}$ (km)
SLy	2.049	11.736	APR3	2.390	12.094
APR4	2.213	11.428	WFF1	2.133	10.414
WFF2	2.198	11.159	ENG	2.240	12.059
MPA1	2.461	12.473	MS1	2.767	14.918
MS1b	2.776	14.583	H4	2.032	13.759
ALF2	2.086	13.188			

Notes. All EOSs have the maximum mass of a NS in GR $M_{\text{NS,max}}^{\text{GR}} > 2M_{\odot}$ and taken in piece-wise polytropic approximation from Read et al. (2009). $R_{1.4}$ is the radius of a $1.4M_{\odot}$ NS.

where we used variables $\beta_O = (G_{\text{pc}} M n / c^3)^{1/3}$, $G_{\text{pc}} = G_*(1 + \alpha_p \alpha_c)$, $m_{\text{tot}} = m_p + m_c$, $X_p = m_p / m_{\text{tot}}$, and $X_c = m_c / m_{\text{tot}} = 1 - X_p$. Moreover, $n = 2\pi / P_b$ is the orbital circular frequency and $v_p = 1/P$ is the pulsar's rotational frequency. The angle i is the inclination of the orbital plane with respect to the plane of sky, while angles λ and η are the polar angles of the spin axis. In the special case of alignment between the orbital and spin axes $\lambda = i$ and $\eta = -\pi/2$ forcing $B = 0$.

The expression for \dot{P}_b is composed from four different contributions (Damour & Esposito-Farèse 1992a)

$$\dot{P}_b^{\varphi, \text{mon}} = -\frac{3\pi X_p X_c}{1 + \alpha_p \alpha_c} \beta_O^5 \frac{e^2 (1 + e^2/4)}{(1 - e^2)^{7/2}} \times \left[\alpha_p + \alpha_c + \frac{2}{3} (\alpha_p X_c + \alpha_c X_p) + \frac{\beta_p \alpha_c + \beta_c \alpha_p}{1 + \alpha_p \alpha_c} \right]^2, \quad (\text{B.10})$$

$$\dot{P}_b^{\varphi, \text{dip}} = -\frac{2\pi X_p X_c}{1 + \alpha_p \alpha_c} \beta_O^3 \frac{1 + e^2/2}{(1 - e^2)^{5/2}} (\alpha_p - \alpha_c)^2 - \frac{4\pi X_p X_c}{1 + \alpha_p \alpha_c} \beta_O^5 \frac{1}{(1 - e^2)^{7/2}} \left[\frac{8}{5} \left(1 + \frac{31e^2}{8} + \frac{19e^4}{32} \right) (\alpha_p - \alpha_c) \times (\alpha_p X_p + \alpha_c X_c) (X_p - X_c) + \left(1 + 3e^2 + \frac{3e^4}{8} \right) \times \frac{(\alpha_p - \alpha_c)(\beta_c \alpha_p X_p - \beta_p \alpha_c X_c)}{1 + \alpha_p \alpha_c} \right], \quad (\text{B.11})$$

$$\dot{P}_b^{\varphi, \text{quad}} = -\frac{32\pi X_p X_c}{5(1 + \alpha_p \alpha_c)} \beta_O^5 \frac{1 + 73e^2/24 + 37e^4/96}{(1 - e^2)^{7/2}} \times (\alpha_p X_c + \alpha_c X_p)^2, \quad (\text{B.12})$$

$$\dot{P}_b^{g, \text{quad}} = -\frac{192\pi X_p X_c}{5(1 + \alpha_p \alpha_c)} \beta_O^5 \frac{1 + 73e^2/24 + 37e^4/96}{(1 - e^2)^{7/2}}, \quad (\text{B.13})$$

where scalar field φ gives rise to a monopolar $\dot{P}_b^{\varphi, \text{mon}}$, a dipolar $\dot{P}_b^{\varphi, \text{dip}}$ and a quadrupolar $\dot{P}_b^{\varphi, \text{quad}}$ contribution. Another quadrupolar term $\dot{P}_b^{g, \text{quad}}$ is associated with metric $g_{\mu\nu}^*$. Following Damour & Esposito-Farèse (1992a), we have ignored a term of order $(\alpha_p - \alpha_c)^2 \beta_O^5$ in Eq. (B.11).

Appendix C: DDSTG implementation into TEMPO

As part of this work, we implement a new DDSTG timing model into an independent version of the standard timing software TEMPO⁷. The new model is based on the DD (Damour & Deruelle 1986) and DDGR (Taylor & Weisberg 1989) models but

Table C.1. Parameters used in the DDSTG timing model in TEMPO.

Name	Description
ALPHA0	α_0 value $\in [-10^{-1}, -10^{-5}] \cup \{0\}$
BETA0	β_0 value $\in [-6, +10]$
EOS	EOS ID, e.g. MPA1
COMP_TYPE	selected from {WD, NS, BH}

modified to work within STG theories. In addition to the parameters present in the DDGR model, the DDSTG model utilises four additional parameters, added to the corresponding parameter input file. New parameters and their possible values are presented in Table C.1.

The first two parameters ‘ALPHA0’ and ‘BETA0’ correspond to the values of two arbitrary DEF gravity parameters $\{\alpha_0, \beta_0\}$. They can be selected to be zero to recover GR. Otherwise, the applicable range of values is determined by the range of the supplementary data files, containing grids of the calculated gravitational form factors and masses. The grids are calculated as part of this work and supplied with the model. The supplied grids cover $\alpha_0 \in [-10^{-1}, -10^{-5}]$ and $\beta_0 \in [-6, +10]$ in the DEF gravity parameter space.

The third parameter ‘EOS’ selects the equation of state to be used to perform the test. In the present work, we generally select the rather stiff EOS MPA1. TEMPO reads the supplied pre-calculated grids of the gravitational form factors and masses corresponding to the selected EOS (see also Table B.1).

The last parameter ‘COMP_TYPE’ is used to select the type of the companion to the pulsar, among a white dwarf (‘WD’), a neutron star (‘NS’) and a black hole (‘BH’). These three compact objects have different properties in terms of gravitational form factors. NSs in DEF gravity can be moderately $|\alpha_{\text{NS}}| \lesssim |\alpha_0|$ or strongly $|\alpha_{\text{NS}}| \sim O(1)$ scalarised depending on the mass m_A and the chosen DEF gravity parameters. White dwarfs are not enough compact to significantly deviate from the weak-field approximation $|\alpha_{\text{WD}}| \lesssim |\alpha_0|$. Thus for WDs we apply $\alpha_{\text{WD}} = \alpha_0$, $\beta_{\text{WD}} = \beta_0$.¹¹ In contrast, BHs in DEF gravity fulfil a no-hair theorem and they are completely descalarised $\alpha_{\text{BH}} = 0$ and $\beta_{\text{BH}} = 0$. The choice of the companion type plays a significant role for the calculated values of PK parameters.

During initialisation, TEMPO reads the parameter file and the file with the TOAs. When the $\{\alpha_0, \beta_0\}$ values and a EOS are provided, it reads the supplementary grids for the selected EOS. The supplied grids are 3D and consist of the masses of NSs m_A and their gravitational form factors $\{\alpha_A, \beta_A, k_A\}$. Two axes are $\{\alpha_0, \beta_0\}$ and the third axis is a fixed range of central pressure (p_c) of a NS.

Bilinear interpolation is performed in the $\{\alpha_0, \beta_0\}$ parameter space for the selected parameters. The result of the interpolation are four 1D arrays for masses and gravitational form factors depending on a range of central pressures. Interpolated 1D arrays are saved in the computer memory and used further without being changed. From this point gravitational form factors are ready to be interpolated in the mass range during the fitting procedure.

TEMPO with the selected DDSTG model fits two masses in addition to spin, astrometric and Keplerian parameters. The masses are iteratively changed during the fitting procedure. Only

¹¹ Since this is the case for any weakly self-gravitating body, the setting WD can also be selected for any non-degenerate companion such as a main-sequence star.

when the masses change TEMPO interpolates the gravitational form factors from the pre-calculated 1D grids. Gravitational form factors are calculated for both the pulsar and the companion and are kept constant for all TOAs within a single iteration.

TEMPO tries to find the companion mass (m_c) and the total mass ($m_{\text{tot}} = m_p + m_c$) corresponding to the best fit. During the fitting, it utilises the derivatives of the timing formula (7) with respect to these two masses. We calculate derivatives separately for each delay in Equations (11 – 14). We used the approximated expressions for the derivatives of the delays with respect to PK parameters provided by Damour & Deruelle (1986). After applying the chain rule for our derivatives we come to expressions depending on the derivatives of PK parameters with respect to the two masses. The model calculates the derivatives of the PK parameters of DEF gravity (Eqs. B.1 – B.9) with respect to the two masses of the timing model,

$$\left(\frac{\partial p_i^{\text{PK}}(m_c, m_{\text{tot}})}{\partial m_c} \right) \Big|_{m_{\text{tot}}}, \quad \left(\frac{\partial p_i^{\text{PK}}(m_c, m_{\text{tot}})}{\partial m_{\text{tot}}} \right) \Big|_{m_c}, \quad (\text{C.1})$$

which are certainly different to those in GR. Moreover, for proper convergence of the model in highly non-linear areas the additional corrections to the derivatives related to variation of gravitational form factors with respect to masses (i.e. $\partial \alpha_A / \partial m_A, \partial \beta_A / \partial m_A$) are required.

The formulae in Eq. (C.1) are large but can be obtained by straightforward differentiation of the PK formulae in Appendix B. The expressions are incorporated in the model and depend on gravitational form factors, masses and Keplerian parameters which are known on each iteration. Once the derivatives of PK parameters are calculated, the model calculates the derivatives of the timing delays with respect to the two masses and proceeds the iteration. The outcome of the DDSTG model in TEMPO is the same as for the DDGR model. The most important result is the χ^2 value, which is used further to perform tests of gravity.

The DDSTG timing model converges well if the eccentricity of the orbit is not too small and TEMPO is supplied with a reasonable initial guess of the orbital parameters. The model can run into difficulties in the convergence for solutions if those parameters are far away from GR or if the orbit is almost circular. The problem can happen if TEMPO changes masses significantly during the iteration and reaches a point where the Shapiro shape parameter calculated with new masses using Eq. (B.4) becomes greater than one, i.e. $s > 1$. In this case the expression of the Shapiro delay from Eq. (13) does not have a physical meaning and TEMPO breaks with an error for TOAs near conjunction, as the argument of the logarithm becomes negative. The convergence can be enforced by applying the special ‘GAIN’ parameter, which however comes at the cost of increased computational time. A GAIN parameter of less than one forces TEMPO to use smaller steps while iterating. Another way which ensures the convergence is to supply TEMPO with a good initial guess for the masses in DEF gravity. The rough estimate of the masses in DEF gravity can be obtained from the traditional PK method. The combination of these two methods helps to converge systems with a small eccentricity and Shapiro shape parameter near one ($s \approx 1$) even in the strongly non-linear part of the $\{\alpha_0, \beta_0\}$ plane.

Appendix D: Mass-mass diagram near the horn

Many binary pulsar tests are not sensitive at high $|\alpha_0|$ values and $\beta_0 \sim -2$. In this region near the ‘horn’, $\alpha_{\text{NS}} \approx \alpha_0$ and the

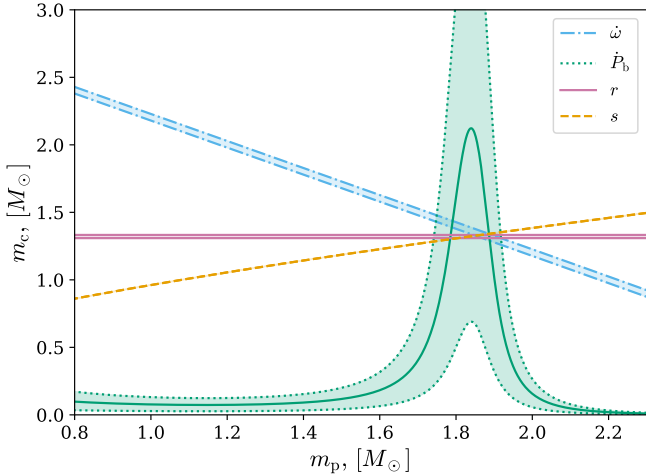


Fig. D.1. Mass-mass diagram in DEF gravity corresponding to a point near the ‘horn’ with $\alpha_0 = -10^{-1}$, $\beta_0 = -1.9$. The performed test is the same as for Fig. 1. The shadowed area is the allowed region at 68% CL limit for a corresponding PK parameter. The solid green line corresponds to the observed value of \dot{P}_b^{int} .

overall \dot{P}_b is close to the GR value, as the dipolar contribution is greatly suppressed. As a consequence, the test is passed. In Fig. D.1 we show a mass-mass diagram in DEF gravity with $\alpha_0 = -0.1, \beta_0 = -1.9$. This point in the DEF gravity parameter space passes the test with timing data but is excluded by Cassini experiment ($|\alpha_0| \leq 3.4 \times 10^{-3}$). The picture is dramatically different from what we see in the zone of scalarisation in Fig. 1.

Appendix E: Proper account of \dot{P}_b^{ext} uncertainty

If the observed \dot{P}_b is measured more precisely than the external contribution \dot{P}_b^{ext} we have to take the uncertainty of the external contribution into account. In most situations \dot{P}_b^{ext} consists of the Shklovskii and the Galactic contribution. The calculation of the Shklovskii contribution is limited by the uncertainty in the distance and the one in the proper motion. The uncertainty in the estimation of the Galactic contribution is determined by the uncertainty in the distance and our imperfect knowledge of the Galactic gravitational potential. The latter is partly systematic in nature and therefore somewhat more difficult to quantify. If the uncertainty $\Delta\dot{P}_b^{\text{ext}}$ is the limiting factor we no longer can have \dot{P}_b^{ext} fixed on one value for the whole experiment (i.e. provide a fixed value for XPBDOT parameter in TEMPO) because its uncertainty can affect the derived limits.

To properly account for the uncertainty in \dot{P}_b^{ext} we follow the method described in Splaver et al. (2002) based on Bayes inference. On the first stage TEMPO calculates χ^2 grid for three independent parameters $\{\alpha_0, \beta_0, \dot{P}_b^{\text{ext}}\}$. Then the shifted value $\Delta\chi^2 = \chi^2 - \chi_{\text{min}}^2$ is described by χ^2 distribution with 3 degrees of freedom. The number of degrees of freedom is important when we calculate confidence level limits. The $\Delta\chi^2$ maps as usual to a Bayesian likelihood function,

$$p(\{t_j\} | \alpha_0, \beta_0, \dot{P}_b^{\text{ext}}) = \frac{1}{2} e^{-\Delta\chi^2/2}, \quad (\text{E.1})$$

where $\{t_j\}$ refers to the used timing data. We treat the obtained probability density $p(\{t_j\} | \alpha_0, \beta_0, \dot{P}_b^{\text{ext}})$ as the likelihood for

Bayes’ theorem

$$p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}} | \{t_j\}) = \frac{p(\{t_j\} | \alpha_0, \beta_0, \dot{P}_b^{\text{ext}})}{p(\{t_j\})} p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}}), \quad (\text{E.2})$$

where $p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}} | \{t_j\})$ is the desired joint posterior probability function and $p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}})$ is the prior. The Bayesian evidence $p(\{t_j\})$ is obtained from normalisation of posterior probability over the whole grid $\{\alpha_0, \beta_0\}$.

In the next stage we select the prior $p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}})$. We assume no prior information about α_0 and β_0 , so their distributions are taken uniformly for $\beta_0 \in (-\infty, +\infty)$ and $\alpha_0 \in (-\infty, 0]$. At this step, we utilise the information about the uncertainties in \dot{P}_b^{ext} . For example, we can select the prior $p(\dot{P}_b^{\text{ext}})$ to be normally distributed with the mean measured value \dot{P}_b^{ext} and its standard deviation $\Delta\dot{P}_b^{\text{ext}}$. Or we can use more elaborate probability distribution accounting for our uncertainty in the measurements of distance, proper motion, and Galactic gravitational potential,

$$p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}}) = p(\alpha_0) \times p(\beta_0) \times p(\dot{P}_b^{\text{ext}}). \quad (\text{E.3})$$

The last step is to marginalise over the \dot{P}_b^{ext} variable using the correct joint posterior probability with desired priors

$$p(\alpha_0, \beta_0 | \{t_j\}) = \int d\dot{P}_b^{\text{ext}} p(\alpha_0, \beta_0, \dot{P}_b^{\text{ext}} | \{t_j\}), \quad (\text{E.4})$$

where the integral is calculated on the grid of \dot{P}_b^{ext} parameter. The marginalised probability then can be used to derive limits, for instance going back to χ^2 representation with 2 degrees of freedom.