Publication Year	2021
Acceptance in OA	2023-01-16T14:12:38Z
Title	Strong-field Gravity Tests with the Double Pulsar
Authors	M. Kramer, I. H. Stairs, R. N. Manchester, N. Wex, A. T. Deller, W. A. Coles, M. Ali, BURGAY, MARTA, F. Camilo, I. Cognard, T. Damour, G. Desvignes, R. D. Ferdman, P. C. C. Freire, S. Grondin, L. Guillemot, G. B. Hobbs, G. Janssen, R. Karuppusamy, D. R. Lorimer, A. G. Lyne, J. W. McKee, M. McLaughlin, L. E. Muench, B. B. P. Perera, N. Pol, POSSENTI, ANDREA, J. Sarkissian, B. W. Stappers, G. Theureau
Publisher's version (DOI)	10.1103/PhysRevX.11.041050
Handle	http://hdl.handle.net/20.500.12386/32866
Journal	PHYSICAL REVIEW. X
Volume	11

Featured in Physics

Strong-Field Gravity Tests with the Double Pulsar

M. Kramer,[1,2,*] I. H. Stairs,[3] R. N. Manchester,[4] N. Wex,[1] A. T. Deller,[5,6] W. A. Coles,[7] M. Ali,[1,8] M. Burgay,[9] F. Camilo,[10] I. Cognard,[11,12] T. Damour,[13] G. Desvignes,[14,1] R. D. Ferdman,[15] P. C. C. Freire,[1] S. Grondin,[3,16] L. Guillemot,[11,12] G. B. Hobbs,[4] G. Janssen,[17,18] R. Karuppusamy,[1] D. R. Lorimer,[19] A. G. Lyne,[2] J. W. McKee,[1,20] M. McLaughlin,[19] L. E. Münch,[1] B. B. P. Perera,[21] N. Pol,[19,22] A. Possenti,[9,23] J. Sarkissian,[4] B. W. Stappers,[2] and G. Theureau[11,12,24]

[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[2]*Jodrell Bank Centre for Astrophysics, The University of Manchester, Oxford M13 9PL, United Kingdom*
[3]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia V6T 1Z1, Canada*
[4]*Australia Telescope National Facility, CSIRO Space and Astronomy, P.O. Box 76, Epping, New South Wales 1710, Australia*
[5]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, Victoria 3122, Australia*
[6]*ARC Centre of Excellence for Gravitational Wave Discovery (OzGrav), P.O. Box 218, Hawthorn, Victoria 3122, Australia*
[7]*Electrical and Computer Engineering, University of California at San Diego, La Jolla, California 92093, USA*
[8]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5*
[9]*INAF-Osservatorio Astronomico di Cagliari, Via della Scienza 5, 09047 Selargius (CA), Italy*
[10]*South African Radio Astronomy Observatory, 2 Fir Street, Black River Park, Observatory 7925, South Africa*
[11]*Laboratoire de Physique et Chimie de l'Environnement et de l'Espace LPC2E CNRS-Université d'Orléans, F-45071 Orléans, France*
[12]*Station de Radioastronomie de Nançay, Observatoire de Paris, CNRS/INSU, F-18330 Nançay, France*
[13]*Institut des Hautes Études Scientifiques, 91440 Bures-sur-Yvette, France*
[14]*LESIA, Observatoire de Paris, Universit PSL, CNRS, Sorbonne Université, Universit de Paris, 5 Place Jules Janssen, 92195, Meudon, France*
[15]*Faculty of Science, University of East Anglia, Norwich NR4 7TJ, United Kingdom*
[16]*David A. Dunlap Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, Ontario M5S 3H4, Canada*
[17]*ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD, Dwingeloo, Netherlands*
[18]*Department of Astrophysics/IMAPP, Radboud University, P.O. Box 9010, 6500 GL, Nijmegen, Netherlands*
[19]*Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, West Virginia 26506, USA*
[20]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada*
[21]*Arecibo Observatory, University of Central Florida, HC3 Box 53995, Arecibo, Puerto Rico 00612, USA*
[22]*Department of Physics and Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, Tennessee 37235, USA*
[23]*Universitá di Cagliari, Dipartimento di Fisica, S.P. Monserrato-Sestu Km 0, 700–09042 Monserrato (CA), Italy*
[24]*Laboratoire Univers et Théories LUTh, Observatoire de Paris, PSL Research University, CNRS/INSU, Université Paris Diderot, 5 place Jules Janssen, 92190 Meudon, France*

(Received 3 May 2021; accepted 25 October 2021; published 13 December 2021)

*mkramer@mpifr-bonn.mpg.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

Continued timing observations of the double pulsar PSR J0737–3039A/B, which consists of two active radio pulsars (A and B) that orbit each other with a period of 2.45 h in a mildly eccentric ($e = 0.088$) binary system, have led to large improvements in the measurement of relativistic effects in this system. With a 16-yr data span, the results enable precision tests of theories of gravity for strongly self-gravitating bodies and also reveal new relativistic effects that have been expected but are now observed for the first time. These include effects of light propagation in strong gravitational fields which are currently not testable by any other method. In particular, we observe the effects of retardation and aberrational light bending that allow determination of the spin direction of the pulsar. In total, we detect seven post-Keplerian parameters in this system, more than for any other known binary pulsar. For some of these effects, the measurement precision is now so high that for the first time we have to take higher-order contributions into account. These include the contribution of the A pulsar's effective mass loss (due to spin-down) to the observed orbital period decay, a relativistic deformation of the orbit, and the effects of the equation of state of superdense matter on the observed post-Keplerian parameters via relativistic spin-orbit coupling. We discuss the implications of our findings, including those for the moment of inertia of neutron stars, and present the currently most precise test of general relativity's quadrupolar description of gravitational waves, validating the prediction of general relativity at a level of 1.3×10^{-4} with 95% confidence. We demonstrate the utility of the double pulsar for tests of alternative theories of gravity by focusing on two specific examples and also discuss some implications of the observations for studies of the interstellar medium and models for the formation of the double pulsar system. Finally, we provide context to other types of related experiments and prospects for the future.

DOI: 10.1103/PhysRevX.11.041050 Subject Areas: Astrophysics, Gravitation

I. INTRODUCTION: PULSARS AS PROBES OF GRAVITATIONAL PHYSICS

The study of gravitational physics currently benefits from a number of experimental advances which provide unprecedented opportunities for constraining the underlying theory. Many of these methods focus on the study of compact objects, namely, neutron stars (NSs) and black holes (BHs), in order to confront the theories to be studied with data obtained under strong-field conditions. A prime example is the availability of ground-based gravitational-wave detectors which provided the first detection of gravitational waves (GWs) with Earth-bound detectors [1]. Almost 40 years earlier, the first evidence for GWs was provided by observations of the first binary pulsar, PSR B1913+16 [2,3]. Pulsar observations provide unique and complementary experimental constraints, especially thanks to the unrivalled precision enabled by radio astronomical observations and the method of pulsar timing (see, e.g., Ref. [4]), which permit us to trace the orbital evolution of a system over long periods of time. Pulsar observations, furthermore, provide valuable information on the structure of NSs, on plasma physics, on asymmetries in the supernova explosion of massive stars, and on the ionized content of the interstellar medium.

A. Pulsars

The discovery of pulsars in 1967 [5] provided astronomers with natural and extraordinarily stable fly-wheel clocks, opening up the prospect of probing the composition of matter at extremely high densities. Pulsars were soon identified as rotating NSs, exceedingly dense and compact stellar remnants formed in supernova explosions [6,7]. Most known pulsars are located within our Galaxy, typically at distances of a few kpc [8] NSs, which typically have a mass of about 1.4 times the mass of the Sun ($M_\odot$) but a radius of only about 12 km, can and, in fact, are observed to spin very rapidly, up to approximately 700 times every second [9]. Precise timing of pulse arrival times at Earth shows that pulsars can have a rotational stability comparable to that of the best atomic clocks. But not all pulsars have highly stable periods. Observations of the rotational instabilities known as "glitches" (in mostly young pulsars) allow a kind of stellar seismology to probe the NS's solid crust, its liquid interior, and the coupling between them (see, e.g., Refs. [10,11]).

Importantly, many pulsars, especially those with pulse periods in the millisecond range, both are stable rotators and are in a binary orbit with another star. Their great timing stability enables tiny variations in the arrival time due to, e.g., relativistic effects, to be accurately measured, thereby allowing tests of gravitational theories and investigations of many other physical phenomena.

The timing of binary pulsars also delivers measurements of NS masses with unprecedented precision, yielding evidence for NSs with maximum masses of at least $2\ M_\odot$ [12–15]. Such large masses are incompatible with a number of theoretical equations of state (EOSs) proposed to describe the properties of superdense matter; the discovery of even more massive NSs would restrict the families of EOSs even further [16]. Other ways of constraining the EOS via pulsar timing include the potential measurement of the moment of inertia (MOI) of a NS.

Soon after the discovery of the first binary pulsar, PSR B1913+16 in 1975 [17], it was understood that measurements of relativistic effects in binary pulsars would allow

investigation of the masses and composition of NSs, with the first attempts at describing relativistic orbital effects as part of a timing model by Ref. [18] and relativistic spin precession (due to spin-orbit coupling) by Refs. [19,20]. To first order, measured orbital perturbations are determined by the line-of-sight motion as the pulsar orbits its companion, meaning that motion in the sky plane is not measurable. Consequently, these measurements permit a range of masses and orbital inclination angles. Multiple relativistic effects must be observed in order to both constrain the masses and begin testing gravitational theories, some of which may depend on the NS composition. Many aspects of gravitational theories are best tested when the strongly self-gravitating NSs are found as pulsars in systems with short orbital periods, nonzero orbital eccentricities, and orbital planes closely aligned to our line of sight.

All of these desirable characteristics are present in the still unique double pulsar system PSR J0737–3039A/B. The 2003 discovery [21] showed it to have a very tight orbit and excellent potential for gravity tests, and shortly thereafter it was shown to be a system with two orbiting active radio pulsars [22]. The system consists of a "recycled" 23-ms pulsar ("A") and the second-born 2.8-s pulsar ("B"). Probably after being "dead" or at least undetectable for a few million years, the A pulsar was spun up and restored to detectability as its companion star, the progenitor of pulsar B, evolved and transferred matter and angular momentum to it. Subsequently, the progenitor of B exploded, leading to two NSs in the highly relativistic, slightly eccentric 2.45-h orbit that we observe today. Further details on the double pulsar system are given in Sec. II.

B. Pulsar timing

Pulsar timing analyses begin with measurement of precise pulse times of arrival (TOAs) at the telescope. While many pulsars emit pulsed signals at high energies (x rays and gamma rays), it is the radio band which is of most interest here. Two orthogonal polarizations of the incoming electromagnetic wave are recorded by the receiver system of the telescope. These measurements are typically made at radio frequencies of hundreds of megahertz (MHz) or several gigahertz (GHz), with signals being Nyquist sampled at twice the receiver bandwidth. This sampling is often preceded by a frequency down-conversion and followed by channelization in frequency using a digital signal-processing system that can differ depending on the telescope. Astronomical radio signals suffer a dispersive delay due to free electrons in the interstellar medium (ISM), parametrized by the dispersion measure (DM), which must be taken into account. Within a given channel bandwidth, e.g., in our observations 1 MHz, the pulsar signals are "coherently dedispersed" to remove this dispersive delay and then folded at the topocentric pulsar period. Again, the methods depend on the telescope and the receiver. The folded profiles are averaged over a "subintegration" interval, which in our observations is 30 s, resulting in a data cube of pulse amplitude versus pulse phase for each frequency channel.

The next step is to convert the observed pulse profiles into TOAs. This is done by comparing each profile, which is time stamped by the observatory's hydrogen-maser atomic clock, with a carefully prepared template, giving both a TOA and the uncertainty in that TOA [23]. The uncertainty due to template matching is independent between TOAs but depends on the strength of the observed pulsar signal, which can vary dramatically due to interstellar scintillation caused by turbulent fluctuations in the ionized ISM density.

However, the uncertainty of the template matching is not the only source of noise. Because of relatively large-scale fluctuations in the ISM electron density, the DM is variable on a timescale of months. Pulsars also display intrinsic "timing noise" due to variations in spin frequency with a similar timescale, and this is then common to all radio frequency channels. Both of these noise sources have steep power-law spectra and become negligible on short timescales such as the orbital period of PSR J0737–3039A/B.

Finally, the observed TOAs are compared with predictions from a timing model. The model describes (and, in fact, counts) the rotations of the pulsar and accounts for physical effects that modify the TOA. The corresponding computations are done using a timing analysis program, either Tempo [24] or Tempo2 [25] in our analysis.

The (usually small) differences between the observed and predicted TOAs are known as "timing residuals." These residuals are basic to pulsar timing analyses, since they reveal effects that are not included in the timing model. The resulting signatures typically have a particular form in the timing residuals, and some of these may result from previously unknown effects.

C. Tests of theories of gravity

The general theory of relativity, or general relativity (GR) [26], has passed its experimental tests with flying colors, so far. Despite its successes, it may not be our final answer in describing gravity on a macroscopic scale. There is a range of parameter space, from the quasistationary weak-field regime of the Solar System to the strong-field regime of compact objects like NSs and BHs, in all of which one may encounter an experiment where the theory could be falsified [27]. It is therefore important to test different aspects of the predictions of GR and alternative theories with different methods. For instance, observations with gravitational-wave detectors are able to test the highly dynamical strong-field regime and radiative aspects of gravity, but they are not able to test aspects of light propagation in strong fields. This, on the other hand, and other aspects can be tested with binary pulsars.

The equations of GR, and indeed of alternative theories of gravity, are nonlinear and must be approximated for comparison with binary pulsar data. Damour and Deruelle [28,29] provide a leading-order pulsar timing model which includes the effects expected in GR, such as the advance of

periastron and Shapiro delay, but which is parametrized in a way that does not assume the validity of GR or any other theory of gravity. Once measured, these "post-Keplerian" (PK) parameters can be used to determine masses (based on an assumed theory) and perform self-consistency tests of theories; this is the approach subsequently taken in timing the Hulse-Taylor pulsar [30] and other systems [27,31–34]. Damour and Taylor [35] expand the formalism to include parameters based on the pulse profile changes expected in relativistic spin precession and simulate measurement timescales for several parameters.

In this framework, any given relativistic theory of gravity provides a description of the PK parameters as functions of the measured Keplerian parameters and the two *a priori* unknown masses of the binary system. Measuring n PK parameters, where $n > 2$, overdetermines this system of equations, providing $n - 2$ independent tests of the studied theory. In the work presented here, by measuring more PK parameters than in any other system and by measuring them also more precisely than usual in any other, we find we need to go well beyond what has been considered in earlier work. As an example, for two of the PK parameters, besides their dependence on the masses and the Keplerian parameters, we need to incorporate their dependence on the MOI of pulsar A in our analyses.

D. A coordinated gravity experiment

In many respects, we have reached a juncture in the application of binary pulsars to tests of gravitational physics. As we demonstrate in this paper, from now on, we have to consider a number of effects that could be neglected in the past but now require attention and the application of new methods. This is true now for the double pulsar but will eventually also apply for other systems in the future.

The high precision of our measurements that we describe in Sec. III forces us to consider the relativistic mass loss of the system due to the pulsar spin-down, while considering the EOS is now essential to interpret our observational results. We show explicitly that relativistic orbital deformation needs to be accounted for in gravity tests based on time dilation (a combination of second-order Doppler effect and gravitational redshift). Higher-order light-propagation effects in strong gravitational fields are also clearly evident in our data. More specifically, we measure an (higher-order) aberration effect due to the deflection of the radio signal in the gravitational field of the companion, which gives us the rotation sense of A relative to the orbital angular momentum. The advance of periastron, meanwhile, requires two important corrections: a second post-Newtonian term and a correction due to relativistic spin-orbit coupling that changes the orientation of the orbit over time [36,37]. The latter is proportional to the pulsar spin and the (generally negligible) companion spin and, therefore, encodes information about the MOI of the pulsar. This offers an opportunity to constrain this NS property, important for determining the EOS, for the first time via pulsar timing. Consequently, we need to develop a new timing model to jointly account for all these effects.

Before we can apply such a model, however, we also need to determine the distance to the pulsar, as this is an important parameter in many applications of pulsar timing, in particular, in tests of the effects of gravitational-wave emission. Indeed, for the Hulse-Taylor pulsar, the limited precision of the known distance (and its acceleration relative to the Solar System) has prevented any improvement as a gravity experiment for a considerable time now [38]. Since the double pulsar is relatively close on a Galactic scale, there is a tiny, but measurable, curvature in the signal wavefront. This results in a small TOA modulation with a period of six months (as we track the source during Earth's motion around the Sun) which allows us to determine the distance to the double pulsar.

Measurement of this "timing parallax" is, unfortunately, hampered by long timescale variations in both the pulsar intrinsic spin and the intervening ISM, the latter causing variations in the observed DM. The effect of DM variations can be corrected but not the intrinsic spin noise. Hence, we cannot constrain the pulsar distance using pulsar timing as tightly as we would prefer. These effects are discussed in Secs. IV B and VIII A.

Fortunately, we have a second independent method of determining the distance to the double pulsar. This method uses high-angular-resolution imaging of the pulsar with continental-scale radio interferometers, a technique known as "very long baseline interferometry" (VLBI). It also depends on Earth's motion around the Sun, detecting small annual modulations in the pulsar's apparent *position* on the sky ("annual geometric parallax"; see, e.g., Ref. [39]). Since the measurements are made in the plane of the sky, as opposed to pulsar timing which is sensitive to line-of-sight changes, such VLBI measurements are subject to different systematic effects and offer complementary information. In our analysis, we adopt the weighted mean of these two independent distance measurements to obtain the distance with a sufficiently small uncertainty.

With the distance measurement in hand, we develop and present the new full-relativistic timing model in Sec. V and present its application to the analysis of our TOAs in Sec. VI. Using the measurement of PK parameters, we obtain precise mass measurements and derive finally a set of independent tests of GR, finding superb internal consistency. Our most precise test is one related to GR's prediction for GW emission, validating this prediction at a level of 1.3×10^{-4} with 95% confidence.

The close agreement of the double pulsar timing with GR, in turn, means that one can use these observations to place tight constraints on various alternatives to GR. In Sec. VII, we demonstrate this on the basis of two well-studied gravity theories. The first such theory is "Damour–Esposito-Farése" (DEF) gravity, a two-parameter monoscalar-tensor gravity [40], containing GR as a limit with the two parameters

$\alpha_0 = \beta_0 = 0$. DEF gravity exhibits the typical effects one would expect if the strong equivalence principle (SEP; see, e.g., Ref. [41]) is violated, for instance, the existence of (scalar) dipolar GWs and a location-dependent gravitational constant, both leading to characteristic modifications of the PK parameters [42]. Moreover, in certain regions of the parameter space, DEF gravity shows genuine nonperturbative strong-field effects, present only in NSs and, therefore, not testable in the weak-field regime of the Solar System. As shown in detail in Sec. VII, with the observations presented here, we limit such deviations from GR, further constraining the $\alpha_0 - \beta_0$ parameter space of DEF gravity.

The second alternative to GR which we confront with our constraints from the double pulsar is Bekenstein's tensor-vector-scalar theory (TeVeS) [43], a MONDian relativistic gravity theory that evades the need for dark matter in galaxies by a modification of GR. In Sec. VII, we show that the TeVeS is practically incompatible with our observations. Despite the fact that the TeVeS is already falsified by the confirmation that tensor modes of GWs travel with the speed of light [44], the double pulsar experiment has its own merits. It tests specifically the scalar sector of the theory and shows that, depending on the details of a MONDian gravity theory, such a theory can be tested by the radiative and strong-field properties of binary pulsars.

As mentioned above, the inhomogeneous turbulent ISM in between the double pulsar and Earth leads to slow variations of the DM. Associated refractive variations affect the apparent pulsar position as measured by VLBI with a corresponding impact on the VLBI distance as discussed in Sec. IV C. However, the study of the ISM can also provide independent information on the pulsar distance. These constraints are discussed in detail in Sec. VIII, where we compare the measured pulsar distance with estimates based on established models of the Galactic ionized gas [45,46]. We also investigate the inferred structure of the ISM and bring the discussion full circle, demonstrating consistency with our preferred pulsar distance. As a last consistency check of our results, in Sec. VIII C, we discuss the system's inferred low space velocity and its implications for formation of the binary system.

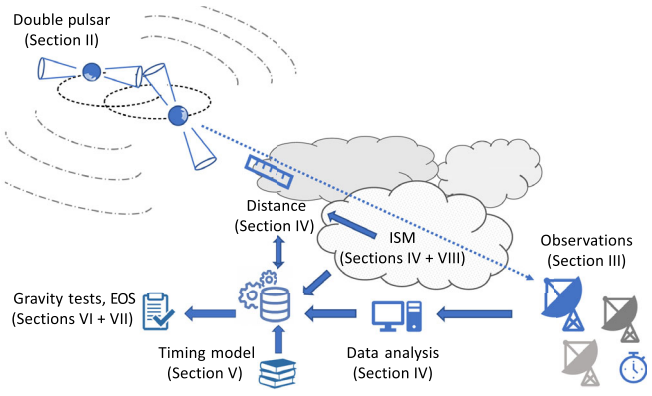

FIG. 1. An overview of the experiment described in this work. The named sections provide an orientation to descriptions of the various components.

In Sec. IX, we discuss the future prospects for tests of gravitational physics with the double pulsar system, and, in Sec. X, we conclude with a summary of our results and how these complement other methods for testing gravity. Figure 1 provides an overview of both our experiment and the descriptions of its various components in this paper.

II. THE DOUBLE PULSAR

In this section, we describe some relevant aspects of the double pulsar system in more detail, including the eclipses of the A pulsar emission by the magnetosphere of B and the resultant constraints on the orientation of the spin vectors of the two NSs relative to the orbital angular momentum vector.

As the original discovery paper [22] shows, the double pulsar system PSR J0737–3039A/B is an eclipsing dual-line binary system. Pulsar B is believed to have been formed in a low-kick supernova event, the details of which are still a matter of debate [47–49]. The relatively low eccentricity ($e = 0.088$), the small system transverse velocity ($v_{\rm trans} \sim 10\ {\rm km\,s^{-1}}$) [31], and the very small misalignment angle of the spin vector of A relative to the total angular momentum vector ($\delta_{\rm A} < 3.2°$) [50,51] are indeed all signposts of a low-kick birth event of the double pulsar. A retrograde solution ($180° - \delta_{\rm A}$) is possible but considered less likely, as it would require a very strong kick with a fine-tuned magnitude and direction. Furthermore, the prograde rotation of A has been confirmed by using emission properties of B [52]. In this work, we confirm this independently using the newly seen relativistic effect of aberrational light bending.

The eclipses of pulsar A by pulsar B were first detected in the B discovery paper [22] and shown to have a duration of just 30 s. The discovery paper already attributed the eclipses to absorption by the magnetosphere of B, and this idea was dramatically confirmed when it was shown that the eclipses are modulated at either the rotation rate of B or twice that rate, depending on the orbital phase [53]. This paper also discusses the likelihood that the properties of the eclipse would be affected by relativistic precession [19] of the B spin axis about the total angular momentum vector.

In contrast to A, B was born with a significant misalignment of its spin axis relative to the total angular momentum (which is dominated by the orbital angular momentum). Detailed modeling of the A eclipse modulation by the magnetosphere of B gives a measurement of B's spin misalignment angle of $\delta_{\rm B} \simeq 50°$ [54]. [55] The rate of the relativistic precession of B's spin vector is also derived from the eclipse modelling: The observed rate given by the PK parameter, $\Omega_{\rm B}^{\rm spin} = 4°.77 \pm 0°.66\ {\rm yr^{-1}}$, is consistent with that predicted by GR (a 70.96-yr precessional period, corresponding to a precessional rate of $5°.073\ {\rm yr^{-1}}$) within

the uncertainty of 13% [54]. The spin precession causes changes in the observed pulse shapes of B [56,57] and ultimately resulted in B's disappearance from our view in March 2008 [58]. When the B pulsar returns to our view depends on the actual beam shape as well as the precessional rate [57,58]. The beam shape is changing with time, as B's magnetosphere is severely distorted by the wind emerging from A to form a magnetotail on the downstream side [22,53]. From the length of the eclipse, presumably caused by synchrotron self-absorption of A's radio emission in B's magnetosphere [59,60], one can estimate that the magnetosphere extends only to about 40% of B's light-cylinder radius. This is consistent with estimates of the dynamic pressure of A's wind [53,61].

The present work builds on the observations and analysis of the double pulsar system presented by Kramer *et al.* [31], together with the measurement of relativistic spin precession by Breton *et al.* [54]. Making use of observations over a 2.5-yr data span, the Kramer *et al.* work presented precise determinations of the relativistic periastron precession of the system, the combined effect of time dilation and gravitational redshift, the Shapiro delay effect due to light propagation in curved spacetime, and the decay of the orbit due to GW emission. These effects are described in more detail in Sec. V.

Combining these observations led to five strong-field tests of gravity. These make use of not only six PK parameters previously measured from relativistic effects, but also of the theory-independent [62] ratio of the two NS masses uniquely available in this system [31,54].

Since then, the system has been studied continuously using a number of radio telescopes, with improved data acquisition systems and better sensitivity, resulting in much improved timing precision over time. Here, we report on new results from the timing of A, extending the data span to 16.2 yr, and describe multiple new phenomena in the system.

III. OBSERVATIONS AND SIGNAL PROCESSING

The vast majority of the results presented in this work are based on an extensive pulse timing experiment made over a 16.2-yr interval made at six observatories around the world: the Parkes 64-m radio telescope ("Murriyang") in New South Wales, Australia, the Green Bank 100-m radio telescope (GBT) in West Virginia, USA, the Westerbork Synthesis Radio Telescope (WSRT) in Netherlands, the Nançay Radio Telescope (NRT) in France, the Effelsberg 100-m radio telescope in Germany, and the Lovell 76-m radio telescope (LT) at Jodrell Bank Observatory in the United Kingdom. In addition, we undertook interferometric imaging observations using the Very Long Baseline Array (VLBA) to obtain complementary information on the pulsar's position in the plane of the sky as a function of time. Details of the datasets from these telescopes and the signal-processing methods are described here.

Timing observations commenced with the confirmation of the A pulsar search candidate in April, 2003. The TOAs included in our timing analysis lie between 2003 May 1, modified Julian day (MJD) 52760, and 2019 July 3, MJD 58667, a total span of 16.2 years. Details of the different datasets used in the analysis are given in Table I. This table lists in order, the observatory and/or telescope used, the common designations of the receiver and signal-processing systems, the center frequency and total bandwidth of the dataset, the number of frequency channels in the processed data and whether or not the channel data are coherently dedispersed, the basic subintegration or TOA sampling time, the data span, the number of TOAs, and the weighted root-mean-square (rms) timing residual of a basic fit to each dataset using Tempo2. As discussed in more detail below (Sec. IV), two related datasets are used for the timing analyses. For the determination of astrometric parameters and DM variations, 4-min-sampled TOAs are analyzed using Tempo2; the penultimate line of Table I summarizes this dataset. Binary and relativistic parameters are determined using a modified version of Tempo which uses a new timing model and modifications not available in Tempo2 (see Sec. V below for more details) with the 30-s-sampled TOAs. The final line in Table I gives the parameters of this combined dataset. Figure 2 illustrates the time and frequency coverage of the datasets from the different observatories.

In many respects, observations and analyses are identical or similar to those used in Ref. [31]. In the following subsections, we describe the summaries and details where required for the data-acquisition systems of the contributing observatories.

A. Parkes Telescope

As summarized in Table I, the Parkes observations were made in three main bands, centered around 700, 1400, and 3100 MHz, respectively, using a variety of receivers and signal processors. As illustrated in Fig. 2, in terms of time coverage, the Parkes 1400-MHz dataset is the most complete of those analyzed. All datasets have a basic time sampling or subintegration length of 30 s.

The 700-MHz and most of the 3100-MHz observations were made using the dual-coaxial 10-cm/50-cm receiver [64], while the 1400-MHz observations primarily use the center beam of the 20-cm 13-beam (MB) receiver [65]. For several months in 2016 while the MB receiver was refurbished, the H-OH receiver which covers a 512-MHz band centered at 1526 MHz was used. Following retirement of the 10-cm/50-cm receiver in 2018 October, the ultra-wideband low (UWL) receiver [66] is used to extend the 1400- and 3100-MHz datasets. All of these receivers have (or had) orthogonal linearly polarized receptors. Signal-processing systems used with these receivers comprise three groups: analog filter banks (AFB), digital filter banks (DFB), and two coherent dedispersion systems used for the

TABLE I. Summary of datasets. Listed are the corresponding observatory, receiver, and instrument used, together with the center frequency (Ctr freq.), bandwidth (BW), subband bandwidth, the dedispersion method (DD), the integration time per TOA, the data span, the number of TOAs, and the root-mean-square residuals (rms Res).

Observatory	Receiver	Instrument	Ctr freq. (MHz)	Total BW (MHz)	Subband BW (MHz)	DD[a]	T_s (s)	Data span (MJD)	No. TOAs	rms Res. (μs)
Parkes	50-cm	AFB	680	64	16	*I*	30	53054–55371	2588	186.1
Parkes	50-cm	CPSR2	732	64	16	*C*	30	53954–55100	3480	110.6
Parkes	50-cm	DFB3	732	64	16	*I*	30	55042–56748	32 595	143.1
Parkes	50-cm	CASPSR[b]	720	200	12.5	*C*	30	55617–58396	2351	103.8
Parkes	MB	AFB	1390	256	64	*I*	30	52760–55371	111 835	96.9
Parkes	MB	DFB1	1369	256	64	*I*	30	54356–54409	3179	94.5
Parkes	MB	DFB2	1369	256	64	*I*	30	54408–54930	9201	83.2
Parkes	MB	DFB3	1369	256	64	*I*	30	54680–56769	46 853	67.9
Parkes	MB	DFB4[b]	1369	256	64	*I*	30	56829–58667	37 217	70.1
Parkes	H-OH	DFB4	1526	512	64	*I*	30	57444–57673	11 817	79.3
Parkes	UWL	DFB4	1369	256	64	*I*	30	58431–58667	12 231	72.3
Parkes	10-cm	AFB	3030	768	768	*I*	30	52987–54619	4262	94.8
Parkes	10-cm	DFB2	3100	1024	512	*I*	30	54376–54526	690	78.6
Parkes	10-cm	DFB3	3100	1024	512	*I*	30	54618–56070	1120	90.8
Parkes	10-cm	DFB4[b]	3100	1024	512	*I*	30	54755–58396	19 342	109.0
Parkes	UWL	DFB4	3100	1024	512	*I*	30	58264–58619	519	105.0
GBT	820-MHz	GASP[c]	820	64	64	*C*	30	53448–55587	21 531	13.6
GBT	1400-MHz	GASP	1404	64	64	*C*	30	53266–55458	7397	22.0
GBT	820-MHz	GUPPI	820	200	25	*I*	30	55002–55431	27 539	18.8
GBT	820-MHz	GUPPI	820	200	25	*C*	30	55607–58140	255 301	18.2
GBT	1400-MHz	GUPPI	1500	800	100	*I*	30	55217–56257	46 521	26.7
GBT	1400-MHz	GUPPI	1500	800	100	*C*	30	56367–58227	76 173	21.3
Nançay	1.4-GHz	NUPPI	1484	512	64	*C*	30	55818–58211	174 173	60.8
Nançay	2.5-GHz	NUPPI[d]	2520	128	128	*C*	30	56191–57958	2594	144.9
WSRT	300-cm	PuMa-II[d]	334.6	70	8.75	*C*	240	54519–57119	16 110	89.2
Effelsberg	20-cm/7-beam	PSRIX[d]	1360	150/250	15.3/15.9	*C*	160/240	55722–58328	3300	40.1
JBO/Lovell	1.4-GHz	PDFB[d]	1516	384	12	*I*	240	55668–58219	33 780	126.9
All 4-min	...	...	...	...	...	...	240	52760–58667	199 913	13.0
All 30-s	...	...	...	...	...	...	30	52760–58640	916 648	26.0

[a]Dedispersion technique: C = coherent, I = incoherent.
[b]Band reference instrument for the DM modeling.
[c]Band reference instrument for the 30-s TOA analysis.
[d]Datasets not included in the 30-s TOA analysis.

700-MHz band, CPSR [67] and CASPSR [68]. The AFB systems had channel bandwidths of 0.5 MHz for the 700-MHz band and 1 MHz for the higher-frequency bands and utilize 1-bit digitizers. The DFB systems used 8-bit digitizers and implement polyphase filter banks with field programmable gate array processors (FPGAs). Both the AFB and DFB system utilized incoherent dedispersion to sum data across subbands. More details of these signal-processing systems can be found in Manchester *et al.* [69].

B. Green Bank Telescope

Observations at the Green Bank Telescope (GBT) are carried out at 820 and 1400 or 1500 MHz. Typically, 5-h observations were conducted with monthly cadence and at alternating frequencies. Twice a year, we usually conducted concentrated observing campaigns in April-May and October-November using observations at 820 MHz on consecutive days to separate short- from long-term orbital changes. The Green Bank Astronomical Signal Processor (GASP [70]) carried out 8-bit Nyquist sampling of the incoming dual-linear-polarization signal, after which it performs coherent dedispersion in software on a Linux-based cluster for each of several 4-MHz channels. Later observations were carried out with the Green Bank Ultimate Pulsar Processing Instrument (GUPPI [71]). Initially, the 8-bit digitized data streams were detected and dedispersed incoherently, while subsequently a coherent dedispersion mode is available. Coherent dedispersion was implemented later at 1500 MHz than at 820 MHz. GASP and GUPPI were operated in parallel during a transition period in order to carefully calibrate an expected small instrumental time offset between the two systems. For those overlapping epochs, in the final analysis we usually choose the GUPPI dataset, unless calibration issues with GUPPI occur. We note that earlier datasets recorded with the previous SPIGOT and BCPM processors are not

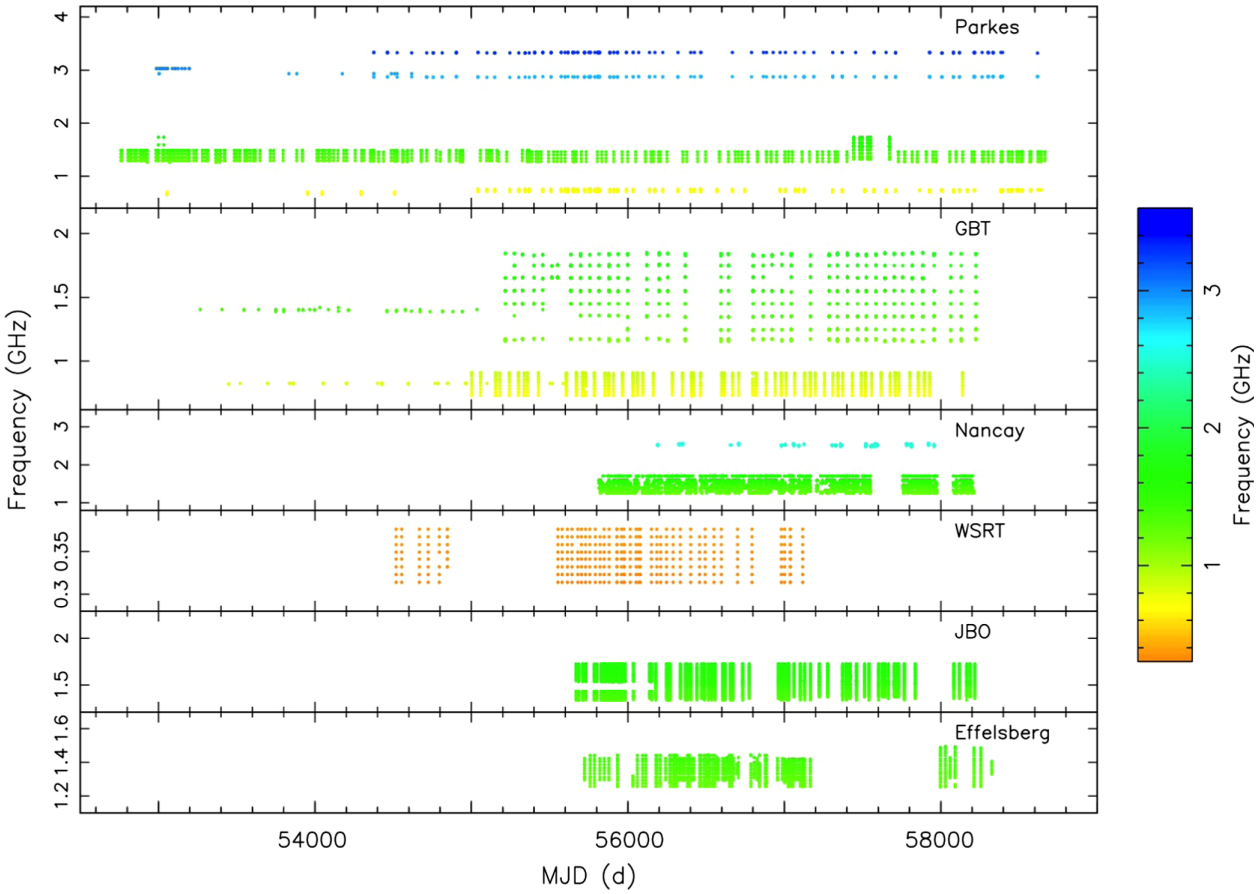

FIG. 2. Time and frequency coverage of the 4-min-sampled TOAs from the contributing observatories.

included in the current analysis (cf. Ref. [31]). A noise diode was used for flux calibration, with observations of the stable continuum source B1442+10 used to calibrate the diode noise power. Flux calibration, interference excision, and subband formation are carried out for GASP using ASPFITSReader [72,73], while for GUPPI they are carried out using PSRCHIVE [74]. The GUPPI interference excision is based on early work carried out by Grandy [75].

C. Nançay Radio Telescope

Our dataset includes the results of high-cadence observations with the Nançay Radio Telescope (NRT), which is a transit telescope of the Kraus-type design with a collecting area equivalent to a 94-m parabolic dish. While observations of the double pulsar commenced late 2004, using an earlier back end, we make use of observations conducted since late 2011 with the Nançay Ultimate Pulsar Processing Instrument (NUPPI). NUPPI is a baseband recording system, similar to GUPPI used at the GBT; see, e.g., Ref. [76] for more details. Data were taken with the 1.4- and 2.5-GHz receivers which have frequency coverage of 1.1–1.8 and 1.7–3.5 GHz, respectively. Most of the observations with the 1.4-GHz receiver were made at a central frequency of 1484 MHz, while those with the 2.5-GHz receiver were generally centered at 2520 MHz. The effective central frequency for a given observation depends on the excision of radio-frequency interference (RFI) which is determined via visual inspection. The data are polarization calibrated using PSRCHIVE's SingleAxis method. Observations of a reference noise diode were made prior to each pulsar observation and the data are calibrated, correcting for differential phase and amplitude between the two polarizations, before being folded into subintegration profiles.

D. Effelsberg 100-m radio telescope

Observations with the 100-m radio telescope at Effelsberg of the Max-Planck-Institut für Radioastronomie were conducted with two different 20-cm receiver systems. For most sessions, the central beam of the seven-beam system with a bandwidth of 250 MHz was used. When this was not available, a single-pixel receiver was used, providing a bandwidth of 150 MHz. All data streams were coherently dedispersed using the ROACH-based PSRIX signal processor (see Ref. [77] for more details). As the double pulsar barely rises above the hills surrounding the telescope, full orbital tracks are not possible, and the system temperature is significantly increased because of spillover. Consequently, longer integration times of 160 and 240 s were used. Standard calibration procedures are applied as described in Ref. [77].

E. Lovell radio telescope

Observations were also conducted with the 76-m Lovell telescope at Jodrell Bank of the University of Manchester. We used a receiver covering a frequency range 1300–1700 MHz, with a maximum usable bandwidth of 400 MHz. Data were acquired using a ROACH-based system that is essentially identical to the Effelsberg PSRIX system (see Ref. [78] for details). Gain and polarization calibration are achieved by monitoring a noise-diode signal and manual adjustments to power levels. Similar to the methods used at other observatories, RFI is excised following visual inspection of the data.

F. Westerbork Synthesis Radio Telescope

The Westerbork Synthesis Radio Telescope (WSRT) consists of 14 25-m dishes arranged in an East-West direction. Observations were made at 334 MHz, the lowest frequency included in our dataset. Dual-polarization signals were acquired with the (nearly real-time) coherently dedispersing PuMaII instrument [79] using a total bandwidth of 70 MHz with 8.75-MHZ-wide subbands. Gain and phase differences between the two polarizations were adjusted during the phased-array calibration of the system (cf. Ref. [80]).

G. Very Long Baseline Array

The VLBA consists of ten 25-m dishes spread across the continental USA, Hawaii, and the Virgin Islands. Dual-polarization observations were taken in standard continuum mode with 256 MHz of observing bandwidth distributed over eight 32-MHz-wide subbands centered at 1.56 GHz. Further details are given in Sec. IV C and Appendix A.

IV. PULSE TIMING AND VLBI DATA ANALYSES

In this section, we describe how the pulse timing observations discussed in Sec. III above are processed to form pulse TOAs, how these TOAs are analyzed to determine the pulsar astrometric parameters and the long-term DM variations, and the results obtained. We also outline the methods used to analyze the VLBA interferometric data and the astrometric results obtained. A key result from these observations is our best estimate of the distance to the double pulsar, an important parameter in the gravitational tests described in Sec. V. Detailed descriptions of the observations, analysis methods, and results are deferred to Appendix A.

For the pulse timing analyses, two sets of TOAs are formed: (a) a set with 30-s spacing that are used for the relativistic-parameter analyses described in Sec. V and (b) a set with 4-min spacing that are used for the timing-based astrometry and the analysis of DM variations described in this section.

Analysis of the timing datasets is carried out in two distinct phases: The first phase uses Tempo2 to derive the pulsar astrometric parameters and the DM variations based on the 4-min-sampled TOAs, whereas the second phase uses Tempo to derive the binary and relativistic parameters based on the 30-s-sampled TOAs. As indicated in Table I, datasets from Westerbork, Effelsberg, and Jodrell Bank observatories are not included in the Tempo analysis, since they have insufficient time resolution to correctly trace the Shapiro delay curve. The Nançay 2.5-GHz dataset is also omitted in the second analysis phase, as the TOA precision is limited and their number is relatively small. For the astrometric and DM analyses, the binary and relativistic parameters are held fixed and the 4-min sampling makes the various Monte Carlo and bootstrap analyses (described in Sec. IV B below) more tractable. The 30-s sampling interval chosen for the binary and relativistic analyses is consistent with that used for the Kramer *et al.* (2006) analysis (Ref. [31]) and represents a good compromise between our ability to resolve orbital effects, signal-to-noise ratio (S/N) considerations (and, hence, TOA uncertainties), and the potential impact of pulse jitter (see, e.g., Ref. [81]).

A. Measurement of TOAs

In general, similar procedures for processing observational data and formation of TOAs are adopted for all telescopes and data acquisition systems. After removal of obvious narrow-band and broadband transient radio-frequency interference (RFI), data are calibrated, in most cases making use of a short observation of an injected noise diode signal that precedes or follows the pulsar observation. The noise diode signal itself is calibrated in flux density units by comparing it to a stable continuum source, normally Hydra A (3C 218) at Parkes and B1442+10 at Green Bank, and in polarization properties by analysis of rise-to-set tracks of the strong millisecond pulsar PSR J0437–4715 at Parkes [82] and the well-studied calibrator PSR B1929+10 at Green Bank. Following calibration, the data are formed into subbands whose width varies according to frequency band as listed in Table I, with wider bandwidths used at higher frequencies (a) because of the reduced DM smearing, (b) to maintain sufficient S/N, and (c) so that the timing program (Tempo or Tempo2) properly accounts for the fact that data at different frequencies received at a given time correspond to different orbital phases at emission because of the differential dispersion delay (see, e.g., Refs. [25,31]). The subbandwidths chosen are sufficiently small to ensure that this effect is not significant.

For the Parkes datasets, TOAs are obtained using the PSRCHIVE routine PAT with the Fourier phase gradient algorithm [23]. Separate pulse profile templates derived are used for the three main receiver bands (50-, 20-, and 10-cm). These are analytic templates aligned in phase and first used in Ref. [31], derived by fitting Gaussian components to pulse profiles that are obtained from averaging a large number ($>100\,000$) of individual pulses.

This helps to avoid "self-mirroring," i.e., using a template based on the same pulse profiles that are to be timed, which leads to potential biases in TOA uncertainties due to correlated noise contributions [31,83]. These same templates are used to derive TOAs for data from the Green Bank GASP instrument and from the WSRT, Effelsberg, and JBO/Lovell telescopes. The wider-bandwidth GUPPI and NUPPI data generally have higher S/N than those from GASP, and use of the analytic templates results in systematics in the residuals that worsen with increasing S/N. For GUPPI and NUPPI, we therefore construct separate templates based on multiple observations aligned and smoothed using ASPFITSReader and PSRCHIVE utilities such as psrsmooth [84].

As mentioned above, 4-min-averaged TOAs are used for the astrometric and DM analyses. The method used to obtain the 4-min TOAs differs for the different observatories. For Parkes, the original 30-s-sampled datasets are directly summed to form 4-min-sampled datasets using standard Psrchive routines and TOAs determined as for the 30-s data. For data from Green Bank and Nançay, the 30-s TOAs are averaged over 4-min intervals using the Tempo2 averageData routine. This routine determines the weighted mean residual and its uncertainty for TOAs in the 4-min sample and applies the necessary pulse phase offset to a central TOA. Data from WSRT, Effelsberg, and JBO/Lovell are used directly.

Because of the short sampling time and the effects of pulse jitter, pulse profiles often vary greatly in shape, e.g., having the interpulse stronger than the main pulse. Consequently, TOAs are sometimes very discrepant in phase. These discrepant TOAs are simply removed from the dataset to produce a "clean" set for each telescope and instrument. Next, we determine scaling factors "EFAC" nd "EQUAD" for each instrument (see, e.g., Ref. [25]). These factors modify the initially derived uncertainty of each TOA, σ_i, according to $\sigma_i' = \mathrm{EFAC} \times \sqrt{\sigma_i^2 + \mathrm{EQUAD}^2}$ such that the reduced χ^2 value obtained after fits of the initial timing model is close to unity. This ensures an appropriate weight for a given dataset in the final combined least-squares fit to our timing model. For well-behaved datasets free of systematic errors or radio interference signals, one expects the TOA uncertainties to reflect the true measurement uncertainties; hence, $\mathrm{EFAC} \sim 1$ and $\mathrm{EQUAD} \sim 0$. For the 30-s dataset, we determine EQUAD values for each observing epoch and exclude those where the resulting EQUADs show large deviation from zero. Then, a single EFAC value is assigned to the each of the datasets listed in Table I. For the 4-min datasets, EFACs and EQUADs are determined for each instrument and each band for each observatory using the efacEquad function of Tempo2. Because of the preliminary "cleaning," the EFACs are generally close to unity and in most cases less than 1.2. EQUADs are in most cases 10 μs or less, although for some systems they are up to 40 μs.

Because of profile frequency evolution and the different templates used for different rf bands, there are systematic offsets between TOAs for different bands. For Parkes data, for each rf band, we determine TOA offsets for each instrument relative to a reference instrument using the Tempo and Tempo2 "JUMPS" facility. The reference instrument for each band is indicated in Table I. In most cases, there is significant overlap between pairs of datasets, which greatly improves the precision of the JUMP determination.

For Green Bank GUPPI data, Nançay 1.4-GHz data, and WSRT data, the fractional bandwidths are much larger, and it is necessary to determine TOA offsets between subbands. In the 4-min TOA analysis, in each case a subband close to the center of the band is chosen as reference. Since these jumps are best determined from individual datasets, they are held fixed in the subsequent analyses. In the 30-s TOA analysis, we determine phase offsets between the analytic templates and the separately derived GUPPI and Nançay templates. We account for these offsets by applying phase shifts to the corresponding TOAs prior to applying the timing model, thereby effectively aligning all templates in phase.

B. Timing astrometry and DM variations

The 4-min-sampled TOAs from all observatories are analyzed using Tempo2 to determine the astrometric parameters for the pulsar and the variations in DM across the total data span. All analyses use the DE436 Solar-System ephemeris [85] to transfer TOAs to the Solar System barycenter (SSB) frame. Studies of the effects of using different Solar-System ephemerides for this transfer (see, e.g., Refs. [86,87]) show differences much below the TOA precision of our datasets.

Initial parameters for the Tempo2 fit are obtained from the Tempo analysis of the 30-s TOAs. Pulsar astrometric parameters (right ascension, declination, proper motions, and annual parallax), the pulsar rotational frequency (ν) and its first time derivative ($\dot{\nu}$), and the secular orbital terms ($\dot{\omega}$ and $\dot{P}_\mathrm{b}$) are fitted. All other binary and relativistic parameters are better determined by the 30-s TOA and, hence, are held fixed and updated in following iterations.

DM and common-mode (i.e., wavelength-independent) variations are fitted for, typically at 100-d intervals in the central part of the dataset and at somewhat longer intervals at the ends where the frequency coverage is poorer, using the methods described by Keith *et al.* [88]. In addition, jumps relative to the Parkes 20-cm dataset are fitted for all observatory bands listed in Table I. DM variations are relative to the reference DM, held at 48.917 208 pc cm^{-3}. The reference solar-wind electron density is set to zero for all fits. Because of the high ecliptic latitude ($-51.2°$), the annual DM variations due to the solar wind are quite smooth, so they are mostly absorbed by the fitted DM offsets. Transient solar events, such as coronal mass ejections, are not absorbed in the fit but are smaller and

short lived. We note that the worst-case solar-wind DM contribution is negligible compared with the errors in estimating DM(t) (see Appendix B). Transient events also have a negligible effect on measured TOAs. Although TOA offsets between observatory bands resulting from DM variations are absorbed by the jumps, variations within each dataset are preserved, and these determine the derived DM variations. Higher-order rotational-frequency derivatives up to the fourth ($\ddot{\nu}$) are held fixed at values determined in the Tempo analysis of the 30-s dataset because of covariances with the common-mode variations. This whole process is iterated as the system parameters are refined.

Astrometric parameters and DM and common-mode variations are initially determined by iterated fits to the entire dataset for the default 100-d sampling of the DM and common-mode variations. Because of the known covariance between astrometric parameters and the sampling of the DM and common-mode variations, we then undertake a Monte Carlo analysis, randomly varying the spacing and phase of DM and common-mode sampling within defined limits, normally 64–256 d, for the spacing and for phase within ± 0.5 of the current spacing. After 4096 iterations, the distribution for each parameter is fitted by a Gaussian function to give estimates of the parameter and its 1σ uncertainty.

To test for systematic effects related to data from a particular observatory, the Monte Carlo analysis is repeated on datasets that (a) exclude the GBT data, (b) exclude the Parkes data, and (c) contain just the GBT and Parkes data. These tests have 2048 iterations, and they are analyzed in the same way as the full dataset.

As an additional check on the reliability of the parameter uncertainties, we perform a bootstrap analysis (with replacement) of the full dataset using the 100-d DM and common-mode sampling. In a similar manner to the Monte Carlo analysis, a Gaussian function is fitted to the parameter distributions after 2048 iterations.

At timescales longer than a few months, steep-spectrum or "red" noise in the intrinsic pulsar spin rate which has a power spectrum of the form f^{α} (where $\alpha \ll 0$) becomes important. In the initial analysis, much of this red noise is removed by fitting for the first four pulse frequency derivatives. However, proper analysis of timing datasets in the presence of red noise requires the estimation of a noise covariance matrix and the use of a generalized least-squares (GLSQ) approach when fitting the timing model [89]. The DM and common-mode analysis does not require GLSQ, because the common mode absorbs all the red noise and the residuals are "white," i.e., flat spectrum. In this case, a weighted least-squares analysis is sufficient. For the GLSQ analysis, the covariance matrix has N_{TOA}^2 elements, and it is not feasible to perform a full Monte Carlo analysis. We choose to estimate the effects of the GLSQ analysis compared to the approximate method described above by forming smaller averaged datasets with 16- and 32-min sampling using the Tempo2 averageData routine. Even for these smaller datasets, it is necessary to modify the indexing of the covariance matrix and eliminate unnecessary matrix inversions in Tempo2. The best-fit value of α is -3.0, but changes of ± 0.3 make very little difference in the parameters or their uncertainties. The results show that the initial analysis underestimates the parallax uncertainty by about 75% but does not significantly bias the value. For the other astrometric parameters, the increase in uncertainty is less than 50% (zero for the position terms), but there are some biases, in all cases less than 50% of the combined uncertainty.

Results for the astrometric parameters from the Monte Carlo analysis of the full 4-min-sampled dataset modified to reflect the biases and uncertainty changes indicated by the GLSQ analyses of the shorter datasets are listed in the second column in Table II. [90] The final three columns in Table II list Monte Carlo results from the partial datasets. All results are within or close to twice the combined uncertainty of the final result, with the parameters from the GBT-omitted dataset differing the most. This shows that the GBT dataset with its high S/N has the greatest influence on the final results. It is notable, though, that the measured parallax is not significantly biased by the dataset from any one observatory.

Results from the bootstrap analysis are given in the third column in Table II. Uncertainties from the bootstrap

TABLE II. Astrometric results from analysis of the 4-min timing datasets [see the text for details of the different datasets and analysis procedures, i.e., bootstrap (boot.) and Monte Carlo (MC)]. Numbers in parentheses are 1σ uncertainties relative to the last quoted digit of the parameter value. For the right ascension, we show only the seconds of $07^{\mathrm{h}}37^{\mathrm{m}}ss.s^{\mathrm{s}}$ and for the declination only the arcseconds of $-30°39'ss.s''$.

Dataset	All data	All data Boot.	All ex. GBT MC	All ex. Pks MC	Pks + GBT MC
Right ascension, α (J2000) (s)	51.248 115(10)	51.248 111(6)	51.248 125(10)	51.248 105(10)	51.248 105(9)
Declination, δ (J2000) ($''$)	40.704 85(17)	40.704 96(11)	40.705 44(21)	40.705 29(23)	40.705 25(18)
Proper motion R.A., μ_α (mas yr^{-1})	−2.567(30)	−2.570(16)	−2.618(15)	−2.581(16)	−2.582(16)
Proper motion Dec., μ_δ (mas yr^{-1})	2.082(38)	2.081(19)	2.178(35)	2.102(35)	2.105(25)
Parallax, π_t (mas)	2.15(48)	2.40(11)	2.41(10)	2.01(22)	2.22(24)
Position epoch (MJD)	55045	55045	55045	55045	55045
Data span (MJD)	52760–58667	52760–58667	52760–58667	53266–58328	52760–58667
N_{TOA}	199 913	199 913	129 787	151 883	118 156

analysis are typically about half those from the Monte Carlo (MC) analysis. The mean values are also offset, although generally within the combined uncertainties. Both of these differences result from the covariance of the astrometric parameters and the sampling of the DM and common-mode variations.

Figure 3 shows the DM variations derived by averaging the Monte Carlo DM-offset results for the 4-min-sampled dataset in 100-d bins across the data span. As noted above, the DM results are unaffected by the GLSQ analysis. Apart from a few points with large uncertainties near the ends of the data span where there is less radio-frequency coverage, the DM shows fluctuations on timescales of a few hundred days with amplitude of up to 0.0015 pc cm^{-3} and a downward trend of approximately -1.41×10^{-4} pc cm^{-3} yr^{-1} over the data span.

For the final Tempo analysis of the 30-s TOAs, the astrometric parameters and DM variations from the final Tempo2 analysis of the 4-min data are applied using a method described in Sec. VI A. The final analysis takes into account the DM variations, which we can describe by fitting a smoothly varying curve that reflects the uncertainties of the DM measurements. This is achieved using a Gaussian learning process as implemented in the PYTHON library George [91]. We explored a number of stationary kernels with very similar results. We finally adopted the Matern32Kernel leading to the modeling of the DM offsets as shown in Fig. 3. Each TOA is modified with a DM offset dithered according to the interpolated DM uncertainty at that epoch.

The observed DM variations imply the presence of transverse phase gradients which will shift the apparent position of the pulsar on a timescale of months. These shifts can be estimated from the measured intensity scintillation

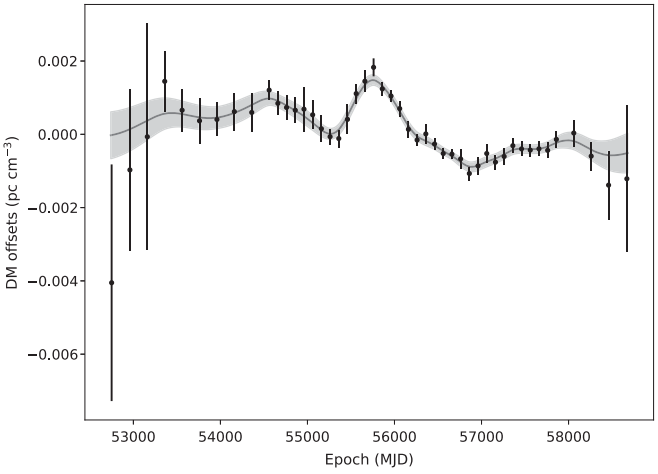

FIG. 3. Variations in DM relative to the reference value, 48.917 208 pc cm^{-3}, derived by averaging over values from the Monte Carlo analysis at 100-d intervals (points with error bars). The solid line is the result of using a Gaussian learning process to describe the data with resulting uncertainties indicated by the gray band.

[92]. They are of the same order as the parallax uncertainty and will affect both timing and VLBI astrometry, which we present below in Sec. IV C. We discuss these effects in Sec. VIII A. We believe that the interstellar electron density fluctuations responsible for both the DM variations and the scintillation occur in or close to the Gum Nebula, as discussed in Sec. VIII B. To avoid the poorly sampled data at the ends of the full data span, this latter discussion is based on a restricted dataset between MJDs of 53500 and 58300.

C. VLBI astrometry

Astrometric terms in the pulsar timing model can be compared to independent estimates for these parameters obtained via radio imaging. PSR J0737–3039A/B was previously targeted by an astrometric program using the Australian Long Baseline Array [93], which measured an annual geometric parallax of 0.87 ± 0.14 mas. This measurement is in tension with the timing parallax presented in Table II. In an attempt to resolve this, we undertook an extended astrometric program with the VLBA. The details of these observations, data reduction, and robustness testing using in-beam reference sources are given in Appendix A, and the resultant astrometric parameters are summarized in Table III.

For this analysis, the proper motion measured by pulsar timing (and listed in Table II) is applied as a prior to the VLBI fit. While some previous studies highlight substantial disagreement between timing and VLBI proper motions (e.g., Ref. [94]), the uncertainty in the timing proper motion for PSR J0737–3039A/B is sufficiently small that even an error exceeding those seen previously for other millisecond pulsars would not materially affect the VLBI parallax uncertainty. Details of the effects of proper motion fitting are included in Appendix A. The timing and VLBI reference positions are consistent when propagated to the same reference epoch, although the precision of the VLBI reference position for PSR J0737–3039A/B is relatively low, for the reasons discussed in Appendix A.

The quantity of most interest, the parallax, lies between the previous VLBI result and the current timing value but is more precise than either. We discuss the difference between the distance implied by a straightforward inversion

TABLE III. VLBA astrometric results. Listed are right ascension (R.A.), declination (Dec.), proper motion in both coordinates, parallax and position epoch.

R.A., α (J2000)	$07^{h}37^{m}51^{s}.247(1)$
Dec, δ (J2000)	$-30°39'40''.68(1)$
Parallax, π_v (mas)	$1.30^{+0.13}_{-0.11}$
Proper motion in R.A., μ_α (mas yr^{-1})[a]	-2.567 ± 0.030
Proper motion in Decl., μ_δ (mas yr^{-1})[a]	2.082 ± 0.038
Position epoch (MJD)	58000

[a]A prior based on the timing proper motion was applied as described in Appendix A.

of the VLBA parallax (770 ± 70 pc) and the timing parallax 465^{+134}_{-85} pc in Sec. VIII A. For the reasons discussed in Appendix A, we adopt the weighted mean of our new VLBA parallax and the timing parallax as our best estimate of the system parallax and use the resultant distance (735 ± 60 pc) and transverse velocity (11.5 ± 1.0 km s^{-1}) for the remainder of the paper. We stress that the precision of the distance measurement does not (yet) affect the gravitational tests presented below.

V. BINARY AND RELATIVISTIC-PARAMETER ANALYSIS METHODS

In this section, we describe the methods used to derive the relativistic terms that form the basis of the tests of gravitational theories, the key results of this paper, from the observational data and results described in Secs. III and IV. The astrometric, pulsar spin, binary, and PK relativistic parameters, derived using a Monte Carlo analysis to allow for the uncertainties in the derived DM variations (Sec. IV), are presented in this section. The results of the gravitational tests are described in the following Sec. VI for tests of GR and in Sec. VII for tests of alternative gravity theories.

A timing model describing the rotational properties of pulsar A, its position and apparent movement on the sky, the propagation delays in the ISM, and the orbital motion and relativistic effects that cause deviations from an arrival time expected for a simple Keplerian orbit is applied to TOA measurements for pulsar A. Before the timing model can be applied, the measured topocentric arrival times need to be transferred to the SSB to account for the varying position of our telescopes during the course of the year and for relativistic effects in the Solar System [4]. Before we present the results of such a procedure, we explain the required relativistic binary timing model. We do this in some detail, as we introduce a number of new considerations, which eventually also become relevant for other systems in the future.

The barycentric arrival time $t_{\rm b}$ of a pulsar signal is related to the proper time T of the pulsar (modulo a constant factor) by a phenomenological timing model that accounts for all relevant contributions related to the orbital motion of the pulsar and the signal propagation in and near the binary system [29,35]:

$$t_{\rm b} - t_0 = \frac{1}{D}[T + \Delta_{\rm R}(T) + \Delta_{\rm E}(T) + \Delta_{\rm S}(T) + \Delta_{\rm A}(T)]. \quad (1)$$

The time t_0 denotes a (chosen) reference epoch. The different Δ_a ($a = {\rm R, E, S, A}$) on the right-hand side denote, respectively, the *Rømer delay* resulting (solely) from the orbital motion of the pulsar, the *Einstein delay* caused by time dilation along the pulsar's worldline, the *signal propagation delay* due to relativistic contributions to the propagation of the radio signal, and the *aberration delay* due to the fact that the radio signals originate from a rotating moving source [18,29,95]. These contributions depend on the proper time T and a set of Keplerian and PK parameters, which we explain in detail below. D is the Doppler factor due to the velocity $\mathbf{v}_{\rm b}$ of the pulsar system with respect to the SSB [29]:

$$D = \frac{1 - \hat{\mathbf{K}}_0 \cdot \mathbf{v}_{\rm b}/c}{\sqrt{1 - \mathbf{v}_{\rm b}^2/c^2}}, \quad (2)$$

which to leading order is dominated by the unknown radial velocity $v_R = \hat{\mathbf{K}}_0 \cdot \mathbf{v}_{\rm b}$ (>0 if the Earth-pulsar distance increases). $\hat{\mathbf{K}}_0$ is the line-of-sight (LOS) unit vector pointing from the SSB to the binary system. Only temporal changes of D are of interest here, and, therefore, $D \equiv 1$ can be chosen at a given epoch (see discussion in Ref. [29]). The proper time of the pulsar is related to the rotational phase ϕ of the pulsar by

$$\frac{\phi(T)}{2\pi} = N_0 + \nu(T - t_0) + \frac{1}{2}\dot{\nu}(T - t_0)^2 + \frac{1}{6}\ddot{\nu}(T - t_0)^3 + \cdots, \quad (3)$$

where N_0 is the pulse number at the reference epoch t_0. In the following, we discuss the individual contributions Δ_a and all significant effects that need to be accounted for in those contributions when analyzing the timing data presented here. In terms of theory-specific interpretation of the different Δ_a and their PK parameters, we primarily focus on GR. As it turns out, for several contributions we need to account for next-to-leading-order (NLO) terms in order to achieve a correct interpretation of our timing data. In the following sections, we therefore discuss the individual contributions to the SSB arrival times in detail.

A. Rømer delay

In terms of the orbital motion, pulsar timing analysis of binary pulsars is based on a particularly compact solution of the first post-Newtonian (PN) two-body problem (Damour-Deruelle solution [28]). As it turns out, this quasi-Keplerian parametrization of the orbital motion is valid for a wide class of alternative theories, in particular, the class of (massless) scalar-tensor theories [29,35]. The *Rømer delay*, $\Delta_{\rm R}$ in Eq. (1), related to the orbital motion of the pulsar, is given by [29]

$$\Delta_{\rm R} = x\left[\sin\omega(\cos u - e_r) + \sqrt{1 - e_\theta^2}\cos\omega\sin u\right], \quad (4)$$

where $x = a_p \sin i/c$ is the projected semimajor axis of the pulsar orbit. The angle ω is the longitude of periastron (measured from the ascending node [96]), and u is the "relativistic" *eccentric anomaly*, related to the proper time of the pulsar via Kepler's equation

$$u - e_T \sin u = 2\pi \left[\left(\frac{T - T_0}{P_b} \right) - \frac{\dot{P}_b}{2} \left(\frac{T - T_0}{P_b} \right)^2 \right]. \quad (5)$$

The $\dot{P}_b$ term accounts for any secular (intrinsic or apparent) change in the orbital period P_b. It therefore accounts for the decay in the orbital period due to GW damping, which enters the GR equations of motion at the 2.5PN level, i.e., the order of v^5/c^5. The parameter T_0 denotes a time of periastron passage. As becomes clear below, for the double pulsar it is important to distinguish between the three different eccentricities e_r, e_θ, and e_T that enter the two equations (4) and (5) and are different at their first PN correction [29]:

$$e_r \equiv e_T(1 + \delta_r), \qquad e_\theta \equiv e_T(1 + \delta_\theta). \quad (6)$$

While the parameter δ_r is not observable in the double pulsar (it is absorbed by higher-order frequency derivatives), the "relativistic deformation of the orbit" δ_θ has to be accounted for in the analysis (Sec. VI B 4). The "time eccentricity" e_T corresponds to the fitting parameter called *"eccentricity of orbit"* in the Tempo and Tempo2 implementation of the Damour-Deruelle (DD) timing model [24,98] and, therefore, can be considered as the observed eccentricity in pulsar timing, a Keplerian parameter.

The longitude of periastron ω that enters the Rømer delay (4) changes over time, due to the relativistic precession of the apsidal line of the double pulsar. In the quasi-Keplerian parametrization of the orbital motion, the longitude of periastron is simply related to the "relativistic" *true anomaly* θ [28], i.e.,

$$\omega = \omega_0 + k\theta, \quad (7)$$

$$\theta = 2 \arctan \left[\sqrt{\frac{1 + e_\theta}{1 - e_\theta}} \tan \left(\frac{u}{2} \right) \right]. \quad (8)$$

By adding or subtracting multiples of 2π, the angle θ is adjusted such that it matches the number of full phase cycles of u, as determined by Eq. (5). The structure of the Rømer delay presented above is expected to be the same for many viable relativistic theories of gravity. What changes is how the PK parameters (k, $\dot{P}_b$, δ_r, and δ_θ) depend on the Keplerian parameters (P_b, x, and e_T) and the inertial masses m_A and m_B of A and B, respectively. Detailed expressions for GR can be found in Refs. [4,29], where one also finds expressions for the PK parameters which we introduce below, in combination with the Einstein and Shapiro delays in the signal propagation.

Instead of the PK parameter k, in pulsar timing one generally quotes the value for the (secular) *periastron advance* parameter defined as $\dot{\omega} \equiv n_b k$, where $n_b \equiv 2\pi/P_b$ is the measured orbital frequency [cf. Eq. (5.8) in Ref. [36]]. Given the high numerical precision of n_b normally obtained in pulsar timing observations, $\dot{\omega}$ and k can be used synonymously. For that reason, whenever comparison with former publications is required, we use $\dot{\omega}$ for convenience.

1. Second post-Newtonian corrections

The high precision achieved for several PK parameters raises the question of whether higher-order corrections need to be taken into account. In terms of the Rømer delay, 2PN corrections to the quasi-Keplerian parametrization are calculated in Refs. [36,99–101]. The result is a generalized quasi-Keplerian description of the orbital motion, in which Kepler's equation and the equation for the true anomaly are augmented by periodic terms of the order of v^4/c^4. These 2PN periodic terms, however, are well below the timing precision so far achieved in the double pulsar system and can be ignored. This also means that we can replace e_θ with e_T in Eq. (8) when calculating the longitude of periastron ω via Eq. (7).

Apart from these additional periodic terms, parameters that are already present in the 1PN solution get extended by 2PN corrections. Of particular importance for the analysis presented here is the secular advance of periastron. The (intrinsic) advance of periastron to 2PN approximation was first calculated within GR in Refs. [36,102]. When expressed as a function of the masses and the (directly observed) Keplerian parameters, it is given by

$$k = k^{1PN} + k^{2PN} + \mathcal{O}(\beta_O^6), \quad (9)$$

where

$$k^{1PN} = \frac{3\beta_O^2}{1 - e_T^2}, \qquad k^{2PN} = \frac{3 f_O \beta_O^4}{1 - e_T^2}, \quad (10)$$

and

$$\beta_O = \frac{(GMn_b)^{1/3}}{c}, \quad (11)$$

$$f_O = \frac{1}{1 - e_T^2} \left(\frac{39}{4} X_A^2 + \frac{27}{4} X_B^2 + 15 X_A X_B \right) - \left(\frac{13}{4} X_A^2 + \frac{1}{4} X_B^2 + \frac{13}{3} X_A X_B \right), \quad (12)$$

with $M \equiv m_A + m_B$, $X_A \equiv m_A/M$, and $X_B \equiv m_B/M$ [cf. Eq. (5.18) in Ref. [36]]. Note that $X_A + X_B = 1$. With the numbers from the timing parameter (Table IV), one finds $\dot{\omega}^{2PN} \equiv n_b k^{2PN} \simeq 4.39 \times 10^{-4}$ deg yr^{-1}, which is about 35 times the measurement error of $\dot{\omega}$, quoted in Table IV. More importantly, other PK parameters, in particular, s and $\dot{P}_b$, have now reached a precision that, for the first time in a binary pulsar, we have to account for k^{2PN} in order to perform, for instance, consistent gravity tests.

The 3PN corrections to k are calculated in Refs. [106,107]. They are of the order of $\beta_O^2 \approx 4 \times 10^{-6}$ times the 2PN terms and, therefore, absolutely negligible.

2. Spin-orbit contribution and equation of state

So far, we have ignored the influence of spin, i.e., the proper rotations of A and B, on the orbital dynamics. Several effects related to spin can lead to significant modifications of the equations of motion of a binary system, in particular, the relativistic spin-orbit and spin-spin coupling and the rotationally induced mass quadrupole moments [37]. For the analysis in this paper, only the coupling between the spin of A and the orbital motion is of any relevance (the spin of B is about 135 times smaller). It is numerically of 2PN order [33,108,109], which is typical for binary pulsars in relativistic double-NS systems [101]. Furthermore, since the spin of A is practically parallel to the orbital angular momentum (see Sec. II), there is only a contribution to the precession of the periastron that is of relevance here. We refer the reader to Refs. [33,110] for more details.

In order to incorporate spin-orbit coupling in our analysis, Eq. (9) needs to be extended by the Lense-Thirring (LT) term, i.e.,

$$k = k^{1PN} + k^{2PN} + k^{LT,A}, \tag{13}$$

where within GR the LT contribution is given by [36,37,111]

$$k^{LT,A} = -\frac{3\beta_O^3 \beta_{S_A}}{1-e_T^2} g^{\parallel}_{S_A}, \tag{14}$$

with

$$\beta_{S_A} = 2\pi\nu \frac{cI_A}{Gm_A^2}, \tag{15}$$

$$g^{\parallel}_{S_A} = \frac{X_A(1+\frac{1}{3}X_A)}{\sqrt{1-e_T^2}}. \tag{16}$$

Apart from the MOI I_A, all quantities in the above equations are known with high precision. I_A depends on the EOS for NS matter, which is still afflicted by considerable uncertainty. Consequently, there is a range in the prediction for $k^{LT,A}$. From Eq. (14), one finds the numerical expression

$$\dot{\omega}^{LT,A} \equiv n_b k^{LT,A} \simeq -3.77 \times 10^{-4} \times I_A^{(45)} \text{ deg yr}^{-1}, \tag{17}$$

where $I_A^{(45)} \equiv I_A/(10^{45}\ \text{g cm}^2)$. Using the multimessenger constraints on the radius that can be inferred from Ref. [112] (probability distribution function of case F in their Fig. 1) [113] in combination with the radius-MOI relation for A given in Ref. [114], we find a range of $I_A^{(45)} \approx 1.15$–1.48 (95% confidence). [115] Alternatively, Eq. (17) can be used to infer limits for the MOI of A purely from the timing observations of the double pulsar, if combined with two other suitable PK parameters [33,110,119]. A corresponding analysis is given in Sec. VI B 3.

3. Proper motion contributions

The proper motion of a binary pulsar leads to a change in its orientation with respect to the observer on Earth. Such a change leads to an apparent change in the longitude of periastron ω and the orbital inclination i [120,121]. The change in ω leads to a proper motion related offset $\dot{\omega}^{pm}$ between the intrinsic and the observed advance of periastron. Using the proper motion and orbital inclination from Table IV, in combination with the longitude of the ascending node obtained from scintillation measurements [92], one finds $\dot{\omega}^{pm} \approx -4 \times 10^{-7}$ deg yr^{-1} (see also Ref. [110]). This is about a factor of 30 smaller than the current measurement error for $\dot{\omega}$ (see Table IV) and can, therefore, be ignored. As a consequence, there is no need to distinguish between the observed and the intrinsic $\dot{\omega}$.

The change in the orbital inclination enters the timing model through a temporal change in the projected semimajor axis of the pulsar orbit, showing up as a $\dot{x}$ in the timing solution, if significant. However, this contribution is even smaller than the contribution to the advance of periastron, since it is greatly suppressed by the fact that i is close to 90° ($\dot{x}^{pm} \propto \cot i \approx 0.01$; see Table IV).

4. Next-to-leading-order contributions in the mass function

The inclination of the binary orbit is linked to the projected semimajor axis x in Eq. (4) via the binary mass function. In Newtonian gravity, one finds (see, e.g., Refs. [4,35])

$$\sin i = \frac{n_b x}{\beta_O X_B}, \tag{18}$$

where n_b and x are both (observable) Keplerian parameters, generally known to very high precision for a binary pulsar. As we discuss later (Sec. V C), the measurement of the Shapiro delay in the double pulsar gives access to $\sin i$, and, therefore, Eq. (18) leads to an additional constraint for the two masses m_A and m_B.

In the 1PN approximation, Kepler's third law, which enters the derivation of Eq. (18), gets modified by an additional term [see Eq. (3.7) in Ref. [28] and Eq. (3.7) in Ref. [35]]. Consequently, Eq. (18) gets modified as well at the 1PN level. Using the 1PN expression for Kepler's third law, one finds

$$\sin i = \frac{n_b x}{\beta_O X_B}\left[1 + \left(3 - \frac{1}{3}X_A X_B\right)\beta_O^2\right]. \tag{19}$$

We use the fact that for the Damour-Deruelle solution the Newtonian relation between the semimajor axis of the pulsar orbit and the semimajor axis of the relative orbit also holds at the 1PN level, i.e., $a_A = (m_B/M)a_R + \mathcal{O}(v^4/c^4)$

TABLE IV. Timing parameters for PSR J0737–3039A in TDB units (see the text). Except for astrometry and DM, the parameters are derived using Tempo with the 30-s TOA dataset. Numbers in parentheses are 1σ uncertainties referred to the last quoted digit. The overall reduced χ^2 is 0.97.

Parameter	Value
Right ascension (R.A.), α (J2000)	$07^h37^m51^s.248115(10)$[a]
Declination (Dec), δ (J2000)	$-30°39'40''.70485(17)$[a]
Proper motion R.A., μ_α (mas yr^{-1})	$-2.567(30)$[a]
Proper motion Dec., μ_δ (mas yr^{-1})	$2.082(38)$[a]
Parallax, π_c (mas)	$1.36(+0.12,-0.10)$[a]
Position epoch (MJD)	55045.0000
Rotational frequency (freq.), ν (Hz)	44.054 068 641 962 81(17)[b]
First freq. derivative, $\dot{\nu}$ (Hz s^{-1})	$-3.4158071(11)\times10^{-15}$[b]
Second freq. derivative, $\ddot{\nu}$ (Hz s^{-2})	$-2.286(29)\times10^{-27}$[b]
Third freq. derivative, $\dddot{\nu}$ (Hz s^{-3})	$1.28(26)\times10^{-36}$[b]
Fourth freq. derivative, $\ddddot{\nu}$ (Hz s^{-4})	$4.580(86)\times10^{-43}$ [b]
Timing epoch, t_0 (MJD)	55700.0
Profile evolution, FD parameter c_1	0.0000180(75)
Profile evolution, FD parameter c_2	−0.0001034(10)
Profile evolution, FD parameter c_3	0.0000474(26)
Dispersion measure, DM (pc cm^{-3})	48.917 208
Orbital period, P_b (day)	0.102 251 559 297 3(10)
Projected semimajor axis, x (s)	1.415 028 603(92)
Eccentricity (Kepler equation), e_T	0.087 777 023(61)
Epoch of periastron, T_0 (MJD)	55 700.233 017 540(13)
Longitude of periastron, ω_0 (deg)	204.753 686(47)
Periastron advance, $\dot{\omega}$ (deg yr^{-1}) [c]	16.899 323(13)
Change of orbital period, $\dot{P}_b$	$-1.247920(78)\times10^{-12}$
Einstein delay amplitude, γ_E (ms)	0.384 045(94)
Logarithmic Shapiro shape, z_s	9.65(15)
Range of Shapiro delay, r (μs)	6.162(21)
NLO factor for signal prop., q_{NLO}	1.15(13)
Relativistic deformation of orbit, δ_θ	$13(13)\times10^{-6}$
Change of proj. semimajor axis, $\dot{x}$	$8(7)\times10^{-16}$
Change of eccentricity, $\dot{e}_T$ (s^{-1})	$3(6)\times10^{-16}$
Derived parameters	
$\sin i = 1-\exp(-z_s)$	$0.999936(+9/-10)$
Orbital inclination, i (deg)	89.35(5) or 90.65(5)
Total mass, M ($M_\odot$)[d]	$2.587052(+9/-7)$
Mass of pulsar A, m_A ($M_\odot$)[d]	$1.338185(+12/-14)$
Mass of pulsar B, m_B ($M_\odot$)[d]	$1.248868(+13/-11)$
Galactic longitude, l (deg)	245.2357
Galactic latitude, b (deg)	−4.5049
Proper motion in l, μ_l (mas yr^{-1})	−3.066(35)
Proper motion in b, μ_b (mas yr^{-1})	−1.233(31)
Distance from π_c, d (pc)	735(60)
Transverse velocity, v_T (km s^{-1})	11.5(10)

[a]See Secs. IV B and IV C for the derivation of these values.
[b]See Ref. [103].
[c]$\dot{\omega}\equiv 2\pi k/P_b$. k is the PK timing parameter in Eq. (7).
[d]See Ref. [104].

(see Refs. [28,29] and the discussion in Ref. [63]). With P_b, x, and the masses from Table IV, one finds for the 1PN correction in Eq. (19) approximately 1.3×10^{-5}, which is only about a factor of 1.3 larger than the error for $\sin i$ in Table IV. For that reason, we use the full 1PN mass function (19). This is the first time that 1PN corrections to the mass function become relevant for any binary pulsar.

5. Secular changes in orbital period

The observed change in the orbital period is a combination of effects intrinsic to the system and apparent changes related to a temporal change in the Doppler factor D in Eq. (1) [122]. For the double pulsar, by far the dominant contribution to $\dot{P}_b$ is the orbital period decay due to the emission of GWs. In GR, GW damping enters at the 2.5PN level in the equations of motion (see Ref. [123] and references therein). The explicit expression for the leading-order changes due to GW emissions for an eccentric orbit are worked out in Ref. [124] (see Ref. [4] for a more pulsar-astronomy-adapted expression). The Hulse-Taylor pulsar is the first binary system where the leading-order GW damping has been tested [2,30].

The NLO correction to the change in the orbital period corresponds to the 3.5PN terms in the equations of motion and, hence, to the 1PN corrections in the radiation reaction force [125–127]. It is calculated in Ref. [128]. For the double pulsar, this contribution amounts to about -1.75×10^{-17} [110]. This is about a factor of 4.5 smaller than the error in $\dot{P}_b$ (see Table IV). Although this higher-order correction is, in principle, still negligible, we include it in our analysis. This is of particular interest for the comparison with the LIGO/Virgo results in Sec. VI B 2. In the near future, however, that contribution will become of importance (see Ref. [110]).

Yet another intrinsic effect that changes the orbital period in the double pulsar is the mass loss related to the spin-down of the pulsars [122]. This mass loss is a result of Einstein's energy-mass equivalence in the sense that here one is seeing the loss of mass associated with the loss of rotational (kinetic) energy of the pulsar. In Ref. [110], these contributions are calculated based on Eq. (4.1) in Ref. [122]. While for B this is negligible, for A one has [110]

$$\dot{P}_b^{\dot{m}_A} = 2.3\times10^{-17}\times I_A^{(45)}. \tag{20}$$

For two reasons, it is important to include this contribution in the analysis below. First, given the range for $I_A^{(45)}$ (see the end of Sec. V A 2), $\dot{P}_b^{\dot{m}_A}$ can be as large as 3.1×10^{-17}, which is a fair fraction of the measurement error of $\dot{P}_b$ (see Table IV). Second, and more importantly, when estimating a MOI constraint based solely on the double pulsar

observations, $\dot{P}_{\rm b}^{\dot{m}_{\rm A}}$ and its dependence on $I_{\rm A}$ has a significant impact on the result (see discussion in Ref. [110]).

Finally, there are the external contributions to the observed $\dot{P}_{\rm b}$, related to a temporal change of D, where $\dot{P}_{\rm b}/P_{\rm b} = \dot{D}/D$ [122]. There is an apparent radial acceleration due to the transverse motion of the double pulsar with respect to the SSB, which leads to the Shklovskii effect [129]:

$$\dot{P}_{\rm b}^{\rm Shk} = \frac{(\mu_\alpha^2 + \mu_\delta^2)d}{c} P_{\rm b}, \tag{21}$$

and a physical radial acceleration due to the Galactic gravitational potential, leading to a $\dot{P}_{\rm b}^{\rm Gal}$. Both effects depend on the distance d, which is the main source of uncertainty in those corrections (see Sec. VIII A below). For the Galactic potential, we use the model provided in Ref. [130]. We also test the analytic correction of Ref. [122] (adapted for $b < 0$) with the latest Galactic parameters from Ref. [131]. The result is practically the same. Numerical details for these contributions are given in Sec. VI B 2.

B. Einstein delay

The Einstein delay $\Delta_{\rm E}$ in Eq. (1) links the proper time of the pulsar T to the coordinate time of the binary system t. To leading order, it can be viewed as a combination of the second-order Doppler effect and the gravitational redshift caused by the companion. If T is renormalized such that the orbital period is the same as measured in t, one finds to leading order the following expression [29,95]:

$$\Delta_{\rm E} = \gamma_{\rm E} \sin u. \tag{22}$$

The *amplitude of the Einstein delay* $\gamma_{\rm E}$ is a PK parameter that in GR can be calculated according to

$$\gamma_{\rm E} = e \frac{P_{\rm b}}{2\pi} X_{\rm B}(1 + X_{\rm B})\beta_{\rm O}^2. \tag{23}$$

In alternative theories of gravity, $\gamma_{\rm E}$ is generally more complicated, since it also includes a periodic variation of the pulsar rotation due to a variation of the pulsar's MOI along the orbit [42,132,133]. The variation of the MOI is caused by a variation of the local gravitational constant along the (eccentric) orbit, a result of a violation of the SEP [41].

NLO contributions are calculated for GR in Refs. [101,134]. Similarly to the periodic 2PN terms in the orbital motion (see Sec. V A 1), these contributions are currently absolutely negligible. They are of the order of 10 ns, which is an order of magnitude below the precision for $\gamma_{\rm E}$ (see Table IV). They are not expected to be of relevance before the full Square Kilometre Array [135] is in operation.

C. Signal propagation delay

For binary pulsars which are sufficiently edge-on, i.e., i close to 90°, the curved spacetime of the companion star has a significant effect on the propagation of the pulsar's radio signal. To leading order, we have the well-known Shapiro delay [136], which for binary pulsars is expressed in the following form [29,95]:

$$\Delta_{\rm S}^{\rm (LO)} = -2r \ln \Lambda_u, \tag{24}$$

$$\Lambda_u = 1 - e_T \cos u - s[\sin\omega(\cos u - e_T) + (1 - e_T^2)^{1/2} \cos\omega \sin u], \tag{25}$$

with the two PK parameters r and s, called the *range* and the *shape* of the Shapiro delay, respectively. The Shapiro shape can quite generally be identified with the sine of the orbital inclination, i.e., $s = \sin i$ [35,137]. The range of the Shapiro delay is linked to the mass of the companion. In GR, one finds $r = Gm_{\rm B}/c^3$ [29,95].

The leading-order expression for the Shapiro delay in a binary pulsar is obtained in Ref. [95], by integrating along a straight line (in an isotropic coordinate system) and by assuming a static mass distribution while the signal propagates away from the pulsar (static limit). Relaxing the first assumption and accounting for the fact that the radio signal propagates along a curved path due to the deflection in the gravitational field of pulsar B leads to a modification of Eq. (24) [138,139]. This *lensing correction* to the Shapiro delay, given the current timing precision in the double pulsar, is, however, not yet observable due to a strong covariance with s. We test this against the real data as well as in mock data simulations [140] and find a shift in s which is yet insignificant (less than 0.5σ). One has to keep in mind that these effects are strongest around conjunction, in a part of the orbit where the timing precision is significantly reduced due to the (partly intermittent) eclipses approximately ±1° around superior conjunction, caused by the rotating plasma-filled magnetosphere of B [53,54,142]. [143] In spite of that, we extend the Shapiro delay implementation in Tempo by an adapted version of the elegant approximation in Eq. (73) in Ref. [144], i.e., $\Lambda_u \to \Lambda_u + \delta\Lambda_u^{\rm len}$ with $\delta\Lambda_u^{\rm len} = 2rc/a_R$, which accounts to good approximation (only few percent error) for the delay related to a curved signal path and, therefore, removes nearly completely the already insignificant bias to s. It is interesting to note that the lensing correction leads to a reduction in the (calculated) propagation time, which is a result of Fermat's principle (see, e.g., Ref. [145]).

A higher-order signal propagation effect that becomes relevant for the double pulsar timing in the meantime, and cannot be absorbed into other timing parameters, is the *retardation effect*, i.e., specifically the velocity-dependent contributions in the Shapiro time delay (1.5PN corrections) [146,147]. It results from the fact that the "lens,"

i.e., pulsar B, moves while the signal of A propagates across the binary system toward Earth. The result derived in Ref. [146] is absolutely sufficient for modeling the TOAs around superior conjunction, [148] and one has

$$\Delta_{\rm S} = -2r\ln(\Lambda_u + \delta\Lambda_u^{\rm len} + \delta\Lambda_u^{\rm ret}), \tag{26}$$

where retardation correction $\delta\Lambda_u^{\rm ret}$ is of the order of $x/P_{\rm b}$ and can directly be taken from Eq. (130) in Ref. [146]. As in Ref. [146], we omit the velocity-dependent term in front of the logarithmic function, since its contribution is negligible (less than about 0.015 μs).

For systems with i close to 90°, like the double pulsar, one often obtains rather asymmetric error bars for the Shapiro shape parameter s. For that reason, in Ref. [149] the *logarithmic Shapiro shape parameter* is introduced:

$$z_s \equiv -\ln(1-s). \tag{27}$$

It not only leads to a more Gaussian distribution for the *fitted* parameter, i.e., z_s instead of s, it also ensures that the *a priori* free parameter s always fulfills $s < 1$, which is important when interpreted as the sine of the inclination angle.

D. Aberration

The term $\Delta_{\rm A}$ in Eq. (1) collectively refers to delays that are related to the fact that the pulsar is acting as a moving "lighthouse" with beamed radio emission. These effects (with one exception) are proportional to the rotational period of the pulsar, $P \equiv 1/\nu$, and would be absent if the variability of the radio signal were caused by an (intrinsic) oscillation or variability instead of the rotation of the NS. For pulsar timing, aberration means that the proper angle of emission of an observed pulse changes along the orbit [18,29]. To leading order, the aberration delay can be written as

$$\Delta_{\rm A}^{\rm (LO)} = \mathcal{A}[\sin\psi + e_T\sin\omega] + \mathcal{B}[\cos\psi + e_T\cos\omega], \tag{28}$$

where $\psi \equiv \omega + \theta$ is the longitude of the pulsar with respect to the plane of the sky. The aberration coefficients $\mathcal{A}$ and $\mathcal{B}$ are proportional to P and depend on the orientation of the pulsar with respect to the orbit and the observer. Since, as discussed in Sec. II, the spin of A is (practically) parallel to the orbital angular momentum, one has $\mathcal{B} = 0$ and

$$\mathcal{A} = \frac{Px}{P_{\rm b}(1-e_T^2)^{1/2}\sin^2 i} \simeq 3.65\ \mu{\rm s}. \tag{29}$$

In practice, $\mathcal{A}$ is a nonobservable parameter. If $\mathcal{A}$ is not provided in the timing model, it gets absorbed by a shift in various timing parameters (see Refs. [29,35] for details). For that reason, we *a priori* add the aberration coefficient $\mathcal{A}$ to our model, as a fixed parameter with the value given in the equation above.

Classical aberration, as expressed in Eq. (28), assumes a flat spacetime in which the signal propagates toward Earth. However, this is only a first approximation, which is no longer sufficient for an adequate description of the double pulsar timing observations, in particular, for TOAs near the superior conjunction of pulsar A. It is necessary to account for the deflection of the radio signal in the gravitational field of pulsar B. This gravitational signal deflection leads to a small change in the proper angle of emission, which, in turn, leads to a lensing correction to the classical ("longitudinal") aberration of Eq. (28). The potential importance of such a correction in close to edge-on binary systems has first been discussed in Ref. [150]. For the specific situation of A (spin aligned with orbital angular momentum), one finds

$$\Delta_{\rm A} = \mathcal{A}[\sin\psi + e_T\sin\omega] + \mathcal{D}\frac{\cos\psi}{\Lambda_u}, \tag{30}$$

where

$$\mathcal{D} = \frac{P}{\pi}\frac{r}{x + x_{\rm B}}. \tag{31}$$

The quantity $x_{\rm B}$ denotes the projected semimajor axis of pulsar B, which is an observed Keplerian parameter in the double pulsar [31]. The NLO contribution in $\Delta_{\rm A}$ contains the same two PK parameters as the Shapiro delay in Sec. V C, i.e., the range r ($\mathcal{D} \propto r$) and the shape s [in Λ_u; see Eq. (25)]. Like in the Shapiro delay (see Sec. V C), also here we account for retardation in our Tempo implementation, i.e., use the position of B when the signal reaches the minimum distance to B. This leads to small corrections in the last term of Eq. (30), i.e., a corresponding shift in ψ defined by the new position of B (and the position of A at emission) and a correction of Λ_u like in Eq. (26). Although these corrections are not yet significant, we implement them along with the retardation in the Shapiro delay, for completeness.

In Refs. [147,151], it is pointed out that the approximation for this lensing correction to the aberration delay, as given in Ref. [150], loses its validity if the impact parameter for the radio signal becomes comparable to the Einstein radius $R_{\rm E} \approx \sqrt{4Gm_{\rm B}a_R/c^2} \approx 2500$ km. [152] Our tests, based on simulated and the real timing data, show, however, that Eq. (30) is absolutely sufficient for the current double pulsar timing analysis. In fact, the main difference between the (retardation-corrected) approximation (30) and the full calculations (cf. Ref. [140]) is below 50 ns (postfit) and confined to a region very close to superior conjunction where, as explained above, the timing precision is anyway significantly compromised. One has to keep in mind that for outside the eclipse region one finds the impact parameter $\gtrsim 7R_{\rm E}$, which for the deflection angle means $\lesssim 0.026°$.

Finally, there is the lensing correction to the "latitudinal" aberration delay, related to a potential change in the observed pulse profile, as the LOS cuts a different part of pulsar A's magnetosphere due to the signal deflection [147,151]. This effect does not directly depend on the rotational period of pulsar A but is linked to A's rotation by having the LOS crossing the emission region of the pulsar. At this stage, it is not clear if there is at all a significant profile evolution across the range of the LOS variation. This is a question currently under investigation, utilizing the improved data quality provided by the MeerKAT telescope [154]. Because of a strong covariance with s and the eclipses near conjunction caused by the magnetosphere of pulsar B, this effect is not (yet) expected to be of any relevance [155].

E. Fitting for NLO contributions in the Shapiro and aberration delay

As it turns out, the NLO contributions in the Shapiro and aberration delays cannot be tested separately in the double pulsar. We have mentioned already in Sec. V C that the lensing correction to the propagation delay is covariant with the Shapiro shape s. Furthermore, the NLO contribution to Δ_S is quite similar to the NLO contribution to Δ_A, which renders a separation of $\delta\Lambda_u^{\rm ret}$ and $\mathcal{D}$, at least under current timing precision, impossible. For that reason, in order to test for the significance of NLO signal propagation contributions in our data, we collectively rescale these contributions with a common factor $q_{\rm NLO}$ that can be fitted for in our new timing model, i.e., for the retardation effect in the Shapiro delay

$$\delta\Lambda_u^{\rm ret} \to \delta\Lambda_u^{\rm ret} \times q_{\rm NLO}, \tag{32}$$

and for effects related to signal deflection

$$\delta\Lambda_u^{\rm len} \to \delta\Lambda_u^{\rm len} \times q_{\rm NLO}, \tag{33}$$

$$\mathcal{D} \to \mathcal{D} \times q_{\rm NLO}. \tag{34}$$

In GR, one has $q_{\rm NLO} = 1$. The NLO corrections above depend on Keplerian parameters and the PK parameters r and s [note that $a_R = c(x + x_B)/s$]. Therefore, $q_{\rm NLO}$ simply measures the relevance of these NLO corrections and is not considered as a new PK parameter.

VI. BINARY AND RELATIVISTIC-PARAMETER RESULTS

In this section, we present key results from this work relating to tests of Einstein's general relativity. These results are based on the methods described in Sec. V above and include a number of previously undetected relativistic effects. All observed effects are shown to be fully consistent with GR, a remarkable confirmation of Einstein's extraordinary intuition and vision. In the first timing parameters subsection, we give a full list of the parameters derived from the analysis of the 30-s sampled dataset. The next subsection presents results depending on our best determined values for the masses of the A and B pulsars, including constraints on the MOI of pulsar A (I_A) and, hence, on the neutron-star EOS, and the newly detected relativistic effects. Using these masses, we can constrain the rate of change of the orbital period due to GW damping $\dot{P}_b^{\rm GW}$ and probe light propagation effects in the gravitational field of a NS, key tests of GR, to unprecedented levels. We compare limits of deviation from GR in GW emission obtained with the double pulsar with limits from the LIGO/Virgo observations of BH-BH and NS-NS coalescences, showing that, at leading PN order (quadrupole formula), the double pulsar is more sensitive by 3 orders of magnitude, whereas at high PN orders, LIGO/Virgo gives more stringent limits, demonstrating the complementarity of these experiments.

A. Timing parameters

For the fit of the timing model to the TOAs, one does not use Eq. (1) in its form $t_b = T + \Delta(T)$ directly. What is needed is the inverted timing model, i.e., something of the form $T = t_b - \tilde{\Delta}(t_b)$. In Ref. [29], an analytic inversion of the timing model is worked out, giving an analytic expression for $\tilde{\Delta}(t_b)$. However, the approximation used in Ref. [29], in particular, with the Shapiro delay, is no longer sufficient for the double pulsar, mainly due to the higher-order signal propagation and aberration delays. For that reason, we modify the standard routine of the DDS model [157] to perform a numerical inversion of model Eq. (1) (for details, see the public Tempo distribution after November 27, 2015).

The results of applying the timing model are shown in Table IV and Fig. 4. In addition to the binary parameters introduced and discussed in detail above, we fit for spin frequency and higher-order derivatives (ν, $\dot{\nu}$, $\ddot{\nu}$, $\dddot{\nu}$, and $\ddddot{\nu}$), FD parameters [158,159], and time offsets of the TOA datasets relative to the reference datasets as indicated in Table I.

FD parameters mitigate the effect of possible unmodeled frequency evolution of the pulse profile by allowing for a time delay between different subbands according to $\Delta t_{\rm FD} = \sum_{i=1}^{i=N} c_i (\log f)^i$ [158]. Here, c_i are the FD parameters and f the subband frequency in GHz, when $\Delta t_{\rm FD}$ is expressed in seconds. Following Ref. [159], we employ an F test to determine that three FD parameters are needed. Similarly, we employ an F test to determine that four frequency derivatives are required to model the spin evolution of pulsar A.

Nonzero FD parameters suggest that future analyses of wideband data of the double pulsar will benefit from deploying smoothly varying, frequency-dependent

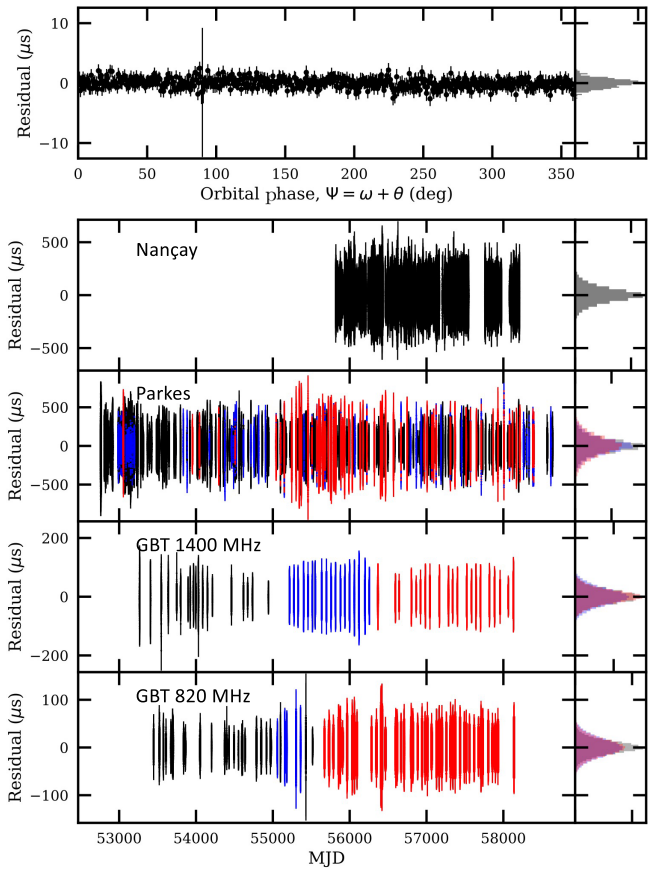

FIG. 4. Residuals obtained after applying the timing model to the 30-s dataset. From bottom to top: GBT 820 MHz (black, GASP; blue, GUPPI incoherent; red, GUPPI coherent), GBT 1400 MHz (same colors as for 820 MHz), Parkes (black, 50 cm; blue, 20 cm; red, 10 cm), and Nançay. The topmost shows residuals as a function of orbital phase, averaged to 15-s-long phase bins. The large uncertainties at $\Psi = 90°$ are the result of the presence of the eclipse. For each, we add a histogram of the distribution of residuals on the right-hand side.

templates [160,161]. But, in general, profile evolution and dispersion measure are highly covariant, with a fit for DM being able to (partly) absorb the effects of profile evolution and vice versa. Indeed, allowing the DM reference value to float in our final fit reduces the number of significant FD parameters to one. We note that the described choice does not have an impact on any of the binary timing parameters. We choose to fix the reference DM to the value given in Table IV and determine the DM variations as described in Sec. IV B.

The parameters shown in Table IV are the result of a large number of MC runs. Before each run, we select a random realization of the astrometry as determined in Sec. IV B, i.e., position, proper motion, and parallax that are consistent with the determined distributions. While these parameters are kept fixed for the given MC run, we use a modified version of Tempo to fit the described timing model to the remainder of the 20 timing parameters (see Table IV) and 22 constant time offsets (JUMPS) between the different datasets (see Sec. IV A). In each run, the DM value of a given TOA epoch is varied according to the DM curve and the estimated (time-varying) uncertainties as determined in Sec. IV B, thereby accounting for the particular DM value and its uncertainty as determined by the Gaussian learning process. The entries in Table IV are the mean of the resulting distributions of each parameter after a few thousand MC runs. The uncertainties are the standard deviation of the distributions or the maximum Tempo error encountered in all MC runs, whichever is larger.

In order to allow direct comparisons with previous publications (especially Ref. [31]), parameters in Table IV are measured within the timescale known as "Barycentric Dynamical Time" (TDB) as implemented in Tempo. TDB runs at a slower rate than the "Barycentric Coordinate Time" (TCB), which is recommended by IAU 2006 Resolution B3 [162]. This choice does not have any consequences for the gravity tests or discussions presented below, as all (dimensionful) parameters determined from TOAs measured using TDB are multiplied by a constant factor, which either drops out or is (still) too small to be relevant for the discussion of masses or PK parameters. In order to transfer from TDB to TCB, parameters with units of time shown in Table IV need to be divided by $\kappa = (1 - 1.550519768 \times 10^{-8})$ and the values adjusted accordingly [25,162,163]. As for the astrometric timing, the transfer of the TOAs from the topocentric to the barycentric reference frame is made using the DE436 Solar-System ephemeris.

While our timing results are perfectly consistent with those based on 2.5 years of timing data presented earlier [31], the increased length and density of our dataset leads to unprecedented precision in the measured parameters. For instance, the orbital period is measured with a precision of 86 ns. Most importantly, the Keplerian and PK parameters reach a precision that leads to a very significant improvement in our ability to conduct precision tests of strong-field gravity, including radiative and light-propagation aspects, as shown in the following sections. The need to fit up to the fourth spin frequency derivative (cf. Sec. V) reflects, on one hand, the exceptional duration and density of our dataset but also indicates a certain degree of timing noise, but at a level that is consistent with other pulsars of this age [103,165].

The observed improvements in PK parameters are in line with the expectation based on our earlier measurements [33], although a detection of the relativistic deformation of the orbit, described by PK parameter δ_θ, has occurred somewhat earlier than predicted. A consequence is a slightly smaller improvement in the precision of PK parameter γ_E, due to a correlation between the two parameters (see Sec. VI B 4). The precision in the measurement of the periastron advance, PK parameter $\dot{\omega}$ (or k), has improved beyond the level of 2PN contributions, making it necessary for the first time to include considerations of the Lense-Thirring effect and EOS of superdense matter in our

analysis (see Sec. VI B 3). Similarly, we have to take higher-order effects into account when studying light propagation in the double pulsar system (see Sec. VI B 5). Here, we point out that the derived orbital inclination angle $i = 89°.35 \pm 0°.05$ (or $180° - i$) is consistent with the value derived in our earlier timing analysis, i.e., $i_{\rm KSM06} = 88^{\circ}.69(+0^{\circ}.50/-0^{\circ}.76)$ [31], but now also in perfect agreement with the value obtained from modeling of the eclipse pattern ($i_{\rm eclipse} = 89°.3 \pm 0°.1$ [54,166]). In contrast, both values are in tension with measurements based on the changes in the intensity scintillation pattern as a function of orbital phase ($i_{\rm scint} = 88°.1 \pm 0°.5$ [92]). The latter method has the potential to resolve the $i \to 180° - i$ ambiguity caused by the Shapiro delay's dependency on $\sin i$. Ongoing MeerKAT measurements promise to improve on the scintillation results to resolve this difference [154].

B. Masses and tests of general relativity

1. Mass measurement

The standard procedure to obtain the masses of a ("clean") relativistic binary pulsar system is the measurement of two PK parameters. Under the assumption of GR, and using the well-measured Keplerian parameters, one can calculate the two *a priori* unknown masses [4]. Alternatively, one can fit the DDGR model [30], which is based on GR and has the companion and total masses as free parameters. For the double pulsar, the situation is more complicated, since the most precisely known PK parameter, $\dot{\omega} \equiv n_{\rm b} k$, has a contribution from the Lense-Thirring effect due to the rotation of pulsar A, which is more than 30 times larger than the measurement error [see Eq. (17) and $\dot{\omega}$ in Table IV]. However, a calculation of the Lense-Thirring contribution requires the knowledge of the MOI of pulsar A, $I_{\rm A}$, which comes with a significant uncertainty due to our imperfect knowledge of the EOS for NS matter (cf. Refs. [119,167]). Figure 5 demonstrates how the masses, determined from k and the Shapiro shape s (the second-most precise PK parameter) depend on $I_{\rm A}$. To constrain $I_{\rm A}$, one can use constraints on the EOS obtained from the double-NS merger GW170817 [168]. As a first approximation, we use the limits on the radius of a NS obtained in Ref. [112] and convert them into a probability distribution for $I_{\rm A}$ with the help of the radius-MOI relation for A given in Ref. [114]. Doing so, one obtains from Eq. (17)

$$\dot{\omega}^{\rm LT,A} = -4.83(+29/-35) \times 10^{-4}\ \text{deg yr}^{-1}, \tag{35}$$

and in combination with the PK parameter s one then gets for the (Doppler-shifted) masses [104]

$$m_{\rm A} = 1.338185(+12/-14)\ \mathrm{M}_{\odot}, \tag{36}$$

$$m_{\rm B} = 1.248868(+13/-11)\ \mathrm{M}_{\odot}, \tag{37}$$

$$M = 2.587052(+9/-7)\ \mathrm{M}_{\odot}. \tag{38}$$

In these calculations, for both of the PK parameters k and s, we account for NLO contributions as given in Sec. V. While for k this 2PN correction is highly significant (approximately 35σ), for s the 1PN term amounts to a correction of only about 1σ. It is interesting to note, while to leading order a measurement of k can directly be converted into a measurement of the total mass M, this is no longer the case once 2PN and Lense-Thirring corrections need to be accounted for [cf. Eqs. (9) and (14)].

Concerning the precision obtained in Eqs. (36)–(38), one has to keep in mind that, strictly speaking, they need to be rescaled with the unknown Doppler factor D of Eq. (1), in order get the actual (intrinsic) inertial masses [29,35]. [170] To calculate D, one would need the radial velocity v_R of the double pulsar system with respect to the SSB, which is unknown. Given the small transverse velocity v_T (see Table IV and also Sec. VIII C), one would naturally expect $v_R \sim 11\ \mathrm{km\,s^{-1}}$ (cf. Fig. 20), which converts into a mass uncertainty of approximately $3 \times 10^{-5}\ \mathrm{M}_{\odot}$. It is important

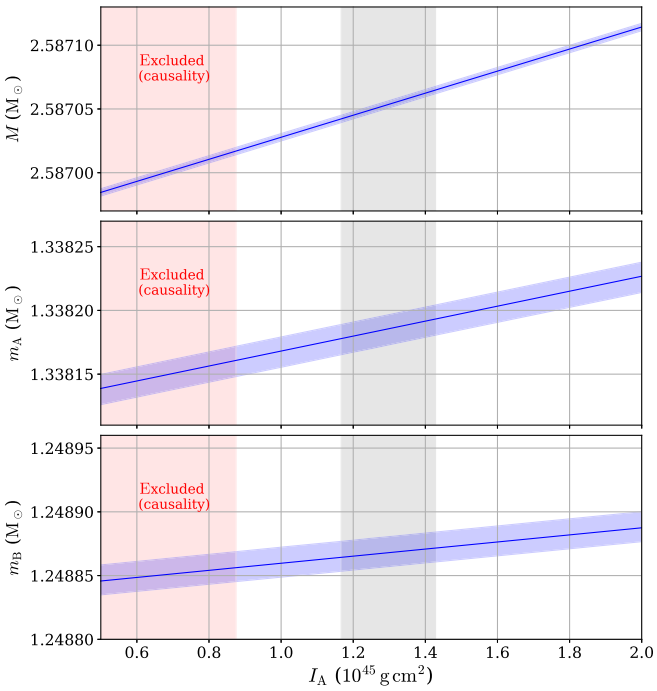

FIG. 5. Dependence on the masses of the double pulsar system on the MOI of pulsar A, $I_{\rm A}$. Mass calculations are based on GR and the two PK parameters $k(m_{\rm A}, m_{\rm B}, I_{\rm A})$ [Eq. (13)] and $s(m_{\rm A}, m_{\rm B})$ [Eq. (19)]. The figure shows the total mass $M = m_{\rm A} + m_{\rm B}$ (upper) and the masses of A (middle) and B (lower) as a function of $I_{\rm A}$. The blue bands indicate the 68.3% confidence range in the masses for each given $I_{\rm A}$. The gray bands indicate the likely MOI range, as suggested by the radius limits of Ref. [112] (90% confidence), when using the radius-MOI relation of Ref. [114]. The red areas are excluded by the causality condition for the EOS [169].

to note that in gravity tests the unknown factor D drops out and is, therefore, irrelevant [29,35].

The PK parameter γ_E (within the approximation required) is not affected by the uncertainty in the EOS. However, a mass determination based on s and γ_E yields considerably larger errors in the masses: $m_A = 1.3393(19)\ M_\odot$ and $m_B = 1.2494(9)\ M_\odot$. The same is true for a mass determination based on s and R [the latter taken from Ref. [31], where $R \equiv m_A/m_B = 1.0714(11)$ is given]: $m_A = 1.3379(29)\ M_\odot$ and $m_B = 1.2487(14)\ M_\odot$. Note that a mass measurement using $\dot{P}_b$ is to some extent also EOS dependent, due to the spin-down mass loss of A [see Eq. (20)], although currently this is still at a negligible level. Combining s and $\dot{P}_b$ (from Sec. VI B 2) gives $m_A = 1.338176(68)\ M_\odot$ and $m_B = 1.248842(34)\ M_\odot$, which is obviously less precise than the masses in Eqs. (36) and (37). In summary, in spite of the uncertainty in the EOS, s and k give by far the best estimate for the double pulsar masses.

2. Gravitational-wave emission

With the (GR) masses (36) and (37), the binary system is fully determined, and any further PK parameter can be used for a test of GR. Of particular importance is the test of GW damping. In their backreaction on the system, GWs extract orbital energy and angular momentum from the binary motion, leading to a secular change of the Keplerian parameters P_b, x, and e. Temporal changes in the latter two are still well below the measurement precision, which in terms of x means that the shrinkage of the orbit due to GW damping is not yet seen directly through a change in the size of the orbit. The (observed) change of the orbital period $\dot{P}_b$, on the other hand, is highly significant, measured with a fractional precision of 6×10^{-5} (see Table IV). This is nearly an order of magnitude better than for the Hulse-Taylor pulsar [172], and, as in that system, it is dominated by the GW damping.

For a GW test, one needs the intrinsic change of the orbital period, $\dot{P}_b^{\rm int}$. Therefore, external contributions $\dot{P}_b^{\rm ext}$, resulting from differential accelerations between the SSB and the double pulsar system, need to be subtracted from the observed orbital period change [122]. The Shklovskii effect [129] leads to an apparent acceleration due to the transverse motion of the double pulsar system and is given by

$$\dot{P}_b^{\rm Shk} = \frac{d(\mu_\alpha^2 + \mu_\delta^2)}{c} P_b = 1.72(15) \times 10^{-16}. \quad (39)$$

The distance d is calculated from the weighted mean probability distribution in Fig. 23. The differential acceleration in the Galactic gravitational potential leads to a correction of similar magnitude. If one takes the potential of Ref. [130] to calculate the Galactic gravitational accelerations of the SSB and the double pulsar, $\vec{g}_{\rm SSB}$ and $\vec{g}_{\rm DP}$, respectively, then one finds

$$\dot{P}_b^{\rm Gal} = \frac{\vec{K}_0 \cdot (\vec{g}_{\rm DP} - \vec{g}_{\rm SSB})}{c} P_b = -3.40(+25/-24) \times 10^{-16}, \quad (40)$$

where $\vec{K}_0$ is a unit vector directed from the SSB to the double pulsar. Instead of the potential in Ref. [130], one can also use a more analytic approach for the Galactic correction, like in Refs. [122,173,174], which in combination with the latest Galactic parameters based on the results in Refs. [131,175,176] gives very similar numbers. Other external contributions to $\dot{P}_b$, as, for instance, discussed in the context of the Hulse-Taylor pulsar in Ref. [122], are still negligible (see also the discussions in Refs. [110,177]). Combining Eqs. (39) and (40), while accounting for correlations, gives

$$\dot{P}_b^{\rm ext} = \dot{P}_b^{\rm Shk} + \dot{P}_b^{\rm Gal} = -1.68(+11/-10) \times 10^{-16} \quad (41)$$

and, consequently,

$$\dot{P}_b^{\rm int} = \dot{P}_b - \dot{P}_b^{\rm ext} = -1.247752(79) \times 10^{-12}. \quad (42)$$

At this point, it is important to note that the error in the calculated $\dot{P}_b^{\rm int}$ is still clearly dominated by the error in the observed orbital period change $\dot{P}_b$ (cf. Table IV). In other words, the uncertainty in our knowledge of the external contributions to the observed change in the orbital period [Eq. (41)] is still significantly smaller than the measurement error of the timing parameter $\dot{P}_b$.

The second-largest contribution to $\dot{P}_b^{\rm int}$, besides GW damping, is related to the mass loss of A as given in Eq. (20). Using the limits of Ref. [112] and the radius-MOI relation in Ref. [114] gives

$$\dot{P}_b^{\dot{m}_A} = 2.9(2) \times 10^{-17}, \quad (43)$$

which is still a factor of 3 smaller than the error of $\dot{P}_b^{\rm int}$. Nevertheless, we include the mass loss contributions in our calculations, in particular, since this is relevant in Sec. VI B 3 below, where instead of using *a priori* constraints on I_A we use $\dot{P}_b^{\rm int}$ to obtain a probability distribution for the MOI of A. The measured change in the orbital period due to GW emission amounts to

$$\dot{P}_b^{\rm GW} = \dot{P}_b^{\rm int} - \dot{P}_b^{\dot{m}_A} = -1.247782(79) \times 10^{-12}, \quad (44)$$

meaning that GW damping in the double pulsar is measured with a fractional precision of 6×10^{-5}. This improves the last published value for the double pulsar [31] by more than a factor of 200. This is also about a factor of 25 more precise than in the Hulse-Taylor pulsar [172], which is presently limited by the uncertainties in the correction for the external effects.

The high precision in Eq. (44) leads to important tests of the radiative properties of gravity theories. In Sec. VII, we use this for constraining two alternatives to GR. In this section, we provide only a test for GR. In GR, to leading order, $\dot{P}_{\rm b}^{\rm GW}$ is given by the quadrupole formula [124], which corresponds to the 2.5PN contributions in the equations of motion. Using the masses in Eqs. (36) and (37), one obtains

$$\dot{P}_{\rm b}^{\rm GW,GR(2.5PN)} = -1.247810(+6/-7) \times 10^{-12}. \quad (45)$$

The next-higher correction (3.5PN in the equations of motion) is calculated in Ref. [128] and amounts to

$$\dot{P}_{\rm b}^{\rm GW,GR(3.5PN)} = -1.75 \times 10^{-17}, \quad (46)$$

which is larger than the error in Eq. (45), and, therefore, we include it in our calculation for the predicted value:

$$\begin{aligned}\dot{P}_{\rm b}^{\rm GW,GR} &= \dot{P}_{\rm b}^{\rm GW,GR(2.5PN)} + \dot{P}_{\rm b}^{\rm GW,GR(3.5PN)} \\ &= -1.247827(+6/-7) \times 10^{-12}. \end{aligned} \quad (47)$$

This value is in perfect agreement with the observed value given in Eq. (44), leading to a GR test of

$$\dot{P}_{\rm b}^{\rm GW}/\dot{P}_{\rm b}^{\rm GW,GR} = 0.999963(63). \quad (48)$$

This is by far the most precise test of GW emission, about a factor of 25 better than in the Hulse-Taylor pulsar [172]. The agreement with GR is further demonstrated in Fig. 6, which shows the cumulative shift in periastron time due to the acceleration in the orbital phase evolution as a result of $\dot{P}_{\rm b}$ [cf. Eq. (5)]. Since this is the most precise of the different tests one obtains from the double pulsar, one can say that GR is in agreement with the double pulsar at a level of 1.3×10^{-4} with 95% confidence. We emphasize that the result (48) is calculated in a Monte Carlo run, where we simultaneously randomize the observed parameters and $I_{\rm A}$ [probability distribution like for Eq. (43)] in order to calculate the masses (from k and s) and all the $\dot{P}_{\rm b}$ contributions.

We also test the dependence of result (48) on our assumptions about the NS EOS. In a separate Monte Carlo run, instead of the limits in Ref. [112] we assume a uniform distribution for the MOI, in the (rather extreme) range of $I_{\rm A}^{(45)} = 0.875$ (causality limit [169]) and $I_{\rm A}^{(45)} = 1.981$ ($R_{\rm A} = 15$ km) [114,167]. The upper limit with its large radius is in clear tension with, e.g., LIGO/Virgo observations [168]. Nevertheless, even for such a distribution for $I_{\rm A}$ we obtain $\dot{P}_{\rm b}^{\rm GW}/\dot{P}_{\rm b}^{\rm GW,GR} = 0.999958(64)$, which shows that Eq. (48) depends only very weakly on the EOS uncertainty.

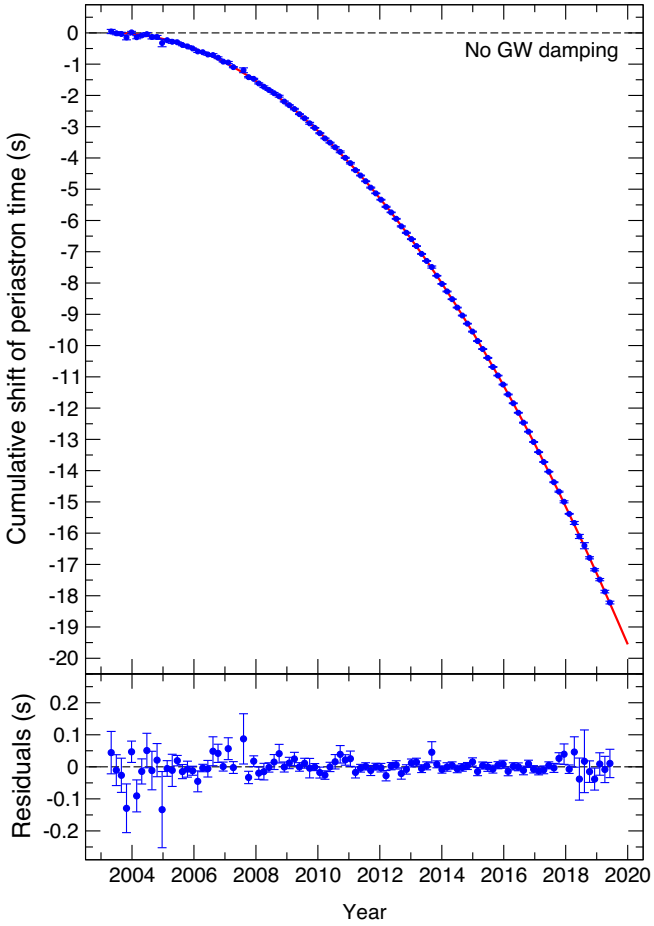

FIG. 6. Cumulative shift of the times of periastron passage relative to a nondissipative model. Each data point covers a time span of 60 d. To each of these subsets, we fit a Keplerian orbit optimizing only for the orbital period and the time and angle of periastron passage. We include the advance of periastron, light-propagation effects, and the Einstein delay in the orbital model but keep the values fixed to those in Table IV. We plot the difference between the measured time of periastron and a periastron time near its discovery (i.e., MJD 52759.89, or 2003.33). The red curve in the top is the GR prediction based on the masses in Sec. VI B 1. The bottom shows the deviation from this prediction, characterized by a normalized $\chi^2 = 0.76$.

In the following, we use the accelerated phase evolution in the orbital motion as shown in Fig. 6 to update the PN parameter test introduced by the LIGO/Virgo Collaboration in Ref. [178]. In particular, it is the tests from the early-inspiral stage, where the phase evolution of the merger of compact objects is analytically described by the PN approximation [123], that can be used in such a comparison with the GW damping measured in the double pulsar. Most importantly, the measurement precision of $\dot{P}_{\rm b}$ has greatly improved compared to Ref. [31], which are the data that are used in Ref. [178]. Figure 7 gives an updated version of Fig. 6 in Ref. [178] (including the new results [179] from Refs. [180,181]). The bounds are plotted for the different

PN levels, allowing for possible GR violations at different PN levels (i.e., different powers of frequency), one at a time. Note that Fig. 7 uses the "relative" PN order in the radiation reaction (i.e., PN order beyond the Einstein quadrupole formula), where the leading order, i.e., 0PN, occurs at the 2.5PN order in the binary equations of motion (see, e.g., Ref. [123] for a detailed discussion). Because of the many orbits since 2003 (approximately 60 000), which can be tracked with high precision in a phase-coherent timing solution, the double pulsar leads to considerably tighter constraints at low PN orders, whereas it becomes very quickly less constraining for higher PN orders, due to its comparatively small velocity ($v \sim 0.002c$).

While Fig. 7 certainly serves as a comparison on how much a given PN parameter of the inspiral phase evolution can (each at a time) deviate from its GR value in the different experiments, that figure has to be taken with a grain of salt when it comes to interpreting these bounds as limits on deviations from GR predicted by alternative theories of gravity. First, such a comparison mixes tests from two different types of compact objects, i.e., NSs and BHs, which might behave quite differently depending on how GR is broken. Hence, constraints from experiments with material bodies might not apply to BH dynamics and vice versa. Particularly obvious cases are alternative theories where BH binaries behave like in GR (e.g., Ref. [183]) or alternative theories where NSs do not carry any scalar charge, while BHs do [184]. Second, the double pulsar tests a different gravity regime (mildly relativistic strong field) compared to the GW merger events (highly relativistic strong field). For instance, the double pulsar test would generally be insensitive to modifications of GR that lead only to short-range effects (e.g., Refs. [185,186]); see also Ref. [187]. Nevertheless, at least to some extent, such a comparison illustrates the complementarity of binary pulsar experiments and merger observations by GW detectors, as long as one keeps in mind the qualitative differences of the various experiments, which are closely linked to the details of a given theory of gravity.

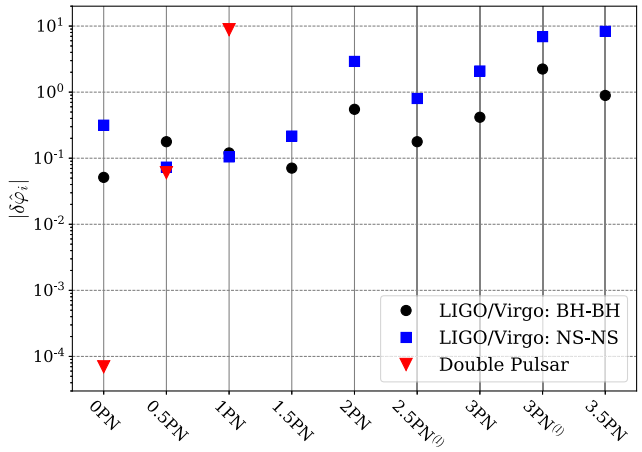

FIG. 7. Update of Fig. 6 in Ref. [178] (including data from Refs. [180,181]), which shows the 90% upper bounds on the absolute magnitude of the GR violation parameters $\delta\hat{\varphi}_i$, from 0PN through 3.5PN ("relative" order) in the inspiral phase (see, e.g., Ref. [182] for the definition of the PN phase coefficients and Ref. [178] for further details on the method). As discussed in Ref. [178], the 0.5PN parameter is zero in GR and, therefore, understood not as a relative but as an absolute shift. Black circles show the combined limits from the double BH mergers, blue squares are the limits from the double-NS merger GW170817, and red triangles give the limits derived from the double pulsar GW test in this paper. The PN order on the x axis is in the GR radiation reaction, where the leading contribution (0PN) corresponds to the dissipative 2.5PN term in the equations of motion. Note that such a comparison of tests with different compact objects (BHs vs NSs) as well as different gravity regimes (mildly relativistic vs highly relativistic strong field) does come with a caveat, which is explained in more detail in the text.

3. *Lense-Thirring effect and equation of state*

In Sec. VI B 1, we use constraints on the MOI of pulsar A, I_A, derived from the multimessenger analysis in Ref. [112], in order to obtain the best mass estimates for the double pulsar, as given in Eqs. (36)–(38). In this section, at first, we ignore any existing constraints on the EOS of NSs and simultaneously determine m_A, m_B, and I_A, following the procedure outlined in Ref. [110]. As in Sec. VI B 1, we assume GR to be the correct theory of gravity and use the three best PK parameters to simultaneously calculate the individual masses of the double pulsar and the MOI of A. From the calculations in Sec. VI B 1, it is already obvious that the combination of the PK parameters k, s, and $\dot{P}_b$ is expected to give by far the best results. In a way, we use s and $\dot{P}_b$ to determine the masses m_A and m_B and then use m_A to extract I_A from the observed advance of periastron $k^{\rm obs}$ (see $\dot{\omega} \equiv n_b k$ in Table IV), a procedure already proposed for the double pulsar in Ref. [33]. In practice, the calculations are slightly more complicated, as $\dot{P}_b$ also has a contribution proportional to I_A [see Eq. (20)]. Although that contribution is still smaller than the error in $\dot{P}_b$, we nevertheless account for it and follow the procedure in Ref. [110], i.e., calculate m_A, m_B, and I_A by simultaneously solving the three equations $k^{\rm obs} = k(m_A, m_B, I_A)$, $s^{\rm obs} = s(m_A, m_B)$, and $\dot{P}_b^{\rm int} = \dot{P}_b(m_A, m_B, I_A)$. By this, we obtain probability distributions for the double pulsar masses and the MOI of pulsar A. For the MOI, we find $I_A < 3.0 \times 10^{45}\ {\rm g\,cm^2}$ with 90% confidence. Figure 8 compares our result with those derived from the GW170817 LIGO/Virgo merger and from NICER x-ray timing. Using a universal relation, like the one in Ref. [114], one can convert the probability distribution of I_A into a probability distribution for A's radius. With 90% confidence, this gives an upper limit for A's radius of 22 km, a value outside any physically valid EOS and clearly exceeding the range used in Ref. [114].

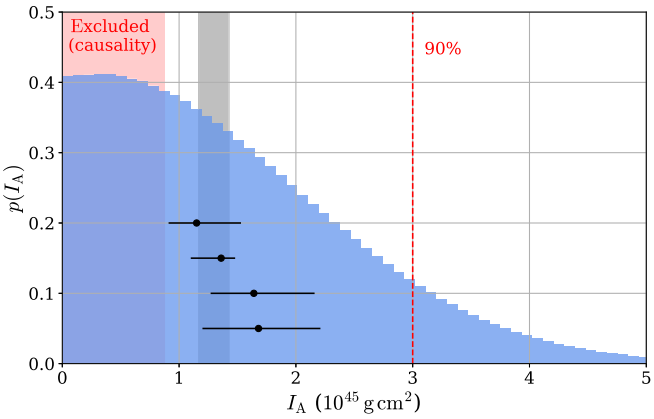

FIG. 8. Probability distribution for the MOI I_A of A derived from the $k - s - \dot{P}_b$ test. The vertical dashed red line indicates the upper bound with 90% confidence. $I_A > 0$ is assumed as a prior. The light gray band shows the 90% credible interval one obtains with the limits from Ref. [112] using the radius-MOI relation of Ref. [114]. As a comparison, the horizontal black lines indicate different 90% ranges derived from (top to bottom) tidal-deformability constraints from GW170817 [117], Bayesian modeling of a range of EOSs [116], and two different constraints from NICER observations [118]. The red area is excluded by the causality condition for the EOS [169].

While these numbers cannot compete with those obtained from LIGO/Virgo and NICER observations (see, e.g., Refs. [112,116–118]), they show that the double pulsar constraints are narrowing in on realistic values for NS radii. This is expected to improve considerably over the next years (see Hu *et al.* [110]). For the masses, we get $m_A = 1.338183(+78/-52)\ M_\odot$ and $m_B = 1.248869(+38/-27)\ M_\odot$. These masses are clearly more uncertain than in Eqs. (36) and (37), but they do not require any assumption about the EOS for matter at supranuclear densities.

Instead of using the Lense-Thirring effect to constrain the MOI, one can conversely use existing constraints on the NS EOS to test the contribution of spin-orbit coupling to the precession of the pulsar orbit (cf. Sec. VII in Ref. [110]). For this, we introduce a scaling factor λ_{LT} for the Lense-Thirring part in Eq. (14), i.e.,

$$k = k^{1PN} + k^{2PN} + \lambda_{LT} k^{LT,A}, \tag{49}$$

For GR, one has $\lambda_{LT} = 1$. Similar to the procedure above, we can now calculate m_A, m_B, and λ_{LT} by simultaneously solving the three equations $k^{obs} = k(m_A, m_B, \lambda_{LT}|I_A)$, $s^{obs} = s(m_A, m_B)$, and $\dot{P}_b^{int} = \dot{P}_b(m_A, m_B|I_A)$, where this time in our Monte Carlo runs I_A is randomly chosen from the probability distribution based on the EOS constraints in Ref. [112] (like in Sec. VI B 1, when determining the masses from k and s). In a sense, we determine the masses of A and B from the PK parameters s and $\dot{P}_b$ to extract the Lense-Thirring contribution form the observed periastron advance k, by assuming a range of values for I_A. We obtain

$$\lambda_{LT} = 0.7 \pm 0.9 \quad (68\%\ \text{C.L.}). \tag{50}$$

This is consistent with GR, but admittedly not very constraining, at least when compared to the weak-field (test-particle-type) frame-dragging experiments in the Solar System (Lense-Thirring precession of satellite orbits [188] and Pugh-Schiff precession of gyroscopes [189]). Moreover, the above result is not generic, in the sense of the parametrized post-Newtonian (PPN) tests in the Solar System. In order to extract the Lense-Thirring contribution to k, we assume that all three, the advance of periastron without Lense-Thirring contribution [i.e., Eq. (9)], the GW damping, and the Shapiro shape, can be calculated from GR. This is generally not expected to be the case in alternative gravity theories. As a result, one would have to implement the equivalent analysis for any other gravity theory, in order to obtain a fully consistent Lense-Thirring test. Given the low precision of the LT test, one may ask if, in view of the other tests with the double pulsar, mostly based on very precisely measured PK parameters, this then yields any additional constraints for the gravity theories under consideration. For the alternative theories discussed in Sec. VII, at least, spin-orbit effects observed in the double pulsar (including Ω_B^{spin}) do not contribute to the constraints. See Ref. [110] for a more detailed discussion, in particular, on this test in the context of near-field modifications of GR.

Finally, from the analysis outlined above, one can also extract a test of the total advance of periastron k, in an s-$\dot{P}_b$-k test. From this, one obtains $k^{obs}/k^{GR} = 1.000015(26)$.

4. Relativistic deformation of the orbit

The relativistic deformation of the orbit discussed in context of the Rømer delay (Sec. V A) is detected in our measurements. Similar to the report of a 1.5σ detection of δ_θ for PSR B1913+16 [172], our value is also formally detected only just above the 1σ level. However, as we demonstrate in Fig. 9, the parameter is, in fact, well constrained and in full agreement with the expectation from GR. The figure also demonstrates a correlation between δ_θ and γ_E. Indeed, due to this correlation, the value derived for γ_E in a timing model ignoring the relativistic deformation, i.e., assuming $\delta_\theta = 0$, would lead to a value for γ_E that would be inconsistent with GR at the 2σ level, while both γ_E and δ_θ are, in fact, in perfect agreement when including δ_θ in our timing model (see the white cross in Fig. 9). In order to illustrate the effect of a nonzero δ_θ value, we compare a deformed orbit ($\delta_\theta > 0$) with an elliptical orbit ($\delta_\theta = 0$) in Fig. 10 for a hugely exaggerated effect.

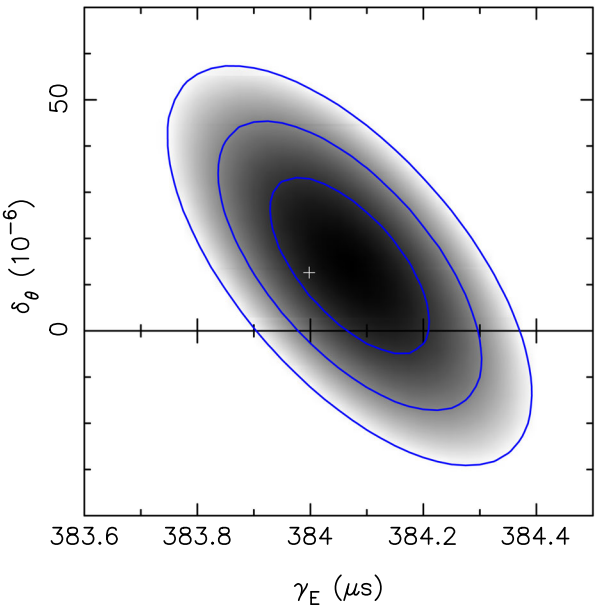

FIG. 9. Map of χ^2 contours demonstrating a correlation between the PK parameters γ_E and δ_θ. In order to produce this map, the two parameters are held fixed at their grid position, while fitting for all remaining parameters in the timing model as described earlier. The contours indicate 68.3%, 95.4%, and 99.7% confidence levels, respectively. The expected GR value is indicated by the cross with uncertainties too small to be visible on this scale: $\gamma_E = 383.997(5)$ μs and $\delta_\theta = 12.60889(4) \times 10^{-6}$ (masses from k and s; cf. Sec. VI B 1).

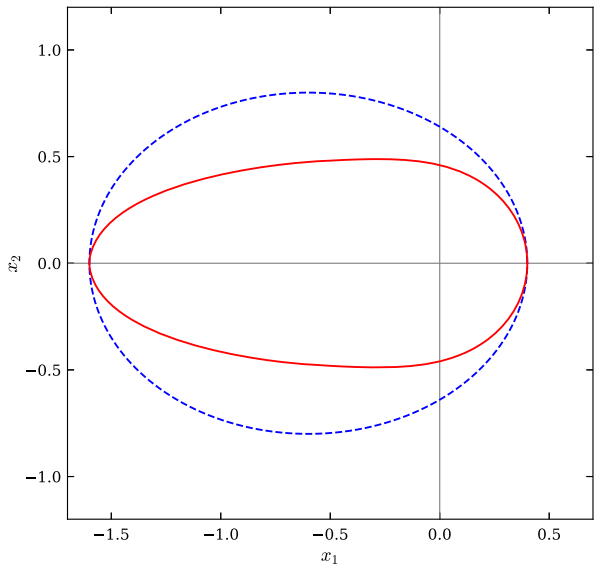

FIG. 10. Hugely exaggerated illustration of the effect of the relativistic deformation δ_θ on the shape of the orbit, which modifies the usually expected elliptical shape. For the red (deformed) orbit, we choose $a = 1$, $e_T = e_r = 0.6$, and $\delta_\theta = 0.5$, while the dashed blue curve has $\delta_\theta = 0$.

5. Signal propagation

Already in the discovery paper of A [21], it was suspected that the double pulsar system is seen nearly edge-on. This was confirmed by the discovery of a prominent Shapiro delay in the TOAs of A, as well as by the observation of eclipses of A near its superior conjunction, caused by the plasma-filled magnetosphere of B [22]. The measurement of the Shapiro delay gave access to two additional PK parameters, i.e., the "Shapiro range" r and the "Shapiro shape" s [29] [see Eq. (24)]. By 2006, these two parameters were measured with high precision, and, in particular s, as part of an $\dot{\omega}$-R-s test, provided a 10^{-3} (95% C.L.) GR test [31]. Unfortunately, the measurement of the mass ratio R has not improved since then, mostly because of the fact that B disappeared from view in early 2008 as a result of relativistic spin precession [58]. In the meantime, the measurement of the Shapiro delay, and, therefore, s, has improved significantly (see Fig. 11), making it the second-most constraining parameter in the mass-mass plane (see Fig. 13 below), which is also obvious from the mass determinations in Sec. VI B 1. Since γ_E by now is more constraining than R, one can test s as part of a $\dot{\omega}$-γ_E-s test. This leads to

$$s^{\rm obs}/s^{\rm GR} = 1.000\,09(18), \tag{51}$$

meaning a 4×10^{-4} (95% C.L.) test in agreement with GR, which is somewhat better than the $\dot{\omega}$-R-s test of Ref. [31], while mainly dominated by the error in γ_E. Moreover, from Eq. (3.15) in Ref. [35], it is clear that, for a large range of alternatives to GR, s is linked only to the effective gravitational constant G_{AB} of the orbital dynamics, as s can be identified with the purely geometric quantity $\sin i$.

A test of the Shapiro range r directly probes the interaction between a strongly self-gravitating NS—in our case, pulsar B—and a photon. This can be easily seen within the modified Einstein-Infeld-Hoffmann (mEIH) formalism [35,137], where $r = (1 + \gamma_{B0})G_{B0}m_B/c^3$ [see, in particular, Eq. (3.14) in Ref. [35] with $\varepsilon_{B0} = 2\gamma_{B0} + 1$]. G_{B0} is the effective gravitational constant and γ_{B0} the strong-field extension of the PPN formalism parameter $\gamma^{\rm PPN}$ (see, e.g., Ref. [41]), which enter the leading-order spacetime metric $g_{\mu\nu}$ of pulsar B. In GR, we have $G_{B0} = G$ and $\gamma_{B0} = 1$. Since r is by far the most uncertain of the PK parameters, apart from δ_θ and $\Omega_B^{\rm spin}$, there are several combinations of PK parameters that can be used in the test. In other words, the agreement of $m_B^{(r)} \equiv rc^3/G = 1.2510(43)\,M_\odot$ (cf. Table IV) with the companion mass determined in Sec. VI B 1 gives

$$r^{\rm obs}/r^{\rm GR} = 1.0016(34), \tag{52}$$

leading to a 7×10^{-3} (95% C.L.) test of r in agreement with GR.

It is interesting to note that, near superior conjunction, the signals of the pulsar propagate through a region with a

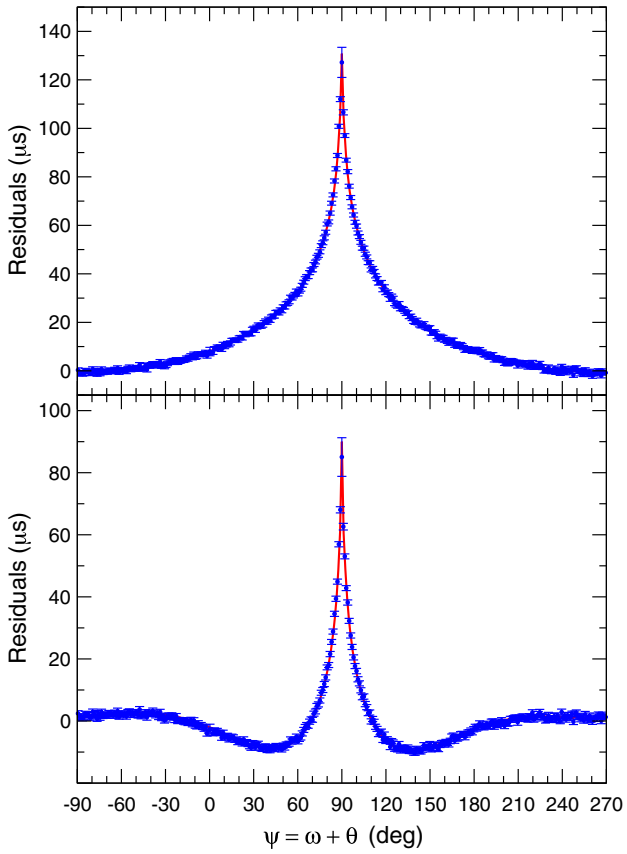

FIG. 11. Signature of the Shapiro delay in the timing residuals, plotted against A's angular orbital position $\psi = \omega + \theta$ with respect to the ascending node. The bottom shows the aggregated residuals as a function of orbital phase when not including the Shapiro delay parameters z_s (or, equivalently, $\sin i$) and r. The shape differs from the expected Shapiro delay curve shown in the top, as some signal power is absorbed in a fit for the orbital parameters, namely, the Rømer delay (see, e.g., Ref. [98]). The expected shape is recovered by computing the residuals when setting the Shapiro delay parameters to zero in the final timing solution. Note that the shown residuals are corrected for a slow variation in the curve due to the relativistic orbital precession. The red curves show the GR expectation using values of $\sin i$ and r derived from $\dot{\omega}$ and $\gamma_{\rm E}$. These are $(\sin i)^{\rm GR} = 0.99988(18)$ and $r^{\rm GR} = 6.1518(11)$ μs, where the uncertainties are determined by those in $\dot{\omega}$ and γ_E. Note that the uncertainty in the prediction of $\sin i$ is correspondingly larger than that of the actually measured $\sin i$.

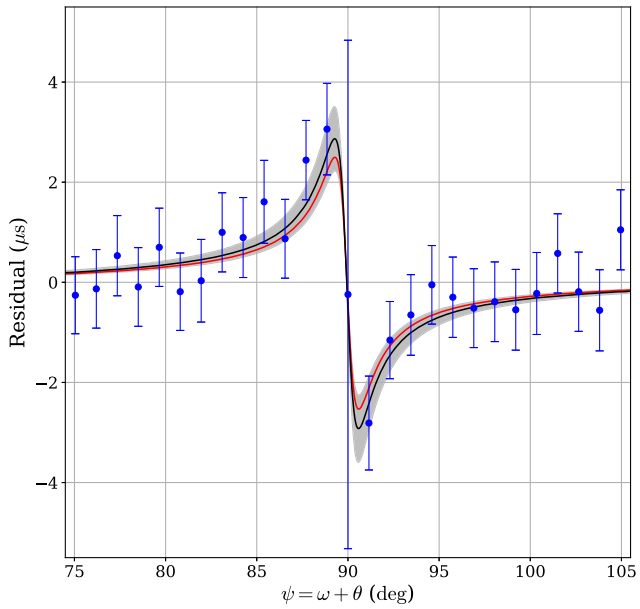

FIG. 12. Aggregated residuals (blue) due to NLO contributions in the Shapiro delay and the aberration, plotted against A's angular orbital position $\psi = \omega + \theta$ with respect to the ascending node. The red curve shows the theoretical prediction, and the black curve corresponds to the fitted $q_{\rm NLO}$ (see Table IV), with a 2σ range indicated by the light gray band. Residuals are rescaled by $(1 + e_T \cos\theta)^{-1}$ to account (to leading order) for a secular variation in the amplitude due to the precession of ω. A data point generally results from approximately 3000 aggregated residuals. The data point at $\psi = 90°$ (i.e., superior conjunction of A) has a much larger uncertainty, as this bin coincides with the 30-s-long eclipse of pulsar A, resulting in both many fewer TOAs available for aggregation and residuals of opposite sign being averaged.

spacetime curvature (as, for instance, measured by the Kretschmann scalar) that is many orders of magnitude stronger than in other experiments that test photon propagation in gravitational fields and, therefore, the coupling between gravitational and the electromagnetic fields. This is of interest, for instance, in the presence of interactions between the electromagnetic field and curvature tensor as studied in Ref. [190].

As it turns out, the leading-order expression (24) no longer provides a complete description of the signal propagation in the double pulsar. Such a model leads to significant residuals near conjunction, as can be seen in Fig. 12. These residuals are fully accounted for by the expected NLO contributions discussed in Secs. V C and V D (red curve in Fig. 12). To test the significance of the NLO corrections in $\Delta_{\rm S}$ and $\Delta_{\rm A}$, we scale them collectively with a factor $q_{\rm NLO}$ [cf. Eqs. (32) and (34)], which is unity in GR. As discussed in Sec. V E, the similarity of the NLO contributions to $\Delta_{\rm S}$ and $\Delta_{\rm A}$ makes it impossible to test these two contributions separately. For the common factor, which also scales the NLO contribution from lensing [see Eq. (33)], we find

$$q_{\rm NLO} = 1.15(13), \tag{53}$$

which can be interpreted as an approximately 8σ measurement of NLO contributions in the signal propagation, in agreement with GR. This clearly shows that the timing of the double pulsar has by now reached a precision where NLO contributions to the Shapiro and aberration delay have to be taken into account.

As as consequence of its definition in Sec. V E, a fit for $q_{\rm NLO}$ combines two aspects of gravity, 1.5PN corrections to

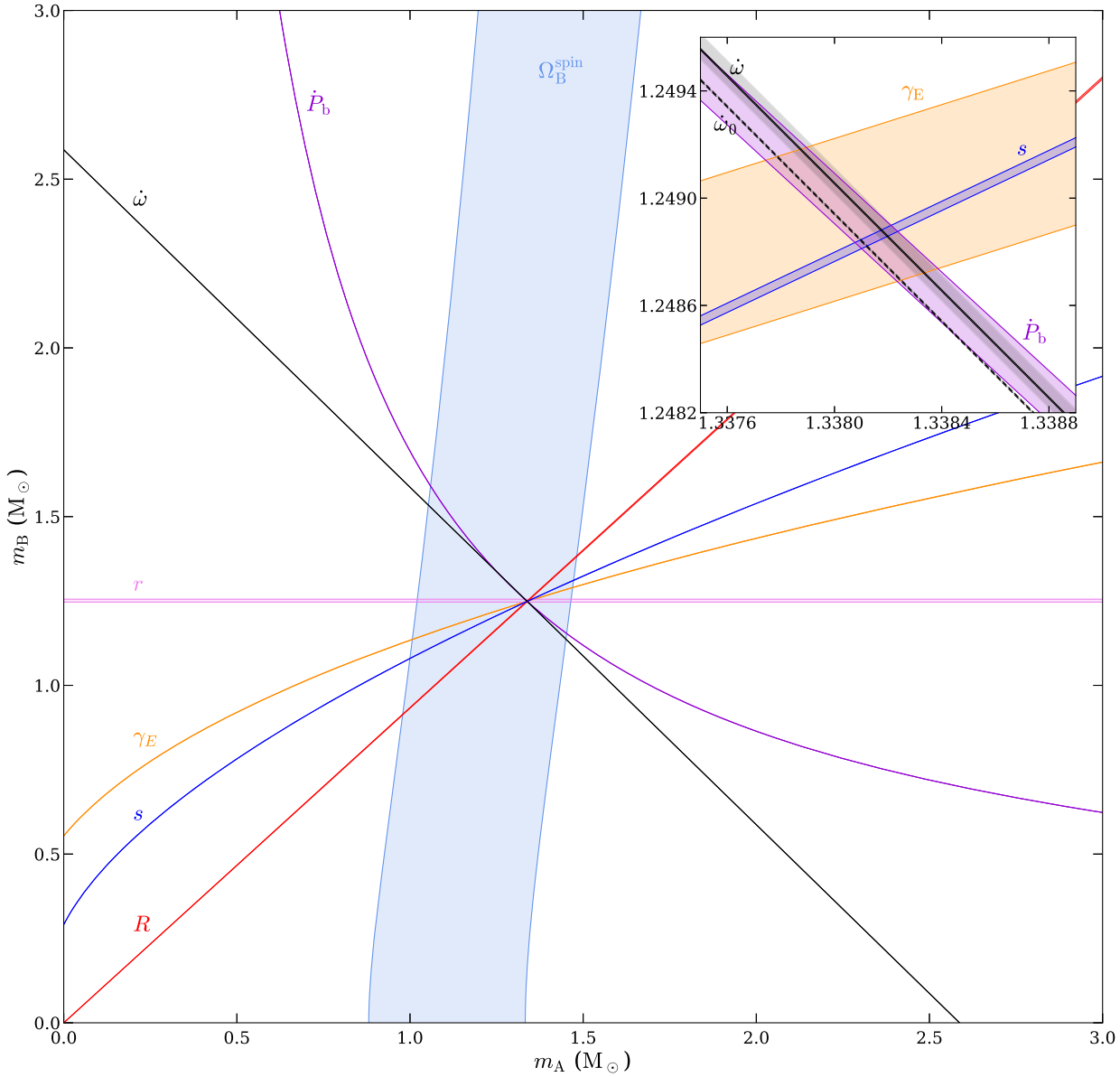

FIG. 13. Mass-mass diagram for the double pulsar based on GR, for the six PK parameters $\dot{\omega} \equiv n_b k$, γ_E, $\dot{P}_b$, r, s, Ω_B^{spin}, and the mass ratio R. The width of each curve indicates the measurement uncertainty of the corresponding parameter. The seventh PK parameter δ_θ is not shown here, since its limits still lie outside the mass ranges shown. For the solid black $\dot{\omega}$ line, $I_A^{(45)} = 1.32$ is used, which corresponds to a NS radius of 12 km (cf. Ref. [112]). The inset is an expanded view of the region of principal interest, where only the four best PK parameters are shown. To illustrate the influence of the LT effect, we draw in the inset a dashed black $\dot{\omega}_0$ line where the LT contribution is ignored [see Eq. (9)]. The gray band indicates the range for the $\dot{\omega}$ line under the variation of I_A, from the causality limit of Ref. [169] (left border) to a NS radius of 15 km (right border). The conversions between radius and MOI are based on the relation in Ref. [114]. There is a small I_A dependence of the $\dot{P}_b$ curves, which is ignored. The intersection of all line pairs is consistent with a small region that corresponds to the masses of A and B.

the Shapiro delay related to the motion of the masses [Eq. (32)] and corrections related to the deflection of the signal beam in the gravitational field of the companion [Eqs. (33) and (34)]. Although, as discussed in Sec. V E, these contributions cannot be tested separately in a simultaneous fit, in a phenomenological approach one can still test for one at a time while keeping the other one fixed. If we apply the scaling factor q_{NLO} only to the deflection-related contributions, one finds

$$q_{NLO}[\text{deflection}] = 1.26(24). \tag{54}$$

This limit comes solely from the NLO aberration contributions [Eq. (34)], since a rescaling of $\delta\Lambda_u^{len}$ is covariant

with s. Given the large uncertainty, the above test is admittedly of little interest in all those theories or frameworks where deviations from GR in the signal deflection (caused by a strongly self-gravitating mass) are already well constrained by the measurement of the Shapiro delay PK parameters r and s, i.e., Eqs. (30) and (31) [191].

Alternatively, one can keep the deflection contributions fixed, i.e., $q_{\rm NLO} = 1$ in Eqs. (33) and (34), and fit only for the retardation effect. In this case, one obtains

$$q_{\rm NLO}[\text{retardation}] = 1.32(28). \tag{55}$$

This result clearly demonstrates the significance of the motion of the companion mass while the signal of A propagates across the binary system on its way toward the observer.

Finally, when assuming that the signal deflection is sufficiently well described by GR, then Eq. (54) provides independent evidence (besides Refs. [52,57]) for the fact that the spin of A is aligned with the orbital angular momentum and not antialigned [$\mathcal{D} \to -\mathcal{D}$ in Eq. (30)] [193].

6. Mass-mass diagram

The PK parameters, once external contributions are removed and a value for $I_{\rm A}$ is chosen, depend only on the *a priori* unknown masses of pulsars A and B and the well-measured (and, hence, determined) Keplerian parameters. Hence, each PK parameter defines a curve in a two-dimensional mass plane displaying their dependence on $m_{\rm A}$ (plotted on the X axis) and $m_{\rm B}$ (plotted on the Y axis) for a given theory of gravity. If the theory of gravity used to describe the dependence of the n curves on the masses is consistent with the experimental data, and if no other effects are present, then all curves will meet in a single point. As two curves each define one intersection point, $n-2$ curves have still the potential to miss the intersection point and to falsify the assumed theory. This "mass-mass" diagram, as shown for GR in Fig. 13, is a graphical representation of the gravity tests using PK parameters described in Sec. II, with $\dot{P}_{\rm b}$ corrected as described in Sec. VI B 2. In the case of the double pulsar, the mass ratio adds yet another (largely theory-independent) curve to this diagram in Fig. 13. The thickness of the lines represents the uncertainty in the measured PK parameters (68% C.L.). For all PK parameters but $\Omega_{\rm B}^{\rm spin}$, which is determined from eclipse modeling (see Sec. II), we have reached a point where we need to enlarge to show the intersection point in more detail to recognize the relative uncertainties. While the large panel in Fig. 13 shows most of the mass range of interest for NSs, the inset shows such an enlarged area, where we choose not to show R, r, and $\Omega_{\rm B}^{\rm spin}$ for clarity. It is obvious that all lines are consistent with a single intersection point given by the masses of Eqs. (36) and (37). The inset, however, highlights the fact that we clearly need to take the Lense-Thirring contribution into account (see discussion in Sec. V A 2). With $n = 7$ constraints (i.e., the six PK parameters measured here and the mass ratio R), we can conduct $n - 2 = 5$ independent tests of theories of gravity, which are all passed with flying colors as shown above [194].

On a final note, in the inset in Fig. 13, one can see that the measured $\gamma_{\rm E}$ nicely agrees with the intersection of the s curve with either the $\dot{\omega}$ curve or the $\dot{P}_{\rm b}$ curve. One can convert this into a test of $\gamma_{\rm E}$:

$$\gamma_{\rm E}/\gamma_{\rm E}^{\rm GR} = 1.000\,12(25). \tag{56}$$

The main interest in such a test lies in the fact that it tests an invariance of the local gravitational constant, which is a specific aspect of local position invariance. We discuss this in more detail in the following section, in the context of two specific alternatives to GR.

VII. ALTERNATIVE GRAVITY THEORIES

In the previous section, we demonstrate that Einstein's GR is in perfect agreement with current observations of the double pulsar. This shows that, in contrast to Solar-System tests which are in the weak-gravity regime, GR also is an accurate description of gravitational interactions for two strongly self-gravitating bodies in a mildly relativistic strong-gravity regime. However, since GR was devised in 1915, many other theories of relativistic gravity have been proposed. In this section, we consider these alternative theories. Since it is impossible to give an even remotely complete discussion, we focus on two examples. The first, known as "Damour–Esposito-Farèse" (DEF) gravity, is a two-parameter class of theories, whereas the second, known as "Tensor-Vector-Scalar (TeVeS)" gravity, represents the class of "Modified Newtonian Dynamics" or MONDian gravitational theories.

These theories are chosen since they are especially suitable for studying the deviations one typically expects when going beyond GR by breaking the SEP and how they are being constrained by the different PK parameters measured in the double pulsar. In this, these theories provide a theory-based approach for putting different GR experiments into context and allow for a better understanding which structures and symmetries of GR are actually being probed by pulsars and, more specifically, the double pulsar (see Sec. V in Ref. [63]).

Before entering the theory-specific discussion, we give a general statement about constraints on deviations from GR in the radiative regime that can come from the precise measurement of the GW emission [see Eq. (48)]. In alternative theories, one often has gravitational dipolar radiation due to a violation of the SEP as a significant contribution to GW damping. In the equations of motion, that contribution enters at the 1.5PN level [$\mathcal{O}(\beta_{\rm O}^3)$; see, e.g., Ref. [183]]. If we parametrize a deviation from GR led by

dipolar GW damping by $\dot{P}_{\rm b}^{\rm GW} = \dot{P}_{\rm b}^{\rm GW,GR}(1 + B_D \beta_{\rm O}^{-2})$, then we find $B_D \lesssim 4 \times 10^{-10}$ (95% C.L.). This is nearly 5 orders of magnitude more constraining than the limit from the double-NS merger GW170817 [195]. Note, however, that the two different limits are coming from different gravity regimes (see the discussion at the end of Sec. VI B 2). How such a limit on B_D converts into actual constraints on a given gravity theory depends on the specifics of that theory. In many theories, it is expected that, due to the similarity of the two masses in the double pulsar, the limit is actually weaker than one would assume directly from the tight limit on B_D.

A violation of the SEP not only gives rise to dipolar GWs, it generally also results in a temporal modification of the (effective) gravitational constant due to the expansion of the Universe (see, e.g., Ref. [196]), by this having an addition effect on the orbital period. The orbital period change due to a time-varying gravitational constant is given by $\dot{P}_{\rm b}^{\dot{G}}/P_{\rm b} = -2(\dot{G}/G)\mathcal{F}_{\rm AB}$, where $\mathcal{F}_{\rm AB}$ accounts for corrections related to the strong gravitational fields of the NSs in the double pulsar system [196–198]. [199] For weakly self-gravitating bodies, as in the Solar System, $\mathcal{F}_{\rm AB} \simeq 1$. $\mathcal{F}_{\rm AB}$ depends on the details of the gravity theory under consideration, as well as on the EOS. While for Jordan-Fierz-Brans-Dicke gravity $\mathcal{F}_{\rm AB} \approx 0.6$, for some theories $\mathcal{F}_{\rm AB}$ can be significantly larger than one (cf. discussion in Ref. [27]). If one assumes that $\dot{P}_{\rm b}^{\dot{G}}$ is the only significant non-GR contribution to the orbital period change, then from Eq. (48) one finds that $|\dot{G}/G| < -0.8(14) \times 10^{-13}\ {\rm yr}^{-1}/\mathcal{F}_{\rm AB}$. This limit is about a factor of 2 better than the currently best pulsar limit [201], although the exact factor is theory dependent. In cases where $\mathcal{F}_{\rm AB}$ is large, this limit can even exceed the best limit obtained in the Solar System (i.e., $|\dot{G}/G| < 4 \times 10^{-14}\ {\rm yr}^{-1}$ [202]). At this stage, these considerations need to be taken with a grain of salt. As stated above, theories with nonvanishing $\dot{G}$ are expected to also predict dipolar GWs, which, in turn, leads to an additional non-GR modification of $\dot{P}_{\rm b}$. As a conclusion, in order to perform a $\dot{G}$ test with the double pulsar, one has to do a fully consistent analysis within a certain class of gravity theories and combine it with other binary pulsars (similar to the method proposed in Ref. [174]). However, such an analysis goes beyond the scope of this paper and will be presented in a future publication.

A. Damour–Esposito-Farèse Gravity

The first alternative to GR which we confront with our double pulsar results is the two-parameter monoscalar-tensor gravity $T_1(\alpha_0, \beta_0)$ introduced in Ref. [40] ("*DEF gravity*"). This two-parameter class of theories exhibits various effects one typically expects from alternatives to GR, including genuine strong-field effects related to the self-gravity of NSs. Hence, $T_1(\alpha_0, \beta_0)$ can be viewed as an ideal minispace of well-studied gravity theories in which GR is embedded ($\alpha_0 = \beta_0 = 0$). They are well suited to contrasting the predictions of GR with the predictions of alternatives and allowing for the comparison of GR tests across different gravity regimes [63]. The constants α_0 and β_0 measure, respectively, the linear and the quadratic coupling strength between matter and the scalar field. For simplicity, we assume that the potential of the scalar field can be ignored on the typical scales for the double pulsar test, while it could still be of importance on Galactic or cosmological scales. We further assume that we can neglect a temporal change in the background scalar field, which among others would lead to a long-term variation of the gravitational constant. For sufficiently negative β_0, DEF gravity shows genuine strong-field effects in binary pulsars, related to "spontaneous scalarization" [40,42]. The presence of the scalar field leads to a modification of the PK parameters in different ways [42,203].

Already at the Newtonian level, the body-dependent effective scalar couplings $\alpha_{\rm A}$ and $\alpha_{\rm B}$ enter via the (effective) gravitational constant $G_{\rm AB} = G_*(1 + \alpha_{\rm A}\alpha_{\rm B})$, where G_* is the bare gravitational constant, linked to Newton's gravitational constant G (as measured in a Cavendish-type experiment) via $G = G_*(1 + \alpha_0^2)$. By this, the acceleration of a body depends on the gravitational binding energy, leading to a violation of the SEP [41]. As a result, one has a strong-field modification of the mass function of the binary system that enters the expression for the Shapiro shape parameter s [cf. Eq. (5.8) in Ref. [42]].

At the first PN level, one has three body-dependent strong-field counterparts of Eddington's weak-field PPN parameters $\gamma^{\rm PPN}$ and $\beta^{\rm PPN}$ [204] and a periodic modification of the MOI due to a violation of local position invariance. All this leads to genuine strong-field modifications of the corresponding PK parameters (see Appendix C for details).

Beyond the first PN level, there are modifications to the conservative dynamics at 2PN [183], which are not (yet) relevant for the test presented here. Hence, it is sufficient to use the corresponding GR terms. However, the presence of the dynamical scalar field φ considerably modifies the loss of energy and angular momentum due to the emission of scalar GWs. As a result, in addition to the quadrupolar tensor GWs, one has monopolar, dipolar, and quadrupolar scalar GWs, entering the equations of motion at the 2.5PN, 1.5PN, and 2.5PN level, respectively. These dissipative contributions to the equations of motion result, most importantly, in a modification of $\dot{P}_{\rm b}$ (see Appendix C for details). The leading term, i.e., scalar dipole radiation, is proportional to $(\alpha_{\rm A} - \alpha_{\rm B})^2$ and, therefore, quite sensitive to the asymmetry in gravitational binding energy in a system. [205] Even though the difference in mass between A and B is only 7%, the small size of the effect is compensated for by the high precision (1.3×10^{-4}) of the GW emission test (see Sec. VI B 2). Consequently, the GW damping test

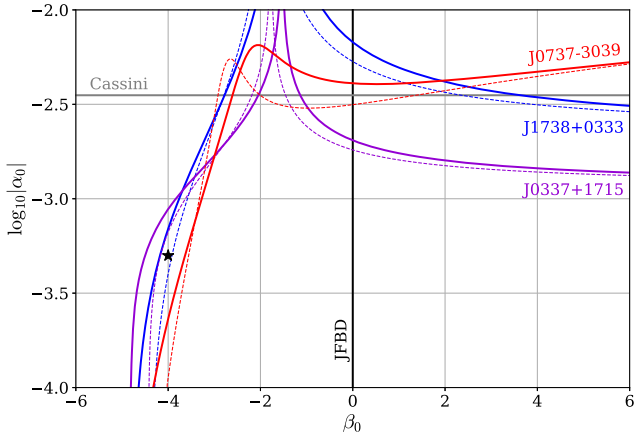

FIG. 14. Constraints on the DEF gravity from different experiments: Shapiro delay with the *Cassini* spacecraft [207], dipolar radiation (J1738+0333) [208], gravitational weak equivalence principle (J0337+1715) [76], and the double pulsar (this paper). Areas above a curve are excluded. (See Refs. [42,206] for details. We use $\chi^2 = 4$ as a conservative limit.) Solid lines are for a comparably stiff EOS (MPA1 in Ref. [167]). The dashed lines are based on a rather soft EOS (WFF1 in Ref. [167]), to illustrate the EOS dependence of the different limits. GR corresponds to $\alpha_0 = \beta_0 = 0$, and Jordan-Fierz-Brans-Dicke (JFBD) theory is along the vertical $\beta_0 = 0$ line with Brans-Dicke parameter $\omega_{\rm BD} = (\alpha_0^{-2} - 3)/2$. The black star indicates the parameters used in Fig. 15.

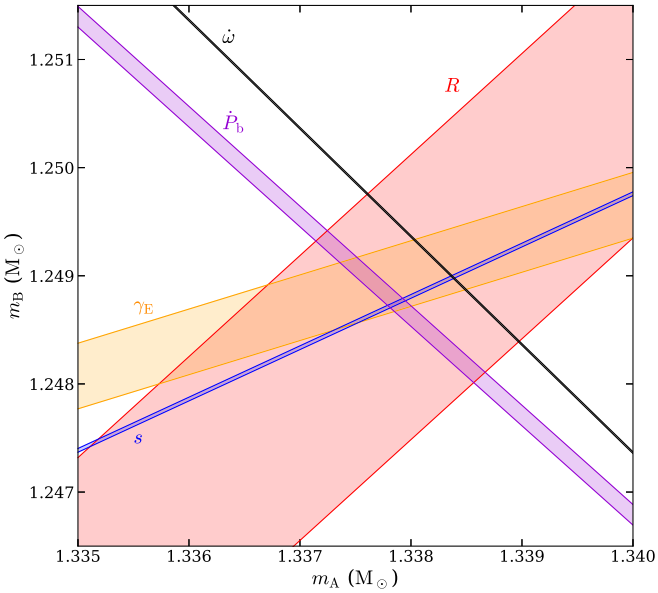

FIG. 15. Mass-mass diagram for DEF gravity with $\alpha_0 = 5 \times 10^{-4}$ and $\beta_0 = -4$. EOS MPA1 is used to calculate the curves. The curves fail to agree on a common region in the mass-mass plane (see, in particular, $\dot{\omega}$ and $\dot{P}_{\rm b}$ curves), meaning that this specific scalar-tensor theory is excluded. The chosen point in the $\alpha_0 - \beta_0$ plane is not excluded by other experiments (see Fig. 14). This figure is merely to illustrate how a specific theory fails the double pulsar test. In this case, it is due to the additional energy loss from scalar GWs, predominantly the dipolar contribution. The $\dot{P}_{\rm b}$ curve is based on Eq. (44). The 2PN and Lense-Thirring contributions to $\dot{\omega}$ are calculated assuming GR, since any corrections from the scalar field at this PN level are insignificant.

provides an important contribution to the constraints on DEF gravity from the double pulsar.

The PK parameters r and δ_θ are not used in the tests presented in this section. Because of their large uncertainty, they are not relevant for any of those tests. The same is the case for effects related to the scaling parameter $q_{\rm NLO}$.

Finally, we note that the expression for the mass ratio remains unchanged, since $R \equiv m_{\rm A}/m_{\rm B} = x_{\rm B}/x + \mathcal{O}(c^{-4})$ [63].

Figure 14 shows the constraints on DEF gravity from the double pulsar, presented as a curve (boundary of allowed region) in the $\alpha_0 - \beta_0$ parameter plane. When calculating the limits, we followed the procedure outlined in Refs. [42,206]. As one can see from Fig. 14, the double pulsar provides the best limit on DEF gravity for $\beta_0 \lesssim -3$. In Fig. 15, for illustration purposes, we plot a mass-mass diagram for a point in the $\alpha_0 - \beta_0$ plane that was previous to this paper not excluded: $\alpha_0 = 5 \times 10^{-4}$ and $\beta_0 = -4$ (marked by $*$ in Fig. 14). As one can see for this specific example, $T_1(5 \times 10^{-4}, -4)$ is falsified by the double pulsar, in particular, by the GW test.

B. Bekenstein's TeVeS

As a second alternative to GR, we discuss Bekenstein's TeVeS [43]. In Ref. [208], TeVeS is extended to a TeVeS-like class of gravity theories with two coupling parameters α_0 and β_0, similar to the class of standard scalar-tensor theories discussed in Sec. VII A. Here, we discuss only Bekenstein's original TeVeS, which corresponds to $\beta_0 = 0$, since this is the case where the double pulsar provides the most important contribution. Although Bekenstein's TeVeS has recently been falsified by the multimessenger observation of the LIGO/Virgo merger event GW170817 [44], we still use it here as an example of a MONDian gravity theory that passes Solar-System tests but can be excluded by a binary pulsar experiment. At the same time, this is a significant update on the result presented in Ref. [208].

The crucial difference between DEF gravity in Sec. VII A and TeVeS is the specific form of the physical metric, which ensures that $\gamma^{\rm PPN} = \beta^{\rm PPN} = 1$, making TeVeS indistinguishable from GR in the weak-field slow-motion regime. Therefore, by design, TeVeS passes Solar-System tests, like the one from the *Cassini* spacecraft [207]. Interestingly, the special form of the physical metric also leads to the fact that $\gamma_{\rm AB} = 1$ in TeVeS-like gravity theories [208], different from Eq. (C2).

As discussed in Ref. [208], the predictions for the Einstein delay amplitude $\gamma_{\rm E}$ and the GW damping $\dot{P}_{\rm b}$ keep the same form as in standard scalar-tensor theories. However, $\alpha_{\rm A} = \alpha_{\rm B} = \alpha_0$, and, therefore, there is no dipolar radiation. As a consequence, dipolar radiation tests like in

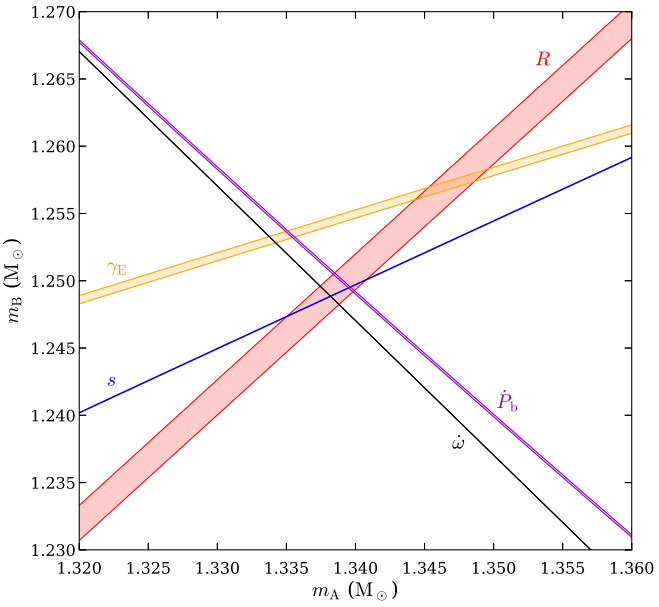

FIG. 16. Mass-mass diagram for Bekenstein's TeVeS with a "natural" transition from the Newtonian to the MONDian regime ($\alpha_0 = 0.04$) [210]. Such a theory is obviously in contradiction with the double pulsar observations. The $\dot{P}_b$ curve is based on Eq. (44). The 2PN and Lense-Thirring contributions to $\dot{\omega}$ are calculated assuming GR, since any TeVeS-related corrections for them are negligible in this test.

Ref. [208] do not provide any constraints on Bekenstein's TeVeS. The same is the case for the universality of free fall test with the pulsar in the stellar triple [76,209]. The double pulsar, on the other hand, is sensitive to the modification of the MOI due to the periodic variation in the local gravitational constant (modification of γ_E) related to the scalar field of TeVeS and has a sufficiently precise GW test that is sensitive to modifications of the quadrupolar GW damping due to the presence of the dynamical scalar field. This is obvious from Fig. 16, which gives the mass-mass diagram for Bekenstein's TeVeS with a coupling parameter α_0 that guarantees a natural transition from the Newtonian to the MONDian regime (see discussion in Ref. [210]). The deviations in γ_E and $\dot{P}_b$ clearly falsify that theory. The current double pulsar observations require $|\alpha_0| < 9 \times 10^{-3}$ (95% C.L.) which requires a highly unnatural behavior of the scalar-field contribution to the gravitational acceleration (cf. Fig. 8 in Ref. [208] and Fig. 3 in Ref. [210]).

TeVeS-like theories that modify the conformal factor between the Einstein and the Jordan frame, as introduced in Ref. [208], are similarly excluded. For β_0 values significantly different from zero, the constraints do primarily come from dipolar radiation and universality of free fall tests, e.g., from limits in Refs. [76,208].

As a final comment, we emphasize that like in Ref. [208] we neglect the contributions of the vector field to the dynamics and the gravitational radiation. We consider this as a conservative approach, since a dynamical vector field is expected to increase the GW damping by extracting additional energy from the orbital dynamics. Furthermore, modifications of TeVeS that modify only the vector part and leave the scalar sector of the theory unchanged should, therefore, generally be excluded by this test as well (see discussion in Ref. [210] on the role of the scalar field to reproduce the MONDian dynamics).

VIII. ASTRONOMICAL IMPLICATIONS

In this section, we complete our analysis by revisiting several important properties of the double pulsar system. We first discuss results relating to the system distance, comparing our best estimate with the DM-based estimates based on the two principal Galactic electron-density models. The observed discrepancies inform how the model parameters of ISM structures in the path to the system should be modified. We next use the observed scintillation properties and DM variations over a 13-yr data span (Sec. IV B) to investigate the structure of the ISM over 6 orders of magnitude in fluctuation scale. Finally, we consider the implications of the observed system proper motion and implied transverse velocity on formation models for the system and summarize the evidence for a prograde rotation sense for the A pulsar (Sec. VI B 5).

A. The distance to the double pulsar

The astrometric results derived from timing (Sec. IV B) and from interferometric data (Sec. IV C) can be compared against each other and with earlier VLBI observations using the Australian Long Baseline Array [93]. A detailed comparison between the two VLBI measurements is presented in Appendix A, in which we demonstrate that our new VLBA results are more reliable than the previous Long Baseline Array results, in part due to a previously unappreciated source of systematic error (refractive image wander). Accordingly, we focus on the VLBA results presented in this work.

The estimated parallax from the VLBA measurements is $\pi_v = 1.30^{+0.13}_{-0.11}$ mas (Sec. IV C). This contrasts with the value estimated from the pulsar timing of $\pi_t = 2.15 \pm 0.48$ mas (Sec. IV B). The implied pulsar distances are, respectively, 770 ± 70 and 465^{+134}_{-85} pc. As described in Appendix A, we adopt a weighted mean of these two distances, 735 ± 60 pc, as our best estimate of the pulsar distance. Note that none of these distance estimates change beyond their uncertainties if one applies the Lutz-Kelker correction (cf. Refs. [211,212]). Refractive image wander also potentially has an effect on the timing-based astrometry. We show in Appendix B 2 that this effect is negligible, changing the values of the astrometric parameters by $\lesssim 0.2\sigma$.

We now consider the impact of the pulsar distance on other observables to see if they can assist in determining the most likely parallax value.

The estimated distance for the pulsar based on the NE2001 Galactic electron-density model [45] is 516 pc and, based on the YMW16 model [46], 1105 pc. We note, however, that the derivation of the YMW16 electron-density model includes the earlier VLBI-based distance [93] (corrected for the Lutz-Kelker bias) of 1100 pc in its set of independently measured distances used to calibrate the model. Removing the double pulsar from the list of independent distances and redetermining the parameters of the YMW16 model has a surprisingly large effect on the estimated distance, changing it from 1105 to 463 pc [213]. Further investigation shows that the reason for this large change is that, in the modified analysis, the radius of the model component representing the Gum Nebula changed from 125 to 128 pc. This has the effect of placing the LOS to the double pulsar just within the Gum Nebula, which is centered at about 450 pc from the Sun [214], rather than just outside it. The resulting additional electron column density is sufficient to place the pulsar within the Nebula at 463 pc. The original YMW16 distance for the double pulsar, 1105 pc, is very close to the adopted independent distance of 1100 pc. This, and the large change when this independent distance is removed from the model, highlights the fact that the number of independent distances used to build the original YMW16 model (viz., 189) is relatively small. In the intervening years, many more independent estimates of pulsar distances have been published, most notably the VLBI pulsar parallax study PSRπ [215]. Consequently, future generations of the YMW16 model will be less susceptible to this problem.

Of course, the real Gum Nebula does not have the relatively sharp and well-defined edge assumed by the YMW16 model. It is a complex region of ionized gas that is believed to be a greatly expanded remnant of a supernova explosion, a fossil H II region, a wind-blown bubble, or some combination of these (see Ref. [214] for an Hα image of the nebula and a summary of its properties). The Vela supernova remnant is superimposed on the Nebula and probably lies within it but is not the source of the bulk of the ionization. The Gum Nebula is roughly circular in shape, with an angular radius of about 23°, corresponding to a physical radius of about 190 pc. It has large extension to the west (i.e., lower right ascension or Galactic longitude) which covers the projected direction of the double pulsar. The identification of this extended nebulosity as the location of the electron density fluctuations responsible for the DM variations and scintillation of the double pulsar is discussed in Sec. VIII B below.

B. ISM properties

In addition to the *average* DM value used above, we can also make use of the observed *variations* of the DM shown in Fig. 3, in order to infer properties of the intervening ISM and to compare these with expectations. These variations are consistent with our LOS moving across electron-density enhancements and deficits in the ISM (see, e.g., Ref. [216]). These density fluctuations are part of the same plasma turbulence that causes angular scattering and intensity scintillation. The diffractive intensity scintillations of pulsar A have been thoroughly analyzed [92], and our observations do not have sufficient frequency resolution to improve on this work. In the following analysis, diffractive parameters are taken from Ref. [92]. They show that the scattering screen is located at about 30% of the distance from Earth to the pulsar or about 220 pc from Earth for the adopted pulsar distance of 735 pc. If the extended nebulosity on the western side were somewhat closer to Earth than the Gum Nebula itself, we could identify the scattering screen with this nebulosity.

We also note that the rotation measure (RM) of the double pulsar is $+120.8 \pm 0.2$ rad m^{-2} [154], which is consistent both with the RM of other pulsars located near the Gum Nebula (in projection) and with the RM of extragalactic sources located behind the western nebulosity near the double pulsar [214]. Both of these are consistent with the double pulsar being located behind the western extension and much of the Faraday rotation occurring in this region.

Structure functions (SFs) are useful here, because the diffractive scattering can be calculated directly from the SF, and stochastic processes with a power-law spectrum have a power-law SF. The SF for DM variations is defined by

$$D_{\mathrm{DM}}(\tau) = \langle [\mathrm{DM}(t+\tau) - \mathrm{DM}(t)]^2 \rangle, \tag{57}$$

where τ is the lag between DM samples (see, e.g., Ref. [217]).

To obtain the most reliable estimates of the DM variations and the corresponding SF, we restrict our analysis to the interval MJD 53500–58300, which has better sampling in both time and frequency compared to earlier and later times (see Fig. 2). We use the results of the Monte Carlo analysis described in Sec. IV B to estimate the DM variations at 24-d intervals. The resulting DM-variation curve is fully consistent with that shown in Fig. 3 for the corresponding interval.

The structure function estimated directly from Eq. (57) is shown at the upper right in Fig. 17. At lags between 100 and 1000 d, the observed SF is approximately power law with an exponent much smaller than the Kolmogorov value of 5/3. [218] A linear fit to the region between 96 and 1008 d is shown as a red line on the plot. The SF is less useful for lags greater than about 2500 d (i.e., half the data span), because there are few independent samples to be averaged.

The LOS velocity (excluding the binary motion) is about 37 km/s [92], so the timescales that are easily measurable in DM(t), 100–1000 d, correspond to spatial scales of approximately 2.0 to 20 A.U. [219] The intensity scintillations, unlike DM(t), have a characteristic spatial scale s_0

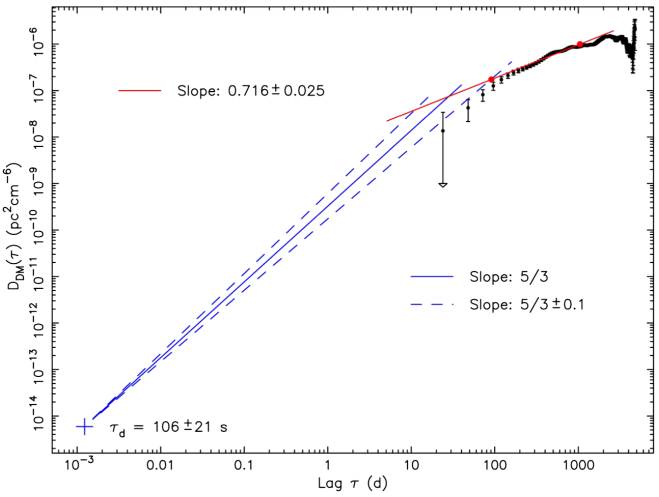

FIG. 17. Structure function of DM from MJD 53500 to 58300 with 24-d sampling. The point marked with a + is derived from diffractive timescale measurements at 820 MHz [92]. The blue solid line shows the extrapolation from the diffractive scale up to the measured lags assuming a Kolmogorov electron-density fluctuation spectrum. The dashed lines indicate ±0.1 in the spectral exponent. The red line shows the result of a weighted fit of a power law to the measured structure function for lags between 96 and 1008 d, as indicated by the red dots.

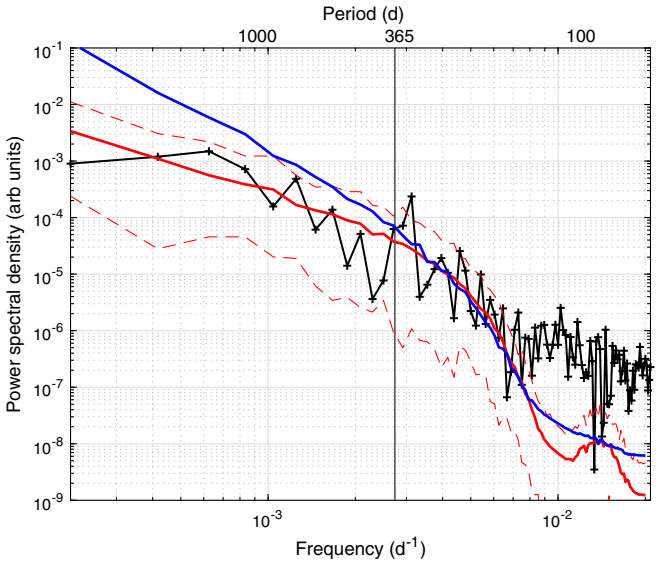

FIG. 18. Power spectral density for the DM variations used to form the SF in Fig. 17 (segmented black line). The blue and red solid lines are means of 100 simulations of the 1D PSD of DM variations given by Tempo2 for fluctuation spectral exponents of $-8/3$ (Kolmogorov) in blue and -1.5 in red. The dashed red lines give the 90% range of the -1.5 simulations. The vertical black line denotes the annual frequency.

which is very much smaller at approximately 3700 km. The spatial autocorrelation function (ACF) of the scintillations is given by $\mathrm{ACF}(s) = \exp[-D_\phi(s)]$, where $D_\phi(s)$ is the structure function of the phase shift ϕ caused by the scattering plasma. The phase delay is the negative of the group delay, so we can obtain the phase directly from DM by computing the group delay and multiplying by $2\pi\nu$. A measurement of s_0 at ν provides $D_\phi(s_0) = 1$, which defines $D_{\mathrm{DM}}(s_0)$. It should be noted that the timescale of the scintillations is extremely variable, because the binary velocity is much higher than the mean LOS velocity. The analysis of Ref. [92] fits the binary velocity to extract s_0 which is not time variable.

In this way, the observed diffractive intensity scintillations can be used to define the point in Fig. 17 marked with a + near lag 110 s. A blue line with the Kolmogorov slope of 5/3 is drawn from the diffractive point to meet the red line at a lag of approximately 30 d. The dashed lines in Fig. 17 illustrate the effect of an uncertainty in the slope of the extrapolation from the diffractive point, indicating a likely range for the intersection lag of between 10 and 100 d, or a spatial scale of 0.2 to 2.0 A.U. using the mean ISM velocity given by Ref. [92].

The location of the spectral break is important, because it can be used to estimate the thickness of the scattering region which is causing both the DM(t) fluctuations and the intensity scintillations. Accordingly, we simulate 100 realizations of a pure power-law process with an SF exponent of +0.5 and apply the convolutions by a 100-d triangle implicit in the Tempo2 analysis and the 24-d binning of the Monte Carlo results. Details of the simulation methods are described in Appendix B 1.

We first adjust the amplitude and spectral exponent of the simulated DM(t) so its power spectral density (PSD) matches that of the observations. [220] The match is much clearer in the PSD than in the SF, because the PSD estimation errors are statistically independent and those of the SF are highly correlated. Figure 18 shows the PSD of the observed DM variations as a segmented black line. The mean PSD of the simulations is overplotted with the 90% confidence limits, and both the level and the spectral exponent are adjusted so the observed PSD fits within these confidence limits. A Kolmogorov PSD simulation is also shown to make it clear that the observations cannot be fitted with a Kolmogorov spectrum. The flatter spectrum with exponent -1.5 agrees well with the observations up to frequencies of approximately 7×10^{-3} d^{-1}. At higher frequencies, the observed PSD is dominated by white noise, which is not included in the simulations.

The observed PSD also shows a statistically significant peak at a frequency close to the annual frequency. This peak arises from the annual motion of Earth moving the LOS to the pulsar across transverse A.U.-scale gradients in the integrated screen density. For further details on both the effect of the Tempo2 convolutions and the annual feature in the PSD, see Appendix B 1.

Figure 19 shows the measured SF from Fig. 17 in black on an expanded scale. The structure functions of the best-fit simulations with a spectral exponent of -1.5 are calculated, and the mean and 90% confidence limits are shown in red.

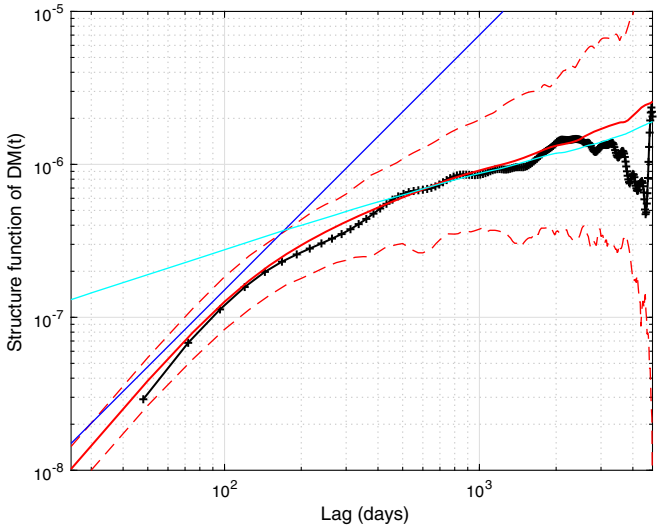

FIG. 19. Comparison of the observed structure function for DM variations with white noise removed (in black) with simulations for an input SF exponent of +0.5. The solid red line is the mean of 100 simulations including the effect of the time-domain convolutions, and the red dashed lines give the 90% range. The cyan line is the mean of the simulations before the time-domain convolutions. The blue line is the Kolmogorov extrapolation (SF exponent 5/3) from the diffractive scale.

The theoretical SF simulated before convolutions required in the data analysis is indicated by the cyan line, and the blue line is the extrapolation from the diffractive timescale, also shown in blue in Fig. 17.

The steepening of the SF toward lags shortward of a few hundred days results from the time-domain convolutions. However, there must be a true break in the SF slope from a value close to the Kolmogorov 5/3 at short lags to a flatter slope at lags greater than a few hundred days. The intersection of the blue and cyan lines at a lag of approximately 170 d, which corresponds to a linear scale of about 3 A.U., probably gives the best estimate of the frequency of spectral break in the ISM fluctuations.

A break in the PSD from a Kolmogorov spectral exponent of $-8/3$ to an exponent of $-5/3$ occurs naturally in turbulence in a thin layer, and the transition frequency is directly related to the thickness of the layer. As discussed by Lay [221], the inverse of the transition frequency for the corresponding PSD corresponds to twice the screen thickness. To relate the SF transition lag and PSD transition frequency, we simulate our observed DM(t) 1000 times with an SF transition at 170 d and compute the average PSD, which we find has a transition scale of 1/800 d. Accordingly, the screen depth must correspond to a lag of about 400 d or a spatial scale of about 8 A.U. The mean electron density integrated through the scattering layer must exceed its rms fluctuation, since it is positive definite and the distribution is roughly symmetric. The rms of DM(t) is about 0.001 pc cm^{-3} (Fig. 3), and, therefore, a lower limit on the mean density in the scattering region is about 25 cm^{-3}.

As mentioned in Sec. VIII A above, the scattering screen can be identified with the nebulosity to the west of the Gum Nebula, which is probably associated with the Nebula but somewhat closer to Earth. From the observed Hα intensity, Purcell *et al.* [214] infer emission measures in the fainter parts of the Nebula that correspond to a mean electron density of about 1.5 cm^{-3}. This estimate is appropriate for the nebulosity in front of J0737–3039A/B. Recognizing that a filament with a transverse scale of even 10 000 A.U. would be invisible on the Hα image (which has a resolution of 6 arcmin), an overdensity by a factor of 20 or even more in the scattering screen of depth 8 A.U. is quite plausible. We, therefore, conclude that the western nebulosity associated with the Gum Nebula is a likely location for the scattering screen. This is consistent with J0737–3039A/B being located at our preferred distance of 735 pc.

The clear flattening of the spectral exponent at a scale of 3 A.U. is very unusual in ISM observations, which usually follow the Kolmogorov exponent. Of the pulsars observed by pulsar timing arrays, which provide DM(t) with the required precision, only PSR J1713+0747 shows similar behavior. It should be noted that the diffractive scintillation of the double pulsar is entirely normal for pulsars with similar DM. Only the flattening of the exponent due to the unusually thin scattering region is abnormal.

C. System velocity and geometry

Using the distance adopted from the weighted mean probability distribution in Fig. 23, i.e., 735 ± 60 pc, and taking the parameters of Ref. [130], in particular, the Solar distance to the Galactic midplane, $z_{\rm SSB} = +14$ pc, [222] we find that the double pulsar is about 43 pc below the Galactic plane.

The proper motion of the double pulsar is extremely well constrained. Its total value 3.304 ± 0.033 mas yr^{-1} converts into a transverse velocity with respect to the SSB of just 11.5 ± 1.0 km s^{-1}, where the uncertainty is dominated by that of the distance measurement. The components in Galactic longitude and latitude are $v_l = -10.7 \pm 0.9$ km s^{-1} and $v_b = -4.3 \pm 0.4$ km s^{-1}, respectively.

The radial velocity v_R with respect to the SSB is unknown, but, for reasonable values for v_R, the double pulsar has a peculiar velocity of $\lesssim 50$ km s^{-1} with respect to its standard of rest (see Fig. 20). In particular, the vertical component of the full Galactic velocity v_z is small ($\lesssim 5$ km s^{-1}), meaning that the double pulsar moves practically along the Galactic plane, information that can be used to constrain possible birthplaces of the double pulsar system.

The small peculiar velocity of the double pulsar is consistent with the suggested low-kick supernova explosion for the birth of B (e.g., Refs. [48,49]), although this interpretation has been challenged [47]. A low-kick

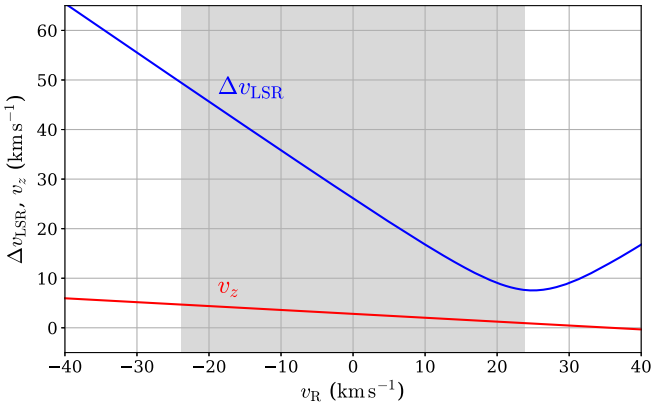

FIG. 20. Velocity of the double pulsar with respect to its standard of rest ($\Delta v_{\rm LSR}$; blue) and its velocity perpendicular to the Galactic plane (v_z; red), as a function of the unknown radial v_R. The gray area shows the 90% probability range for the radial velocity, assuming an (*a priori*) uniform probability distribution for the direction of the 3D velocity vector with respect to the SSB (cf. Ref. [48]). Calculations are based on the parameters and Galactic potential in Ref. [130], including the components for the peculiar velocity of the Sun $(U_\odot, V_\odot, W_\odot) = (8.6, 13.9, 7.1)\ {\rm km\,s^{-1}}$.

scenario is, however, also supported by the small misalignment angle $\delta_{\rm A}$ between A's spin and the total angular momentum vector, which is based on the lack of profile changes in A [50,51]. Visible secular changes in the pulse profile of A are still lacking, even with the doubling of the observation data span, leading to even tighter constraints that will be presented elsewhere.

The previous argument that the prograde solution $\delta_{\rm A} < 3.2°$ is more likely than the retrograde solution ($176.5° < \delta_{\rm A} < 180°$) is based on the justifiable notion that a large and fortuitous kick would be required for the observed alignment [51]. Here, we use our measurement of the relativistic effect of aberrational light bending to show unequivocally that A is rotating prograde (see Sec. VI B 5 for details), supporting the earlier findings of Pol *et al.* [52] using a fully independent method based on the emission properties of pulsar B. We note that B is also rotating prograde [57] but with a significant misalignment angle of $\delta_{\rm B} \simeq 50°$ [54,57].

We conclude that our results are fully consistent with theoretical considerations for the formation of the double pulsar system, as well as with the observed stability of the mean pulse profile.

IX. PROSPECTS

The results presented in the previous sections verify our past expectations on the development of the precision in the timing parameters (cf. Ref. [33]). This indicates that not only has the instrumentation improved as expected, but also that possible, as yet undetected, effects have not yet impacted on our timing observations or their interpretation. In particular, this is true for relativistic spin precession, which could lead to a changing pulse profile, hence making the timing more difficult and potentially less accurate (see, e.g., Refs. [224,225]). In fact, spin precession rendered B temporarily undetectable [58,226], so that the precision of the (direct) mass ratio measurement [22,31], which requires timing of both A and B, cannot easily be improved. We refer to the recent work by Noutsos *et al.* [57] for a detailed study of possible future improvements in this additional constraint.

The results presented here are not affected by the disappearance of B. The contrasting lack of any profile changes in A further strengthens the notion that the spin direction of A is closely aligned with the total angular momentum vector of the system (see Sec. VIII C). In this case, A will continue to be visible and available for precision timing for the foreseeable future, enabling improvements on all parameters and effects studied here. This includes the new effects presented in this work, which will become even more relevant in the future. For instance, in the future, the system's acceleration relative to the SSB will become a crucial parameter, while we now also have to take into account the relativistic mass loss due to the pulsar spin down and the NS structure. This will allow us eventually to convert our current constraints on the MOI and the NS radius into concrete measurements. Corresponding detailed computations and simulations are presented by Hu *et al.* [110], who show that an MOI measurement with 11% accuracy by 2030 is likely. Combination with results from other sources, such as GW emission from NS-NS mergers, to determine the EOS sufficiently well, would allow for a 7% test of Lense-Thirring precession or, alternatively, provide a 3σ measurement of the NLO GW damping in GR.

Obviously, future studies of the double pulsar will benefit greatly from more sensitive telescopes. For example, with observations such as those with MeerKAT [110,154], we can study possible profile changes of A due to signal deflection at B near superior conjunction. Moreover, the much-improved sensitivity should also result in a better timing performance during the eclipse, in general, allowing us to probe this important part of the Shapiro delay curve in more detail for improved NLO tests. Detailed eclipse studies will allow us to track the rotations of B even without seeing its radio signal, also helping to improve the mass-ratio measurement [54,154]. Finally, continued studies (timing, eclipse, and profile studies) of A will improve the precision of all relativistic effects

Future studies will also include a Bayesian-based analysis of the timing data. Currently, the number of TOAs to be studied typically overwhelms the available algorithms and computing power. Optimization work, including studies to deploy smoothly varying, frequency-dependent templates [160,161], is underway, with initial results on a subset of data being fully consistent with the results shown here.

In addition to timing observations, information obtained via different routes will further aid the exploitation of the system. First, further VLBA astrometric imaging observations can reduce the current uncertainty of the VLBI parallax and, hence, of the distance. A factor-of-2 reduction in the VLBI parallax uncertainty is within reach of an extended campaign, which would lead to a distance precision of 5%. Second, as mentioned in Sec. VI A, scintillation measurements can result in an independent measurement of the orbital inclination angle. This is one of the goals of ongoing MeerKAT observations [154].

X. SUMMARY AND CONCLUSIONS

We have presented results from a timing campaign for the double pulsar with a duration of more than 16 years, using data from six different telescopes. In order to analyze the data adequately, we have introduced a number of new methods. To facilitate analyses of the system astrometric parameters and DM variations, we formed a dataset with 4-min integration time per TOA, which provided the widest possible coverage in terms of time and observing frequency. Analysis of the binary parameters, including the important PK parameters, requires a second dataset with TOAs based on 30-s subintegrations in order to optimally resolve the fast compact orbit of PSR J0737–3039A. The analysis of this latter dataset required also the development of a new timing model and an improved implementation in Tempo. This takes into account higher-order effects, including velocity-dependent terms in the Shapiro delay. With these efforts, our results not only improve on precision of the previously measured parameters [31], but also reveal newly measured effects.

Using a Monte Carlo analysis, we obtained a probability distribution for the weighted mean of the VLBI and pulse timing annual parallaxes to obtain our best estimate of $1.36^{+0.12}_{-0.10}$ mas for the parallax and 735 ± 60 pc for the distance of the double pulsar. We emphasize again that, because of the fortunate arrangement of the secular (Shklovskii) and the Galactic acceleration being of opposite sign for the double pulsar, the impact of distance uncertainties is minimal for our current GW emission test [177]. This will eventually change, but continued VLBI and timing observations should converge onto a higher-precision distance well before any limitations are reached for the GW emission test.

We then discussed in detail the contributions to the timing model that need to be considered given the much-improved precision of our measurements. Here, we considered, in particular, higher-order timing effects, extrinsic modifications to our observed timing parameters, and also spin contributions. The latter requires us, for the first time, to pay attention to the EOS of NSs when interpreting pulsar timing data.

In this work, we obtained more independent tests of GR than are possible in any other system. We summarize the six PK parameters measured in this work in Table V, additionally including the test of relativistic spin precession. The measurements allow us to test conservative aspects of the orbital dynamics of two strongly self-gravitating masses up to 2PN order, including an approximately 1σ constraint on the Lense-Thirring contribution, which, in turn, could be used to constrain the MOI of pulsar A under the assumption of GR. A signal propagation test resulted in a confirmation of GR at the 4×10^{-4} level for the propagation of photons in the gravitational field of a strongly self-gravitating (material) body (see Sec. VI B 5). Moreover, NLO contributions are clearly present in the timing data, confirmed with a precision of about 10%. The most precise GR test available with our data probes the radiative aspects of GR, yielding a test at 2.5PN level in the equations of motion with a precision of 1.3×10^{-4} (95% C.L.; see Sec. VI B 2). In terms of overall fractional precision, this is the most precise test of GR's predictions for GW emission currently available.

TABLE V. Summary of the relativistic effects measured in our analysis and list of the resulting independent strong-field tests of GR. For each test, the remaining PK parameters and the mass ratio are used to determine the masses of pulsars A and B as input for GR predictions. In addition, parameters that test the significance of specific higher-order contributions in the advance of periastron and the signal propagation are given.

Relativistic effect	Parameter	Obs./GR pred.
Shapiro delay shape	s	1.00009(18)
Shapiro delay range	r	1.0016(34)
Time dilation	$\gamma_{\rm E}$	1.00012(25)
Periastron advance	$\dot{\omega} \equiv n_{\rm b} k$	1.000015(26)
GW emission	$\dot{P}_{\rm b}$	0.999963(63)
Orbital deformation	δ_θ	1.3(13)
Spin precession	$\Omega_{\rm B}^{\rm spin}$	0.94(13)[a]
Tests of higher-order contributions		
Lense-Thirring contrib. to k	$\lambda_{\rm LT}$	0.7(9)
NLO signal propagation	$q_{\rm NLO}$[total]	1.15(13)
… from signal deflection	$q_{\rm NLO}$[deflect]	1.26(24)
… from signal retardation	$q_{\rm NLO}$[retard]	1.32(24)

[a]Determined in Ref. [54].

The new effects that we detected include a relativistic deformation of the orbit and higher-order contributions to the Shapiro and aberration delay. The latter allow us to infer the spin direction of A as prograde, confirming earlier results that require a low-kick supernova as the formation process for pulsar B. We can determine the masses of pulsars A and B with a precision of 10^{-5} (modulo an unknown Doppler factor, which is expected to deviate by less than 10^{-4} from unity), which we can combine with the first constraints on the NS MOI ever obtained from pulsar timing. With the double pulsar, we are able to greatly improve the measurement of orbital period decay resulting from GW damping compared to the current best

measurement from the Hulse-Taylor binary system. The second-largest contribution to the observed orbital period decay is related to the effective mass loss of pulsar A, which now also has to be taken into account. Improving the precision of the double pulsar tests even further will, from now on, be constrained by our ability to correct for kinematic effects.

We pointed out that a direct comparison of tests of PN inspiral phase coefficients with different compact objects (BHs vs NSs) as well as different gravity regimes (mildly relativistic vs highly relativistic strong field) comes with certain caveats. Stating this, it is nevertheless obvious that our high-precision timing tests superbly complement the LIGO/Virgo tests, which currently are less precise at low-GW orders but allow us to probe higher-order contributions to the GW emission. This can be seen more clearly in approaches based on different theories of gravitation. For instance, in DEF gravity (Sec. VII A), the constraints from the double pulsar are considerably tighter than those from the double-NS merger GW170817 (binary BH mergers do not provide any constraints for DEF gravity). Section VII describes two alternatives to GR, namely, DEF gravity and Bekenstein's TeVeS, in some detail. We show that the new double pulsar results constrain effects that one would typically expect from certain modifications of GR including dipolar radiation and a periodic change of the MOI due to a varying local gravitational constant along the orbit of pulsar A. The double pulsar observations presented here lead to further constraints in the two-parameter space of DEF gravity and result in an additional falsification of TeVeS, which is qualitatively different to that from GW170817 [44], by improving on the limits in Ref. [208]. We also show that the double pulsar can be used to improve constraints on a time-varying gravitational constant, in particular, for effects related to strong gravitational fields. However, a more complete analysis is still needed.

The double pulsar system is, so far, the only double-NS system where the orientations of both spins relative to the total angular momentum vector can be measured. Combined with a superb knowledge about the systemic velocity and the very high precision measurements of both NS masses, the system is a unique laboratory to study binary evolution, the mechanisms of core-collapse supernovae, and the formation and structure of NSs. Obviously, the double pulsar is a rich laboratory for a wide variety of physics and astrophysics but especially for testing GR and its alternatives. For the latter application, past and future observations of the system provide constraints that are highly complementary to other methods, such as observations with GW detectors, the study of (supermassive) BHs via VLBI imaging [227], or orbits of stars [228] and flares [229]. Figure 21 demonstrates one way of using a two-parameter space to illustrate the complementarity of different gravity experiments and put the experiments presented in this paper into context. The first parameter is the potential of the gravitational interaction Φ. This is typically the (external) potential probed by a photon or a mass (e.g., a pulsar) in the gravitational field of another mass. Φ/c^2 represents a quantity that typically enters the PN and post-Minkowskian approximation schemes. As a second parameter, we have chosen the maximum spacetime curvature in the system, $\xi_{\max}$ defined as the square root of the Kretschmann scalar. Among material bodies, $\xi_{\max}$ allows distinguishing between weakly and strongly self-gravitating masses. For a given EOS, a NS shows a monotonic relation between $\xi_{\max}$ and, e.g., its surface potential Φ_s. For NSs, one finds $\Phi_s/c^2 \sim -0.2$, which underlines the strong-field aspect of pulsar tests. For BHs, the maximum spacetime curvature (causally connected to its environment) is the one at the horizon, which is a measure for the size and mass of the BH ($\xi_{\max} \propto M_{\rm BH}^{-2}$, for a nonrotating BH). This is, for instance, of relevance for BHs in alternatives to GR, where in certain cases the magnitude of the modification decreases with increasing mass and, therefore, with decreasing $\xi_{\max}$ (see, e.g., Ref. [230]).

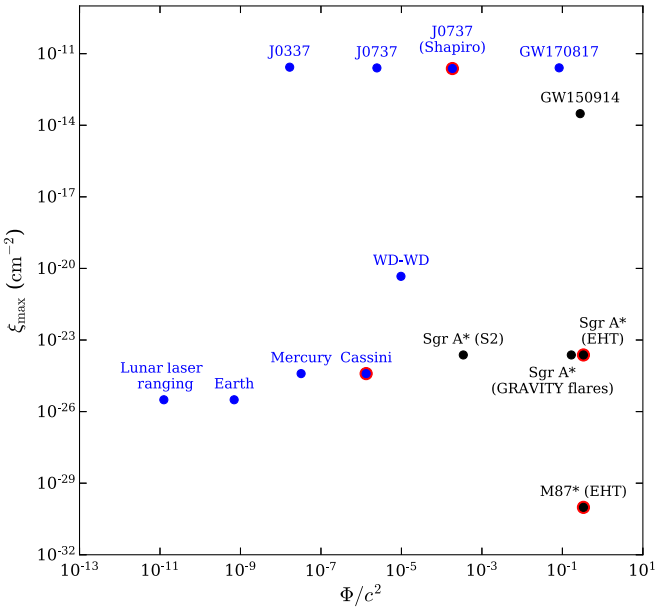

FIG. 21. Parameter space for putting the experiments of this paper into context with other gravity experiments. Φ is the potential of the gravitational interaction, and $\xi_{\max}$ is the maximum spacetime curvature in the system. Experiments involving BHs are indicated by black filled circles, whereas other experiments are reported in blue. For pulsar experiments we show the Double Pulsar ('J0737' and 'J0737 (Shapiro)') and the universality-of-free-fall test with the triple system pulsar ('J0337') [76,209]. Besides these, the following experiments are shown: solar system (cf. [41]), LIGO/Virgo mergers (cf. Section VI B 2), experiments with the supermassive BHs Sgr A* and M87* [227–229,233–235], and compact double white dwarf systems ('WD-WD'; e.g., [236]). Experiments that probe the propagation of photons in curved spacetime are highlighted by red circles. Details can be found in the text.

Figure 21 illustrates well how the double pulsar with its strongly self-gravitating components probes the mildly relativistic strong field regime. The double pulsar appears twice in Fig. 21, once for the experiments related to orbital dynamics, like GW damping (Sec. VI B 2) and periastron precession (Sec. VI B 3), and a second entry (with label "Shapiro") for the test related to photon propagation (Sec. VI B 5). It is evident that, in terms of coupling between gravitational and electromagnetic fields, the Shapiro delay test in Sec. VI B 5 is the precision timing experiment which probes the strongest spacetime curvature. When comparing the different gravity experiments in the parameter space in Fig. 21, one has to keep in mind the qualitative difference between them, for instance, BHs vs material bodies (cf. the discussion at the end of Sec. VI B 2), stationary vs dynamical or radiative situations, etc. Needless to say, a two-parameter plot cannot capture all quantities relevant for characterizing gravity tests and, therefore, always gives an incomplete comparison (for alternative parameter spaces, see, e.g., Refs. [34,228,231,232]). In conclusion, it is clear that studies of the double pulsar will continue to be extremely useful, with new applications undoubtedly awaiting us. We believe that this work presents an important milestone in this endeavor.

ACKNOWLEDGMENTS

We acknowledge the contribution of our dear colleague and collaborator Nichi D'Amico, who sadly passed away far too early. We are indebted to a number of colleagues, who have contributed to our discussions or helped with the observations. These include in alphabetical order: David Champion, Gilles Esposito-Farèse, Aditya Parthasarathy, Nataliya Porayko, Roman Rafikov, Gerhard Schäfer, and Vivek Venkatraman Krishnan. We are especially grateful to Jumei Yao for exploring the YMW16 electron density model for our new astrometric constraints, to Michael Keith for his contributions to optimization of Tempo2 for operation with large datasets, to Victoria Grandy for calibrations of much of the GBT data, and to Lijing Shao for carefully reading the manuscript and many valuable comments that helped to improve it. We thank the staff of all observatories for their continued help over the years. The Green Bank Observatory is a facility of the National Science Foundation (NSF) operated under cooperative agreement by Associated Universities, Inc. The National Radio Astronomy Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. The Parkes radio telescope is part of the Australia Telescope National Facility (grid.421683.a) which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. The Nançay Radio Observatory is operated by the Paris Observatory, associated with the French Centre National de la Recherche Scientifique (CNRS). This publication is also based on observations with the 100-m telescope of the Max-Planck-Institut für Radioastronomie at Effelsberg. The Westerbork Synthesis Radio Telescope is operated by the Netherlands Institute for Radio Astronomy (ASTRON) with support from The Netherlands Foundation for Scientific Research (NWO). Pulsar research and access to the Lovell telescope is supported by consolidated grants from the Science and Technology Facilities Council in the United Kingdom. We acknowledge financial support from the Action Fédératrice PhyFOG funded by Paris Observatory and from the "Programme National Gravitation, Références, Astronomie, Métrologie" (PNGRAM) funded by CNRS/INSU and CNES, France. Pulsar research at U. B. C. is supported by an NSERC Discovery Grant and by the Canadian Institute for Advanced Research. J. W. M. is supported by a CITA Postdoctoral Fellowship. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference No. CITA 490888-16. A. T. D. is supported by an Australian Research Council Future Fellowship (FT150100415). D. R. L. acknowledges support from the NSF Grants No. AAG-1616042 and No. OIA-1458952. P. C. C. F. gratefully acknowledges financial support by the European Research Council, under the European Unions Seventh Framework Program (FP/2007-2013) Grant Agreement No. 279702 (BEACON), and continuing support from the Max Planck Society. This work is supported by the Max-Planck Society as part of the "LEGACY" collaboration on low-frequency gravitational wave astronomy. M. K., N. W., and G. D. acknowledge financial support by the European Research Council for the ERC Synergy Grant BlackHoleCam (ERC-2013-SyG, Grant Agreement No. 610058). Finally, we thank three anonymous referees for their kind comments and suggestions, which helped us to improve the manuscript.

APPENDIX A: VLBI OBSERVATIONS, DATA REDUCTION, AND VERIFICATION

A total of 18 observations were made using the VLBA between 2016 October and 2018 May (VLBA project code BD193). Each observation is one hour in duration and records 256 MHz of dual-polarization data with a center frequency of 1.56 GHz. A single scan per observation on the source ICRF J074533.0+101112 is utilized for calibrating the instrumental bandpass. The source ICRF J073038.2–320820 is used as the primary phase reference calibrator, with NVSS J073709–302710 observed in beam and contemporaneously with the pulsar to refine the calibration solutions and to define a relative position reference point. Four other background radio sources are also in beam and observed simultaneously and are subsequently used as astrometric check sources. NVSS J073709–302710 is separated by 15′ from PSR J0737–3039A/B on the sky and by 5′ to 15′ from the other background sources.

The calibration procedure we employ makes use of the pipeline described in detail in Ref. [215]. Because of the southerly location of PSR J0737–3039A/B, we relax

TABLE VI. VLBI astrometric results for PSR J0737–3039A/B and in-beam background sources. Listed are reference position in right ascension (R.A.) and declination (Dec.), proper motion in each coordinate, parallax, and reference epoch. The top half of the table shows results when proper motion is fixed to expected values, while the second half shows results when proper motion is freely fitted. See the text for details.

	J0737–3039A/B	J0736–3027	J0736–3037	J0737–3018	J0738–3025
RA (J2000)	$07^{h}37^{m}51^{s}.247(1)$	$07^{h}36^{m}45^{s}.378(1)$	$07^{h}36^{m}51^{s}.234(1)$	$07^{h}37^{m}09^{s}.378(1)$	$07^{h}38^{m}19^{s}.793(1)$
Dec (J2000)	$-30°39'40''.68(1)$	$-30°27'18''.77(1)$	$-30°37'46''.16(1)$	$-30°18'52''.68(1)$	$-30°25'04''.92(1)$
Parallax, π_v (mas)	$1.30^{+0.13}_{-0.11}$	$0.02^{+0.06}_{-0.07}$	$-0.24^{+0.18}_{-0.21}$	$-0.03^{+0.11}_{-0.11}$	$-0.07^{+0.11}_{-0.11}$
μ_α (mas yr^{-1}, fixed)	-2.567 ± 0.030	0.0	0.0	0.0	0.0
μ_δ (mas yr^{-1}, fixed)	2.082 ± 0.038	0.0	0.0	0.0	0.0
Position epoch (MJD)	58000	58000	58000	58000	58000
RA (J2000)	$07^{h}37^{m}51^{s}.247(1)$	$07^{h}36^{m}45^{s}.378(1)$	$07^{h}36^{m}51^{s}.234(1)$	$07^{h}37^{m}09^{s}.378(1)$	$07^{h}38^{m}19^{s}.793(1)$
Dec (J2000)	$-30°39'40''.68(1)$	$-30°27'18''.77(1)$	$-30°37'46''.16(1)$	$-30°18'52''.68(1)$	$-30°25'04''.92(1)$
Parallax, π_v (mas)	$1.43^{+0.14}_{-0.13}$	$-0.05^{+0.07}_{-0.08}$	$-0.35^{+0.27}_{-0.34}$	$-0.12^{+0.17}_{-0.17}$	$-0.10^{+0.13}_{-0.14}$
μ_α (mas yr^{-1}, fitted)	$-2.23^{+0.22}_{-0.21}$	$-0.20^{+0.15}_{-0.14}$	$-0.32^{+0.42}_{-0.48}$	$-0.25^{+0.26}_{-0.25}$	$-0.12^{+0.21}_{-0.22}$
μ_δ (mas yr^{-1}, fitted)	$2.79^{+0.45}_{-0.43}$	$0.01^{+0.11}_{-0.11}$	$0.06^{+0.61}_{-0.57}$	$-0.50^{+0.30}_{-0.32}$	$-0.03^{+0.48}_{-0.46}$
Position epoch (MJD)	58000	58000	58000	58000	58000

the elevation-based flagging to times when a source is below an elevation of 15°, although we test cutoff elevations in the range 12°–18° and find minimal variation. After the application of flagging and calibration, the calibrated visibility dataset for each of the in-beam sources is divided by a source model (derived using the concatenated observations from all 18 epochs), and the source position and uncertainty at that epoch are estimated with an image-plane fit to the resultant dataset. The division by a source model removes any source structure from the image to be fitted, leaving a pointlike source whose centroid position is unaffected by changes in resolution due to the absence of different VLBA antennas in some epochs. In addition to the statistical uncertainty recorded from the image-plane fit, an estimate for the systematic position uncertainty is made using Eq. (1) of Ref. [215], and this uncertainty is added in quadrature to form a total estimated positional uncertainty for each source at every epoch. As discussed in Ref. [215], this empirical estimator of systematic position uncertainties encapsulates likely contributions such as the differential troposphere and ionosphere between the sight lines to the pulsar and the nearby calibrators.

Again following Ref. [215], we estimate the astrometric parameters for PSR J0737–3039A/B and the in-beam background sources using a bootstrap method. The orbital reflex motion of the double pulsar is negligibly small and is not included in the fit. We repeat this process twice—once utilizing the proper motion as a free parameter in the fit and once fixing the proper motion to the expected value (zero for the background sources, and the pulsar timing result reported in Table IV for the pulsar). In this latter case, a proper motion value for right ascension and declination is randomly drawn for each bootstrap trial using the corresponding mean and standard deviation from the pulsar timing fit; i.e., the uncertainty in the pulsar timing proper motion is accounted for.

The best-fit VLBI results for PSR J0737–3039A/B and the background radio sources used as astrometric checks are shown in Table VI. Given that the background radio sources are expected to be distant radio-emitting active galactic nuclei, the parallax and proper motion measured for these sources are expected to be consistent with zero. Uncertainties for proper motion and parallax are derived from the bootstrap procedure described above. The uncertainties on the reference position are dominated by the uncertainty in the absolute position of the in-beam reference source NVSS J073709–302710. This quantity is relatively poorly constrained, being derived purely from phase referencing from the primary calibrator ICRF J073038.2–320820, and we estimate a nominal uncertainty of 10 mas based on the approximately 2° angular separation between these two sources. Accordingly, we assign an uncertainty of 10 mas to the reference position for all sources in Table VI. While we focus on the proper motion and parallax (which are unaffected by errors in the absolute position that are constant in time) from this point forward, we do note that comparing the reference positions for PSR J0737–3039A/B obtained from timing and VLBI at a common epoch shows agreement within this nominal 10-mas uncertainty.

As our reference case, we utilize the fit in which the proper motion is constrained based on the timing value (top section of Table VI) but note that the best-fit value is consistent to within 1σ regardless of whether the proper motion is fixed or not. The results are similarly insensitive to other potential choices in the data reduction, such as a scaling factor applied to the estimated systematic uncertainty added to the position measurements or the elevation flagging limit. As an illustration of the results, we show the offset in the position in right ascension as a function of time for both PSR J0737–3039A/B (where the parallax signature is clearly seen) and the in-beam background source NVSS J073709–301853 (which shows no parallax signature, as

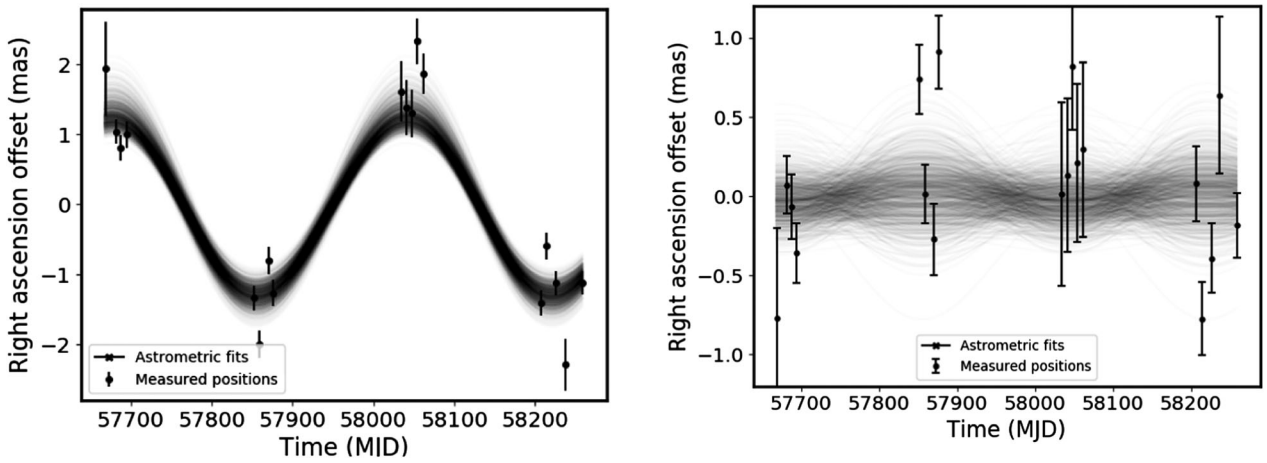

FIG. 22. Bootstrap fit plots of right ascension vs time, with the fixed proper motion subtracted. Each line shows the result of a single bootstrap trial. The left panel shows PSR J0737–3039A/B, and the right panel shows the in-beam background source NVSS J073709–301853, which shows a parallax consistent with zero as expected.

expected) after the subtraction of proper motion for our reference case in Fig. 22. Right ascension is shown for three reasons: first, because the elongated synthesized VLBA beam leads to greater precision in the right ascension axis than in declination; second, the parallax signature is larger in right ascension than declination; and third, because of the first two reasons, observations are scheduled at times of maximum parallax displacement in right ascension (and, hence, minimal parallax displacement in declination).

As we are primarily concerned with the parallax (and the corresponding uncertainty) for PSR J0737–3039A/B, the consistency of the astrometric results for the contemporaneously observed background radio sources with expectations is of considerable interest. As shown in Table VI, the fitted parallax and proper motion are consistent with expectations at the 1σ level in 75% of cases and at the 2σ level in 100% of cases. This offers a high degree of reassurance that the parallax uncertainty for PSR J0737–3039A/B is well estimated. It is, however, necessary to note that the background sources (and PSR J07373–3039A/B) span a range of angular separations from $5'$ to $15'$ to the in-beam calibrator, and at $15'$ the angular separation between PSR J0737–3039A/B and the in-beam calibrator is the equal highest of all source pairs considered. While systematic errors are likely to scale with angular separation [215], the bootstrap technique employed both here and in the study made by Deller *et al.* [215] offers a high degree of robustness in incorporating these into the estimated uncertainties.

The parallax probability distributions resulting from the bootstrap fits described above are not perfectly Gaussian, exhibiting a slight skew with a tail toward larger values.

We now consider the potential effects of refractive image wander on the VLBI results. Refraction caused by large-scale gradients in the ionized ISM can cause the apparent position of a radio source to shift with a characteristic timescale usually on the order of months to years, depending on the observing frequency and relative pulsar-ISM-Earth velocity [237]. For most radio sources, this astrometric offset would be uncorrectable but, as discussed in Sec. VIII B, for radio pulsars the time variability of the pulsar DM can be related back to gradients in the ionized ISM and, hence, used to estimate (a component of) the necessary correction. Figure 24 shows both the DM variations and the derived image wander (see Appendix B 2 for details) over the entire observing span.

We apply these corrections to the VLBA positions presented above and refit the dataset for parallax and reference position. When doing so, the best-fit parallax (from our reference-case fit with the proper motion constrained by the timing values) becomes $1.27^{+0.11}_{-0.10}$ mas—a negligible change. This result is unsurprising when considering Fig. 24, in which the derived offsets can be seen to be small at the times sampled by our VLBA observations (MJD 57670–58260.)

Since the refractive wander corrections that we can derive based on DM variations are necessarily incomplete, we also estimate a "worst-case" impact of refractive wander by sampling the corrections that would have been applied had the VLBA observations began in earlier years. We sample corrections from the VLBA observation dates shifted by an integer number of years in the past, generating a total of ten potential sets of corrections. Across these ten sets of potential corrections, the mean change to parallax is 0.03 mas, with a maximum change of 0.09 mas—smaller in all cases than the currently estimated uncertainty.

These refractive wander corrections can also be applied to historical VLBI datasets, in particular, to the Long Baseline Array work from Ref. [93]. As can be seen from Fig. 24, the impact during the time of these observations (MJD 53970–54500) is considerably greater than during

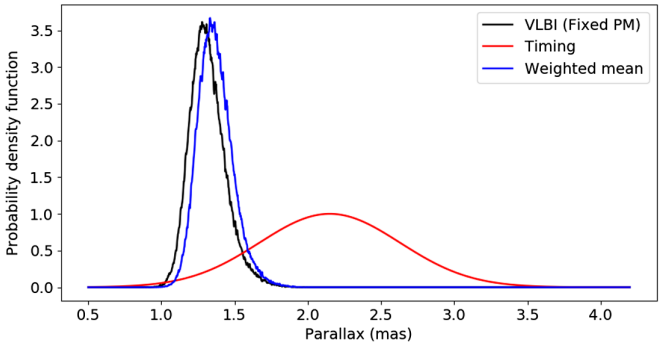

FIG. 23. Probability distribution for parallax resulting from VLBI with proper motion fixed to the timing value (black), from pulsar timing (red), and the weighted mean of the two (blue).

our VLBA observations. Moreover, the LBA parallax fit is particularly sensitive to the observation at MJD 54308 (which is the sole point taken near the parallax maximum in declination), and the estimated refractive shift is particularly pronounced at this time. After correcting the estimated refractive offsets, the best-fit parallax from the LBA dataset changes appreciably, from 0.87 to 0.96 mas. We also note that the original fit to the LBA data [93] did not have the advantage of the now well-constrained proper motion and uses a least-squares fit to estimate the parallax uncertainty rather than a bootstrap fit as we do for the newer VLBA data. Repeating the analysis of the LBA data using the same approach taken here yields a 68% confidence interval for the parallax of 0.61–1.17 mas, which becomes consistent at the approximately 1σ level with the VLBA results (but still favors a larger distance, in contrast to the timing, which favors a smaller distance). We note that the increased number of observations (by a factor of almost 3), a more compact primary calibrator source, and the presence of in-beam background sources that can be used to check for unmodeled systematic astrometric offsets all act to add extra confidence in the accuracy of the uncertainty estimation for the VLBA dataset, compared to the LBA dataset.

Finally, we combine the VLBA parallax π_v with the timing parallax π_t into a single estimate π_c. As π_v and π_t are independent, a weighted mean, with weights equal to the inverse variance of each measurement, is the best linear unbiased estimator of π_c. The variances are well estimated in both cases. In a weighted mean, one would normally compute a χ^2 value as a goodness of fit measure. A large χ^2 would prompt a search for an error. In this case, we have only two estimates to combine, so we simply take the difference $D = \pi_v - \pi_t = 0.85$. The variance of D is the sum of the variances of π_v and π_t, which gives a standard deviation of 0.51. Thus, the measured D represents a 1.7σ event—not improbable enough to cause a search for errors in π_v and π_t. We can put it as $\chi^2(1) = (0.85/0.51)^2$ and use the $\chi^2(1)$ distribution to show that the probability of $D > 0.85 = 10\%$, confirming that the two independent estimates are adequately consistent.

The probability distribution for π_c is derived using a Monte Carlo approach, in which we randomly draw a timing parallax and a VLBI parallax from their distributions and then take a weighted average to form a combined sample. This process is repeated 100 000 times to form the combined parallax probability distribution. The final result for the parallax is $\pi_c = 1.36^{+0.12}_{-0.10}$ mas (68% confidence levels), a shift of approximately 0.5σ from the VLBA-only estimate. For each iteration, we also compute the distance. Because of the compensating effects of the tail on the high side of the parallax distribution and the bias resulting from the parallax inversion, the distance probability distribution is close to symmetric, giving our best estimate for the distance of 735 ± 60 pc. We show the parallax probability distributions from our VLBA observations, from pulsar timing, and for the weighted mean in Fig. 23.

APPENDIX B: ANALYSIS OF ISM ELECTRON-DENSITY FLUCTUATIONS

1. PSD and SF simulations

The measured fluctuations DM(t) (Sec. IV B) are distorted by the fitting technique in two ways. First, the Tempo2 analysis fits a piecewise linear model of DM(t) with a sampling of 100 d. This is equivalent to convolving the data with a triangle of baseline ± 100 d. Second, the Monte Carlo scheme adjusts the phase of the DM(t) sampling and then bins it in 24-d bins. This is equivalent to convolving the data by a 24-d rectangle. To estimate the spectral exponent of the original DM(t) variations, we simulate various pure power-law processes, convolve them by the two measurement effects, and compare their power spectra with that of the observations.

We make the comparison in the spectral domain (rather than the SF), because the errors in the spectral harmonics are independent and we can adjust the level and the exponent to minimize the number of harmonics in the observed spectrum that exceed the 90% confidence limits. The samples of the SF, like those of a covariance function, are highly correlated, and this makes them appear smoother than the errors would suggest. The mean and 90% confidence limits of the best match are shown in Fig. 18 in red. We also show the mean of the expected Kolmogorov exponent in blue. Note that the spectra are relatively steep, so the PSD must be computed with prewhitening and postdarkening to avoid a bias due to spectral leakage. We prewhiten the time series with a first difference, compute the power spectrum, and correct the spectrum with the Fourier transform of the difference operator [89].

The match is good from 2×10^{-4} d^{-1} up to about 7×10^{-3} d^{-1}, where white noise begins to dominate the spectrum. It is very clear that a Kolmogorov spectrum cannot fit the data. The PSD also shows a peak near the annual frequency (at a period of 335 d). This is probably

due to Earth's orbital motion through phase gradients in the ISM. As the ISM is also moving, these phase gradients can essentially Doppler shift the annual motion to a higher frequency. The same simulation is used to compare with the observed structure function in more detail. This is shown in Fig. 19.

2. Refractive image wander

The basic phenomenon causing intensity scintillations is multipath propagation, also referred to as scattering. This both increases the apparent angular diameter of the pulsar and broadens its pulse width. Interference between the scattered rays causes the intensity scintillation. The scattering contribution to the pulse width can be estimated directly from the bandwidth of the intensity scintillations [92] as $\tau_p = 1/(2\pi\nu_{0.5}) \sim 3\ \mu$s at 820 MHz for the double pulsar. Both diffractive and refractive intensity scintillations cause fluctuations in the TOAs of the pulses on short timescales, and phase gradients cause fluctuations of the TOAs on longer scales. However, all of these fluctuations are comparable with, but less than, the pulse width [238], so we do not correct for them in this work. Phase gradients do have an important effect, because they displace the apparent position of the pulsar by $\theta_r = \nabla\phi/k$, where $k = 2\pi/\lambda$ is the propagation constant. This displacement can be comparable with the uncertainty in parallax measurements, so we examine the effect of refractive displacement in detail.

The phase gradients for PSR J0737–3039A/B are estimated directly from intensity scintillations [92]. Unfortunately, few of our observations have sufficient frequency resolution to measure the dynamic spectra of intensity scintillations accurately. However, we can use a daily interpolation of the observed DM variations to estimate their temporal gradient, which can be scaled to phase by

$$\phi = 2\pi \times 10^6 \mathrm{DM}/2.41 \times 10^{-4} \nu_{\mathrm{MHz}}. \qquad \text{(B1)}$$

We then scale the temporal gradient to a spatial gradient in the direction of the velocity by dividing by the velocity $\mathbf{v}(t)$. We must assume that there is a comparable gradient perpendicular to the velocity, since the turbulence is known to be roughly isotropic [92]; consequently, the rms gradient should be direction independent. In Fig. 24, we show the displacement θ_r at 1.56 GHz in the direction of $\mathbf{v}$. The $\mathrm{rms}(\theta_r) = 0.125$ mas, so the total rms image wander should be $\sqrt{2}$ times greater, approximately 0.177 mas.

The dates of the VLBI observations used to estimate the parallax (and, thus, the pulsar distance) are marked as blue x symbols in Fig. 24. One can see that the later VLBI observations, which are used in this paper, are at a time of relatively low phase gradients (at least in the direction of the velocity). To estimate the effect on parallax, we resolve the vector displacement θ_r into components parallel to the R.A. and Dec. axes. The actual effects of these displacements are already discussed in Appendix A.

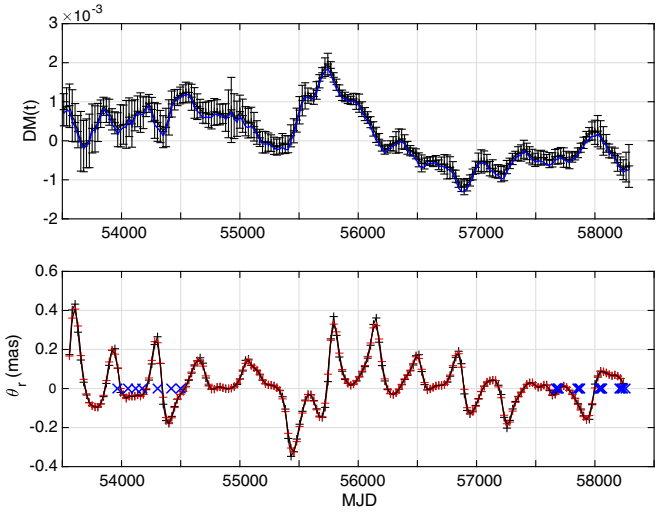

FIG. 24. Top: dispersion measure of the double pulsar as a function of time. The black line shows the interpolated DM variations, and the blue line shows these interpolated variations with the worst-case solar-wind contribution subtracted. Bottom: the image wander in the direction of the velocity θ_r (mas) at 1.56 GHz. The times of VLBI observations are marked with the blue x symbols, with the early observations being made with the LBA and the more recent observations with the VLBA and presented in this paper.

Refractive image wander also potentially has an effect on the astrometric parameters derived from pulsar timing. The timing signature of a position offset is an annual sinusoid, with the phase determined by the ratio of the perpendicular components of the offset (in R.A. and Dec., for example). For proper motion the signature is an annual sinusoid of uniformly increasing amplitude, and for parallax it is a six-month sinusoid. Since the refractive image wander (Fig. 24) has a strong annual component, attributed to the annual motion of the LOS over a gradient in the screen electron density (Appendix B 1), there is a possibility of a significant covariance between image wander and the astrometric parameters.

To assess the significance of this effect, we take the variations in refractive angle shown in Fig. 24, interpolated to daily values, and transform these to daily values of TOA offset using

$$\Delta t_{\mathrm{refr}} = 0.5\theta_r^2 \frac{d}{c}\frac{s}{1-s}, \qquad \text{(B2)}$$

where θ_r is the refractive angle in radians, d is the pulsar distance (735 pc), c is the velocity of light, and s is the

distance of the scattering screen from Earth as a fraction of the pulsar distance. These daily TOA offsets are linearly interpolated to the MJD of each TOA in the 4-min dataset and then scaled by the ratio $(f_{\rm TOA}/f_{\rm Ref})^{-2}$, where $f_{\rm TOA}$ is the radio frequency for the TOA and $f_{\rm Ref}$ is the reference frequency for the refractive angles shown in Fig. 24, 1.56 GHz. These time offsets are then added to the corresponding TOA. Comparison of the results of a standard Tempo2 analysis solving for the astrometric parameters with and without the refractive delay corrections shows that the uncertainties in the astrometric parameters are unaffected by the refractive image wander and that the changes in the values are all less than or about 0.2σ.

In this discussion, we assume that DM(t) is entirely due to the ISM. However, the observed DM(t) includes a solar-wind component. The worst-case solar-wind component would be a spherical solar wind with $n_e \propto R_{\rm Sun}^{-2}$ and $n_e = 10\ {\rm cm}^{-3}$ at Earth. We subtract this from the observations and plot the result as a blue line in Fig. 24. Evidently, the worst-case solar-wind contribution is negligible, and our assumption is justified.

APPENDIX C: PK PARAMETERS IN SCALAR-TENSOR GRAVITY

In this section, we provide expressions for the PK parameters valid for a wide class of scalar-tensor gravity (STG), including DEF gravity [40,42] used in Sec. VII A of this paper. More details can be found in reviews like Refs. [34,35,41,63,203], and the corresponding references given therein. Here, we closely follow the representation used in Ref. [42]. We restrict the discussion to PK parameters that we actually use for the tests in Sec. VII, i.e., k, $\gamma_{\rm E}$, s, and $\dot{P}_{\rm b}$.

1. Advance of periastron: k

For the advance of periastron in the presence of two strongly self-gravitating bodies, one finds in fully conservative, boost-invariant gravity theories

$$k = (2 + 2\gamma_{\rm AB} - X_{\rm B}\beta^{\rm A}_{\rm BB} - X_{\rm A}\beta^{\rm B}_{\rm AA})\frac{\hat{\beta}_{\rm O}^2}{1-e^2} + \mathcal{O}(\hat{\beta}_{\rm O}^4), \quad \text{(C1)}$$

where $\hat{\beta}_{\rm O} \equiv (G_{\rm AB} M n_{\rm b})^{1/3}/c$. $G_{\rm AB}$ is the effective gravitational constant, entering the orbital dynamics of the binary system already at the Newtonian level. In GR, $\hat{\beta}_{\rm O}$ is equal to $\beta_{\rm O}$ of Eq. (11). In STG, $G_{\rm AB} = G_*(1 + \alpha_{\rm A}\alpha_{\rm B})$, where G_* is the bare gravitational constant and α_a ($a = {\rm A}, {\rm B}$) denotes the effective scalar coupling of body a. α_a is a measure for the change of the inertial mass of body a with respect to the (asymptotic) scalar field φ_0, i.e., $\alpha_a = \partial \ln m_a/\partial\varphi_0$, where the number of baryons of the NS is kept fixed when taking the partial derivative. The parameters $\gamma_{\rm AB}$, $\beta^{\rm A}_{\rm BB}$, and $\beta^{\rm B}_{\rm AA}$ represent three body-dependent strong-field generalizations of Eddington's two weak-field PPN parameters $\gamma^{\rm PPN}$ and $\beta^{\rm PPN}$ [204]. They enter the modified Einstein-Infeld-Hoffmann equations of motion for a binary system consisting of strongly self-gravitating bodies at the 1PN level [63,137]. In GR, as a result of the effacement principle [239], $G_{\rm AB} = G$ and $\gamma_{\rm AB} = \beta^{\rm A}_{\rm BB} = \beta^{\rm B}_{\rm AA} = 1$. In STG, they are given by [63]

$$\gamma_{\rm AB} = 1 - \frac{2\alpha_{\rm A}\alpha_{\rm B}}{1 + \alpha_{\rm A}\alpha_{\rm B}}, \quad \text{(C2)}$$

$$\beta^{\rm A}_{\rm BB} = 1 + \frac{\beta_{\rm A}\alpha_{\rm B}^2}{2(1 + \alpha_{\rm A}\alpha_{\rm B})^2}, \quad \text{(C3)}$$

$$\beta^{\rm B}_{\rm AA} = 1 + \frac{\beta_{\rm B}\alpha_{\rm A}^2}{2(1 + \alpha_{\rm A}\alpha_{\rm B})^2}. \quad \text{(C4)}$$

The quantity β_a ($a = {\rm A}, {\rm B}$) is calculated according to $\beta_a = \partial\alpha_a/\partial\varphi_0$, where the number of baryons of the NS is kept fixed when taking the partial derivative with respect to the asymptotic scalar field φ_0. Depending on the coupling parameters α_0 and β_0 of DEF gravity, β_a can assume rather large values. More generally, in a system consisting of two NSs, these generalized PPN parameters can be very different from the corresponding weak-field PPN parameters. Even if there are only very small (or no) deviations from GR in the weak field, one can have order unity deviations in a double-NS system (see, e.g., Ref. [40]).

As discussed in Sec. VII A, although the $\mathcal{O}(\hat{\beta}_{\rm O}^4)$ corrections in Eq. (C1) do become relevant in the analysis of the double pulsar, presently it is still sufficient to use the corresponding GR expressions in the gravity tests presented here.

2. Amplitude of the Einstein delay: $\gamma_{\rm E}$

In STG, the amplitude $\gamma_{\rm E}$ of the time dilation (Einstein delay) gets modified as well. Apart from changes in the expressions for the time dilation, there is a modulation of the spin period of the pulsar caused by a periodic variation of the local gravitational constant, at the location of the pulsar as it moves around its companion. This effect has the same orbital dependence as the time dilation and, hence, can be absorbed in $\gamma_{\rm E}$. The resulting (total) expression for A is [42]

$$\gamma_{\rm E} = \frac{e}{n_{\rm b}} X_{\rm B}\left(X_{\rm B} + \frac{1 + k_{\rm A}\alpha_{\rm B}}{1 + \alpha_{\rm A}\alpha_{\rm B}}\right)\hat{\beta}_{\rm O}^2, \quad \text{(C5)}$$

where $k_{\rm A} \equiv -\partial \ln I_{\rm A}/\partial\varphi_0$. Here again, the number of baryons of A is kept fixed when taking the partial derivative. Note that $\alpha_{\rm A} = \alpha_{\rm B} = k_{\rm A} = 0$ in GR.

3. Shapiro shape parameter: s

The Shapiro shape parameter s can simply be identified with $\sin i$. When expressed as a function of the masses of the binary system, one uses the mass function, which is linked to the theory-dependent third Kepler law. To leading order, one finds in STG [42]

$$s = \frac{x n_{\rm b}}{X_{\rm B}\hat{\beta}_{\rm O}}. \tag{C6}$$

NLO corrections to the mass function depend on $\gamma_{\rm AB}$ [cf. Eq. (3.9) in Ref. [35]], where $\varepsilon = 2\gamma_{\rm AB} + 1$. Since such NLO corrections are at the one-sigma level of the measured s, it is sufficient to use the corresponding GR expression in Eq. (19) in the PK-parameter test.

4. GW damping: $\dot{P}_{\rm b}$

Concerning GW damping, the (dynamical) scalar field leads to various modifications of the expression for $\dot{P}_{\rm b}$. In terms of PN order, the leading contribution comes from scalar dipole radiation and enters already at the 1.5PN level, i.e., $\hat{\beta}^3_{\rm O}$. In STG, the leading-order term for dipolar GW damping is given by

$$\dot{P}^{\rm dipole}_{\rm b} = -2\pi X_{\rm A} X_{\rm B} \hat{\beta}^3_{\rm O} \frac{1+\frac{1}{2}e^2}{(1-e^2)^{5/2}} \frac{(\alpha_{\rm A}-\alpha_{\rm B})^2}{1+\alpha_{\rm A}\alpha_{\rm B}}. \tag{C7}$$

In our test, we also use those NLO corrections which are of the order of $(\alpha_{\rm A}-\alpha_{\rm B})\times\hat{\beta}^5_{\rm O}$, even though $(\alpha_{\rm A}-\alpha_{\rm B})$ has to be small, given the tight agreement with GR as seen, e.g., in Eq. (48). The rather lengthy expression for these NLO corrections can be found in Eq. (6.52b) in Ref. [203]. There are also terms of the order of $(\alpha_{\rm A}-\alpha_{\rm B})^2\times\hat{\beta}^5_{\rm O}$, which can be safely ignored.

At the 2.5PN order, one finds a modified quadrupole formula that combines quadrupolar contributions from the tensor and the scalar field. For STG, it is given by [203]

$$\dot{P}^{\rm quadrupole}_{\rm b} = -\frac{192\pi}{5} X_{\rm A} X_{\rm B} \hat{\beta}^5_{\rm O} \frac{1+\frac{73}{24}e^2+\frac{37}{96}e^4}{(1-e^2)^{7/2}} \times \frac{1+\frac{1}{6}(\alpha_{\rm A}X_{\rm B}+\alpha_{\rm B}X_{\rm A})^2}{1+\alpha_{\rm A}\alpha_{\rm B}}. \tag{C8}$$

In addition, the leading contribution from the scalar monopole radiation also enters at the 2.5PN order and in STG is given by [203]

$$\dot{P}^{\rm monopole}_{\rm b} = -3\pi X_{\rm A} X_{\rm B} \hat{\beta}^5_{\rm O} \frac{e^2(1+\frac{1}{4}e^2)}{(1-e^2)^{7/2}} \times \frac{[\alpha_{\rm A}(1+\frac{2}{3}X_{\rm B})+\alpha_{\rm B}(1+\frac{2}{3}X_{\rm A})+\frac{\beta_{\rm A}\alpha_{\rm B}+\beta_{\rm B}\alpha_{\rm A}}{1+\alpha_{\rm A}\alpha_{\rm B}}]^2}{1+\alpha_{\rm A}\alpha_{\rm B}}. \tag{C9}$$

Note that in the double pulsar this contribution is greatly suppressed, compared to the other $\mathcal{O}(\hat{\beta}^5_{\rm O})$ contributions, as it is proportional to $e^2 \approx 0.008$.

APPENDIX D: GLOSSARY OF FREQUENTLY USED SYMBOLS AND ABBREVIATIONS

For the convenience of the reader, we provide a list of frequently used symbols and their meaning. We also list abbreviations frequently used in the paper.

1. Symbols

α:	Right ascension (RA) in equatorial coordinates; spectral index
δ:	declination (DEC) in equatorial coordinates
l:	Galactic longitude
b:	Galactic latitude
μ_α:	proper motion in RA
μ_δ:	proper motion in DEC
μ_l:	proper motion in l
μ_b:	proper motion in b
v_T:	transverse velocity
π:	annual parallax
ν:	pulsar spin frequency
$\dot{\nu}$:	first spin frequency derivative
$\ddot{\nu}$:	second spin frequency derivative
$\dddot{\nu}$:	third spin frequency derivative
$\ddddot{\nu}$:	fourth spin frequency derivative
ϕ:	rotational phase
N_0:	pulse number at a reference epoch t_0
c_i:	profile evolution "FD parameters"
D:	Doppler factor between SSB and pulsar system
$P_{\rm b}$:	orbital period
$n_{\rm b}$:	orbital angular frequency, $2\pi/P_{\rm b}$
x:	projected semimajor axis
e_T:	eccentricity (Kepler equation)
e_θ, e_r:	additional eccentricities [see Eqs. (4) and (6)]
T_0:	epoch of periastron
ω, ω_0:	longitude of periastron, at T_0
$\Delta_{\rm R}$:	Rømer delay
$\Delta_{\rm E}$:	Einstein delay
$\Delta_{\rm S}$:	signal propagation delay
$\Delta_{\rm A}$:	aberration delay
u:	relativistic eccentric anomaly
θ:	true anomaly
β_O:	dimensionless orbital velocity $(GMn_{\rm b})^{1/3}/c$
k:	periastron advance parameter
$\dot{\omega}$:	secular periastron advance ($\dot{\omega} \equiv n_{\rm b}k$)
$\dot{P}_{\rm b}$:	change of orbital period
$\gamma_{\rm E}$:	Einstein delay amplitude
r:	range parameter of the Shapiro delay
s:	shape parameter of the Shapiro delay; scintillation scale; fractional scattering screen distance
t:	time
z_s:	logarithmic Shapiro shape, $z_s \equiv -\ln(1-s)$

Λ_u: argument of the Shapiro delay [see Eq. (24)]
$\lambda_{\rm LT}$: scaling parameter of Lense-Thirring precession
$q_{\rm NLO}$: NLO factor for signal propagation
δ_θ: relativistic deformation of orbit
$\mathcal{A}$, $\mathcal{B}$: aberration coefficients [see Eq. (28)]
$\dot{x}$: change of projected semimajor axis
$\dot{e}_T$: change of eccentricity
$\Omega_{\rm B}^{\rm spin}$: rate of relativistic spin precession of B
i: orbital inclination
M: total system mass
$m_{\rm A}$: mass of pulsar A
$m_{\rm B}$: mass of pulsar B
$X_{\rm A,B}$: $m_{\rm A,B}/M$
R: mass ratio $m_{\rm A}/m_{\rm B}$
$I_{\rm A}$: MOI of pulsar A
α_0, β_0: coupling strength parameters in STG
$T_1(\alpha_0,\beta_0)$: Damour–Esposito-Farèse STG
α_a: effective scalar coupling of body a ($a = $ A, B)
β_a: derivative of α_a with respect to scalar field
$k_{\rm A}$: derivative of $-\ln I_{\rm A}$ with respect to scalar field
$\gamma^{\rm PPN}$: spatial-curvature parameter of PPN formalism
$\beta^{\rm PPN}$: nonlinearity parameter of PPN formalism
γ_{ab}: body-dependent strong-field generalization of $\gamma^{\rm PPN}$ $(a, b \in \{\rm A, B, 0\})$
β^c_{ab}: body-dependent strong-field generalization of $\beta^{\rm PPN}$ $(a, b, c \in \{\rm A, B\})$

2. Abbreviations

BH: black hole
DM: dispersion measure
EOS: equation of state
GLSQ: generalized least-squares timing method
GR: general relativity
GW: gravitational wave
ISM: interstellar medium
LOS: line of sight
LT: Lense-Thirring
MJD: modified Julian date
MOI: moment of inertia
NLO: next-to-leading-order
NS: neutron star
PK: post-Keplerian
PN: post-Newtonian
PPN: parametrized post-Newtonian
PSD: power spectral density
SEP: strong equivalence principle
SSB: Solar System barycenter
SF: structure function
S/N: signal-to-noise ratio
STG: scalar-tensor gravity
TOA: time of arrival
VLBA: Very Long Baseline Array
VLBI: Very Long Baseline Interferometry

[1] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, K. Agatsuma, N. Aggarwal *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, 061102 (2016).

[2] J. H. Taylor, L. A. Fowler, and P. M. McCulloch, *Measurements of General Relativistic Effects in the Binary Pulsar PSR* 1913+16, Nature (London) **277**, 437 (1979).

[3] J. H. Taylor and J. M. Weisberg, *A New Test of General Relativity: Gravitational Radiation and the Binary Pulsar PSR* 1913+16, Astrophys. J. **253**, 908 (1982).

[4] D. R. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy* (Cambridge University Press, Cambridge, England, 2004), Vol. 4.

[5] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins, *Observation of a Rapidly Pulsating Radio Source*, Nature (London) **217**, 709 (1968).

[6] T. Gold, *Rotating Neutron Stars as the Origin of the Pulsating Radio Sources*, Nature (London) **218**, 731 (1968).

[7] M. I. Large, A. E. Vaughan, and B. Y. Mills, *A Pulsar Supernova Association*, Nature (London) **220**, 340 (1968).

[8] One parsec (pc) is approximately 3.086×10^{16} m, One kpc = 1000 pc.

[9] J. W. T. Hessels, S. M. Ransom, I. H. Stairs, P. C. C. Freire, V. M. Kaspi, and F. Camilo, *A Radio Pulsar Spinning at 716 Hz*, Science **311**, 1901 (2006).

[10] A. Lyne and F. Graham-Smith, *Pulsar Astronomy* (Cambridge University Press, Cambridge, England, 2012).

[11] N. Andersson, B. Haskell, G. L. Comer, and L. Samuelsson, *The Dynamics of Neutron Star Crusts: Lagrangian Perturbation Theory for a Relativistic Superfluid-Elastic System*, Classical Quantum Gravity **36**, 105004 (2019).

[12] P. B. Demorest, T. Pennucci, S. M. Ransom, M. S. E. Roberts, and J. W. T. Hessels, *A Two-Solar-Mass Neutron Star Measured Using Shapiro Delay*, Nature (London) **467**, 1081 (2010).

[13] J. Antoniadis *et al.*, *A Massive Pulsar in a Compact Relativistic Binary*, Science **340**, 448 (2013).

[14] H. T. Cromartie *et al.*, *Relativistic Shapiro Delay Measurements of an Extremely Massive Millisecond Pulsar*, Nat. Astron. **4**, 72 (2020).

[15] E. Fonseca *et al.*, *Refined Mass and Geometric Measurements of the High-Mass PSR* J0740+6620, Astrophys. J. Lett. **915**, L12 (2021).

[16] S. K. Greif, K. Hebeler, J. M. Lattimer, C. J. Pethick, and A. Schwenk, *Equation of State Constraints from Nuclear Physics, Neutron Star Masses, and Future Moment of Inertia Measurements*, Astrophys. J. **901**, 155 (2020).

[17] R. A. Hulse and J. H. Taylor, *Discovery of a Pulsar in a Binary System*, Astrophys. J. **195**, L51 (1975).

[18] L. L. Smarr and R. Blandford, *The Binary Pulsar: Physical Processes, Possible Companions and Evolutionary Histories*, Astrophys. J. **207**, 574 (1976).

[19] T. Damour and R. Ruffini, *Certain New Verifications of General Relativity Made Possible by the Discovery of a*

Pulsar Belonging to a Binary System, Acad. Sci. Paris C. R. Ser. Sci. Math. **279**, 971 (1974).
[20] B. M. Barker and R. F. O'Connell, *Relativistic Effects in the Binary Pulsar PSR* 1913+16, Astrophys. J. **199**, L25 (1975).
[21] M. Burgay, N. D'Amico, A. Possenti, R. N. Manchester, A. G. Lyne, B. C. Joshi, M. A. McLaughlin, M. Kramer, J. M. Sarkissian, F. Camilo, V. Kalogera, C. Kim, and D. R. Lorimer, *An Increased Estimate of the Merger Rate of Double Neutron Stars from Observations of a Highly Relativistic System*, Nature (London) **426**, 531 (2003).
[22] A. G. Lyne, M. Burgay, M. Kramer, A. Possenti, R. N. Manchester, F. Camilo, M. A. McLaughlin, D. R. Lorimer, N. D'Amico, B. C. Joshi, J. Reynolds, and P. C. C. Freire, *A Double-Pulsar System: A Rare Laboratory for Relativistic Gravity and Plasma Physics*, Science **303**, 1153 (2004).
[23] J. H. Taylor, *Pulsar Timing and Relativistic Gravity*, Phil. Trans. R. Soc. A **341**, 117 (1992).
[24] D. Nice, P. Demorest, I. Stairs, R. Manchester, J. Taylor, W. Peters, J. Weisberg, A. Irwin, N. Wex, and Y. Huang, *Tempo: Pulsar Timing Data Analysis*, 2015, ascl:1509.002, https://ui.adsabs.harvard.edu/abs/2015ascl.soft09002N.
[25] G. B. Hobbs, R. T. Edwards, and R. N. Manchester, *Tempo2, a New Pulsar-Timing Package—I. An Overview*, Mon. Not. R. Astron. Soc. **369**, 655 (2006); http://ascl.net/1210.015.
[26] A. Einstein, *Erklärung der Perihelbewegung des Merkur aus der Allgemeinen Relativitätstheorie*, Sitzungsber. K. Preuß. Akad. Wiss. **XLVII**, 831 (1915).
[27] N. Wex, in *Frontiers in Relativistic Celestial Mechanics: Applications and Experiments*, Vol. 2, edited by S. M. Kopeikin (Walter de Gruyter GmbH, Berlin, 2014), Chap. 2, p. 39.
[28] T. Damour and N. Deruelle, *General Relativistic Celestial Mechanics of Binary Systems. I. The Post-Newtonian Motion*, Ann. Inst. Henri Poincaré, A **43**, 107 (1985).
[29] T. Damour and N. Deruelle, *General Relativistic Celestial Mechanics of Binary Systems. II. The Post-Newtonian Timing Formula*, Ann. Inst. Henri Poincaré, A **44**, 263 (1986).
[30] J. H. Taylor and J. M. Weisberg, *Further Experimental Tests of Relativistic Gravity Using the Binary Pulsar PSR* 1913+16, Astrophys. J. **345**, 434 (1989).
[31] M. Kramer, I. H. Stairs, R. N. Manchester, M. A. McLaughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possenti, N. D'Amico, J. M. Sarkissian, G. B. Hobbs, J. E. Reynolds, P. C. C. Freire, and F. Camilo, *Tests of General Relativity from Timing the Double Pulsar*, Science **314**, 97 (2006).
[32] M. Kramer and I. H. Stairs, *The Double Pulsar*, Annu. Rev. Astron. Astrophys. **46**, 541 (2008).
[33] M. Kramer and N. Wex, *TOPICAL REVIEW: The Double Pulsar System: A Unique Laboratory for Gravity*, Classical Quantum Gravity **26**, 073001 (2009).
[34] N. Wex and M. Kramer, *Gravity Tests with Radio Pulsars*, Universe **6**, 156 (2020).
[35] T. Damour and J. H. Taylor, *Strong-Field Tests of Relativistic Gravity and Binary Pulsars*, Phys. Rev. D **45**, 1840 (1992).
[36] T. Damour and G. Schäfer, *Higher-Order Relativistic Periastron Advances and Binary Pulsars*, Nuovo Cimento Soc. Ital. Fis. **101B**, 127 (1988).
[37] B. M. Barker and R. F. O'Connell, *Gravitational Two-Body Problem with Arbitrary Masses, Spins, and Quadrupole Moments*, Phys. Rev. D **12**, 329 (1975).
[38] A. T. Deller, J. M. Weisberg, D. J. Nice, and S. Chatterjee, *A VLBI Distance and Transverse Velocity for PSR* B1913+16, Astrophys. J. **862**, 139 (2018).
[39] M. J. Reid and M. Honma, *Microarcsecond Radio Astrometry*, Annu. Rev. Astron. Astrophys. **52**, 339 (2014).
[40] T. Damour and G. Esposito-Farèse, *Nonperturbative Strong-Field Effects in Tensor-Scalar Theories of Gravitation*, Phys. Rev. Lett. **70**, 2220 (1993).
[41] C. M. Will, *Theory and Experiment in Gravitational Physics*, second ed. (Cambridge University Press, Cambridge, England, 2018).
[42] T. Damour and G. Esposito-Farèse, *Tensor-Scalar Gravity and Binary-Pulsar Experiments*, Phys. Rev. D **54**, 1474 (1996).
[43] J. D. Bekenstein, *Relativistic Gravitation Theory for the Modified Newtonian Dynamics Paradigm*, Phys. Rev. D **70**, 083509 (2004).
[44] S. Boran, S. Desai, E. O. Kahya, and R. P. Woodard, *GW170817 Falsifies Dark Matter Emulators*, Phys. Rev. D **97**, 041501(R) (2018).
[45] J. M. Cordes and T. J. W. Lazio, *NE2001.I. A New Model for the Galactic Distribution of Free Electrons and Its Fluctuations*, arXiv:astro-ph/0207156.
[46] J. M. Yao, R. N. Manchester, and N. Wang, *A New Electron-Density Model for Estimation of Pulsar and FRB Distances*, Astrophys. J. **835**, 29 (2017).
[47] B. Willems, J. Kaplan, T. Fragos, V. Kalogera, and K. Belczynski, *Formation and Progenitor of PSR J0737–3039: New Constraints on the Supernova Explosion Forming Pulsar B*, Phys. Rev. D **74**, 043003 (2006).
[48] I. H. Stairs, S. E. Thorsett, R. J. Dewey, M. Kramer, and C. A. McPhee, *The Formation of the Double Pulsar PSR J0737–3039A/B*, Mon. Not. R. Astron. Soc. **373**, L50 (2006).
[49] T. M. Tauris, M. Kramer, P. C. C. Freire, N. Wex, H. T. Janka, N. Langer, P. Podsiadlowski, E. Bozzo, S. Chaty, M. U. Kruckow, E. P. J. van den Heuvel, J. Antoniadis, R. P. Breton, and D. J. Champion, *Formation of Double Neutron Star Systems*, Astrophys. J. **846**, 170 (2017).
[50] R. N. Manchester, M. Kramer, A. Possenti, A. G. Lyne, M. Burgay, I. H. Stairs, A. W. Hotan, M. A. McLaughlin, D. R. Lorimer, G. B. Hobbs, J. M. Sarkissian, N. D'Amico, F. Camilo, B. C. Joshi, and P. C. C. Freire, *The Mean Pulse Profile of PSR J0737–3039A*, Astrophys. J. **621**, L49 (2005).
[51] R. D. Ferdman, I. H. Stairs, M. Kramer, R. P. Breton, M. A. McLaughlin, P. C. C. Freire, A. Possenti, B. W. Stappers, V. M. Kaspi, R. N. Manchester, and A. G. Lyne, *The Double Pulsar: Evidence for Neutron Star Formation without an Iron Core-Collapse Supernova*, Astrophys. J. **767**, 85 (2013).
[52] N. Pol, M. McLaughlin, M. Kramer, I. Stairs, B. B. P. Perera, and A. Possenti, *A Direct Measurement of Sense of Rotation of PSR J0737–3039A*, Astrophys. J. **853**, 73 (2018).

[53] M. A. McLaughlin, A. G. Lyne, D. R. Lorimer, A. Possenti, R. N. Manchester, F. Camilo, I. H. Stairs, M. Kramer, M. Burgay, N. D'Amico, P. C. C. Freire, B. C. Joshi, and N. D. R. Bhat, *The Double Pulsar System J0737–3039: Modulation of A by B at Eclipse*, Astrophys. J. **616**, L131 (2004).

[54] R. P. Breton, V. M. Kaspi, M. Kramer, M. A. McLaughlin, M. Lyutikov, S. M. Ransom, I. H. Stairs, R. D. Ferdman, F. Camilo, and A. Possenti, *Relativistic Spin Precession in the Double Pulsar*, Science **321**, 104 (2008).

[55] Note that the angle θ in Ref. [54] is (to a good approximation) $180° - \delta_B$, since in the coordinate system of Ref. [54] the orbital angular momentum is closely aligned with the negative z direction.

[56] M. Burgay, A. Possenti, R. N. Manchester, M. Kramer, M. A. McLaughlin, D. R. Lorimer, I. H. Stairs, B. C. Joshi, A. G. Lyne, F. Camilo, N. D'Amico, P. C. C. Freire, J. M. Sarkissian, A. W. Hotan, and G. B. Hobbs, *Long-Term Variations in the Pulse Emission from PSR J0737–3039B*, Astrophys. J. **624**, L113 (2005).

[57] A. Noutsos, G. Desvignes, M. Kramer, N. Wex, P. C. C. Freire, I. H. Stairs, M. A. McLaughlin, R. N. Manchester, A. Possenti, M. Burgay, A. G. Lyne, R. P. Breton, B. B. P. Perera, and R. D. Ferdman, *Understanding and Improving the Timing of PSR J0737–3039B*, Astron. Astrophys. **643**, A143 (2020).

[58] B. B. P. Perera, M. A. McLaughlin, M. Kramer, I. H. Stairs, R. D. Ferdman, P. C. C. Freire, A. Possenti, R. P. Breton, R. N. Manchester, M. Burgay, A. G. Lyne, and F. Camilo, *The Evolution of PSR J0737–3039B and a Model for Relativistic Spin Precession*, Astrophys. J. **721**, 1193 (2010).

[59] M. Lyutikov and C. Thompson, *Magnetospheric Eclipses in the Double Pulsar System PSR J0737–3039*, Astrophys. J. **634**, 1223 (2005).

[60] R. P. Breton, V. M. Kaspi, M. A. McLaughlin, M. Lyutikov, M. Kramer, I. H. Stairs, S. M. Ransom, R. D. Ferdman, F. Camilo, and A. Possenti, *The Double Pulsar Eclipses. I. Phenomenology and Multi-frequency Analysis*, Astrophys. J. **747**, 89 (2012).

[61] M. Lyutikov, *On the Nature of Eclipses in Binary Pulsar J0737–3039*, Mon. Not. R. Astron. Soc. **353**, 1095 (2004).

[62] At least up to 1PN order for boost-invariant gravity theories [63].

[63] T. Damour, in *Physics of Relativistic Objects in Compact Binaries: From Birth to Coalescence*, Astrophysics and Space Science Library Vol. 359, edited by M. Colpi, P. Casella, V. Gorini, U. Moschella, and A. Possenti (Springer, New York, 2009), pp. 1–41.

[64] C. Granet, H. Z. Zhang, A. R. Forsyth, G. R. Graves, P. Doherty, K. J. Greene, G. L. James, P. Sykes, T. S. Bird, M. W. Sinclair, G. Moorey, and R. N. Manchester, *The Designing, Manufacturing, and Testing of a Dual-Band Feed System for the Parkes Radio Telescope*, IEEE Trans. Antennas Propag. **47**, 13 (2005).

[65] L. Staveley-Smith, W. E. Wilson, T. S. Bird, M. J. Disney, R. D. Ekers, K. C. Freeman, R. F. Haynes, M. W. Sinclair, R. A. Vaile, R. L. Webster, and A. E. Wright, *The Parkes 21 cm Multibeam Receiver*, Pub. Astron. Soc. Aust. **13**, 243 (1996).

[66] G. Hobbs *et al.*, *An Ultra-wide Bandwidth (704 to 4032 MHz) Receiver for the Parkes Radio Telescope*, Pub. Astron. Soc. Aust. **37**, e012 (2020).

[67] A. Hotan, *High-Precision Observations of Relativistic Binary and Millisecond Pulsars*, Ph.D. thesis, Swinburne University of Technology, 2007.

[68] W. van Straten and M. Bailes, *DSPSR: Digital Signal Processing Software for Pulsar Astronomy*, Pub. Astron. Soc. Aust. **28**, 1 (2011).

[69] R. N. Manchester *et al.*, *The Parkes Pulsar Timing Array Project*, Pub. Astron. Soc. Aust. **30**, e017 (2013).

[70] P. B. Demorest, *Measuring the Gravitational Wave Background Using Precision Pulsar Timing*, Ph.D. thesis, University of California, Berkeley, 2007.

[71] R. DuPlain, S. Ransom, P. Demorest, P. Brandt, J. Ford, and A. L. Shelton, in *Advanced Software and Control for Astronomy II*, SPIE Proceedings Vol. 7019 (SPIE-International Society for Optical Engineering, Bellingham, WA, 2008), p. 70191D.

[72] R. D. Ferdman, *Binary Pulsars: Evolution and Fundamental Physics*, Ph.D. thesis, University of British Columbia, 2008.

[73] http://ascl.net/1509.003.

[74] W. van Straten, P. Demorest, and S. Oslowski, *Pulsar Data Analysis with PSRCHIVE*, Astron. Res. Technol. **9**, 237 (2012).

[75] V. R. Grandy, *Wideband Timing of the Double Pulsar (J0737–3039a)*, Bachelor's thesis, University of British Columbia, 2008.

[76] G. Voisin, I. Cognard, P. C. C. Freire, N. Wex, L. Guillemot, G. Desvignes, M. Kramer, and G. Theureau, *An Improved Test of the Strong Equivalence Principle with the Pulsar in a Triple Star System*, Astron. Astrophys. **638**, A24 (2020).

[77] P. Lazarus, R. Karuppusamy, E. Graikou, R. N. Caballero, D. J. Champion, K. J. Lee, J. P. W. Verbiest, and M. Kramer, *Prospects for High-Precision Pulsar Timing with the New Effelsberg PSRIX Backend*, Mon. Not. R. Astron. Soc. **458**, 868 (2016).

[78] C. G. Bassa, G. H. Janssen, R. Karuppusamy, M. Kramer, K. J. Lee, K. Liu, J. McKee, D. Perrodin, M. Purver, S. Sanidas, R. Smits, and B. W. Stappers, *LEAP: The Large European Array for Pulsars*, Mon. Not. R. Astron. Soc. **456**, 2196 (2016).

[79] R. Karuppusamy, B. Stappers, and W. van Straten, *PuMa-II: A Wide Band Pulsar Machine for the Westerbork Synthesis Radio Telescope*, Publ. Astron. Soc. Pac. **120**, 191 (2008).

[80] G. Desvignes *et al.*, *High-Precision Timing of 42 Millisecond Pulsars with the European Pulsar Timing Array*, Mon. Not. R. Astron. Soc. **458**, 3341 (2016).

[81] R. M. Shannon, S. Osłowski, S. Dai, M. Bailes, G. Hobbs, R. N. Manchester, W. van Straten, C. A. Raithel, V. Ravi, L. Toomey, N. D. R. Bhat, S. Burke-Spolaor, W. A. Coles, M. J. Keith, M. Kerr, Y. Levin, J. M. Sarkissian, J.-B. Wang, L. Wen, and X.-J. Zhu, *Limitations in Timing Precision due to Single-Pulse Shape Variability in Millisecond Pulsars*, Mon. Not. R. Astron. Soc. **443**, 1463 (2014).

[82] W. van Straten, *Radio Astronomical Polarimetry and Point-Source Calibration*, Astrophys. J. Suppl. Ser. **152**, 129 (2004).

[83] A. W. Hotan, M. Bailes, and S. M. Ord, *Geodetic Precession in PSR J1141–6545*, Astrophys. J. **624**, 906 (2005).

[84] P. B. Demorest *et al.*, *Limits on the Stochastic Gravitational Wave Background from the North American Nanohertz Observatory for Gravitational Waves*, Astrophys. J. **762**, 94 (2013).

[85] https://naif.jpl.nasa.gov/pub/naif/JUNO/kernels/spk/de436s.bsp.lbl.

[86] M. Vallisneri *et al.*, *Modeling the Uncertainties of Solar System Ephemerides for Robust Gravitational-Wave Searches with Pulsar-Timing Arrays*, Astrophys. J. **893**, 112 (2020).

[87] R. N. Caballero *et al.*, *Studying the Solar System with the International Pulsar Timing Array*, Mon. Not. R. Astron. Soc. **481**, 5501 (2018).

[88] M. J. Keith, W. Coles, R. M. Shannon, G. B. Hobbs, R. N. Manchester, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, D. J. Champion, A. Chaudhary, A. W. Hotan, J. Khoo, J. Kocz, S. Osowski, V. Ravi, J. E. Reynolds, J. Sarkissian, W. van Straten, and D. R. B. Yardley, *Measurement and Correction of Variations in Interstellar Dispersion in High-Precision Pulsar Timing*, Mon. Not. R. Astron. Soc. **429**, 2161 (2013).

[89] W. Coles, G. Hobbs, D. J. Champion, R. N. Manchester, and J. P. W. Verbiest, *Pulsar Timing Analysis in the Presence of Correlated Noise*, Mon. Not. R. Astron. Soc. **418**, 561 (2011).

[90] We note that we have also explored alternative methods, e.g., using the "DMX" method implemented in Tempo, in order to determine the astrometry and its uncertainties. Using simulated timing data with parallaxes of known sizes injected into our datasets, we determined that the method implemented here appears to be the most reliable in recovering the correct astrometry with appropriate uncertainties.

[91] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT, Cambridge, MA, 2006).

[92] B. J. Rickett, W. A. Coles, C. F. Nava, M. A. McLaughlin, S. M. Ransom, F. Camilo, R. D. Ferdman, P. C. C. Freire, M. Kramer, A. G. Lyne, and I. H. Stairs, *Interstellar Scintillation of the Double Pulsar J0737–3039*, Astrophys. J. **787**, 161 (2014).

[93] A. T. Deller, M. Bailes, and S. J. Tingay, *Implications of a VLBI Distance to the Double Pulsar J0737–3039A/B*, Science **323**, 1327 (2009).

[94] A. T. Deller, S. J. Vigeland, D. L. Kaplan, W. M. Goss, W. F. Brisken, S. Chatterjee, J. M. Cordes, G. H. Janssen, T. J. W. Lazio, L. Petrov, B. W. Stappers, and A. Lyne, *Microarcsecond VLBI Pulsar Astrometry with* PSRπ. *I. Two Binary Millisecond Pulsars with White Dwarf Companions*, Astrophys. J. **828**, 8 (2016).

[95] R. Blandford and S. A. Teukolsky, *Arrival-Time Analysis for a Pulsar in a Binary System*, Astrophys. J. **205**, 580 (1976).

[96] Depending on the literature, this is called the *argument of periastron*. We follow the naming that is more common in pulsar astronomy and, for instance, also used in Ref. [97] (see their Fig. 3.2).

[97] E. Poisson and C. M. Will, *Gravity* (Cambridge University Press, Cambridge, England, 2014).

[98] R. T. Edwards, G. B. Hobbs, and R. N. Manchester, *Tempo2, A New Pulsar Timing Package—II. The Timing Model and Precision Estimates*, Mon. Not. R. Astron. Soc. **372**, 1549 (2006).

[99] G. Schäfer and N. Wex, *Second Post-Newtonian Motion of Compact Binaries*, Phys. Lett. A **174**, 196 (1993).

[100] G. Schäfer and N. Wex, *Second Post-Newtonian Motion of Compact Binaries*, Phys. Lett. A **177**, 461 (1993).

[101] N. Wex, *The Second Post-Newtonian Motion of Compact Binary-Star Systems with Spin*, Classical Quantum Gravity **12**, 983 (1995).

[102] T. Damour and G. Schaefer, *The Two-Body Problem in General Relativity*, Acad. Sci. Paris C. R. Ser. B **305**, 839 (1987).

[103] The rotational spin frequency parameters are estimated with a standard Tempo analysis. As discussed in Ref. [89], this procedure underestimates the true uncertainty in the presence of unmodeled red timing noise. Comparison of the standard analysis with the results of a full generalized least-squares analysis shows that the true uncertainty in the spin frequency parameters is about an order of magnitude larger than the values quoted in Table IV. We emphasize that the unmodeled red noise has no effect on the orbital parameters as expected because of their much shorter timescale.

[104] Mass determination from PK parameters yields the product Gm_i $(i = \mathrm{A}, \mathrm{B})$. Consequently, whenever we express the masses in terms of solar masses $\mathrm{M}_\odot$, the numerical value is calculated via $Gm_i/(\mathcal{GM})^{\mathrm{N}}_\odot$, where $(\mathcal{GM})^{\mathrm{N}}_\odot \equiv 1.3271244 \times 10^{26}\ \mathrm{cm^3\,s^{-2}}$ is the nominal solar mass parameter as defined by the IAU 2015 Resolution B3 [105].

[105] E. E. Mamajek *et al.*, *IAU 2015 Resolution B3 on Recommended Nominal Conversion Constants for Selected Solar and Planetary Properties*, arXiv:1510.07674.

[106] T. Damour, P. Jaranowski, and G. Schäfer, *Dynamical Invariants for General Relativistic Two-Body Systems at the Third Post-Newtonian Approximation*, Phys. Rev. D **62**, 044024 (2000).

[107] R.-M. Memmesheimer, A. Gopakumar, and G. Schäfer, *Third Post-Newtonian Accurate Generalized Quasi-Keplerian Parametrization for Compact Binaries in Eccentric Orbits*, Phys. Rev. D **70**, 104011 (2004).

[108] C. Königsdörffer and A. Gopakumar, *Post-Newtonian Accurate Parametric Solution to the Dynamics of Spinning Compact Binaries in Eccentric Orbits: The Leading Order Spin-Orbit Interaction*, Phys. Rev. D **71**, 024039 (2005).

[109] L. Iorio, *Prospects for Measuring the Moment of Inertia of Pulsar J0737–3039A*, New Astron. **14**, 40 (2009).

[110] H. Hu, M. Kramer, N. Wex, D. J. Champion, and M. S. Kehl, *Constraining the Dense Matter Equation-of-State with Radio Pulsars*, Mon. Not. R. Astron. Soc. **497**, 3118 (2020).

[111] For the precision needed, we can safely use $g^{\parallel}_{S_\mathrm{A}}$ instead of g_{S_A} in Ref. [36], since $\delta_\mathrm{A} < 3.2°$, while $g_{S_\mathrm{A}} \simeq g^{\parallel}_{S_\mathrm{A}} \cos\delta_\mathrm{A}$ for the double pulsar (cf. Ref. [33]).

[112] T. Dietrich, M. W. Coughlin, P. T. H. Pang, M. Bulla, J. Heinzel, L. Issa, I. Tews, and S. Antier, *Multimessenger Constraints on the Neutron-Star Equation of State and the Hubble Constant*, Science **370**, 1450 (2020).

[113] Although the mass of A ($1.34\ \mathrm{M}_\odot$) is slightly below $1.4\ \mathrm{M}_\odot$, within the accuracy needed here, the radius constraints in Ref. [112] for a $1.4 - \mathrm{M}_\odot$ NS can also be applied to pulsar A.

[114] J. M. Lattimer, *Neutron Star Mass and Radius Measurements*, Universe **5**, 159 (2019).
[115] There are other distributions that have been derived for the MOI of pulsar A; see, e.g., Refs. [116–118]. However, for the results of this paper, this does not make any difference.
[116] Y. Lim, J. W. Holt, and R. J. Stahulak, *Predicting the Moment of Inertia of Pulsar J0737–3039A from Bayesian Modeling of the Nuclear Equation of State*, Phys. Rev. C **100**, 035802 (2019).
[117] P. Landry and B. Kumar, *Constraints on the Moment of Inertia of PSR J0737–3039A from GW170817*, Astrophys. J. **868**, L22 (2018).
[118] H. O. Silva, A. M. Holgado, A. Cárdenas-Avendaño, and N. Yunes, *Astrophysical and Theoretical Physics Implications from Multimessenger Neutron Star Observations*, Phys. Rev. Lett. **126**, 181101 (2021).
[119] J. M. Lattimer and B. F. Schutz, *Constraining the Equation of State with Moment of Inertia Measurements*, Astrophys. J. **629**, 979 (2005).
[120] Z. Arzoumanian, K. Joshi, F. Rasio, and S. E. Thorsett, in *Pulsars: Problems and Progress, IAU Colloquium 160*, edited by S. Johnston, M. A. Walker, and M. Bailes (Astronomical Society of the Pacific, San Francisco, 1996), pp. 525–530.
[121] S. M. Kopeikin, *Proper Motion of Binary Pulsars as a Source of Secular Variations of Orbital Parameters*, Astrophys. J. Lett. **467**, L93 (1996).
[122] T. Damour and J. H. Taylor, *On the Orbital Period Change of the Binary Pulsar PSR* 1913+16, Astrophys. J. **366**, 501 (1991).
[123] L. Blanchet, *Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*, Living Rev. Relativity **17**, 2 (2014).
[124] P. C. Peters, *Gravitational Radiation and the Motion of Two Point Masses*, Phys. Rev. **136**, B1224 (1964).
[125] B. R. Iyer and C. M. Will, *Post-Newtonian Gravitational Radiation Reaction for Two-Body Systems: Nonspinning Bodies*, Phys. Rev. D **52**, 6882 (1995).
[126] M. E. Pati and C. M. Will, *Post-Newtonian Gravitational Radiation and Equations of Motion via Direct Integration of the Relaxed Einstein Equations. II. Two-Body Equations of Motion to Second Post-Newtonian Order, and Radiation Reaction to 3.5 Post-Newtonian Order*, Phys. Rev. D **65**, 104008 (2002).
[127] C. Königsdörffer, G. Faye, and G. Schäfer, *Binary Black-Hole Dynamics at the Third-and-a-Half Post-Newtonian Order in the ADM Formalism*, Phys. Rev. D **68**, 044004 (2003).
[128] L. Blanchet and G. Schäfer, *Higher Order Gravitational Radiation Losses in Binary Systems*, Mon. Not. R. Astron. Soc. **239**, 845 (1989).
[129] I. S. Shklovskii, *Possible Causes of the Secular Increase in Pulsar Periods*, Sov. Astron. **13**, 562 (1970).
[130] P. J. McMillan, *The Mass Distribution and Gravitational Potential of the Milky Way*, Mon. Not. R. Astron. Soc. **465**, 76 (2017).
[131] R. Abuter, A. Amorim, M. Bauböck, J. P. Berger, H. Bonnet, W. Brand ner, Y. Clénet, V. C. Du Foresto, P. T. de Zeeuw, J. Dexter, G. Duvert *et al.* (Gravity Collaboration), *A Geometric Distance Measurement to the Galactic Center Black Hole with 0.3% Uncertainty*, Astron. Astrophys. **625**, L10 (2019).
[132] D. M. Eardley, Astrophys. J. **196**, L59 (1975).
[133] C. M. Will and H. W. Zaglauer, *Gravitational Radiation, Close Binary Systems, and the Brans-Dicke Theory of Gravity*, Astrophys. J. **346**, 366 (1989).
[134] N. Wex and R. Rieth, in *Proceedings of the 2nd Gauss Symposium, Conference A: Mathematics and Theoretical Physics, Munich, Germany, 1993*, edited by M. Behara, R. Fritsch, and R. G. Lintz (Walter de Gruyter, Berlin, 1995), pp. 681–694.
[135] See http://www.skatelescope.org.
[136] I. I. Shapiro, Fourth Test of General Relativity, Phys. Rev. Lett. **13**, 789 (1964).
[137] C. M. Will, *Theory and Experiment in Gravitational Physics* (Cambridge University Press, Cambridge, England, 1993).
[138] J. Schneider, *Gravitational Beam Deviation Effects in the Timing Formula of Binary Pulsars*, Astron. Astrophys. **232**, 62 (1990).
[139] D. Lai and R. R. Rafikov, *Effects of Gravitational Lensing in the Double Pulsar System J0737–3039*, Astrophys. J. **621**, L41 (2005).
[140] Simulations are based on the integration of null geodesics in a Schwarzschild spacetime, comoving with pulsar B. Since additive constant terms can be ignored, the use of Eddington-Finkelstein coordinates turns out to be of great advantage in the numerical integration for the propagation time [141].
[141] M. Ali, *Relativistic Propagation Effects in Binary Pulsar Signals for Next Generation Radio Telescopes*, Master's thesis, Friedrich-Wilhelms-Universität Bonn, Bonn, Germany, 2011.
[142] V. M. Kaspi, S. M. Ransom, D. C. Backer, R. Ramachandran, P. Demorest, J. Arons, and A. Spitkovsky, *Green Bank Telescope Observations of the Eclipse of Pulsar "A" in the Double Pulsar Binary PSR J0737–3039*, Astrophys. J. **613**, L137 (2004).
[143] The (intermittent) eclipses of A by the rotating magnetosphere of B lead to a significant reduction in the number of pulses available for the integrated pulse profile. This, in turn, reduces its S/N, leading to significantly fewer viable TOAs, which at the same time have considerably larger errors.
[144] S. A. Klioner and S. Zschocke, *Numerical versus Analytical Accuracy of the Formulae for Light Propagation*, Classical Quantum Gravity **27**, 075015 (2010).
[145] V. Perlick, *Gravitational Lensing from a Spacetime Perspective*, Living Rev. Relativity **7**, 9 (2004).
[146] S. M. Kopeikin and G. Schäfer, *Lorentz Covariant Theory of Light Propagation in Gravitational Fields of Arbitrary-Moving Bodies*, Phys. Rev. D **60**, 124002 (1999).
[147] R. R. Rafikov and D. Lai, *Effects of Gravitational Lensing and Companion Motion on the Binary Pulsar Timing*, Phys. Rev. D **73**, 063003 (2006).
[148] Lensing corrections to the retardation effect as discussed in Ref. [147] are totally negligible. The corresponding postfit residuals are below 0.02 μs, restricted to a narrow range around conjunction where, because of the eclipse, the timing precision is anyway greatly compromised.

[149] M. Kramer, I. H. Stairs, R. N. Manchester, M. A. McLaughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possenti, N. D'Amico, J. Sarkissian, B. C. Joshi, P. C. C. Freire, and F. Camilo, *Strong-Field Tests of Gravity with the Double Pulsar*, Ann. Phys. (Berlin) **15**, 34 (2006).

[150] O. V. Doroshenko and S. M. Kopeikin, *Relativistic Effect of Gravitational Deflection of Light in Binary Pulsars*, Mon. Not. R. Astron. Soc. **274**, 1029 (1995).

[151] R. R. Rafikov and D. Lai, *Effects of Pulsar Rotation on Timing Measurements of the Double Pulsar System J0737–3039*, Astrophys. J. **641**, 438 (2006).

[152] Note that the expressions in Refs. [147,151] are based on the lensing equation, which makes the assumption of small angles (cf., e.g., Ref. [153]), and, therefore, they are strictly speaking valid only near conjunction, which also limits their usability in a Tempo implementation.

[153] P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses* (Springer, New York, 1992).

[154] M. Kramer *et al.*, *The Relativistic Binary Programme on MeerKAT: Science Objectives and First Results*, Mon. Not. R. Astron. Soc. **504**, 2094 (2021).

[155] Periodic variations in the pulsar profile due to a latitudinal shift in the LOS caused by aberration are observed in PSR B1534+12 [156].

[156] I. H. Stairs, S. E. Thorsett, and Z. Arzoumanian, *Measurement of Gravitational Spin-Orbit Coupling in a Binary-Pulsar System*, Phys. Rev. Lett. **93**, 141101 (2004).

[157] The DDS timing model is a modification of the DD model [29], where the Shapiro shape parameter s is replaced by the logarithmic Shapiro shape parameter z_s via $s \equiv 1 - \exp(-z_s)$ (see Ref. [149] for details). It is part of the Tempo standard distribution.

[158] W. W. Zhu, I. H. Stairs, P. B. Demorest, D. J. Nice, J. A. Ellis, S. M. Ransom, Z. Arzoumanian, K. Crowter, T. Dolch, R. D. Ferdman, E. Fonseca, M. E. Gonzalez, G. Jones, M. L. Jones, M. T. Lam, L. Levin, M. A. McLaughlin, T. Pennucci, K. Stovall, and J. Swiggum, *Testing Theories of Gravitation Using 21-Year Timing of Pulsar Binary* J1713+0747, Astrophys. J. **809**, 41 (2015).

[159] Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. Chamberlin, S. Chatterjee, B. Christy, J. M. Cordes *et al.* (NANOGrav Collaboration), *The NANOGrav Nine-Year Data Set: Observations, Arrival Time Measurements, and Analysis of 37 Millisecond Pulsars*, Astrophys. J. **813**, 65 (2015).

[160] K. Liu, G. Desvignes, I. Cognard, B. W. Stappers, J. P. W. Verbiest, K. J. Lee, D. J. Champion, M. Kramer, P. C. C. Freire, and R. Karuppusamy, *Measuring Pulse Times of Arrival from Broad-band Pulsar Observations*, Mon. Not. R. Astron. Soc. **443**, 3752 (2014).

[161] T. T. Pennucci, *Frequency-Dependent Template Profiles for High-Precision Pulsar Timing*, Astrophys. J. **871**, 34 (2019).

[162] N. Capitaine, in Journées Systèmes de Référence Spatio-temporels 2007, edited by N. Capitaine (2008), p. 61, https://ui.adsabs.harvard.edu/abs/2008jsrs.meet...61C.

[163] Even though the factor $(1 - \kappa)$ is small, such correction is required before using the listed timing results for predictions of the folding parameters with Tempo2, which implements TCB. According to IAU Resolution 2006 B3, the conversion takes place as $\mathrm{TDB} = \mathrm{TCB} - (1 - \kappa) \times (\mathrm{JD_{TCB}} - 2443144.5003725) \times 86400 - 6.55 \times 10^{-5})$, where $\mathrm{JD_{TCB}}$ is the relevant Julian date [164]. Note that this described conversion needs to be applied as given to the quoted epoch of periastron passage (T_0) value before it is used with Tempo2.

[164] See https://www.iau.org/static/resolutions/IAU2006_Resol3.pdf.

[165] G. Hobbs, A. G. Lyne, and M. Kramer, *An Analysis of the Timing Irregularities for 366 Pulsars*, Mon. Not. R. Astron. Soc. **402**, 1027 (2010).

[166] R. P. Breton, *Radio Pulsars in Binary Systems*, Ph.D. thesis, McGill University, 2009.

[167] J. M. Lattimer and M. Prakash, *Neutron Star Structure and the Equation of State*, Astrophys. J. **550**, 426 (2001).

[168] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, Phys. Rev. Lett. **119**, 161101 (2017).

[169] J. M. Lattimer, in *Proceedings of the Xiamen-CUSTIPEN Workshop on the Equation of State of Dense Neutron-Rich Matter in the Era of Gravitational Wave Astronomy*, American Institute of Physics Conference Series Vol. 2127 (AIP, New York, 2019) p. 020001.

[170] This Doppler-shift rescaling, $m^{\mathrm{measured}} = m^{\mathrm{intrinsic}}/D$, between the pulsar-timing measured masses and the intrinsic ones [29] is due to the same effect which causes the masses directly deduced from gravitational-wave-template fitting in LIGO/Virgo data analysis to differ from the intrinsic, source-frame masses by a redshift factor: $m^{\mathrm{obs}} = (1 + z)m^{\mathrm{source}}$ [171].

[171] A. Krolak and B. F. Schutz, *Coalescing Binaries—Probe of the Universe*, Gen. Relativ. Gravit. **19**, 1163 (1987).

[172] J. M. Weisberg and Y. Huang, *Relativistic Measurements from Timing the Binary Pulsar PSR* B1913+16, Astrophys. J. **829**, 55 (2016).

[173] D. J. Nice and J. H. Taylor, *PSRs* J2019+2425 *and* J2322+2057 *and the Proper Motions of Millisecond Pulsars*, Astrophys. J. **441**, 429 (1995).

[174] K. Lazaridis, N. Wex, A. Jessner, M. Kramer, B. W. Stappers, G. H. Janssen, G. Desvignes, M. B. Purver, I. Cognard, G. Theureau, A. G. Lyne, C. A. Jordan, and J. A. Zensus, *Generic Tests of the Existence of the Gravitational Dipole Radiation and the Variation of the Gravitational Constant*, Mon. Not. R. Astron. Soc. **400**, 805 (2009).

[175] M. J. Reid and A. Brunthaler, *The Proper Motion of Sagittarius A*. III. The Case for a Supermassive Black Hole*, Astrophys. J. **892**, 39 (2020).

[176] R. Abuter *et al.* (Gravity Collaboration), *Improved GRAVITY Astrometric Accuracy from Modeling Optical Aberrations*, Astron. Astrophys. **647**, A59 (2021).

[177] M. S. Kehl, N. Wex, M. Kramer, and K. Liu, in *Proceedings of the Fourteenth Marcel Grossmann Meeting—MG14*, edited by M. Bianchi, R. T. Jansen, and R. Ruffini (World Scientific, Singapore, 2018), pp. 1860–1865.

[178] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso *et al.* (LIGO Scientific and Virgo Collaborations), *Tests of*

General Relativity with GW150914, Phys. Rev. Lett. **116**, 221101 (2016).

[179] We take the combined results of these references, obtained from the IMRPhenomPv2 waveform family.

[180] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Catalog GWTC-1*, Phys. Rev. D **100**, 104036 (2019).

[181] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Tests of General Relativity with Binary Black Holes from the Second LIGO-Virgo Gravitational-Wave Transient Catalog*, Phys. Rev. D **103**, 122002 (2021).

[182] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, *Frequency-Domain Gravitational Waves from Nonprecessing Black-Hole Binaries. II. A Phenomenological Model for the Advanced Detector Era*, Phys. Rev. D **93**, 044007 (2016).

[183] S. Mirshekari and C. M. Will, *Compact Binary Systems in Scalar-Tensor Gravity: Equations of Motion to 2.5 Post-Newtonian Order*, Phys. Rev. D **87**, 084070 (2013).

[184] K. Yagi, L. C. Stein, and N. Yunes, *Challenging the Presence of Scalar Charge and Dipolar Radiation in Binary Pulsars*, Phys. Rev. D **93**, 024010 (2016).

[185] E. Barausse, C. Palenzuela, M. Ponce, and L. Lehner, *Neutron-Star Mergers in Scalar-Tensor Theories of Gravity*, Phys. Rev. D **87**, 081506(R) (2013).

[186] J. Alsing, E. Berti, C. M. Will, and H. Zaglauer, *Gravitational Radiation from Compact Binary Systems in the Massive Brans-Dicke Theory of Gravity*, Phys. Rev. D **85**, 064041 (2012).

[187] Short-range modifications of GR that significantly affect the MOI could, in principle, show up in the Lense-Thirring test of Eq. (50); see discussion in Ref. [110].

[188] I. Ciufolini, A. Paolozzi, E. C. Pavlis, G. Sindoni, J. Ries, R. Matzner, R. Koenig, C. Paris, V. Gurzadyan, and R. Penrose, *An Improved Test of the General Relativistic Effect of Frame-Dragging Using the LARES and LAGEOS Satellites*, Eur. Phys. J. C **79**, 872 (2019).

[189] C. W. F. Everitt *et al.*, *Gravity Probe B: Final Results of a Space Experiment to Test General Relativity*, Phys. Rev. Lett. **106**, 221101 (2011).

[190] A. Ritz and J. Ward, *Weyl Corrections to Holographic Conductivity*, Phys. Rev. D **79**, 066003 (2009).

[191] This is, for instance, the case in the mEIH formalism, where (to leading order) the deflection angle is proportional to $(1+\gamma_{B0})G_{B0}$ (see e.g., Appendix C in Ref. [192]), which is already tightly constrained by the test of the Shapiro range parameter r. More generally, the NLO contribution to Δ_A is parametrized by r and s according to Eqs. (30) and (31).

[192] T. Damour and G. Esposito-Farèse, *Testing Gravity to Second Post-Newtonian Order: A Field-Theory Approach*, Phys. Rev. D **53**, 5541 (1996).

[193] It follows from the analysis in Ref. [51] that the rotational axis of A is (practically) perpendicular to the orbital plane. See also Sec. VIII C.

[194] An additional test is provided by the PK parameter $\Omega_B^{\rm spin}$; see Table V.

[195] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, Adhikari *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Tests of General Relativity with GW170817*, Phys. Rev. Lett. **123**, 011102 (2019).

[196] K. Nordtvedt, *$\dot{G}/G$ and a Cosmological Acceleration of Gravitationally Compact Bodies*, Phys. Rev. Lett. **65**, 953 (1990).

[197] T. Damour, G. W. Gibbons, and J. H. Taylor, *Limits on the Variability of G Using Binary-Pulsar Data*, Phys. Rev. Lett. **61**, 1151 (1988).

[198] K. Nordtvedt, *Cosmological Time Evolution in the Newtonian Dynamics of Compact Celestial Bodies*, Astrophys. J. **407**, 5 (1993).

[199] There are special cases, like Barker's constant-G theory [200], where $\dot{P}_b^{\dot{G}}$ cannot be written in this form.

[200] B. M. Barker, *General Scalar-Tensor Theory of Gravity with Constant g*, Astrophys. J. **219**, 5 (1978).

[201] W. W. Zhu *et al.*, *Tests of Gravitational Symmetries with Pulsar Binary* J1713+0747, Mon. Not. R. Astron. Soc. **482**, 3249 (2019).

[202] A. Genova, E. Mazarico, S. Goossens, F. G. Lemoine, G. A. Neumann, D. E. Smith, and M. T. Zuber, *Solar System Expansion and Strong Equivalence Principle as Seen by the NASA MESSENGER Mission*, Nat. Commun. **9**, 289 (2018).

[203] T. Damour and G. Esposito-Farèse, *Tensor-Multi-scalar Theories of Gravitation*, Classical Quantum Gravity **9**, 2093 (1992).

[204] A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge University Press, Cambridge, England, 1922).

[205] To leading order, one finds in DEF gravity $B_D \simeq (5/96)\times(\alpha_A-\alpha_B)^2[1+\mathcal{O}(\alpha_A,\alpha_B)][1+\mathcal{O}(e^2)]$ (cf. Refs. [203,206]).

[206] T. Damour and G. Esposito-Farèse, *Gravitational-Wave versus Binary-Pulsar Tests of Strong-Field Gravity*, Phys. Rev. D **58**, 042001 (1998).

[207] B. Bertotti, L. Iess, and P. Tortora, *A Test of General Relativity Using Radio Links with the Cassini Spacecraft*, Nature (London) **425**, 374 (2003).

[208] P. C. C. Freire, N. Wex, G. Esposito-Farèse, J. P. W. Verbiest, M. Bailes, B. A. Jacoby, M. Kramer, I. H. Stairs, J. Antoniadis, and G. H. Janssen, *The Relativistic Pulsar-White Dwarf Binary PSR* J1738+0333—*II. The Most Stringent Test of Scalar-Tensor Gravity*, Mon. Not. R. Astron. Soc. **423**, 3328 (2012).

[209] A. M. Archibald, N. V. Gusinskaia, J. W. T. Hessels, A. T. Deller, D. L. Kaplan, D. R. Lorimer, R. S. Lynch, S. M. Ransom, and I. H. Stairs, *Universality of Free Fall from the Orbital Motion of a Pulsar in a Stellar Triple System*, Nature (London) **559**, 73 (2018).

[210] J.-P. Bruneton and G. Esposito-Farèse, *Field-Theoretical Formulations of MOND-like Gravity*, Phys. Rev. D **76**, 124012 (2007).

[211] J. P. W. Verbiest, D. R. Lorimer, and M. A. McLaughlin, *Lutz-Kelker Bias in Pulsar Parallax Measurements*, Mon. Not. R. Astron. Soc. **405**, 564 (2010).

[212] J. P. W. Verbiest, J. M. Weisberg, A. A. Chael, K. J. Lee, and D. R. Lorimer, *On Pulsar Distance Measurements and Their Uncertainties*, Astrophys. J. **755**, 39 (2012).

[213] J. Yao (private communication).
[214] C. R. Purcell, B. M. Gaensler, X. H. Sun, E. Carretti, G. Bernardi, M. Haverkorn, M. J. Kesteven, S. Poppi, D. H. F. M. Schnitzeler, and L. Staveley-Smith, *A Radio-Polarisation and Rotation Measure Study of the Gum Nebula and Its Environment*, Astrophys. J. **804**, 22 (2015).
[215] A. T. Deller, W. M. Goss, W. F. Brisken, S. Chatterjee, J. M. Cordes, G. H. Janssen, Y. Y. Kovalev, T. J. W. Lazio, L. Petrov, B. W. Stappers, and A. Lyne, *Microarcsecond VLBI Pulsar Astrometry with* PSRπ *II. Parallax Distances for 57 Pulsars*, Astrophys. J. **875**, 100 (2019).
[216] M. T. Lam, J. M. Cordes, S. Chatterjee, and T. Dolch, *Pulsar Timing Errors from Asynchronous Multi-frequency Sampling of Dispersion Measure Variations*, Astrophys. J. **801**, 130 (2015).
[217] X. P. You, G. Hobbs, W. A. Coles, R. N. Manchester, R. Edwards, M. Bailes, J. Sarkissian, J. P. W. Verbiest, W. van Straten, A. Hotan, S. Ord, F. Jenet, N. D. R. Bhat, and A. Teoh, *Dispersion Measure Variations and Their Effect on Precision Pulsar Timing*, Mon. Not. R. Astron. Soc. **378**, 493 (2007).
[218] Since the DM is the column integral of the electron density along the line-of-sight (z) axis, its 2D spatial spectrum is the same as that of the 3D density with the wave number in z set to zero. Therefore, the 2D spectral exponent of DM (β) is the same as that of the 3D electron-density fluctuations, i.e., $-11/3$ for a Kolmogorov process. The corresponding 2D structure function has a spectral exponent $\alpha = -(\beta + 2) = +5/3$ for Kolmogorov.
[219] One astronomical unit (A.U.) is approximately 1.496×10^8 km.
[220] The fluctuations in DM(t) are those of a 1D cut across the 2D spatial distribution. Therefore, the 1D PSD of DM(t) is the 2D spatial spectrum of DM (scaled by the velocity) with the perpendicular wave number integrated out. This reduction to 1D reduces the Kolmogorov spectral exponent from $-11/3$ to $-8/3$.
[221] O. P. Lay, *The Temporal Power Spectrum of Atmospheric Fluctuations due to Water Vapor*, Astron. Astrophys. Suppl. Ser. **122**, 535 (1997).
[222] This value agrees well with a pulsar-based estimate [223]. There is some spread in measured values for z_{SSB} (see, e.g., Table 1 in Ref. [223]), but this has little impact on our conclusions.
[223] J. M. Yao, R. N. Manchester, and N. Wang, *Determination of the Sun's Offset from the Galactic Plane Using Pulsars*, Mon. Not. R. Astron. Soc. **468**, 3289 (2017).
[224] R. N. Manchester, M. Kramer, I. H. Stairs, M. Burgay, F. Camilo, G. B. Hobbs, D. R. Lorimer, A. G. Lyne, M. A. McLaughlin, C. A. McPhee, A. Possenti, J. E. Reynolds, and W. van Straten, *Observations and Modeling of Relativistic Spin Precession in PSR J1141–6545*, Astrophys. J. **710**, 1694 (2010).
[225] V. V. Krishnan, M. Bailes, W. van Straten, E. F. Keane, M. Kramer, N. D. R. Bhat, C. Flynn, and S. Osłowski, *Relativistic Spin Precession in the Binary PSR J1141–6545*, Astrophys. J. **873**, L15 (2019).
[226] M. Burgay, A. Possenti, R. N. Manchester, M. Kramer, M. A. McLaughlin, D. R. Lorimer, I. H. Stairs, B. C. Joshi, A. G. Lyne, F. Camilo, N. D'Amico, P. C. C. Freire, J. M. Sarkissian, A. W. Hotan, and G. B. Hobbs, *Long-Term Variations in the Pulse Emission from PSR J0737–3039B*, Astrophys. J. **624**, L113 (2005).
[227] K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett, D. Bintley *et al.* (Event Horizon Telescope Collaboration), *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, Astrophys. J. **875**, L1 (2019).
[228] R. Abuter, A. Amorim, M. Bauböck, J. P. Berger, H. Bonnet, W. Brandner, V. Cardoso, Y. Clénet, P. T. de Zeeuw, J. Dexter *et al.* (Gravity Collaboration), *Detection of the Schwarzschild Precession in the Orbit of the Star S2 near the Galactic Centre Massive Black Hole*, Astron. Astrophys. **636**, L5 (2020).
[229] R. Abuter, A. Amorim, M. Bauböck, J. P. Berger, H. Bonnet, W. Brand ner, Y. Clénet, V. C. Du Foresto, P. T. de Zeeuw, C. Deen, J. Dexter, G. Duvert *et al.* (Gravity Collaboration), *Detection of Orbital Motions near the Last Stable Circular Orbit of the Massive Black Hole SgrA**, Astron. Astrophys. **618**, L10 (2018).
[230] N. Yunes, K. Yagi, and F. Pretorius, *Theoretical Physics Implications of the Binary Black-Hole Mergers GW150914 and GW151226*, Phys. Rev. D **94**, 084002 (2016).
[231] T. Baker, D. Psaltis, and C. Skordis, *Linking Tests of Gravity on All Scales: From the Strong-Field Regime to Cosmology*, Astrophys. J. **802**, 63 (2015).
[232] M. Kramer, *Probing Einstein's Universe and Its Physics*, Astron. Geophys. **58**, 3.31(2017).
[233] R. Abuter, A. Amorim, N. Anugu, M. Bauböck, M. Benisty, J. P. Berger, N. Blind, H. Bonnet, W. Brandner, A. Buron, C. Collin, F. Chapron, Y. Clénet, V. C. Du Foresto, P. T. de Zeeuw, C. Deen, F. Delplancke-Ströbele, R. Dembet, J. Dexter, G. Duvert *et al.* (Gravity Collaboration), *Detection of the Gravitational Redshift in the Orbit of the Star S2 near the Galactic Centre Massive Black Hole*, Astron. Astrophys. **615**, L15 (2018).
[234] T. Do *et al.*, *Relativistic Redshift of the Star S0-2 Orbiting the Galactic Center Supermassive Black Hole*, Science **365**, 664 (2019).
[235] D. Psaltis, L. Medeiros, P. Christian, F. Ozel, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, D. Ball *et al.* (The EHT Collaboration), *Gravitational Test beyond the First Post-Newtonian Order with the Shadow of the M87 Black Hole*, Phys. Rev. Lett. **125**, 141104 (2020).
[236] J. J. Hermes, M. Kilic, W. R. Brown, D. E. Winget, C. A. Prieto, A. Gianninas, A. S. Mukadam, A. Cabrera-Lavers, and S. J. Kenyon, *Rapid Orbital Decay in the 12.75-Minute Binary White Dwarf* J0651+2844, Astrophys. J. **757**, L21 (2012).
[237] B. J. Rickett, *Radio Propagation through the Turbulent Interstellar Plasma*, Annu. Rev. Astron. Astrophys. **28**, 561 (1990).
[238] W. A. Coles, B. J. Rickett, J. J. Gao, G. Hobbs, and J. P. W. Verbiest, *Scattering of Pulsar Radio Emission by the Interstellar Plasma*, Astrophys. J. **717**, 1206 (2010).
[239] T. Damour, in *Three Hundred Years of Gravitation*, edited by S. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987), pp. 128–198.