Extended Reduced-order surrogate models for scalar-tensor gravity in the strong field and applications to binary pulsars and gravitational wave*

Minghao Guo, Junjie Zhao, and Lijing Shao†
Peking University, Beijing 100871, China
(MUSO Collaboration)

(Dated: November 20, 2020)

We investigate the scalar-tensor gravity of Damour and Esposito-Farèse (DEF) with spontaneous scalarization phenomena developed for neutron stars. We construct reduced-order surrogate model for the derived quantities and integrate the model into a python package pySTGROMX that speeds up calculations at two order-of-magnitude yet still keeps accuracy, compared with the previous method. The timing of binary pulsars allows us to place some of the tightest constraints on modified theories of gravity. We apply pySTGROMX to constrain the parameters of the DEF theory with well-timed binary pulsars.

I. INTRODUCTION

Albert Einstein's theory of general relativity (GR) has been tested in many cases, e.g., the Solar System, the timing of binary pulsars, and the gravitational-wave (GW) observation of coalescing binary black holes (BBHs) and binary neutron stars (BNSs).

Gravitational test has a long history.

In this paper, we design and develop a method for computing derived quantities in the scalar-tensor gravity of Damour and Esposito-Farèse (DEF) with spontaneous scalarization phenomena developed for neutron stars. We construct reduced-order surrogate model for the derived quantities and integrate the model into a python package pySTGROMX that speeds up calculations at two order-of-magnitude yet still keeps accuracy, compared with the previous method. The timing of binary pulsars allows us to place some of the tightest constraints on modified theories of gravity. We apply pySTGROMX to constrain the parameters of the DEF theory with well-timed binary pulsars.

II. SPONTANEOUS SCALARIZATION IN THE DEF THEORY

In this section, we study the DEF theory, which is defined by the following general action in *Einstein frame* [1, 2],

$$\begin{aligned} S =& \frac{c^4}{16\pi G_\star}\int \frac{\mathrm{d}^4x}{c}\sqrt{-g_\star}\left[R_\star - 2g_\star^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi)\right] \\ &+ S_m\left[\psi_m; A^2(\varphi)g^\star_{\mu\nu}\right]. \end{aligned} \quad (1)$$

Here, $G_\star$ denotes the bare gravitational constant, $g_\star \equiv \det g^\star_{\mu\nu}$ is the determinant of Einstein metric $g^\star_{\mu\nu}$, $R_\star$ is the Ricci curvature scalar of $g^\star_{\mu\nu}$, and φ is a dynamical scalar field. In the last term of Eq. (1), ψ_m denotes matter fields collectively, and the conformal coupling factor $A(\varphi)$ describes how φ couples to ψ_m in Einstein frame. Varying the action (1) yields the field equations,

$$R^\star_{\mu\nu} = \partial_\mu\varphi\partial_\nu\varphi + \frac{8\pi G_\star}{c^4}\left(T^\star_{\mu\nu} - \frac{1}{2}T^\star g^\star_{\mu\nu}\right), \quad (2)$$

$$\Box_{g^\star}\varphi = -\frac{4\pi G_\star}{c^4}\alpha(\varphi)T_\star\,, \quad (3)$$

where $T_\star^{\mu\nu} \equiv 2c(-g_\star)^{-1/2}\delta S_m/\delta g^\star_{\mu\nu}$ denotes the matter stress-energy tensor, and $T^\star \equiv g^\star_{\mu\nu}T_\star^{\mu\nu}$ is the trace. In Eq. (3), the quantity $\alpha(\varphi)$ is defined as the logarithmic derivative of $A(\varphi)$,

$$\alpha(\varphi) \equiv \frac{\partial \ln A(\varphi)}{\partial\varphi}\,, \quad (4)$$

which indicates the coupling strength between the scalar field and matters.

In the DEF theory [2], $\ln A(\varphi)$ is designated as

$$\ln A(\varphi) = \frac{1}{2}\beta_0\varphi^2\,. \quad (5)$$

Then $\alpha(\varphi) = \partial \ln A(\varphi)/\partial\varphi = \beta_0\varphi$. We designate $\alpha_0 \equiv \beta_0\varphi_0$, where φ_0 is the asymptotic scalar field value of φ at spatial infinity. Note that we have $\alpha_0 = \beta_0 = 0$ in GR.

For NSs, nonperturbative scalarization phenomena develop when [1, 3]

$$\beta_0 \equiv \left.\frac{\partial^2 \ln A(\varphi)}{\partial\varphi^2}\right|_{\varphi=\varphi_0} \lesssim -4\,. \quad (6)$$

Generally, a more negative β_0 means more manifest spontaneous scalarization in the strong-field regime. In such case, the *effective scalar coupling* for a NS "A" with Arnowitt–Deser–Misner (ADM) mass m_A is

$$\alpha_A \equiv \left.\frac{\partial \ln m_A(\varphi)}{\partial\varphi}\right|_{\varphi=\varphi_0}\,, \quad (7)$$

which measures the coupling strength between the scalar field and the NS.

Now we consider a scalarized NS in a binary pulsar system. For a NS binary system with the pulsar labeled "A" and its companion labeled "B", the quantities α_A and α_B contribute to the secular change of the orbital period decay $\dot{P}_b$ [2].Should I wirte the exact formula here? Correspondingly, we define

$$\beta_A \equiv \left.\frac{\partial \alpha_A}{\partial\varphi}\right|_{\varphi=\varphi_0}\,, \quad (8)$$

---

* A footnote to the article title
† lshao@pku.edu.cn

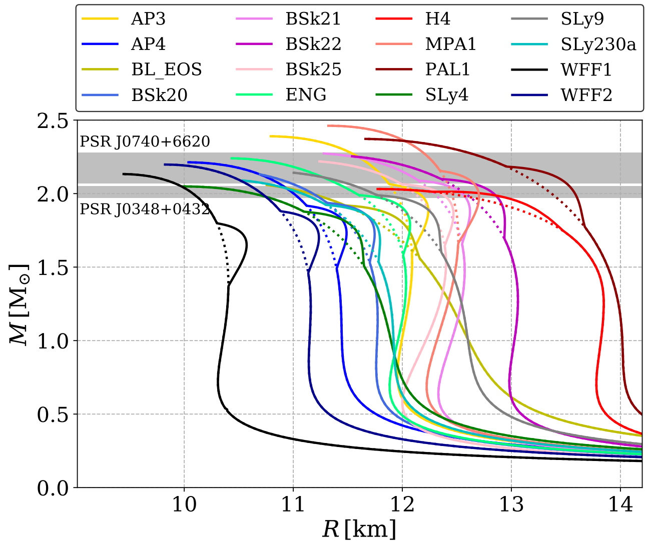

FIG. 1. Mass-radius relations of NSs for the EOSs that we adopt in this study. The mass-radius relations are derived from GR (dashed lines) and from a DEF theory with $\log_{10}|\alpha_0| = -5.0$ and $\beta_0 = -4.5$ (solid lines). The masses from PSRs J0740+6620 and J0348+0432 are overlaid in grey. The "bumps" show the deviation of the DEF theory from GR.

which is the strong-field analogue of the quantity β_0. Then the theoretical prediction for the periastron advance rate is [2]

$$\dot{\omega}^{\text{th}}(m_A, m_B) \equiv \frac{3n_b}{1-e^2}\left(\frac{G_{AB}(m_A+m_B)n_b}{c^3}\right)^{2/3} \times\left[\frac{1-\frac{1}{3}\alpha_A\alpha_B}{1+\alpha_A\alpha_B} - \frac{X_A\beta_B\alpha_A^2 + X_B\beta_A\alpha_B^2}{6(1+\alpha_A\alpha_B)^2}\right], \tag{9}$$

where $n_b \equiv 2\pi/P_b$, $G_{AB} \equiv G_\star(1+\alpha_A\alpha_B)$, and $X_A \equiv m_A/(m_A + m_B) \equiv 1 - X_B$. Finally, consider a NS with inertia moment (in Einstein units) I_A. We denote

$$k_A \equiv \left.\frac{\partial \ln I_A}{\partial\varphi}\right|_{\varphi=\varphi_0} \tag{10}$$

as the "coupling factor" of inertia moment. The theoretical prediction of the Einstein delay parameter is [2],

$$\gamma \equiv \gamma^{\text{th}}(m_A, m_B) = \frac{e}{n_b}\frac{X_B}{1+\alpha_A\alpha_B}\left(\frac{G_{AB}(m_A+m_B)n_b}{c^3}\right)^{2/3} \times[X_B(1+\alpha_A\alpha_B) + 1 + K_A^B], \tag{11}$$

where $K_A^B \equiv -\alpha_B(m_B)k_A(m_A)$ describes the contribution from the variation of I_A under the influence of the companion B.

Maybe add a discussion about what kind of system (like binary NS) should be applied?

III. METHODOLOGY

We here turn our attention to the calculation of the quantities in strong field. For a specific nuclear EOS of NSs, given

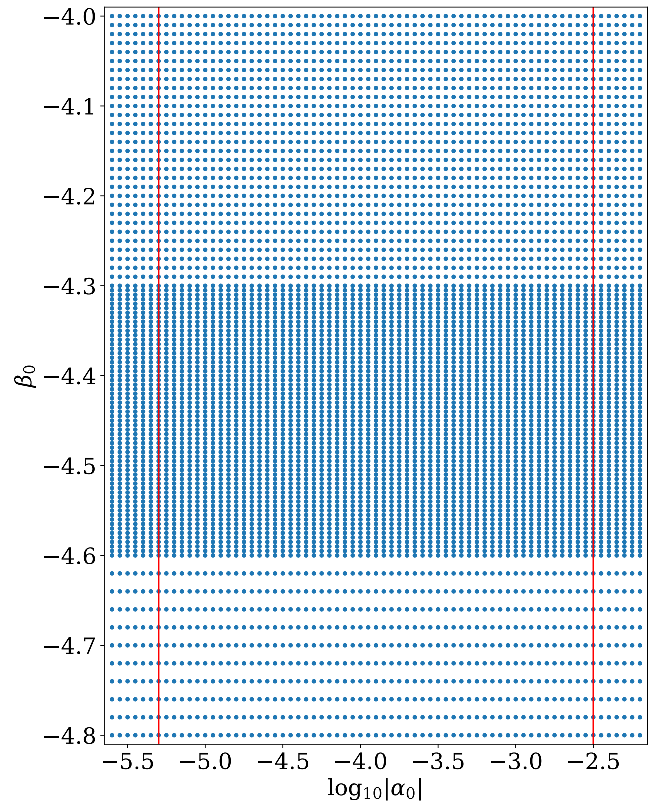

FIG. 2. An uneven grid in the parameter space ($\log_{10}|\alpha_0|, -\beta_0$) for calculating β_A and k_A and building ROMs. We generate a set of $69 \times 101 = 6969$ parameter pairs as the training data. The region between red lines corresponds to the data we use in later calculation.

the center mass density ρ_c and the parameters of the theory (namely, α_0, β_0), we can obtain macroscopic quantities of a NS (e.g, R, m_A, α_A and I_A), by solving the modified TOV equations with the shooting method (see Ref. [4] for details). In Fig. 1 we show mass-radius relation of NSs in the DEF theory with $\log_{10}|\alpha_0| = -5.0$ and $\beta_0 = -4.5$ for the EOSs we adopt in this study. It shows clearly that the spontaneous scalarization phenomena develop for NSs with certain masses, and lager radii are predicted in this range. However, to determine quantities β_A and k_A, we have to calculate the derivatives from Eqs. (8) and (10) for a fixed form of the conformal coupling factor $A(\varphi)$ (i.e, with a fixed β_0) and a fixed baryonic mass $\bar{m}_A$. This requires the data with different φ_0's (or equivalently, α_0's). In order to do so, we calculate the derivatives on a grid. Should I show the PSR data in Fig.1? Also probably some data about the radius of NS.

In practice, for each EOS, we choose the range of ρ_c so that $m_A \in (1\,\mathrm{M}_\odot, m_A^{\max})$ with the maximum NS mass $m_A^{\max}$ being EOS-dependent. Then we generate an uneven gird of $[\log_{10}|\alpha_0|, \beta_0] \in [-5.6, -2.2] \times [-4.8, -4.0]$, as shown in Fig. 2. The number of nodes in grid is set to $N_{\alpha_0} \times N_{\beta_0} = 69 \times 101 = 6969$. We calculate β_A and k_A on each node with a reasonable differential step. Finally, we use the data of $\log_{10}|\alpha_0| \in [-5.3, -2.5]$ for further calculation to avoid the inaccuracy of derivatives at boundaries. The boundary value

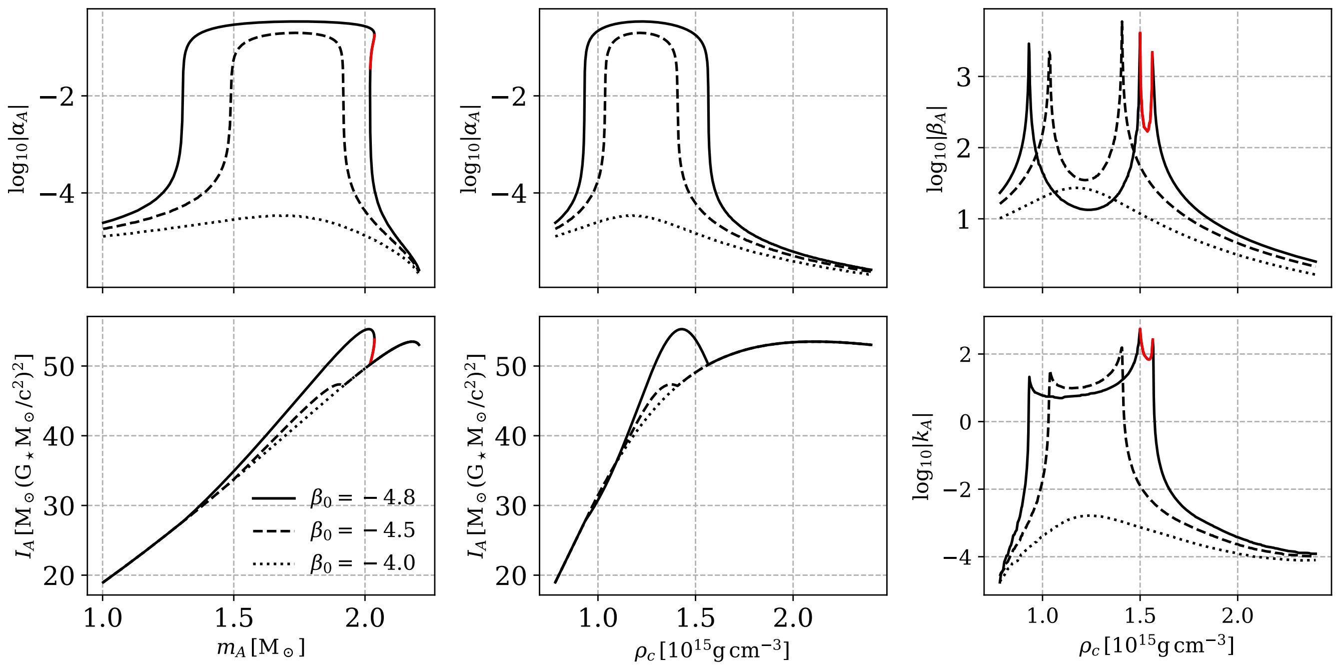

FIG. 3. Pathological phenomena occur when integrating the modified TOV equations for the EOS AP4. The calculation assumes the DEF parameters $\log_{10}|\alpha_0| = -5.3$ and $\beta_0 = -4.8$ (solid lines), -4.5 (dashed lines) and -4.0 (dotted lines). For $\log_{10}|\alpha_0| = -5.3$, the scalar field is weak for $\beta_0 = -4.0$, strong for $\beta_0 = -4.5$, and this causes the pathological phenomena for $\beta_0 = -4.8$. The red lines mark the pathological region. In this region, β_A and k_A are negative.

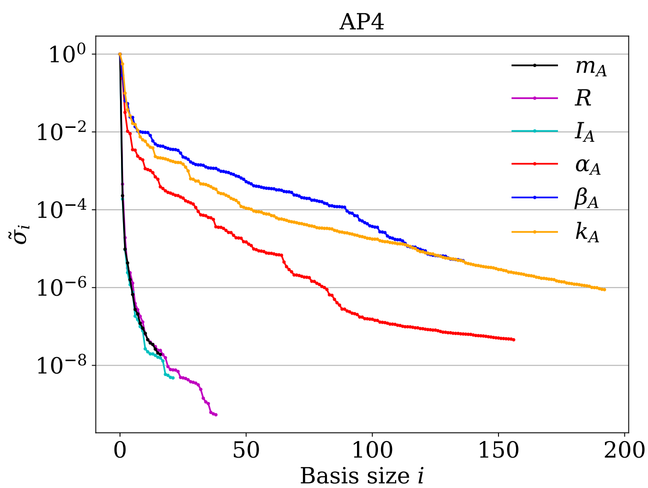

FIG. 4. Relative maximum projection error, $\tilde{\sigma}_i$, in building the ROMs for the EOS AP4. We set $\Sigma = 10^{-7}$ for m_A, R and I_A, $\Sigma = 10^{-5}$ for α_A, and $\Sigma = 10^{-4}$ for β_A and k_A.

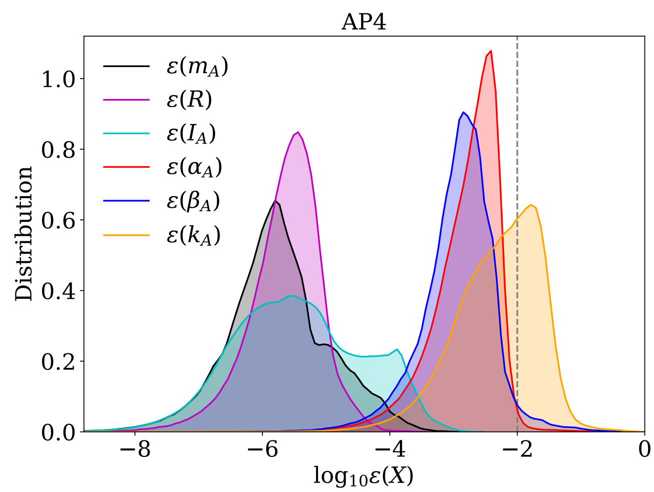

FIG. 5. Kernel density estimation (KDE) distribution of the relative error $\varepsilon(X)$, where $X \in \{m_A, R, I_A, \alpha_A, \beta_A, k_A\}$. The dashed line shows the relative tolerable error in the TOV integration ($\lesssim 1\%$).

$\alpha_0 \approx 10^{-2.5}$ is the upper limit given by the Cassini spacecraft [5], and $\beta_0 \lesssim -4.0$ corresponds to values where spontaneous scalarization happens in the DEF theory.

We have to point it out that in practice it is difficult to calculate k_A when the scalar field is weak. In this case, a change in I_A due to the weak field is comparable to the random noises during the integration in solving the modified TOV equations. The calculation of k_A is therefore not accurate. Here we propose a reasonable approximation that $k_A \sim \varphi_0^2$ when the spontaneous scalarization is not excited. Based on this assumption, we choose a large differential step and calculate $k_A = 2\varphi\partial \ln I_A/\partial\varphi^2$ to reduce the influence of numerical noises.

Due to the time-consuming computation of the TOV inte-

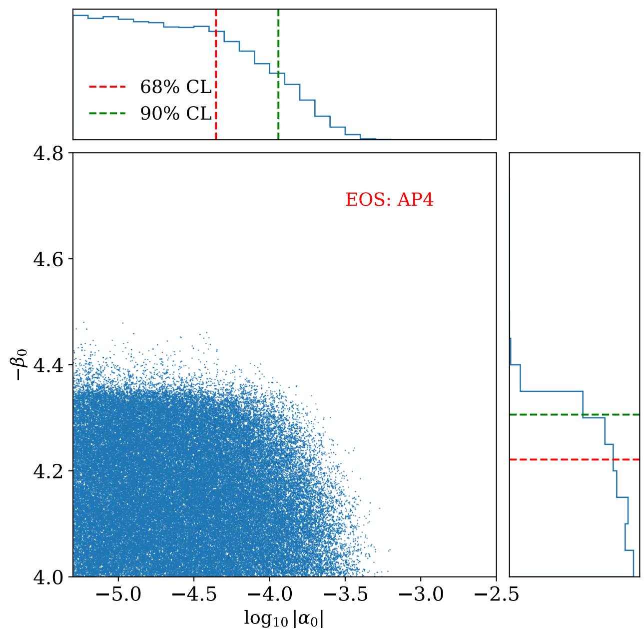

FIG. 6. Kernel density estimation (KDE) distribution of the relative error $\varepsilon(X)$, where $X \in \{m_A, R, I_A, \alpha_A, \beta_A, k_A\}$. The dashed line shows the relative tolerable error in the TOV integration ($\lesssim 1\%$).

gration and the shooting method for large-scale calculations, such as the parameter estimation with the MCMC approach, we build ROMs for the quantities to improve the efficiency. In brief, to generate a ROM for a curve $h(t;\lambda)$ with parameters λ, one provides a training space of data $\mathbf{V} \equiv \{h(t;\lambda_i)\}$ on a given grid of parameters and select a certain number (denoted as m) of bases as a chosen space $\mathbf{RV} = \{e_i\}_{i=1}^{m}$ with the reduced basis (RB) method. In practice, given the starting RB ($i = 0$), one iteratively seeks for m orthonormal RBs by iterating the Gram-Schmidt orthogonalization algorithm with greedy selection to minimize the maximum projection error, [LS: references needed]

$$\sigma_i \equiv \max_{h\in\mathbf{V}} \left\| h(\cdot;\lambda) - \mathcal{P}_i h(\cdot;\lambda) \right\|^2, \tag{12}$$

where $\mathcal{P}$ describes the projection of $h(t;\lambda)$ onto the span of the first i RBs. The process terminates when $\sigma_{m-1} \lesssim \Sigma$, a user-specified error bound. Then every curve in the training space is well approximated by

$$h(t;\lambda) \approx \sum_{i=1}^{m} c_i(\lambda) e_i(t) \approx \sum_{i=1}^{m} \langle h(\cdot;\lambda), e_i(\cdot)\rangle e_i(t), \tag{13}$$

where $c_i(\lambda)$ is the coefficient to be used for the ROM. Finally, one performs a fit to the parameter space, $\{\lambda_i\}$, and complete the construction of ROM. More details can be found in Ref. [4] where ROMs of α_A were built.

Extending the work by Zhao *et al.* [4], we build ROMs for six quantities, R, m_A, I_A, α_A, β_A and k_A, as functions of the central mass density ρ_c, with specialized parameters $\lambda = (\alpha_0, \beta_0)$.[1] We choose the implicit parameter ρ_c as an independent variable to avoid the the multivalued relations between m_A and R, as well as α_A and I_A [4]. We show this phenomena in Fig. 3. Due to the multivalued relations, β_A and k_A are negative when the α_A-m_A and I_A-m_A curve are bent backwards.

In balancing the computation cost and the accuracy of ROMs, we set the error bound $\Sigma = 10^{-7}$ for m_A, R and I_A, $\Sigma = 10^{-5}$ for α_A, and $\Sigma = 10^{-4}$ for β_A and k_A. The relative projection error $\tilde{\sigma}_i \equiv \sigma_i/\sigma_0$ as a function of the basis size is shown in Fig. 4. To achieve the desired projection error, the basis size is $\sim$ 20-40 for m_A, R and I_A, but $\sim$ 150-200 for α_A, β_A and k_A. This is due to the fact that there are more features in the latter set of parameters. Considering the error involved in the shooting method and the calculation of derivatives, which is $\sim$ 1%, the precision loss in ROM building is negligible. But, $\varepsilon(k_A)$ is much larger than 0.01, maybe we should point it out. About $\varepsilon(k_A)$, since we build ROM for $\ln|k_A + k_0|$, is it reasonable to build assess the accuracy using Eq. 14?

To assess the accuracy of the ROMs, we define

$$\varepsilon(X) = \left| \frac{X_{\rm ROM} - X_{\rm mTOV}}{X_{\rm ROM} + X_{\rm mTOV}} \right|, \tag{14}$$

where $X \in \{m_A, R, I_A, \alpha_A, \beta_A, k_A\}$, to indicate the fractional accuracy of the ROMs. In Eq. (14), we denote $X_{\rm ROM}$ as the prediction of ROM, and $X_{\rm mTOV}$ as the value from the shooting algorithm and derivatives on the grid. Explain why choosing grid here. To calculate the derivatives, instead of randomly generating parameters, we choose another grid as the test space which is shifted from the training space for α_0, β_0 and ρ_c, and calculate the quantities in the same way. We should explain why we test the parameters shifted in training space instead of random space The test space has sparser distribution of β_0. The distributions of $\varepsilon(X)$ are shown in Fig. 5. The relative errors of m_A, R and I_A are $\lesssim 10^{-5}$. On the contrary, relative errors of α_A, β_A and k_A is mostly smaller than 1%. Although this error is larger than those of R and m_A, in most cases, the error is still small enough to be neglected compared with the error from the shooting method and the calculation of derivatives. About the error: For k_A, due to the additional error from the method in calculating the derivatives, a small fraction of prediction have the error in the range $\sim 1 - 10\%$.

IV. CONSTRAINTS FROM BINARY PULSARS

In this section, we apply our ROMs to various scenarios, and discuss the improvement in deriving NS properties.

To be finished...

In Table I, we show

[1] In practice, we use $\ln|I_A|$, $\ln|\alpha_A|$, $\ln|\beta_A|$, and $\ln|k_0 + k_A|$—instead of β_A and k_A—for a better numerical performance, where k_0 is an EOS-dependent constant to avoid negative values of k_A in the weak field. Generally we have $k_0 \lesssim 0.1$.

TABLE I. parameters of binary pulsars.

Name	J0348+0432	J1012+5307	J1738+0333	J1909-3744	J2222-0137
Orbital period, P_b (d)	0.102424062722(7)	0.60467271355(3)	0.3547907398724(13)	1.533449474305(5)	2.44576454(18)
Eccentricity, e	0.0000026(9)	0.0000012(3)	0.00000034(11)	0.000000115(7)	0.00038096(4)
Observed $\dot{P}_b$, $\dot{P}_b^{\rm obs}$ (fs s^{-1})	-273(45)	50(14)	-17.0(31)	-510.87(13)	200(90)
Intrinsic $\dot{P}_b$, $\dot{P}_b^{\rm int}$ (fs s^{-1})	-274(45)	-5(9)	-27.72(64)	-4.4(79)	-60(90)
Periastron advance, $\dot{\omega}$ (deg yr^{-1})	—	—	—	—	0.1001(35)
Einstein delay γ (ms)	—	—	—	—	—
Pulsar mass, m_p (M$_\odot$)	2.01(4)	—	—	1.492(14)	1.76(6)
Companion mass, m_c (M$_\odot$)	$0.1715^{+0.0045}_{-0.0030}$	0.174(7)	$0.1817^{+0.0073}_{-0.0054}$	0.209(1)	1.293(25)
Mass ratio, $q \equiv m_p/m_c$	11.70(13)	10.5(5)	8.1(2)	—	—

TABLE II. parameters of binary pulsars.

Name	B1913+16	J0737-3039A	J1757-1854	B1534+12
Orbital period, P_b (d)	0.322997448918(3)	0.10225156248(5)	0.18353783587(5)	0.420737298879(2)
Eccentricity, e	0.6171340(4)	0.0877775(9)	0.6058142(10)	0.27367752(7)
Observed $\dot{P}_b$, $\dot{P}_b^{\rm obs}$ (fs s^{-1})	-2423(1)	-1252(17)	-5300(200)	-136.6(3)
Intrinsic $\dot{P}_b$, $\dot{P}_b^{\rm int}$ (fs s^{-1})	-2398(4)	-1252(17)	-5300(240)	—
Periastron advance, $\dot{\omega}$ (deg yr^{-1})	4.226585(4)	16.89947(68)	10.3651(2)	1.7557950(19)
Einstein delay γ (ms)	4.307(4)	0.3856(26)	3.587(12)	2.0708(5)
Pulsar mass, m_p (M$_\odot$)	1.438(1)	1.3381(7)	1.3384(9)	1.3330(2)
Companion mass, m_c (M$_\odot$)	1.390(1)	1.2489(7)	1.3946(9)	1.3455(2)
Mass ratio, $q \equiv m_p/m_c$	—	—	—	—

In Table II, we show three double neutron stars.

$$P(\alpha_0,\beta_0|\mathcal{D},\mathcal{H},\mathcal{I}) = \int \frac{P(\mathcal{D}|\alpha_0,\beta_0\Xi,\mathcal{H},\mathcal{I})P(\alpha_0,\beta_0|,\Xi|\mathcal{H},\mathcal{I})}{P(\mathcal{D}|\mathcal{H},\mathcal{I})}\mathrm{d}\Xi\,, \tag{15}$$

[1] T. Damour and G. Esposito-Farèse, Phys. Rev. Lett. **70**, 2220 (1993).

[2] T. Damour and G. Esposito-Farèse, Phys. Rev. D **54**, 1474 (1996).

[3] E. Barausse, C. Palenzuela, M. Ponce, and L. Lehner, Phys. Rev. D **87**, 081506 (2013).

[4] J. Zhao, L. Shao, Z. Cao, and B.-Q. Ma, Phys. Rev. D **100**, 064034 (2019).

[5] B. Bertotti, L. Iess, and P. Tortora, Nature **425**, 374 (2003).