Violation of the weak energy condition: Is it generic of spontaneous scalarization?

Marcelo Salgado* and Daniel Sudarsky†
Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Apartado Postal 70-543 México 04510 D.F., México

Ulises Nucamendi‡
Instituto de Física y Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, Edificio C-3, Ciudad Universitaria, Morelia, Michoacán, C.P. 58040, México
(Received 19 March 2004; published 15 October 2004)

It was recently shown by Whinnett and Torres [A.W. Whinnett and D. F. Torres, Astrophys. J. **603**, L133 (2004).] that the phenomenon of spontaneous scalarization in compact objects (polytropes) was accompanied also by a *spontaneous violation of the weak energy condition* (WEC). Notably, by the encounter of negative-energy densities as measured by a static observer at several points of the star. Here we argue that such a situation is not generic of scalar-tensor theories of gravity (STT). We support this conclusion by numerical results within a class of STT and by using three realistic models of dense matter. However, we show that the "angular parts" of the additional conditions needed for the WEC to hold $\rho_{\rm eff} + T_i^{i({\rm eff})} \geq 0$ tend to be "slightly violated" at the outskirts of the star.

DOI: 10.1103/PhysRevD.70.084027 PACS numbers: 04.50.+h, 04.25.Dm, 97.60.Jd

I. INTRODUCTION

Scalar-tensor theories of gravity (STT) are one of the simplest alternatives to Einstein's general relativity (GR) [1]. A particular example of these is the Brans-Dicke theory. The general feature of STT is the introduction of a new fundamental scalar field which couples to gravity in a nonminimal fashion (NMC). Nonetheless, the NMC is such that it preserves the equivalence principle, because the scalar field couples to the ordinary matter only through the spacetime metric.

The NMC in STT unlike simple Einstein-Higgs systems gives rise to field equations that can differ substantially from the usual Einstein's equations. Therefore, when applied to the astrophysical and cosmological settings these theories can produce severe deviations from the general relativity predictions. The fact that STT introduce in general only a few new parameters makes them in principle easily testable. Perhaps the most notable probes in this direction are those connected with the binary pulsar, the primordial nucleosynthesis, the cosmic microwave background (CMB), and the luminosity distance with type II supernova. The success of GR lies in the fact that with suitable matter models, the theory is compatible with the available information in the above systems. Modifications of the gravitational sector in the way of STT lead thus to a serious change in the geometry-matter relation that can upset the compatibility of the theory with the observations. The stringent bounds on STT make it difficult to accommodate these theories to all the observations while at the same time predicting new observable phenomena. One of the potentially observable new phenomenon is precisely the phenomenon of spontaneous scalarization (SC) arising in compact objects. As was first shown by Damour and Esposito-Farèse [2], STT can induce nonperturbative effects in neutron stars (NS) which, roughly speaking, consists of a phase transition that endows a compact star with a new global quantity, the *scalar charge*. This "charge" is the analogous of the *magnetization* in ferromagnets at low temperatures and the central-energy density (or equivalently the baryon mass) of the object plays in turn the role of the temperature in spontaneous magnetization. Damour and Esposito-Farèse [2,3] also show that the consequences of SC can be dramatic even if the boundary conditions (asymptotic conditions) for the scalar field are chosen so as to satisfy the solar-system bounds. In particular, they showed that the binary-pulsar dynamics is very sensitive to the new phenomenon, and therefore, that STT can be further constrained even if they successfully pass the local and the cosmological tests.

In this paper we will not be concerned with the confrontation of STT with the experiments, but rather with a result obtained by Whinnett and Torres [4] who presented evidence that the phenomenon of SC in NS is linked to a violation of the weak energy condition (WEC). Notably, Whinnett and Torres [4] found that the effective-energy density of STT as measured by a static observer and defined as $\rho_{\rm eff} := n^a n^b G_{ab}/(8\pi)$ (where G_{ab} is the Einstein tensor and n^a a unit timelike vector parallel to the static Killing vector ξ_t^a) is negative and suggested that such a finding might be independent of the class of STT used (although depending on the chosen value of the NMC constant of the theory). A negative $\rho_{\rm eff}$ is a sufficient condition for the violation of the WEC. The authors argue that the violation of the WEC in NS might lead to their instability (even for masses quite below their corresponding Landau-Tolman-Volkoff-Oppenheimer limit),

*Electronic address: marcelo@nuclecu.unam.mx
†Electronic address: sudarsky@nuclecu.unam.mx
‡Electronic address: ulises@fis.cinvestav.mx

1550-7998/2004/70(8)/084027(7)$22.50

a very undesirable feature that would presumably rule out the STT as an alternative for a viable spacetime theory.

While it is true that the effective-energy density $\rho_{\rm eff}$ of STT as defined above is not automatically positive definite due to the presence of second order derivatives of the scalar field, the main goal of this paper is to show that there are classes of STT that do not violate the condition $\rho_{\rm eff} \geq 0$ for the problem at hand. As we shall argue in Sec. III, the violation of the condition $\rho_{\rm eff} \geq 0$ as found by Whinnett and Torres [4] is linked to the use of a parametrization *à la* Brans-Dicke which prejudices the choice of the specific theory which can potentially violate such a condition. In our case, we use a parametrization of the STT where the effective gravitational constant can be chosen to be positive definite. In such a case, as shown in Sec. III, the theories parametrized in both ways may be very different (the transformation that would be necessary to change from one parametrization to the other might not be well defined).

Although the class of STT analyzed by us satisfy the condition $\rho_{\rm eff} \geq 0$ for the NS of interest, they are prone to violate "slightly" the angular parts of the additional conditions $\rho_{\rm eff} + T_i^{i({\rm eff})} \geq 0$ for the WEC to hold. Since this violation is rather weak and occurs when the energy density and pressure of the star fluid is very low (near the star surface), we argue that the stability of the neutron star is not in jeopardy, and therefore that a large class of STT cannot be ruled out as viable theories by this argument alone.

II. THE MODEL

The general action for a STT with a single scalar field is given by

$$S = \int \left\{ \frac{1}{16\pi} F(\phi) R - \left[\frac{1}{2} (\nabla\phi)^2 + V(\phi) \right] \right\} \sqrt{-g} d^4x + S_{\rm matt}, \tag{1}$$

where the matter sector will be represented by a perfect fluid.

The field equations obtained from the above action are

$$G_{ab} = 8\pi T_{ab}^{\rm eff}, \tag{2}$$

$$T_{ab}^{\rm eff} = G_{\rm eff}(T_{ab}^{F} + T_{ab}^{\phi} + T_{ab}^{\rm matt}), \tag{3}$$

$$T_{ab}^{F} = \frac{1}{8\pi}[\nabla_a(\partial_\phi F \nabla_b \phi) - g_{ab}\nabla_c(\partial_\phi F \nabla^c \phi)], \tag{4}$$

$$T_{ab}^{\phi} = (\nabla_a\phi)(\nabla_b\phi) - g_{ab}[\tfrac{1}{2}(\nabla\phi)^2 + V(\phi)], \tag{5}$$

$$T_{ab}^{\rm matt} = (\rho + p)u_a u_b + g_{ab} p, \tag{6}$$

$$\Box\phi = \frac{F\partial_\phi V - 2(\partial_\phi F)V + \frac{1}{2}(\partial_\phi F)[T_{\rm matt} - (1 + \frac{3\partial^2_{\phi\phi}F}{8\pi})(\nabla\phi)^2]}{F + \frac{3(\partial_\phi F)^2}{16\pi}}, \tag{7}$$

where $T_{\rm matt}$ stands for the trace of $T^{ab}_{\rm matt}$, $G_{ab} = R_{ab} - \frac{1}{2} g_{ab} R$, and $G_{\rm eff} = \frac{1}{F}$.

It is to be emphasized that the action (1) corresponds to STT in the physical (Jordan) frame, as opposed to the one written in terms of an unphysical (conformally transformed) metric (the Einstein frame). Hereafter we consider $V(\phi) \equiv 0$.

In a previous work [5] (hereafter SSN), we analyzed the SC phenomenon in detail for the class of STT given by $F(\phi) = 1 + 16\pi\xi\phi^2$ (here ξ stands for the NMC constant) and for spherically symmetric (static) neutron stars. In that analysis, we exhibited the SC phenomenon for the cases $\xi = 2, 6$ and employed three representative realistic equations of state (EOS) for the nuclear matter: the model of Pandharipande [6] (hereafter PandN, representative of a "soft" EOS), the model II of Díaz-Alonso [7] (hereafter DiazII, representative of a "medium" EOS), and the model "0.17" of Haensel *et al.* [8] (hereafter HKP, representative of a "stiff" EOS). Our conclusion in SSN was that the SC is a phenomenon that ensues independently of the details of the EOS, but that strongly depends on the compactness of the star (high central-energy density). Moreover, we provided a heuristic analysis that clarifies why NS configurations with SC are energetically preferred over those with a trivial scalar field, despite the unintuitive expectation coming from the fact that the effective gravitational constant $G_{\rm eff} = 1/F$ decreases as the scalar field departs from the trivial configuration, which in turn leads to a reduction of the absolute value of the star's negative binding energy.

In order to analyze the SC phenomenon we computed hundreds of NS configurations. Now, to inquire about the status of the WEC, we take representative data of our catalog and recompute the energy-density profiles $\rho_{\rm eff}(r)$ as well as the other two relevant quantities $\rho_{\rm eff} + T_r^{r({\rm eff})}$ and $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ (where the spherical symmetry implies $T_\phi^{\phi({\rm eff})} = T_\theta^{\theta({\rm eff})}$). The WEC will be violated if any of the three above quantities is negative.

For the static and spherically symmetric metric $g_{ab} = {\rm diag}(-N^2, A^2, r^2, r^2\sin^2\theta)$, the effective-energy density $\rho_{\rm eff} := n^a n^b T_{ab}^{\rm eff}$, and the quantities $\rho_{\rm eff} + T_r^{r({\rm eff})}$ and $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$, respectively [where $T_{ab}^{\rm eff}$ is given by Eq. (3), and n^a stands for the unit timelike vector associated with a static observer] we have

$$\rho_{\rm eff} = \frac{G_{\rm eff}}{1+192\pi\xi^2\phi^2G_{\rm eff}}\Big[-\frac{4\xi\phi(\partial_r\phi)(\partial_r\tilde{N})}{\tilde{N}A^2} \times(1+192\pi\xi^2\phi^2G_{\rm eff}) + \frac{1}{2A^2}(\partial_r\phi)^2 \times(1+8\xi+64\pi\xi^2\phi^2G_{\rm eff}) + \rho + 64\pi\xi^2\phi^2G_{\rm eff} \times(2\rho+3p)\Big], \quad (8)$$

$$\rho_{\rm eff} + T_r^{r({\rm eff})} = \frac{G_{\rm eff}}{1+192\pi\xi^2\phi^2G_{\rm eff}}\Big[-\frac{8\xi\phi\partial_r\phi}{A^2} \times\Big(\frac{1}{r}+\frac{\partial_r\tilde{N}}{\tilde{N}}\Big)(1+192\pi\xi^2\phi^2G_{\rm eff}) + \frac{1}{A^2} \times(\partial_r\phi)^2(1+4\xi+128\pi\xi^2\phi^2G_{\rm eff}) + \rho + p + 128\pi\xi^2\phi^2G_{\rm eff}(\rho+3p)\Big], \quad (9)$$

$$\rho_{\rm eff} + T_\theta^{\theta({\rm eff})} = G_{\rm eff}\Big[\frac{4\xi\phi\partial_r\phi}{A^2}\Big(\frac{1}{r}-\frac{\partial_r\tilde{N}}{\tilde{N}}\Big) + \rho + p\Big], \quad (10)$$

which explicitly shows the usual contributions of the perfect fluid (ρ and p) and the scalar field plus those arising due to the NMC. Here $\tilde{N} = N/N_0$ is the renormalized lapse with respect to its value at $r = 0$. Clearly when $\phi(r) \equiv 0$, one simply recovers $\rho_{\rm eff} = \rho$, and $\rho_{\rm eff} + T_r^{r({\rm eff})} = \rho + p = \rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$.

We have integrated the system of equations with adequate regularity and boundary conditions (leading to asymptotically flat spacetimes) and with $\phi(r \to \infty) \sim Q_s/r$ where Q_s is the scalar charge. For a fixed ξ, all the NS models are parametrized by the central-energy density of the fluid ρ_0. Therefore a critical value $\rho_0^{\rm crit}$ (or equivalently $M_{\rm bar}^{\rm crit}$) characterizes the onset of SC. Figures 1–3 depict $\rho_{\rm eff}$ (left panels), $\rho_{\rm eff} + T_r^{r({\rm eff})}$ (right panels; solid curves), and $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ (right panels; dashed curves). The lower mass ($M_{\rm bar}^{\rm crit}$) curves of Figs. 1–3 indicate the critical configurations at the transition to SC while the rest of the curves correspond to NS endowed with a "scalar charge" Q_s.

The important point to note is that none of the curves of Figs. 1–3 show a violation of the condition $\rho_{\rm eff} \geq 0$. From the figures we appreciate that the condition $\rho_{\rm eff} + T_r^{r({\rm eff})} \geq 0$ is also satisfied. However, the quantity $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ starts becoming slightly negative near the surface of the star.

As we have stressed above, $\rho_{\rm eff}$ has not an *a priori* definite sign, and in fact, the term linear in $\partial_r\phi$ could be negative within the NS. Moreover, one can define an energy density that is useful to understand the appearance of the phenomenon of SC, and is the one given as follows [5]:

$$\rho^\xi = \frac{G_{\rm eff}}{1+192\pi\xi^2\phi^2G_{\rm eff}} \times[\rho + 64\pi\xi^2\phi^2G_{\rm eff}(2\rho+3p)] - \rho, \quad (11)$$

which heuristically encompasses the contribution of the NMC to the energy density of the perfect fluid. Note, however, that $\rho_{\rm eff} = \rho + \rho^\xi + \cdots$ with $\rho + \rho^\xi > 0$. Therefore $\rho_{\rm eff} > 0$, since typically $|-\phi(\partial_r\phi) \times (\partial_r\tilde{N})/(\tilde{N}A^2)| \ll \rho$ within the NS. In our heuristic analysis (SSN) which was also confirmed by the numerical solutions, we showed that the negative contribution ρ^ξ more than compensates the decrease in the absolute value of the binding energy due to the reduction of $G_{\rm eff}$, resulting in an overall lower total energy [the Arnowitt-Deser-Misner (ADM) energy that is obtained from the integration of Eq. (8)] as compared with the energy of the

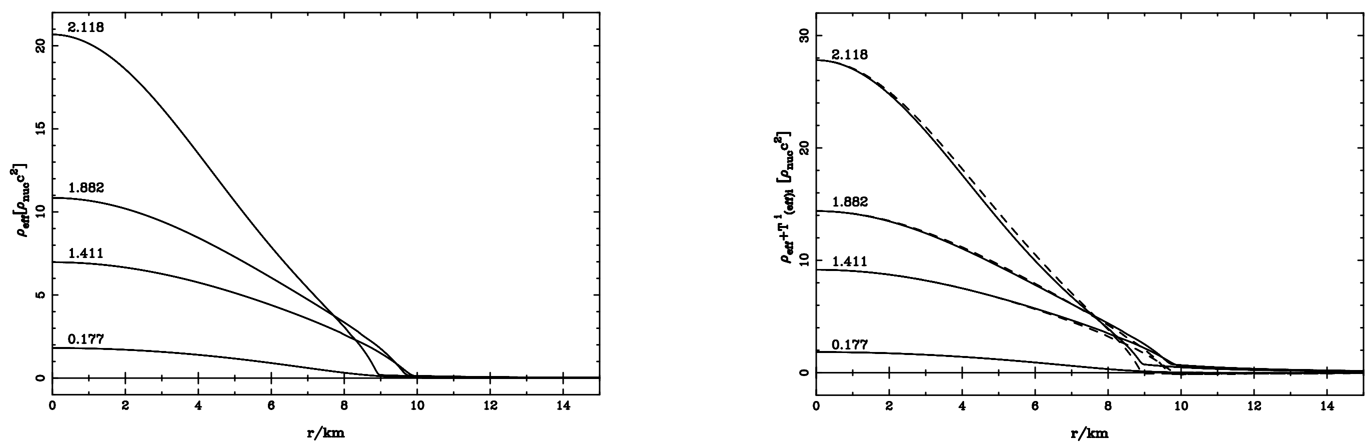

FIG. 1. Neutron star models constructed with the EOS PandN within the scalar-tensor theory $F(\phi) = 1 + 16\pi\xi\phi^2$, with $\xi = 6$. Left panel: effective-energy densities $\rho_{\rm eff}$. Right panel: $\rho_{\rm eff} + T_r^{r({\rm eff})}$ (solid lines) and $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ (dashed lines). In both panels the curves are labeled by the corresponding baryon mass (in solar mass units). The lowest baryon-mass configuration marks the onset of spontaneous scalarization. The figure also shows the maximum mass models. Here $\rho_{\rm nuc} = 1.66 \times 10^{17}$ kg m^{-3}.

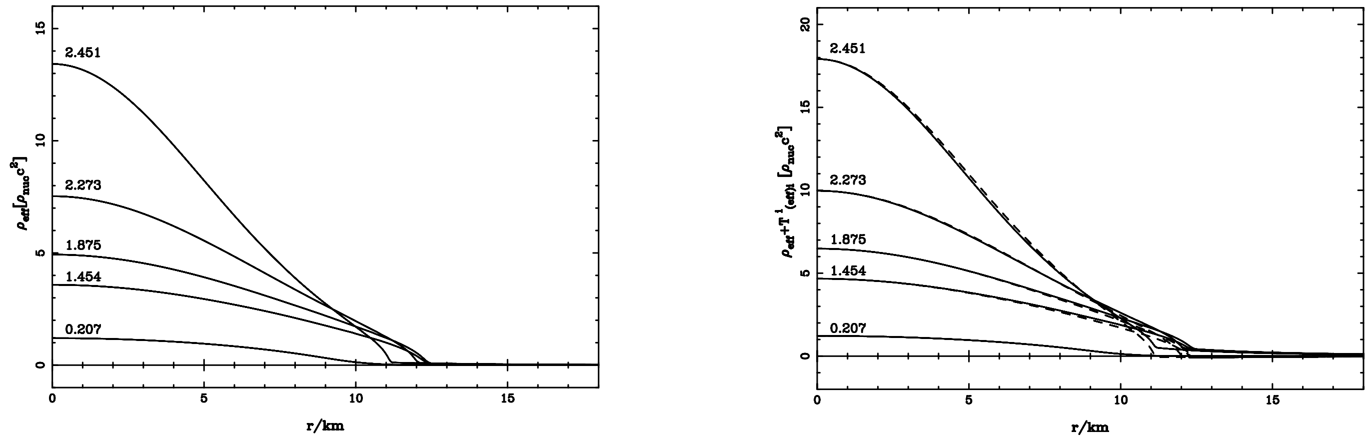

FIG. 2. Same as Fig. 1 for the EOS DiazII.

configuration with the same baryon mass but with a null scalar field (a configuration in pure GR).

III. DISCUSSION

In order to give an insight for the non-negativeness of the effective-energy density let us consider Eq. (3) without specifying $F(\phi)$ and compute $\rho_{\rm eff} = n^a n^b T^{\rm eff}_{ab}$:

$$\rho_{\rm eff} = \frac{1}{F}\bigg[\frac{\partial_\phi F}{8\pi} n^a n^b \nabla_a \nabla_b \phi + \frac{\frac{(\nabla\phi)^2}{2}\big(1 + \frac{(\partial_\phi F)^2}{16\pi F} + \frac{\partial^2_{\phi\phi} F}{4\pi}\big) + \rho + \frac{(\partial_\phi F)^2}{16\pi F}(2\rho + 3p)}{1 + \frac{3(\partial_\phi F)^2}{16\pi F}}\bigg]. \tag{12}$$

This reduces to Eq. (8) for $F(\phi) = 1 + 16\pi\xi\phi^2$ and for the static and spherically symmetric case.

For usual nuclear matter ($\rho > 0$ and $p > 0$) and for $F > 0$, $\partial^2_{\phi\phi} F > 0$, only the first term is not positive semidefinite. Whinnett and Torres [4] found a negative $\rho_{\rm eff}$ at various star positions, notably at the center of the star ($r = 0$). So let us first focus on $\rho_{\rm eff}$ at $r = 0$. By regularity at the origin, $\partial_r \phi|_{r=0} = 0$, therefore

$$\rho^0_{\rm eff} = \frac{1}{F_0}\bigg[\frac{\rho_0 + \frac{(\partial_\phi F)^2_0}{16\pi F_0}(2\rho_0 + 3p_0)}{1 + \frac{3(\partial_\phi F)^2_0}{16\pi F_0}}\bigg]. \tag{13}$$

We can now investigate in which situations $\rho^0_{\rm eff}$ can be negative [9].

(a) STT with $F(\phi) > 0$ (in particular $F_0 > 0$). Clearly for this case, $\rho^0_{\rm eff} > 0$ (since we consider only $\rho_0 > 0$ and $p_0 > 0$). This includes the class of STT we analyzed numerically $F(\phi) = 1 + 16\pi\xi\phi^2$, $\xi > 0$, for which we did not encounter negative effective central-energy densities [cf. Figs. 1–3 (left panels)].

(b) STT with $F(\phi) < 0$ for some ϕ (in particular $F_0 < 0$). Clearly $F_0 < 0$ is a necessary condition for $\rho^0_{\rm eff} < 0$. One can then obtain $\rho^0_{\rm eff} < 0$ if in addition the following inequalities hold: $\rho_0 + [(\partial_\phi F)^2_0/16\pi F_0](2\rho_0 + 3p_0) > 0$ and $1 + [3(\partial_\phi F)^2_0/16\pi F_0] > 0$. From this, we have the following two possibilities: (1) If $\rho_0 - 3p_0 \geq 0$ then the two inequalities hold provided $16\pi|F_0| > 3(\partial_\phi F)^2_0$; (2) if $3p_0 - \rho_0 \geq 0$ then the two inequalities hold provided $16\pi|F_0| > C(\partial_\phi F)^2_0$ where $C = 2 + 3p_0/\rho_0$.

According to (b), it is in principle possible to obtain $\rho^0_{\rm eff} < 0$ for $F_0 < 0$. However, this case is completely

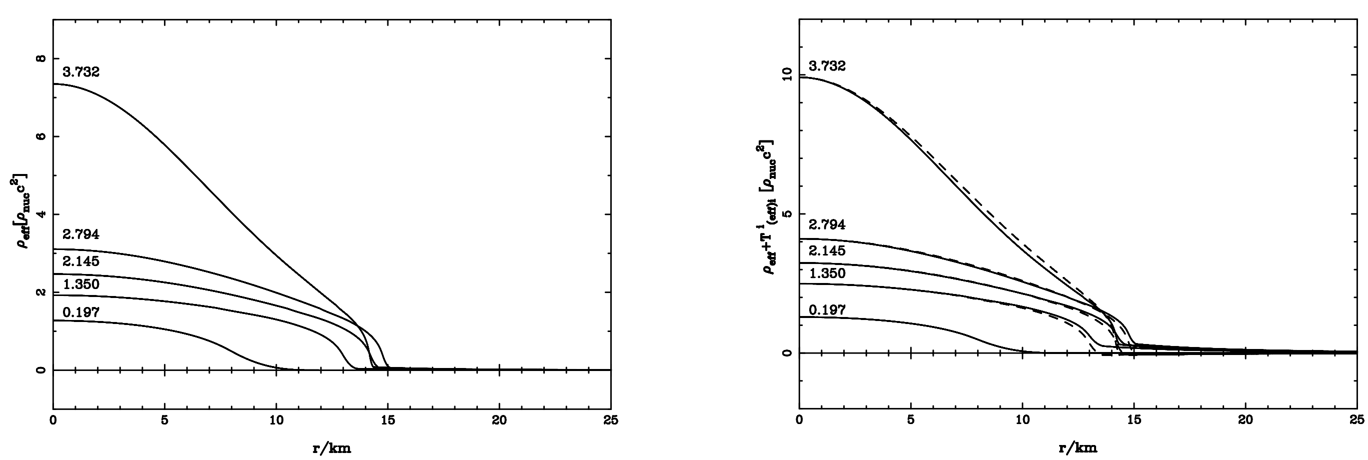

FIG. 3. Same as Fig. 1 for the EOS HKP.

pathological: on one hand at the center of the star F_0 is negative (which in turn implies a negative effective gravitational constant $G^0_{\rm eff} = 1/F_0 < 0$). On the other hand, asymptotically $F = G^{-1}_{\rm eff} \sim 1 > 0$ (i.e., $G_{\rm eff}$ approaches the Newtonian value) in order to pass successfully the solar-system bounds. Therefore $F(r) = F[\phi(r)]$ must interpolate between the negative and the positive value, implying that at some spherical shell (inside or outside the star) $F = 0$ and thus $G_{\rm eff} \to \infty$, where the equations become singular.

In order to avoid these kinds of pathologies within a neutron star model one should consider only the cases $F(\phi) > 0$ which lead us to the conclusion that any physically viable STT will not produce a negative central effective-energy density [10].

Now, although we have just proven rigorously that $F_0 > 0$ implies $\rho^0_{\rm eff} > 0$, we still have to prove the same for $\rho_{\rm eff}(r)$ given by Eq. (12). Although this could be a difficult task, the empirical numerical evidence shows that if $\rho^0_{\rm eff} > 0$ and for a class of STT (e.g., a class of STT with $F > 0$ and $\partial^2_{\phi\phi} F > 0$) then $\rho_{\rm eff}(r) \geq 0$ for the same class due to the fact that $|(\partial_\phi F) n^a n^b \nabla_a \nabla_b \phi| \ll \rho$ within a NS.

The above heuristic analysis which is supported by our numerical results leads us to the following conjecture:

Conjecture.—Consider a STT with $F(\phi) > 0$ and $\partial^2_{\phi\phi} F > 0$ (in particular $F_0 > 0$ at the center of a neutron star which implies $\rho^0_{\rm eff} > 0$), such that asymptotically $F \sim 1$ and with $V(\phi) \equiv 0$, then $\rho_{\rm eff}(r) \geq 0$ for all the static NS configurations, including all those endowed with Q_s (where the equality holds only outside the star and for a trivial scalar field) [11].

Now, in order to give some insight about the violation of the WEC found by Whinnett and Torres [4], let us consider their Eq. (5) (where $\rho_{\rm eff}$ is denoted by μ) and focus on the central-energy density (where the regularity conditions and the spherical symmetry implies that the gradient contributions become null):

$$\mu_0 = \frac{1}{\Phi_0}\left[\rho_0 + \frac{(3p_0 - \rho_0)}{2\omega_0(\Phi_0) + 3}\right]. \tag{14}$$

This expression is in fact equivalent to our Eq. (13), which we showed to be positive definite for $F_0 > 0$ and for usual neutron matter ($\rho_0 > 0$ and $p_0 > 0$). The translation from the Brans-Dicke parametrization used in Eq. (14) to ours is provided by

$$\Phi = F(\phi), \qquad \omega(\Phi) = \frac{8\pi F}{(\partial_\phi F)^2}. \tag{15}$$

On the other hand μ_0 might not look as having a definite sign. We have explicitly

$$\mu_0 = \frac{1}{\Phi_0}[\rho_0(1 - \sigma_0) + 3\sigma_0 p_0], \tag{16}$$

where $\sigma(\Phi) = 1/(3 + 2\omega)$. Then, clearly due to the Brans-Dicke parametrization one could consider situations where $\sigma_0 > 1$ or $\sigma_0 < 0$, which depending on the values of ρ_0 and p_0 can lead to $\mu_0 < 0$. However, we note that the condition $F(\phi) > 0$ (in particular $F_0 > 0$) leading to $\rho^0_{\rm eff} > 0$ (or equivalently $\mu_0 > 0$) implies $\Phi > 0$ and also $0 \leq \sigma < 1/3$: In fact, from $\sigma(\Phi) = 1/(3 + 2\omega)$ and Eqs. (15) one obtains $\sigma(\Phi) = (\partial_\phi F)^2/[3(\partial_\phi F)^2 + 16\pi F]$ [thus $F(\phi) > 0$ implies $0 \leq \sigma < 1/3$]. Moreover, since $\Phi = F(\phi)$ [assuming $F(\phi) \neq$ const] the inverse mapping is $\phi = \int \sqrt{[1 - 3\sigma(\Phi)]/16\pi\Phi\sigma(\Phi)}\, d\Phi$. Therefore the conditions $\Phi > 0$ and $0 < \sigma < 1/3$ are required for this mapping $\Phi \to \phi$ to be well defined [the separate case $\sigma \equiv 0$ corresponds to $F(\phi) =$ const from which one simply recovers GR].

Hence, the restrictions on σ bound in turn the choice of the Brans-Dicke function $\omega(\Phi)$. Any choice of $\omega(\Phi)$ leading to a violation of the condition $0 \leq \sigma < 1/3$ with $\Phi > 0$ will imply a loss of the connection with the original theory. It is in this sense that the use of the Brans-Dicke parametrization prejudices the kind of STT (those STT which allow values for σ to be outside the bounds $0 \leq \sigma < 1/3$) that can lead to violations of the positiveness of $\rho_{\rm eff}$.

One could then use the SC phenomenon in NS to classify the STT in two types. Type I: STT with $\sigma(\Phi)$ such that $0 \leq \sigma < 1/3$ and $\Phi > 0$ (which, in particular, are valid at the star's center and therefore that coincide with the case (a) analyzed above). Type II: STT with $\sigma(\Phi)$ (in particular σ_0) not necessarily limited in the range $[0, 1/3)$ and with $\Phi_0 > 0$. The first kind of STT can be put in correspondence with the original parametrization ($F(\phi) > 0$) and these theories lead to $0 < \rho^0_{\rm eff} \equiv \mu_0$, regardless of the EOS for nuclear matter (assuming ρ and p non-negative). Type II is a STT which cannot be put in correspondence with the original parametrization (in terms of ϕ) since the transformation needed is not well defined. The type II theories can lead to situations where $\mu_0 < 0$ but are not even related to the theories which in the original parametrization give rise to $\rho^0_{\rm eff} < 0$ [case (b) above] which require $F_0 < 0$. For instance, it is possible to have $\mu_0 < 0$ with $\Phi_0 > 0$ with a suitable σ_0, as in the following situations: (A) $\rho_0/(\rho_0 - 3p_0) < \sigma_0$ with $3p_0 < \rho_0$; (B) $\sigma_0 < 0$ with $\rho_0/(3p_0 - \rho_0) < |\sigma_0|$ and $\rho_0 < 3p_0$. Whinnett and Torres [4] considered two classes of STT. The first one with $\sigma = 2\kappa \ln\Phi$ and $\kappa \geq 3$. This example can belong to type IIA (and produce $\mu_0 < 0$) if for instance $\sigma_0 \approx 2$, $p_0 < \rho_0/6$, and $1 < \Phi_0 \lesssim 1.4$. These conditions could be met in neutron stars with a low baryon mass like some of the ones computed by Whinnett and Torres [4].

The second class of STT used by those authors is one with $\sigma = \lambda(\Phi - 1)$ and $\lambda \geq 5.35$. This class is of type IIA or type IIB for $\Phi_0 > 1$ or $0 < \Phi_0 < 1$, respectively. As mentioned above, the condition $3p_0 < \rho_0$ in type IIA is met usually in NS with a low baryon mass. Since for this second class of STT, $\mu_0 < 0$ was found by Whinnett and Torres [4] in low baryon-mass configurations, it would seem that we must have $\Phi_0 > 1$ there.

We have thus given a heuristic explanation for a negative μ_0 in both of their classes of STT.

Of course, STT of type II can lead to positive μ_0's without implying that the WEC will not be violated somewhere within the star (see the figures of Whinnett and Torres [4]). On the other hand, the fact that STT's of type I lead to positive definite μ_0 does not imply that the WEC is not violated either. However, our numerical analysis shows that this type of STT (theories with $\mu_0 > 0$) indeed satisfy the condition $\mu(r) \geq 0$ (where the equality holds only outside the star and for a trivial scalar field) and this is the motivation for our conjecture above. Along the same line of reasoning one could further conjecture (conjecture 2) that given a fixed EOS, if a STT of type II violates the condition $\mu(r) \geq 0$ somewhere within a star (not necessarily at the center) then the STT will violate the WEC at the center for some other ρ_0 (defined as ρ_0^-). That is, the second conjecture states that $\mu_0 = f(\rho_0)$ has always a negative branch $[f(\rho_0^-) < 0]$ in the type II theories. The first conjecture on the other hand roughly states that a STT (with the given assumptions) that does not violate $\mu_0 > 0$ (type I with $\mu_0(\rho_0) > 0$) will not violate it elsewhere within the NS [i.e., $\mu(r) \geq 0$ within the star].

We can now turn the attention to the condition $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ which the numerical analysis shows to be negative near the star surface. From Eq. (10) we appreciate that beyond the compact region containing the nuclear fluid, $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})} \sim G_{\rm eff}\{(4\xi\phi\partial_r\phi/A^2)[(1/r) - (\partial_r\tilde{N}/\tilde{N})]\}$. If we roughly approximate $A^{-1} \approx N \approx (1 - 2M/r)^{1/2}$ for $r \geq R$ (where R and M stand for the star radius and total mass) then the factor $[(1/r) - (\partial_r\tilde{N}/\tilde{N})] \approx (1/r) \times [(1 - 3M/r)/(1 - 2M/r)]$ which is positive even for the maximum mass configurations. Therefore $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})} \sim G_{\rm eff}[4\xi\phi\partial_r\phi(1 - 3M/r)/r]$. Since the SC configurations have $\partial_r\phi < 0$ near and beyond R [$\phi(r) \sim Q_s/r$ asymptotically, with $Q_s > 0$], then $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ is negative (for $\xi > 0$) near and beyond the star's surface. Note from Eq. (9) that the term $\rho_{\rm eff} + T_r^{r({\rm eff})}$ [unlike $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$; cf. Eq. (10)] does not become negative because of the sign difference in the term with $\xi\phi\partial_r\phi$, and the remaining terms involving the scalar field are positive definite (for $\xi > 0$).

In summary, there exists classes of STT that give rise to the phenomenon of spontaneous scalarization with and without a violation of the condition $\rho_{\rm eff} \geq 0$. Examples of STT like the ones we analyzed numerically do not violate that condition. However they do violate the WEC slightly because $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ becomes slightly negative near and beyond the star's surface. On the other hand, there are STT like the ones analyzed by Whinnett and Torres [4] which do violate the condition $\rho_{\rm eff} \geq 0$ which straightforwardly leads to a violation of the WEC. In our case it is unclear whether the small violation of the WEC by the STT considered here will necessarily lead to the instability of the NS. In Ref. [5], we have shown that the energetically preferred NS configurations (with a fixed total baryon mass) are those with lower ADM mass which correspond to scalarized NS as opposed to the corresponding NS in pure GR. The curves of the ADM as a function of the central-energy density ρ_0 suggest that configurations with ADM mass smaller than the maximum mass models (corresponding to $\rho_0^{\rm max}$) are stable with respect to radial linear perturbations [5]. Moreover, numerical nonlinear analysis strongly suggest that ordinary NS (with weak or zero scalar charge) can decay to their corresponding more stable strong scalarized configurations [12] and furthermore that among these scalarized NS only the configurations with $\rho_0 > \rho_0^{\rm max}$ are unstable and can collapse into a Schwarszschild black hole like ordinary NS, by radiating away all the scalar field [13].

If it were found that only the scalarized NS which violate the WEC through the violation of $\rho_{\rm eff} \geq 0$ condition within the star are unstable, then only the STT allowing such violations would be ruled out by this argument. On the other hand, it is very possible that "small" violations of the WEC through the term $\rho_{\rm eff} + T_\theta^{\theta({\rm eff})}$ in the outskirts of the NS while respecting $\rho_{\rm eff} \geq 0$ within the star might not lead to instabilities and therefore that the corresponding STT are viable candidates for a spacetime theory.

IV. CONCLUSION

Our main conclusion is that the large violations of the WEC deep inside the neutron stars as an additional feature to the phenomenon of spontaneous scalarization is not generic of all the classes of scalar-tensor theories of gravity as the work of Torres and Whinnett might be thought to suggest. We have argued that there are classes of STT [at least a subset with $F(\phi) > 0$ and $\partial^2_{\phi\phi}F > 0$] where $\rho_{\rm eff} \geq 0$ and where the violations of the WEC in neutron stars occur only near and beyond the surface of the star via the angular parts of $\rho_{\rm eff} + T_i^{i({\rm eff})}$. Strong numerical evidence supports this conclusion.

ACKNOWLEDGMENTS

M. S. and D. S. were partially supported by DGAPA-UNAM Grants No. IN112401, No. IN122002, and No. IN108103. U. N. acknowledges partial support from SNI, and Grants No. 4.8 CIC-UMSNH, No. PROMEP PTC-61, and No. CONACYT 42949-F.

[1] T. Damour and G. Esposito-Farèse, Classical Quantum Gravity **9**, 2093 (1992).
[2] T. Damour and G. Esposito-Farèse, Phys. Rev. Lett. **70**, 2220 (1993); Phys. Rev. D **54**, 1474 (1996).
[3] T. Damour and G. Esposito-Farèse, Phys. Rev. D **58**, 042001 (1998).
[4] A.W. Whinnett and D. F. Torres, Astrophys. J. **603**, L133 (2004).
[5] M. Salgado, D. Sudarsky, and U. Nucamendi, Phys. Rev. D **58**, 124003 (1998) (referred to as SSN in the text).
[6] V. R. Pandharipande, Nucl. Phys. **A174**, 641 (1971).
[7] J. Díaz-Alonso, Phys. Rev. D **31**, 1315 (1985).
[8] P. Haensel, M. Proszynski, and M. Kutschera, Astron. Astrophys. **102**, 299 (1981).
[9] It is important to note that $\phi(r) = \text{const} \neq 0$ cannot solve the Klein-Gordon equation (7) within a neutron star. Therefore this excludes the possibility of a neutron star configuration with a nonzero but homogeneous scalar field. As a consequence, once a $\phi_0 \neq 0$ ensues at the center of the star (this depends on the value of the central-energy density) the only possibility is an inhomogeneous scalar field which interpolates between ϕ_0 and the asymptotic value. This is another way to view the onset of the spontaneous scalarization
[10] Clearly this conclusion could dramatically change with the inclusion of a potential $V(\phi)$.
[11] A class of STT that can jeopardize this conjecture is with $F(\phi) = \chi e^{\xi\phi^2}$ $(\chi, \xi > 0)$, since $\rho_{\nabla\nabla} := (\partial_\phi F) n^a n^b \nabla_a \nabla_b \phi \sim -\xi e^{\xi\phi^2} \phi(\partial_r \phi)(\partial_r \tilde{N})/(\tilde{N} A^2)$. So for ξ large enough and with $\phi(\partial_r\phi)(\partial_r\tilde{N}) > 0$ in some region within the star, it is possible to imagine a scenario with $\rho_{\nabla\nabla} < 0$ and $|\rho_{\nabla\nabla}| > \rho$ so that $\rho_{\rm eff}$ has regions with negative values. In this case the neutron star would probably be unstable or even such class of STT with large ξ would be ruled out by cosmological arguments (such as primordial nucleosynthesis, galaxy formation, CMB, etc.). Needless to say one requires further numerical studies to explore a concrete realization of this possibility.
[12] J. Novak, Phys. Rev. D **57**, 4789 (1998).
[13] J. Novak, Phys. Rev. D **58**, 064019 (1998).