Spontaneous scalarization

Daniela. D. Doneva

Theoretical Astrophysics, Eberhard Karls University of Tübingen, 72076 Tübingen, Germany
INRNE - Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria

Fethi M. Ramazanoğlu

Department of Physics, Koç University, Rumelifeneri Yolu, 34450 Sariyer, Istanbul, Turkey

Hector O. Silva

Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, D-14476 Potsdam, Germany

Thomas P. Sotiriou

Nottingham Centre of Gravity and School of Mathematical Sciences and School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom

Stoytcho S. Yazadjiev

Theoretical Astrophysics, Eberhard Karls University of Tübingen, 72076 Tübingen, Germany
Department of Theoretical Physics, Faculty of Physics, Sofia University, 1164 Sofia, Bulgaria
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Academy Georgi Bonchev Street 8, 1113 Sofia, Bulgaria

Scalarization is a mechanism that endows strongly self-gravitating bodies, such as neutron stars and black holes, with a scalar-field configuration. It resembles a phase transition in that the scalar configuration appears only when a certain quantity that characterizes the compact object, for example, its compactness or spin, is beyond a threshold. A critical and comprehensive review of scalarization, including the mechanism itself, theories that exhibit it, its manifestation in neutron stars, black holes and their binaries, potential extension to other fields, and a thorough discussion of future perspectives, is provided.

CONTENTS

I. Introduction 2

II. Theoretical background: mechanism and theories
   A. Spontaneous scalarization mechanism
      1. Tachyonic instability and nonlinear quenching 3
      2. Tachyonic instability in curved spacetime 3
      3. Scalarization and gravity 3
      4. Strong-field phase transitions and weak-field screening 4
   B. Models of scalarization
      1. Tachyonic instability and the minimal action 5
      2. Damour–Esposito–Farèse Model 7
      3. Scalar–Gauss–Bonnet gravity 8
      4. The Ricci–Gauss–Bonnet model 9
      5. Tensor-multiscalar theories 9
   C. Types of scalarization
      1. Induced by compactness 10
      2. Induced by spin 10
      3. Induced by matter or coupling to other fields 10
      4. Dynamical scalarization 11
      5. Beyond scalarization 11
   D. Quantum aspects and classical analogs 11

III. Neutron star scalarization
   A. Equilibrium neutron stars in the Damour–Esposito–Farèse model 12
      1. The original Damour–Esposito–Farèse model and binary-pulsar constraints 12
      2. Massive scalar field 16
      3. Incorporating further physics 17
   B. Dynamics of scalarized neutron stars and binary mergers 18
      1. Linearized dynamics 18
      2. Nonlinear stability and collapse to a black hole 20
      3. Stellar core collapse 21
      4. Dynamical scalarization and neutron star mergers 22
   C. Astrophysical implications of scalarized neutron stars in the Damour–Esposito–Farèse model 23
   D. Extended scalar-tensor theories beyond the Damour–Esposito–Farèse model 26

IV. Black-hole scalarization
   A. Black-hole scalarization: Vacuum spacetimes 28
      1. Scalarized black holes 29
      2. Stability of scalarized black holes and implications for model building 30
      3. Spin-induced scalarization 32
      4. Scalarized black holes in binary systems 33
   B. Black-hole scalarization in the presence of matter 34
   C. Variations of the curvature-induced scalarization model 35

V. Generalizations of scalarization to other fields and instabilities
   A. Spontaneous vectorization 36
   B. Vectorization and ghosts 38
   C. Spontaneous tensorization 40
I. INTRODUCTION

Exploring the nature of gravity in the strong curvature regime has seen a recent surge of interest. This is expected to intensify, as it is driven by current and future observations of compact objects: black holes (BHs) and neutron stars (NSs). In particular, gravitational waves (GWs) produced by coalescing compact binaries have by now been routinely detected by the LIGO-Virgo-Kagra (LVK) Collaboration (Abbott et al., 2019, 2021a,b). These observations enabled us to probe the highly dynamical and strong-field regime of general relativity (GR) for the first time. They have enabled one to perform new tests of GR and to constrain modifications thereof in a hitherto unexplored regime. Future spaceborne and ground-based GW observatories have testing GR and the standard model (SM) among their key priorities (Arun et al., 2022; Barausse et al., 2020; Kalogera et al., 2021; Sathyaprakash et al., 2019). At the same time, there is a new suite of electromagnetic observations that probe NSs with unprecedented sensitivity and timing resolution (Arzoumanian et al., 2014; Gendreau and Arzoumanian, 2017; Gendreau et al., 2012). On other fronts, the precision timing of binary pulsars has continually improved (Kramer et al. (2021)), measurements of the motion of stars at the Galactic Center are becoming more precise (Abuter et al., 2018, 2020; Do et al., 2019), and we have witnessed breakthroughs in supermassive BH imaging (Akiyama et al., 2019).

An interesting prospect is that these observations may reveal the existence of some otherwise elusive new fundamental fields, which could be an ingredient of new physics beyond the SM or beyond GR; see Clifton et al. (2012), Yunes and Siemens (2013), Berti et al. (2015), and Barack et al. (2019). For such fields to have remained undetected, there has to exist a mechanism to suppress them when gravity is weak. For scalar fields, which are ubiquitous in extensions of the SM and of GR, a possible realization of such a mechanism was first proposed by Damour and Esposito-Farèse (1993) and named spontaneous scalarization. They showed that a specific type of nonminimal coupling between scalar field and gravity (or matter, after a field redefinition) leads to a theory that is indistinguishable from GR in weak-field gravitational experiments and yet predicts order unity deviations from general-relativistic expectations in the strong-gravity regime of NSs.

As today, the first model of spontaneous scalarization came about at a time in which gravitational experiments were producing new data from a then unexplored regime of gravity: the slow-velocity but strong-field regime of binary pulsars discovered by Hulse and Taylor (1975); see Damour (2015). This discovery inaugurated a new arena to test GR and its contenders (Damour and Taylor, 1992; Taylor and Weisberg, 1982; Taylor et al., 1993). Meanwhile, slow velocities and weak-field Solar System tests had reached an accuracy that made it questionable whether viable theories that predict deviations from GR that are measurable with binary pulsars can exist (Will, 2014). Spontaneous scalarization settled this question and provided further motivation for the use of binary-pulsar observations to test GR. By now these observations have ruled out the Damour-Esposito-Farèse (DEF) scalarization model (Antoniadis et al., 2013; Kramer et al., 2021; Zhao et al., 2022).

In recent years, however, spontaneous scalarization has received renewed interest. This is due to the realization that vacuum BH solutions of GR can also scalarize when the scalar field or fields couple suitably to the spacetime curvature (Doneva and Yazadjiev, 2018b; Silva et al., 2018). This development also showed that the earlier DEF model is part of a much broader class of theories (Andreou et al., 2019) that exhibit what resembles a phase transition in the strong field: once a quantity that describes a compact object, such as its compactness (Damour and Esposito-Farèse, 1993; Doneva and Yazadjiev, 2018b; Silva et al., 2018) or spin (Dima et al., 2020), exceeds a certain threshold, the scalar field switches from a trivial constant configuration to a nontrivial one, and large deviations from GR appear. Conversely, one can think of this as deviations from GR getting severely “screened” as soon as one crosses the same threshold in the opposite direction. It is this phase transition behavior that distinguishes scalarization from other models in which deviation from GR is induced and controlled by coupling to curvature, such as more general scalar-tensor theories with linear (Sotiriou and Zhou, 2014a,b; Yunes and Stein, 2011) or exponential (Kanti et al., 1996) couplings to the Gauss-Bonnet invariant.

Since the advent of GW astronomy, the broader class of theories that exhibit spontaneous scalarization have played a role similar to the one that the DEF model has played for binary-pulsar observations. These theories provide a putative explanation of why we have not detected new fundamental fields with existing observations, but we might still uncover them with high precision observations of astrophysical systems with specific characteristics.

The aim of this review is to summarize, in a unified manner, the current status of this field. In Sec. II, we start by providing the theoretical background of the scalarization mechanism and its various subcases, following a
pedagogical, rather than historical, approach. Next we discuss in more detail the literature on scalarization, first of NSs in Sec. III and then of BHs in Sec. IV. In so doing, we present the state of the art of our understanding of the consequences of scalarization for various situations of observational interest. In Sec. V, we discuss attempts to generalize scalarization to other field types. In Sec. VI, we outline open issues and summarize future perspectives.

II. THEORETICAL BACKGROUND: MECHANISM AND THEORIES

A. Spontaneous scalarization mechanism

1. Tachyonic instability and nonlinear quenching

Before we discuss spontaneous scalarization in the context of gravity, it is instructive to review the dynamics of a real scalar field $\varphi$ with a quartic self-interaction in Minkowski spacetime. The Lagrangian for this field is

$$\mathcal{L} = \frac{1}{2} \eta^{\mu\nu} \partial_{\mu} \varphi \partial_{\nu} \varphi + V(\varphi),$$

where $\eta_{\mu\nu}$ is the Minkowski metric

$$V(\varphi) = \frac{1}{2} \mu^2 \varphi^2 + \frac{1}{4} \lambda \varphi^4,$$

$\mu$ is the bare mass, and $\lambda$ is a coupling constant. The scalar then satisfies the following field equation:

$$\Box \varphi - \mu^2 \varphi - \lambda \varphi^3 = 0,$$

where $\Box = \eta^{\mu\nu} \partial_{\mu} \partial_{\nu}$ is the flat-spacetime d'Alembertian. $\varphi = 0$ is a solution of this equation. Consider now small perturbations $\delta \varphi$ around $\varphi = 0$. By linearizing Eq. (3), we find that $\delta \varphi$ obeys

$$\Box \delta \varphi - \mu^2 \delta \varphi = 0.$$

The corresponding dispersion relation is $\omega^2 = k^2 + \mu^2$, where $\omega$ is the frequency and $k$ is the wavenumber. If $\mu^2 > 0$, the relevant solutions to this equation are plane waves and the perturbations decay. If instead $\mu^2 < 0$, one encounters a tachyonic instability and the perturbations with small wave number exhibit exponential growth.

This exponential growth seems catastrophic at first sight, but it does not have to be. As $\varphi$ grows, the previously used linear approximation will quickly become invalid, and the nonlinear self-interaction $\lambda \varphi^3$ will become important. It will be this interaction that will determine the end point of the instability. Assume that $\lambda > 0$ (and $\mu^2 < 0$), in which case the potential has the well-known "Mexican hat" shape. Equation (3) will then admit a second solution with constant $\varphi$, which we denote as $\varphi_{\text{min}}$, as it will correspond to the minimum of the potential. Equation (3) implies that $\varphi_{\text{min}}^2 = -\mu^2/\lambda$. Thus, the tachyonic instability simply drives the scalar field away from the unstable local maximum of the potential and toward a stable minimum. This is sometimes referred to as tachyon condensation and is associated to a phase transition of the system.

The key message from this simple example is that linearized perturbations around the unstable maximum capture the onset of the tachyonic instability but they are oblivious to the shape of the rest of the potential and hence cannot determine the end point. Nonlinear interactions, represented in this specific case by the $\varphi^4$ term in the potential (2), eventually quench the instability and drive the field to a different, stable configuration.

2. Tachyonic instability in curved spacetime

In Sec. II.A.1 we considered a scalar field that exhibited a tachyonic instability in flat spacetime. The generalization to curved spacetime is simple. If we promote the Minkowski metric to some general curved background described by a metric $g_{\mu\nu}$, Eq. (4) becomes

$$\Box \delta \varphi - \mu^2 \delta \varphi = 0,$$

where $\Box = g^{\mu\nu} \nabla_{\mu} \nabla_{\nu}$, with $\nabla_{\mu}$ the covariant derivative. The key difference here is that in curved space $\mu^2 < 0$ is no longer sufficient for having a tachyonic instability.

To see this, we take $g_{\mu\nu}$ to be the Schwarzschild metric, which can describe either a nonrotating BH or the exterior spacetime of a nonrotating NS in GR. The line element is

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2,$$

where $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$ and $M$ is the mass of the compact object. Because the spacetime is static and spherically symmetric, we can decompose the scalar perturbation $\delta \varphi$ into spherical harmonics $Y_{\ell m}(\theta, \phi)$ and assume a harmonic time dependence

$$\delta \varphi = \sum_{\ell m} \frac{\psi_{\ell m}(r)}{r} Y_{\ell m}(\theta, \phi) e^{-i\omega t},$$

and by substitution into Eq. (5) we obtain a Schrödinger-like equation

$$\frac{d^2 \psi_{\ell m}}{dr_s^2} + \left[\omega^2 - V_{\text{eff}}(r)\right] \psi_{\ell m} = 0,$$

where we introduced the tortoise coordinate $r_s$ defined as $dr/dr_s = 1 - 2M/r$ and $V_{\text{eff}}$ is an effective potential given by

$$V_{\text{eff}} = \left(1 - \frac{2M}{r}\right) \left[\frac{\ell(\ell + 1)}{r^2} + \frac{2M}{r^3} + \mu^2\right].$$
which encodes information about the background, curved spacetime. To have an instability in a Schwarzschild BH spacetime it is sufficient but not necessary that
\[ \int_{-\infty}^{\infty} dr_* V_{\text{eff}}(r) \leq 0, \tag{10} \]
where \( r_* = -\infty \) corresponds to the horizon radius in tortoise coordinates. For \( M = 0 \) (flat spacetime) this condition is always satisfied when \( \mu^2 < 0 \) and Eq. (8) yields the same dispersion relation that we discussed in Sec. II.A. The situation is different for \( M \neq 0 \). Although it is not immediately obvious from upon inspection of Eq. (9), it turns out that \( \mu^2 \) would have to be sufficiently negative for the tachyonic instability to occur. The main lesson here is that, in curved spacetime, the threshold for the tachyonic instability to happen depends on the spacetime; we will return to how one can determine it later. Note that, although we previously used Schwarzschild spacetime as an example, one can rederive Eq. (8) for a general static, spherically symmetric background provided that \( r_* \) is chosen appropriately.

3. Scalarization and gravity

We have thus far considered a scalar field with a negative bare mass squared, which is not well motivated. However, fields can acquire an effective mass squared \( \mu^2_{\text{eff}} \) in specific situations due to their coupling to other fields. As an example, consider a scalar field that is nonminimally coupled to gravity and a term \( \varphi^2 R \) is present in the action, where \( R \) is the Ricci scalar. We then expect a contribution proportional to \( \varphi R \) to the scalar’s field equation; hence, the Ricci scalar contributes to the field’s effective mass, that is, \( \mu^2_{\text{eff}} \propto \mu^2 + R \). Assume that the scalar field has no bare mass (i.e., \( \mu = 0 \)) and that the coupling to \( R \) is the only contributor to its effective mass \( \mu^2_{\text{eff}} \). In flat spacetime, scalar-field perturbations would then be massless, whereas in curved spacetimes (with \( R \neq 0 \)) they would be massive and in general also position dependent. Moreover, the sign of \( \mu^2_{\text{eff}} \) would be controlled by the sign of \( R \) in this case. Hence, in some situations it would be possible for \( \mu^2_{\text{eff}} \) to become sufficiently negative in some spacetime region and cause the scalar field to become tachyonically unstable despite this being impossible in flat spacetime.

Just as in the flat-spacetime example of the scalar with negative \( \mu^2 \) and quartic interactions in Sec. II.A.1, this tachyonic instability does not have to be catastrophic. It can simply signal that the scalar needs to transition to a different configuration once curvature exceeds some threshold. The instability implies that the scalar field will grow, nonlinearities will become important, and, if they can quench the instability, then one can end up with a stable, different configuration, for both the scalar field and the spacetime.

This is precisely the idea behind spontaneous scalarization, first proposed in Damour and Esposito-Farèse (1993). In a given generalized scalar-tensor theory, a configuration with a constant scalar field and a metric that solves Einstein’s equations describes all gravitating systems except some that exhibit strong gravity. In the latter case, curvature becomes significant enough to render the constant scalar configurations tachyonically unstable. The tachyonic instability is eventually quenched by nonlinear effects, and there is a stable configuration with a nontrivial scalar and a spacetime that is no longer a solution to Einstein’s equations.

4. Strong-field phase transitions and weak-field screening

We have not yet shown that the mechanism of spontaneous scalarization, as described heuristically earlier, can be at play within a consistent gravity theory: we do so in Sec. II.B. Nonetheless, assuming that the proposal can be successfully implemented in some model, the following key observations can already be made:

- Scalarization is a sharp transition to a new configuration that can differ significantly from the GR configuration for the same object, even when one is very near the threshold of the tachyonic instability. This is intuitive when one thinks of scalarization as a linear instability quenched by nonlinearity; even for a mild instability (large timescale) the scalar has to grow significantly for nonlinear effects to become important and manage to stop further growth.

- The onset of instability can be controlled by curvature couplings. In Sec. II.A.3, we considered as an example a coupling between the scalar field and the Ricci scalar \( R \), but one can envisage couplings with other curvature invariants, as we see in in Sec. II.B. Hence, there can be models in which scalarization will occur only in the strong-field regime (where curvature can become large), while objects that exhibit weak gravity will show no deviation from GR (because the curvature is small). Combined with the previous point, this suggests that spontaneous scalarization can be thought of as a strong-field phase transition, whereby a field that is dormant in the weak field transitions to a nontrivial configuration in the strong field. Alternatively, one can think of scalarization in the reverse way: as a screening mechanism that forces a scalar field to transition to a trivial configuration in the weak field and hence explain why this field has managed to remain undetected so far.

1 To our knowledge, the expression “spontaneous scalarization” was first used in print by Damour and Esposito-Farèse (1996a).
The previous argument is based on the rather naive expectation that curvature invariants are a good measure of how strong the gravitational interaction is. As an example of the failure of this expectation, recall that for a Schwarzschild BH the Ricci scalar is zero; however, other curvature scalars, such as the Kretschmann scalar $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$, are nonzero. In general, curvature invariants can have a complicated dependence on the properties of compact objects. Hence, more work is needed to understand what controls the threshold of the tachyonic instability and the onset of scalarization. This is addressed in Sec. II.C.

B. Models of scalarization

1. Tachyonic instability and the minimal action

We now turn our attention to the gravity theories that can exhibit spontaneous scalarization. As discussed, at the perturbative level the hallmark of spontaneous scalarization is a tachyonic instability. This begs the following question: Can we construct a minimal gravity theory which can have scalar-field perturbations that are tachyonomically unstable? To do so, consider a gravity theory with a metric $g_{\mu\nu}$ and a scalar field $\phi$. Assume that the theory is such that the following hold:

A.1 Spacetimes that are solutions of Einstein’s equations, potentially with a cosmological constant, and a constant scalar field are admissible solutions of this theory as well.

A.2 All terms in the action are at least quadratic in $\phi$.

A.3 The field equations are second-order partial differential equations.

Under these requirements, the equation governing the dynamics of scalar perturbations $\delta \phi$ on GR spacetimes can be cast in the form

$$g_{\text{eff}}^{\mu\nu}\nabla_\mu^{(0)}\nabla_\nu^{(0)}\delta \phi - \mu_{\text{eff}}^2 \delta \phi + \text{NLC} = 0,$$

where $g_{\text{eff}}^{\mu\nu}$, $\nabla^{(0)}$ and $\mu_{\text{eff}}^2$ are all computed in the background spacetime $g_0^{\mu\nu}$, and NLC denotes nonlinear corrections. In Eq. (11) $g_{\text{eff}}^{\mu\nu}$ is an effective metric that can differ from $g_0^{\mu\nu}$ for certain types of nonminimal couplings between the metric and the scalar field and $\mu_{\text{eff}}^2$ may contain not only a bare mass term but also other contributions.

If one neglects nonlinearities and assumes that $g_{\text{eff}}$ is nondegenerate and has a Lorentzian signature, Eq. (11) becomes a curved-spacetime version of Eq. (4). This means that one can identify all theories with a single scalar field that are expected to lead to spontaneous scalarization by considering which couplings between a scalar and the metric can contribute to $g_{\text{eff}}$ and $\mu_{\text{eff}}^2$ while still satisfying the aforementioned assumptions A.1–A.3.

The benefit of taking into account all possible such terms is that it would allow one to fully explore the onset of scalarization and identify a class of gravity theories that result in a scalarized spacetime.

Assumption A.2 appears to be essential to avoid having a source term in Eq. (11). We return to this shortly and show that it is a redundant assumption. But we first consider assumption A.3. It ensures that there are no unwanted degrees of freedom, as would generically be the case if the equations contained higher-order derivatives (exceptions can exist, most notably in cases where field redefinitions can reduce the differential order of the equations). This assumption does limit the possibilities of the terms that one can consider. For example, a coupling term of the type $\phi^2 R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma}$ in the action would contribute to $\mu_{\text{eff}}^2$ but leads to higher-order field equations (in the absence of suitable counterterms). To deal with this potential pitfall, one can follow the lines of Andreou et al. (2019) and start with the Horndeski action (Deffayet et al., 2009; Horndeski, 1974), also known as generalized scalar-tensor theory,

$$S = \frac{1}{16\pi G} \sum_{i=2}^{5} \int d^4x \sqrt{-\tilde{g}} \mathcal{L}_i + S_m[\Psi_m; g_{\mu\nu}]$$

and

$$\mathcal{L}_2 = G_2(\phi, X),$$

$$\mathcal{L}_3 = -G_3(\phi, X) \Box \phi,$$

$$\mathcal{L}_4 = G_4(\phi, X) R + G_{4X} \left[ (\Box \phi)^2 - (\phi_{\mu\nu})^2 \right],$$

$$\mathcal{L}_5 = G_5(\phi, X) G^{\mu\nu} \phi_{\mu\nu} - \frac{1}{6} G_{5X} \left[ (\Box \phi)^3 - 3 \Box (\phi_{\mu\nu})^2 \right.$$

$$+ 2 (\phi_{\mu\nu})^3 \right],$$

where $G_2$, $G_3$, $G_4$, and $G_5$ are arbitrary functions of the scalar field $\phi$ and its kinetic term $X = -\Box \phi / \phi / 2$. In Eq. (13) $G_{iX} = \partial G_i / \partial X$ ($i = 4$ and 5). $G^{\mu\nu}$ is the Einstein tensor, and the notation $\varphi_{\mu\nu} = \nabla_\mu \nabla_\nu \phi$ was introduced so, for example, $(\phi_{\mu\nu})^2 = \phi_{\mu\nu} \phi^{\mu\nu} = \nabla_\mu \nabla_\nu \phi \nabla^\mu \nabla^\nu \phi$. Finally, $S_m$ is the matter action, with matter fields collectively denoted by $\Psi_m$. This is the most general action for a metric and a scalar field that leads to second-order field equations in four dimensions upon direct variation; see Kobayashi (2019) for a review. We assume for the moment that matter couples minimally to the metric only. This means that the choice of fields $g_{\mu\nu}$ and $\phi$ correspond to the so-called Jordan frame; we return to this issue later.

Imposing assumptions A.1 and A.2 on the action (12) places restrictions on the $G_i$ functions, as we later see.

---

2 See also Motohashi and Minamitsuji (2018) for a classification of a broader class of scalar-tensor theories according to their BH solutions, including those of GR.
We refer to Andreou et al. (2019) for a detailed discussion. For our purposes, it is sufficient to say that, by perturbing around an arbitrary spacetime that is assumed to be a solution of Einstein’s equations with a constant scalar field, we can identify all of the terms that contribute, at the linear level, to $g_{\mu\nu}^{\text{eff}}$ and $\mu_{\text{eff}}^2$, as defined in Eq. (11). These terms amount to the following action:

$$S_{\text{min}} = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - \frac{1}{2} (\gamma_1 + \gamma_2 R) \nabla_\mu \varphi \nabla^\mu \varphi + \gamma_2 R_{\mu\nu} \nabla^\mu \varphi \nabla^\nu \varphi - \frac{1}{4} \mu_{\text{eff}}^2 \varphi^2 - \frac{1}{4} \beta \varphi^2 R + \frac{1}{2} \alpha \varphi^2 \eta \right] - 2\Lambda \right] + S_m[\Psi_m; g_{\mu\nu}],$$

(14)

where $\mathcal{G}$ is the Gauss-Bonnet invariant, which is defined in terms of the Riemann tensor and its familiar contractions as

$$\mathcal{G} = R^{\mu\nu\rho\sigma} R_{\mu\nu\rho\sigma} - 4 R^{\mu\nu} R_{\mu\nu} + R^2,$$

(15)

and where $\alpha$, $\beta$, $\gamma_i$, and $\mu_{\text{eff}}^2$ can be expressed in terms of the $G_i$ functions and their derivatives evaluated in the background configuration (Andreou et al., 2019).\textsuperscript{3} We refer to this action, in a slight abuse of terminology, as scalarization, and to determine the relevant instability onset of contains all of the terms that contribute to the action.\textsuperscript{4}

Before we go further, we examine what happens if we decide to drop assumption A.2 altogether, but still impose assumptions A.1 and A.3. Working along the same lines as before, one arrives at a different set of terms composing the action

$$S''_{\text{min}} = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - \frac{1}{2} \gamma_1 R \nabla_\mu \varphi \nabla^\mu \varphi + \gamma_2 R_{\mu\nu} \nabla^\mu \varphi \nabla^\nu \varphi + \frac{1}{2} \mu_{\text{eff}}^2 \varphi^2 + \tau \varphi + \eta \varphi R + \lambda \varphi \eta \right] - 2\Lambda \right] + S_m[\Psi_m; g_{\mu\nu}],$$

(16)

As before, $\tau$, $\eta$, $\lambda$, and $\gamma'_i$ can be expressed in terms of the $G_i$ functions and their derivatives evaluated in the background configuration. It might seem counterintuitive that an action containing terms linear in $\varphi$ leads to perturbation equations that are in the form of Eq. (11), which has no source terms. This is due to the presence of terms that are nonanalytic in $\varphi$.

The first impression may be that abandoning assumption A.2 has given rise to a second minimal action for scalarization. However, action (16) is just a field redefinition away from action (14). Indeed, one can start with Eq. (16), introduce the redefinition $\varphi \rightarrow \varphi^2$, and obtain action (14), with the following mapping of parameters: $\gamma_1 = 4 \gamma'_1$, $\mu_{\text{eff}}^2 = -4 \tau$, $\beta = -\eta$, and $\alpha = \lambda$ (Andreou et al., 2019). This equivalence demonstrates (i) that assumption A.2 is redundant and (ii) that up to field redefinitions Eq. (14) is sufficient to capture all terms that contribute to the onset of scalarization and satisfy assumptions A.1 and A.3.

One can see by inspection that the $\gamma_i$ terms in Eq. (14) will contribute to $g_{\mu\nu}^{\text{eff}}$, while the rest of the terms will contribute to $\mu_{\text{eff}}^2$. Hence, if the latter vanish, the former cannot trigger scalarization by themselves, as the effective mass would vanish. Nevertheless, the $\gamma_i$ terms will affect the threshold of the tachyonic instability we associate with scalarization; cf. the discussion about the tachyonic instability in curved spacetime in Sec. II.A.2. Additionally, $\mu_{\text{eff}}^2$ corresponds to the bare mass of the scalar field, so it is expected to be positive. We then conclude that the terms that are expected to trigger scalarization in the strong-field regime are only the couplings of $\varphi$ to $R$ and $\mathcal{G}$. In fact, we see shortly that these are indeed the terms present in the known models of scalarization.

To summarize, the minimal action (14) can be used to study the onset of spontaneous scalarization triggered by a nonminimal coupling to gravity. It contains all possible terms that contribute to the associated tachyonic instability at the linearized level, so it could be used to study the threshold and onset of this instability in full generality. As previously discussed, as the instability progresses it is expected to be quenched nonlinearly, and the end point will be a scalarized configuration. The terms in Eq. (14) can contribute nonlinearly as well, but one could add a plethora of other nonlinear interactions, ranging from scalar self-interactions, for example, $\varphi^4$, to nonminimal coupling terms that do not contribute to linear perturbations around curved spacetime with constant scalar, for instance, $\varphi^4 \mathcal{G}$. That is, one can start with Eq. (14), or even a subset of terms therein, and construct different scalarization models. Models that differ only by terms that are not in Eq. (14) will have the same behavior regard to the onset of scalarization, and hence the configurations that one expects to not scalarize, but they can differ in the properties of scalarized solutions (Andreou et al., 2019; Macedo et al., 2019; Minamitsuji and Ikeda, 2019b; Silva et al., 2019). This is further discussed along with specific examples in Secs. IV.A.1 and IV.A.2.

\textsuperscript{3} We are not following the notation of Andreou et al. (2019), but have instead adapted it to match that of some of the specific models that we later study. While expected, it is nontrivial to show how the Gauss-Bonnet invariant emerges from Eq. (12). This was first shown by Kobayashi et al. (2011) at the level of the field equations, and by Langlois et al. (2022) at the level of the action.

\textsuperscript{4} Here and elsewhere in the text, when two actions are related by a field redefinition we do not relabel the field in order to keep the notation lighter.
Before proceeding to discuss more specific models, we return to the issue of the coupling to matter. We have thus far assumed that matter couples minimally to the metric only. This assumption is sufficient to ensure that the theory is compatible with the weak equivalence principle (WEP) (Will, 2018). To satisfy the WEP it is sufficient to have matter couple minimally to some metric, but this does not need to be the same metric (or choice of other fields) for which the theory has second-order field equations. However, it is known that a disformal transformation (Bekenstein, 1993) of the form

$$g_{\mu\nu} \to C(\phi) [g_{\mu\nu} + D(\phi) \nabla_\mu \phi \nabla_\nu \phi] \quad (17)$$

leaves the Horndeski action (12) formally invariant (Bettoni and Liberati, 2013; Zumalacárregui and García-Bellido, 2014). It was shown by Andreou et al. (2019) that coupling matter minimally to a metric that is related to $g_{\mu\nu}$ by such a disformal transformation as done, Minamitsuji and Silva (2016), amounts to a redefinition of $\gamma_2$ in the linearized theory around spacetimes that are solutions of Einstein’s equations. Hence, such a coupling would be redundant when the onset of scalarization is studied using the minimal action (14). One could also entertain the idea of coupling matter to some composite metric $\bar{g}_{\mu\nu}$ that depends on both $g_{\mu\nu}$ and $\varphi$ in a different manner than in Eq. (17). In such a case, it is likely that assumption A.3 would be violated and one would have to start the analysis presented here with a generalization of the action (14).

2. Damour–Esposito-Farèse Model

As mentioned, the concept of scalarization as we described it was first discussed by Damour and Esposito-Farèse (1993). They considered the theory

$$S = \frac{1}{16\pi G_*} \int d^4x \sqrt{-g} \left[ R - 2\nabla_\mu \varphi \nabla^\mu \varphi \right] + S_m[\Psi_m; \mathcal{A}^2(\varphi)g_{\mu\nu}] \quad (18)$$

Equation (18) is said to be written in the Einstein frame, which means that, contrary to our previous assumptions and conventions, the scalar field is coupled minimally to gravity and has a canonical kinetic term. The coupling with matter field $\Psi_m$ is through the function $\mathcal{A}^2(\varphi)$. In Eq. (18) $G_*$ carries a subscript, as it is not generally equal to $G$ used thus far. Variation of the action (18) with respect to $\varphi$ yields the field equation

$$\Box \varphi = -4\pi G_* \alpha(\varphi) T, \quad (19)$$

where

$$\alpha(\varphi) = \frac{d \ln \mathcal{A}(\varphi)}{d \varphi} \quad (20)$$

and $T = g_{\mu\nu} T^{\mu\nu}$ is the trace of the matter energy-momentum tensor in the Einstein frame defined as $T^{\mu\nu} = 2(-g)^{-1/2} \delta S_m / \delta g_{\mu\nu}$. We see that $\alpha(\varphi)$ controls the coupling strength between the scalar field and matter.

If $\alpha(\varphi_0) = 0$ for some constant scalar-field value $\varphi_0$, the constant scalar configuration with $T \neq 0$ will be an admissible solution of the theory. It then follows from the generalized Einstein’s equations

$$R_{\mu\nu} = 2\nabla_\mu \varphi \nabla_\nu \varphi + 8\pi G_*(T_{\mu\nu} - 1/2 g_{\mu\nu} T) \quad (21)$$

that these will be solutions of GR since the first term on the right-hand side vanishes.

At the same time, if we perturb Eq. (19) linearly in $\varphi$ in a fixed background metric that is a solution of GR and compare with Eq. (11) we find that $\beta(\varphi_0) = (\delta \alpha / \delta \varphi)_{\varphi=\varphi_0}$ and $T$ determine the value and sign of the effective mass square of the perturbations, namely,

$$\mu_{\text{eff}}^2 = -4\pi G_* \beta(\varphi_0) T. \quad (22)$$

For stars, one generally has $T < 0$. Hence, for a negative sign of $\beta(\varphi_0)$ and the right magnitude of both quantities, the scalar can develop a tachyonic instability around a spacetime that describes stars in GR, as previously discussed and as studied in detail by Harada (1997).\footnote{Exceptions exist [see Mendes et al. (2014a)], as discussed in Sec. III.A.}

It was shown by Novak (1998a) that this instability is quenched by nonlinearities and that the outcome is a NS with a nontrivial scalar-field configuration. These scalarized NSs were shown by Damour and Esposito-Farèse (1993) to have properties, such as their masses $M$ and radii $R$, that can be dramatically different from their GR counterparts.

In much of the literature considering scalarization, the function $\alpha(\varphi)$ is taken to have the form\footnote{Damour and Esposito-Farèse (1993) also studied the case in which $\mathcal{A} = \cos(\sqrt{\beta} \varphi)$ and hence $\alpha = -\sqrt{\beta} \tan(\sqrt{\beta} \varphi) \approx -6\varphi + 12\varphi^3 + \ldots$, which includes higher powers in the scalar-field-matter interaction series (23) for $\alpha_0 = 0$ and $\beta_0 = -6$.}

$$\alpha = \alpha_0 + \beta_0 \varphi \equiv \alpha_{\text{DEF}}, \quad (23)$$

where $\alpha_0$ and $\beta_0$ are dimensionless constants. Sometimes this choice, rather than the more general action of Eq. (18), is referred to as the DEF model. The constant $\alpha_0$ is then assumed to vanish to allow for constant $\varphi$ solutions [cf. Eq. (19)], or it is assumed to be small. In the latter case, all stars will carry some nontrivial scalar field, but by tuning down $\alpha_0$ any deviation from GR would be undetectable until scalarization kicks in. In its original formulation, the DEF model did not include a bare mass or self-interactions for the scalar, but a potential $V(\varphi)$ can be added to the action and this option has been considered in the literature by Chen et al. (2015); Popchev (2015); and Ramazanoglu and Pretorius (2016), as we see in Sec. III.
Thus far it appears that the DEF model is not covered by our minimal action (14), because of our earlier assumption that the scalar does not couple to the matter. However, by defining $A^2(\phi)g_{\mu\nu}$ as a new metric and rewriting the action (18) in terms of that new metric, the scalar field is no longer coupled to matter. This is referred to as the Jordan frame. It was shown by Andreou et al. (2019) that, at linearized level and after a suitable scalar-field redefinition, the DEF model is equivalent to the action

$$
S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[ \left( 1 - \frac{1}{4} \delta_0 \phi^2 \right) R - \frac{1}{2} \nabla_\mu \phi \nabla^\mu \phi \right] + S_m[\Psi_m; g_{\mu\nu}],
$$

Equation (24) is indeed a particular case of the action (14) in which $\gamma_1 = 1$, $\gamma_2 = \alpha = \Lambda = 0$, and $\beta = \beta_0$. Hence, the DEF model, with what regards to the onset of the tachyonic instability that leads to scalarization, is captured by the minimal action (14) and corresponds to one of the two couplings to curvature that can trigger scalarization.

Before moving on, we mention that the original formulation of the DEF model, which leads directly to Eq. (19), suggests that it is the coupling to matter that controls and triggers scalarization. Indeed, when $T = 0$, as is the case for BHs, Eq. (19) becomes $\Box \phi = 0$ and admits only constant $\phi$ solutions for stationary and asymptotically flat configurations by virtue of a no-hair theorem by Hawking (1972) [this remains true when one includes a potential; see Sotiriou and Faraoni (2012)]. However, our previous analysis and the correspondence between the DEF model and action (24) at the linearized level makes it clear that the DEF model is part of a broader class of theories in which scalarization is present and controlled by the couplings to curvature, rather than matter, and this observation has been crucial for the development of models that exhibit BH scalarization. It is the fact that $R$ and $T$ are related through the trace of the theory’s generalized Einstein equation [cf. Eq. (21)] that allows for both interpretations in the DEF model.

3. Scalar-Gauss-Bonnet gravity

It was first shown by Doneva and Yazadjiev (2018b) and Silva et al. (2018) that theories described by the scalar-Gauss-Bonnet (sGB) action

$$
S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[ R - \frac{1}{2} \nabla_\mu \phi \nabla^\mu \phi + f(\phi) \mathcal{G} \right] + S_m[\Psi_m; g_{\mu\nu}],
$$

can exhibit BH scalarization provided that $(df/d\phi)_{\phi=\phi_0} = 0$ for some constant $\phi_0$. This is an existence condition for constant $\phi$ configurations that are solutions of GR. As proven by Silva et al. (2018), BH solutions of GR are unique solutions to the theory (25) provided that $(d^2f/d\phi^2)_{\phi=\phi_0} \mathcal{G} < 0$. To understand this, we can proceed as follows. By varying the action with respect to $\phi$, we find

$$
\Box \phi + f_\phi(\phi) \mathcal{G} = 0, \quad f_\phi(\phi) = df/d\phi.
$$

Once again, we can consider linear perturbation of $\phi$ on a fixed background and compare with Eq. (11). We find that $(d^2f/d\phi^2)_{\phi=\phi_0} \mathcal{G}$ plays the role of an effective mass square for the scalar perturbations,

$$
\mu_{\text{eff}}^2 = -(d^2f/d\phi^2)_{\phi=\phi_0} \mathcal{G}.
$$

Hence, violating the condition $(d^2f/d\phi^2)_{\phi=\phi_0} \mathcal{G} < 0$ is necessary, but not sufficient, to develop a tachyonic instability that can lead to scalarization.

Doneva and Yazadjiev (2018b) chose $f(\phi)$ to be $f(\phi) = \lambda^2(1 - e^{-3/2\phi^2})/12$, whereas Silva et al. (2018) focused on $f(\phi) = \alpha \phi^2/2$. Note that, in the linearized theory around $\phi = \phi_0$ and provided that the condition $(df/d\phi)_{\phi=\phi_0} = 0$ is satisfied, any choice of $f(\phi)$ is equivalent to $f(\phi) = \alpha \phi^2/2$. Hence, all scalarization models described by action (25) are captured by the minimal action (14) and, in particular, by the coupling between the scalar and the Gauss-Bonnet invariant in regard to the onset of scalarization. However, different choices of $f(\phi)$ will exhibit different behavior in the nonlinear regime and hence scalarized BHs will generally have different properties.

Indeed, it was shown by Blázquez-Salcedo et al. (2018) that the static, spherically symmetric scalarized BHs that were found by Silva et al. (2018) for $f(\phi) = \alpha \phi^2/2$ are unstable against radial perturbations, unlike their counterparts for $f(\phi) = \lambda^2(1 - e^{-3/2\phi^2})/12$ found by Doneva and Yazadjiev (2018b). Silva et al. (2019) later demonstrated by examining the case $f(\phi) = \alpha \phi^2/2 + \xi \phi^4$ that it is indeed the nonlinearity in $\phi$ that controls the stability of scalarized BHs. It was further shown by Macedo et al. (2019) and Minamitsuji and Ikeda (2019b) that a quartic scalar-field self-interaction would be sufficient to make scalarized BHs stable for the $f(\phi) = \alpha \phi^2$ case. These results, which are discussed in Sec. IV, are a clear demonstration that although the onset of scalarization can be described fully using action (14), the end point of the tachyonic instability and the properties of the scalarized configurations will depend to the nonlinear interaction between the scalar and curvature and are thus model dependent.

We remark that models described by action (25) lead to scalarization of compact stars as well for certain regions of their parameter spaces; see Doneva and Yazadjiev (2018a) and Silva et al. (2018).
4. The Ricci-Gauss-Bonnet model

As we have seen, the DEF model and the sGB models of scalarization correspond respectively, to the $\varphi^2 R$ and $\varphi^2 \mathcal{G}$ terms (in addition to $R$ and the canonical kinetic term for the scalar field) in the minimal action (14). We have also argued that these two terms are the only terms that can trigger scalarization as a tachyonic instability around a spacetime that is a solution of GR. These facts together suggest considering the following action (Antoniou et al., 2021a,b):

$$S = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - \frac{1}{2} \nabla_{\mu} \varphi \nabla^{\mu} \varphi - \frac{1}{2} \beta \varphi^2 R + \frac{1}{2} \alpha \varphi^2 \mathcal{G} \right] + S_m[\Psi_m; g_{\mu\nu}].$$

(28)

This theory is interesting from the perspective of an effective field theory (EFT). The previously considered terms can be seen as part of an EFT in which the scalar field enjoys reflection symmetry (i.e., invariance under $\varphi \to -\varphi$), while shift symmetry (i.e., invariance under $\varphi \to \varphi + \text{const.}$) is broken only by the coupling to the curvature scalars. This theory is not a complete EFT, as there are other operators compatible with these symmetries, such as $\varphi^4 R$ and $G^{\mu\nu} \nabla_{\mu} \varphi \nabla_{\nu} \varphi$. Nonetheless, the theory is phenomenologically interesting for various reasons.

To begin, it has GR with a constant scalar field as a late-time cosmic attractor for $\beta > 0$ (Antoniou et al., 2021a). To appreciate why this is important, recall that one can think of scalarization in terms of a tachyonic instability of compact object configurations that are solutions of GR with $\varphi = \varphi_0$. Below the threshold of this instability, these configurations are expected to be stable and exhibit no deviation from GR. The attractive feature of scalarization is that weakly gravitating systems will belong in this category and hence scalarization can be a form of weak-field screening of the scalar field. However, this argument assumes that $\varphi = \varphi_0$ everywhere in the Universe and deviates from this value only due to scalarization. If there were another reason for $\varphi$ to evolve away from $\varphi_0$, this would make even weakly gravitating objects develop a nontrivial scalar configuration. Cosmic evolution can indeed cause such evolution, as shown by Damour and Nordtvedt (1993), to the extent that weakly gravitating systems would become sufficiently scalarized to make the DEF model (Anderson et al. (2016)) and sGB scalarization models (Anson et al. (2019a) and Franchini and Sotiriou (2020)) fail weak-field and cosmological tests of gravity without severely fine-tuning the initial conditions for cosmic evolution.

The theory described by action (28) provides an elegant solution to this problem (Antoniou et al., 2021a). As pointed out by Damour and Nordtvedt (1993), the DEF model has GR as a cosmic attractor for $\beta_0 > 0$, whereas scalarization requires $\beta_0$ to be sufficiently negative. It was already mentioned that, at the linearized level around a GR background, the DEF model is equivalent to action (24), so it is reasonable to expect that the $\varphi^2 R$ term in action (28) would tend to drive the scalar field to a constant in late-time cosmology. Moreover, this term should be dominant over the $\varphi^2 \mathcal{G}$ term at low curvatures, while the latter should dominate at high curvatures and trigger scalarization. It was indeed shown by Antoniou et al. (2021a) that cosmic evolution in the model of action (28) tracks the cosmic evolution of GR from radiation domination onward, and that $\varphi$ is driven to $\varphi_0$ rapidly during matter domination.

It was also shown by Antoniou et al. (2021b) and Ventaggi et al. (2021) that for $\beta > 0$ one can have a range of values for $\alpha$ in which BHs scalarize but NSs do not. This is interesting because the strongest constraints on scalarization thus far, and specifically the DEF model, are based on binary pulsars (Kramer et al., 2021; Zhao et al., 2022). These constraints can be evaded if scalarization is limited to BHs. Antoniou et al. (2021b) also provided strong indication that scalarized BHs should be radially stable for $\beta > 0$, which is not the case for $\beta = 0$. Both of these issues are discussed in Sec. IV.

In summary, adding a $\beta \varphi^2 R$ term with $\beta > 0$ to the simplest sGB scalarization model addresses a series of concerns. This term would be there anyway in an EFT, as it has lower mass dimensions than $\varphi^2 \mathcal{G}$. These considerations should act as a reminder that scalarization theories are currently still toy models in need of a completion.

5. Tensor-multiscalar theories

Thus we have considered theories with a single field, but one could also study scalarization in models with multiple scalar fields. These tensor-multiscalar theories were studied by Damour and Esposito-Farèse (1992) and in their simplest form they are described by the action

$$S = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - 2\gamma_{ab}(\varphi) \nabla_{\mu} \varphi^{a} \nabla^{\mu} \varphi^{b} - 4V(\varphi) \right] + S_m[\Psi_m; A^{2}(\varphi) g_{\mu\nu}],$$

(29)

where $\varphi$ denotes a multiplet of $N$ scalar fields $\{\varphi_1, \ldots, \varphi_N\}$ and $\gamma_{ab}(\varphi)$ is a target-space metric that mixes their kinetic terms. Equation 29 can be seen as a generalization of action (18) to multiple fields. We return to this theory in Sec. III, as it was studied in the context of scalarized NSs by Horbatsch et al. (2015) and Doneva and Yazadjiev (2020a). However, we emphasize that one could consider further generalizations that involve, for example, couplings to the Gauss-Bonnet invariant or further derivative interactions between the scalar fields.
C. Types of scalarization

One can classify types of scalarization based on which property of the compact object triggers scalarization and controls the threshold of the tachyonic instability. As we saw in our discussion of the model (14), it is the couplings between the scalar and curvature invariants that control the onset of scalarization. Therefore, if one thinks of the onset of scalarization as an instability around a spacetime of GR, what controls the onset reduces to how curvature invariants depend on the properties of the object that curves spacetime.

1. Induced by compactness

For static, spherically symmetric BHs, there is a straightforward answer: in GR the exterior is described by the Schwarzschild spacetime, so $R = 0$ and $\mathcal{G} = 48M^2/r^6$, where $M$ is the mass of the BH and $r$ is the areal radial coordinate. Hence, for a given $M$, $\mathcal{G}$ scales monotonically with $r$ and whether the scalar will develop a negative enough effective mass square outside the horizon is controlled by $M$, with lighter BHs being susceptible to the scalarization instability. The opposite trend is present in NS scalarization in the DEF model: the trace of the energy-momentum tensor of the star is what controls the onset, and for most equations of state (EOS) its magnitude shows a steady increase as the radius decreases. The tachyonic instability is stronger at the center of the star, where the density is higher, and this in turn correlates with the total mass of the star, with heavier stars scalarizing.

To unify these two pictures, one can observe that BHs get more compact, in the sense of average mean density, as their mass decreases, whereas for NSs it is the other way around for most EOS. Hence, how compact an object is plays a key role in scalarization and can determine whether the object will be affected by the tachyonic instability that triggers scalarization.

Although we colloquially refer to compactness as one of the triggers of scalarization in the rest of the review, there are some caveats that we ought to mention and on which we further elaborate. First, compactness for BHs is often defined in the literature as the mass over the horizon radius. By that definition, unlike with the more colloquial meaning of the term, Schwarzschild BHs of any mass have the same compactness (but not the same curvature near the horizon). Second, for NSs the relation between compactness and scalarization is highly dependent on the EOS, and there are cases where the tachyonic instability is not strongest at the center of the star. Finally, as we next discuss, other properties of compact objects also affect the value of curvature invariants, and hence can also control scalarization.

2. Induced by spin

Rapidly rotating NSs in the context of scalarization were first studied by Doneva et al. (2013). They showed that for DEF-like models rapid rotation can enhance scalarization by increasing the parameter space where scalarization can occur. Conversely, it is also known that larger spin leads to weaker scalarization for BHs in sGB gravity with $\text{d}^2f/\text{d}q^2 < 0$ (Cunha et al., 2019).

In suitable theories, spin can in itself induce a tachyonic instability that triggers scalarization (Dima et al., 2020), and spinning scalarized BHs in these theories were subsequently constructed explicitly by Herdeiro et al. (2021b) and Berti et al. (2021). The fact that spin can trigger BH scalarization (rather than just controlling the scalar charge) is important because it opens up the possibility that only rapidly rotating BHs carry scalar charge, irrespective of their mass.

These results, as well as further work on scalarized rotating compact objects that is discussed in Secs. III and IV, show that spin, a ubiquitous property of astrophysical objects, can play an important role in the amount of scalar charge that a scalarized object carries.

3. Induced by matter or coupling to other fields

Thus far we have linked scalarization to a tachyonic instability at the perturbative level (although it is fundamentally a nonperturbative effect), and we have focused on models in which that instability is controlled by the nonminimal coupling to gravity. However, as we saw, in the DEF model one can think of $\mu_\text{eff}^2$ as being controlled either by the trace of the stress-energy tensor of matter $T$ [see Eq. (22)] or by the Ricci scalar $R$ [see the action (24)]. The latter interpretation has the advantage of providing a unified framework of scalarization of BHs and stars as linked to a nonminimal coupling to gravity, which is the perspective that we followed. However, the former interpretation highlights that $\mu_\text{eff}^2$ in Eq. (11) could instead be attributed to any type of coupling between $\varphi$ and another matter field.

For example, Stefanov et al. (2008) considered a scalar field coupled to nonlinear electrodynamics, while Herdeiro et al. (2018) focused on Einstein-Maxwell-scalar theory with the addition of the coupling $e^{-\lambda\varphi^2}F_{\mu\nu}F^{\mu\nu}$, where $F_{\mu\nu}$ is the usual Faraday tensor. In both cases it was shown that electrically charged BHs can develop scalar hair through scalarization. Further work in this direction is summarized in Sec. IV.C.

When $\mu_\text{eff}^2$ in Eq. (11) is thought of as being introduced by a coupling to matter, one is also led to consider whether surrounding matter, such as an accretion disk, a companion, or the Galaxy, could scalarize a BH even in models where BHs cannot scalarize in vacuum. It was shown by Cardoso et al. (2013a,b) that this can indeed...
occur in the DEF model.

4. Dynamical scalarization

Thus far we have discussed the scalarization of isolated compact objects, but what happens when they form a binary? As we alluded to, when we embed a NS in an ambient scalar-field environment [which is the case for the nonvanishing $\alpha_0$ in Eq. (23)], stars will always carry some small scalar charge and they can still nonperturbatively develop large charge values when they scalarize. This is called induced scalarization (Salgado et al., 1998), and binaries provide a natural scenario for it to occur. Imagine two NSs, each with its own compactness, such that one is scalarized and the other is not. As the system inspirals, at some point the nonscalarized NS will start experiencing the presence of the scalar field sourced by its companion, and induced scalarization will then take place (Barausse et al., 2013).

Another, perhaps more dramatic, scenario is that of dynamical scalarization. In this case, two nonscalarized NSs can become scalarized once their orbital separation becomes sufficiently small. Qualitatively this can be quantified by some measure of an “effective compactness” of the binary that, only for an isolated NS, can trigger scalarization once it reaches a certain threshold (Palenzuela et al., 2014; Shibata et al., 2014; Taniguchi et al., 2015). In a quasicircular binary this effective compactness only increases (it scales inversely with the orbital separation), but that does not have to be the case in an eccentric orbit. In such cases, the effective compactness oscillates in time, being largest when the NSs are closest. This leads to a transient dynamical scalarization of the system, whereby the two NSs continuously scalarize and descalarize as the system inspirals.

What about BHs? In sGB models the scalar field is sourced by the Gauss-Bonnet invariant, and therefore a binary composed of two scalarized BHs would in general result in an unscalarized BH remnant since the latter has a larger mass and therefore a smaller spacetime curvature. That is, the system descalsarizes (Silva et al., 2021b). However, depending on the initial nonscalarized BHs’ spins and masses, one can have cases where the remnant scalarizes due to its large spin, i.e., there can be dynamical spin-induced scalarization (Elley et al., 2022).

As we have seen, compact binaries lead to new manifestations of scalarization. We further discuss these in Secs. III and IV.

5. Beyond scalarization

Thus far we have discussed scalarizations as a linear, tachyonic instability for a scalar field that is then quenched nonlinearly and leads to a nontrivial scalar configuration. There are many ways to extend this paradigm and yet keep the key outcome: to have fields that undergo what resembles a phase transition — a sharp change from a trivial to a nontrivial configuration) in the strong-field regime.

One direction is to generalize the mechanism to different fields, such as vectors, tensors, and spinors, generating models of spontaneous vectorization, tensorization or spinorization (Ramazanoğlu, 2017, 2018a,b,c). Another approach would be to construct models in which the transition is triggered not by a tachyonic instability but by some other linear instability (Ramazanoğlu, 2018a). Both of these directions are discussed in Sec. V. A third direction comes from the possibility that the transition might not be triggered by a linear instability, but might instead be a fully nonlinear effect (Doneva and Yazadjiev, 2022). It was shown that if one chooses $\phi = \exp(\beta \phi^4)$ in the action (25), one obtains a theory in which scalar perturbations are massless around Kerr BHs, and hence there cannot be any linear tachyonic instability. Yet, stable scalarized solutions still exist for certain masses and spins.

D. Quantum aspects and classical analogs

While most of this review deals with scalarization from a classical field theory perspective, it is noteworthy to comment on the quantum aspects of this phenomenon. In particular, Lima et al. (2010) studied a quantum scalar field nonminimally coupled to gravity living in the classical background of a compact star spacetime (i.e., they worked in the semiclassical approximation). They showed that some field modes can go through an exponential growth causing the vacuum expectation value of the field operator $\hat{\phi}^2$ to grow, ultimately causing the vacuum expectation value of the energy-momentum tensor $T_{\mu\nu}$ of the field itself to grow in the same manner; see also Lima and Vanzella (2010). In this sense, one can say that classical curved spacetimes can “awake the vacuum state” of quantum fields. But what is the connection with spontaneous scalarization? The model studied by Lima et al. (2010) is simply Eq. (24) for a massive scalar field: the Jordan-frame equivalent of the DEF model for a massive scalar. In the parameter space spanned by the scalar-field–Ricci-scalar coupling constant and stellar compactness relevant to our discussion, the regions where the instability occurs agree with the classical prediction to where scalarization should happen, as shown by Pani et al. (2011b). Thus, one can think of the classical “scalar-field perturbations,” which we have so often spoken about to be seeded by quantum field fluctuations (Landulfo et al., 2015). We refer the interested reader to Landulfo et al. (2012); Lima et al. (2013); Mendes et al. (2014a,b); and Santiago et al. (2016) for other works on the semiclassical approach.

Finally, while our review focuses on astrophysical implications of scalarization, we remark that the realization
of this effect in condensed matter systems has also been studied. More specifically Ribeiro and Vanzella (2020) devised a classical analog that exhibits this phenomenon based on the nonlinear optics of metamaterials, and this could in principle be observed experimentally.

III. NEUTRON STAR SCALARIZATION

Scalarization of compact objects was first considered in the context of NSs by Damour and Esposito-Farèse (1993). As discussed in Sec. II, in this case the nonzero trace of the energy-momentum tensor of the nuclear matter acts as a source of the scalar field and evades the no-hair theorems existing for vacuum BHs in certain scalar-tensor theories. Since then, the scalarized NSs in the DEF model have attracted significant attention, and they are perhaps the most studied compact objects beyond GR. Over the years, NS scalarization has also been examined in other scalar-tensor theories beyond the DEF model.

In this section, we present the developments in the field over the past three decades. In Sec. III.A, we start with the DEF model and its generalizations, placing a special emphasis on the constraints coming from binary-pulsar observations. In Sec. III.B, we proceed to the dynamics of such compact objects, both isolated ones and those in binaries, discussing their stability in addition to GW emission. In Sec. III.C, we discuss the various astrophysical implications of scalarized NSs. Finally, in Sec. III.D we review NSs scalarization beyond the DEF model.

A. Equilibrium neutron stars in the Damour–Esposito-Farèse model

In this section, we review the equilibrium properties of NSs in the DEF theory, both in its original form and in extensions of the theory. The strongest constraints on the former come from binary-pulsar constraints, which we also discuss here.

1. The original Damour–Esposito-Farèse model and binary-pulsar constraints

a. Static neutron stars: As we saw in Sec. II.B.2, the effective mass squared of scalar-field perturbations in the DEF model is given by Eq. (22), namely,

$$\mu^2_{\text{eff}} = -4\pi G \beta(\phi_0) T,$$

where we recall that $\beta(\phi_0)$ is the derivative of the scalar-matter coupling $\alpha(\phi)$ evaluated at a constant background scalar-field value $\phi_0$. Thus, the condition for scalarization $\mu^2_{\text{eff}} < 0$ can be satisfied when the trace of the energy-momentum tensor $T$ and $\beta(\phi_0)$ have the same sign. For realistic NSs, which are modeled as a perfect fluid with pressure $p$ and energy density $\varepsilon$, we normally have

$$T = 3p - \varepsilon < 0,$$

and hence scalarization can happen when $\beta(\phi_0) < 0$. In particular, for the coupling function (23) studied by Damour and Esposito-Farèse (1993), we have $\beta(\phi_0) = \beta_0$ being a constant. Thus, most studies focus on the case where both $T$ and $\beta_0$ are negative, but we remark that at sufficiently high densities some EOSs predict a positive sign of $T$, resulting in scalarization when $\beta_0 > 0$; see e.g., Mendes (2015) and Podkowka et al. (2018). For most NS models, $p$ and $\varepsilon$ are related by a barotropic EOS $\varepsilon = \varepsilon(p)$.

In Fig. 1 we show some properties of nonrotating scalarized NSs for $\alpha_0 = 0$ and some illustrative values of $\beta_0 < 0$. Here in Sec. III.A.1, we consider the case of a vanishing scalar-field potential as in the original DEF model. The influence of a nonzero potential is further discussed later. The EOS is taken to be that of Akmal, Pandharipande, and Ravenhall (APR) (Akmal et al., 1998). In the left panel of Fig. 1, we show the Einstein-frame Arnol’d-Deser-Misner (ADM) mass of a sequence of NS solutions, parametrized by the energy density at the center of the star $\varepsilon_c$.

When $\alpha_0 = 0$, NS solutions of GR are also solutions of the DEF model; they are indicated by solid black lines. These solutions are characterized by having a zero scalar field. We see that when a specific critical central energy-density value is reached (say, $\varepsilon_{c,1}$), the GR sequence becomes unstable and a new branch of stable solutions with a nontrivial scalar field (i.e., scalarized stars) bifurcates from it. In our example, the value of $\varepsilon_{c,1}$ depends only on $\beta_0$ and on the EOS. The scalarized branch merges again with the GR branch at a second bifurcation point at a larger energy density $\varepsilon_{c,2}$. Hence, scalarized NSs exist only in the range $\varepsilon_c \in [\varepsilon_{c,1}, \varepsilon_{c,2}]$. We see that the larger $|\beta_0|$ becomes, the more dramatic the deviations in the NS mass relative to GR are. In addition, the range of $\varepsilon_c$ in which scalarization can happen increases. This is also shown in the middle panel of Fig. 1, where we plot the ADM mass as a function of the radius $R$, and in the right panel of Fig. 1, where we show the scalar charge $D$ as a function of the ADM mass. The scalar charge is defined in terms of an expansion at spatial infinity of the scalar field $\varphi$, namely,

$$\varphi = \varphi_0 + D/r + O\left(r^{-2}\right),$$

where $\varphi_0$ is the cosmological background value of $\varphi$, often assumed to be zero for simplicity. The right panel of Fig. 1 shows how $D$ has a small magnitude near the bifurcation point where scalarization kicks in, grows monotonically with $M$, and then approaches zero again once $M$ approaches the mass corresponding to the second bifurcation point at $\varepsilon_{c,2}$. 
For which values of $\beta_0$ does scalarization occur? Recall that the effective potential $V_{\text{eff}}$ needs to be sufficiently negative in order to support at least one tachyonic mode; cf. Eqs. (9) and (10). For typical NS densities, Eq. (22) implies that scalarization exists for $\beta_0 \lesssim -4$. Damour and Esposito-Farèse (1993) made this estimate using a Newtonian approximation and confirmed it by integrating the fully relativistic equations of stellar equilibrium. Subsequent works refined the threshold for scalarization to $\beta_0 \lesssim -4.35$ and also showed that this bound is not very sensitive to the EOS (Altaha Motahar et al., 2017; Novak, 1998b; Silva et al., 2015). We note that, similar to the GR branch of NSs, scalarized solutions in the original DEF model are stable up to the maximum mass of the corresponding branch, and the stable solutions are generally energetically favorable over the GR NSs. This is further discussed in Sec. III.B.1.

Let us briefly comment on the exact definition of mass in scalar-tensor theories (shown in Fig. 1 and elsewhere). In contrast to GR, the definition of mass in scalar-tensor theories is subtle, due to the fact that these theories violate the strong equivalence principle. This results in the appearance of different possible masses as a measure of the total energy of the star (Lee, 1974; Scheel et al., 1995a,b; Whinnett, 1999; Yazadjiev, 1999). These works showed that only the so-called tensor mass has natural energylike properties. For example, the tensor mass is positive definite, it decreases monotonically by the emission of GWs and it is well defined even in dynamical spacetimes (Lee, 1974; Scheel et al., 1995a,b). In addition, only the tensor mass leads to a physically acceptable picture since it peaks at the same point as the particle number, a property crucial for the stability of the static stars (Whinnett, 1999; Yazadjiev, 1999). Therefore, the tensor mass, which is defined as the ADM mass in the Einstein frame, should be taken as the physical mass. As a matter of fact, though, for most of the commonly used coupling functions the Jordan-frame and the Einstein-frame NS masses are identical.

After the discovery of the phenomenon, scalarization was examined for a larger set of parameters and in more detail by Damour and Esposito-Farèse (1996a). They showed that the presence of some externally imposed scalar-field background $\varphi_0$, as well as considering $\alpha_0 \neq 0$, smoothens the transition to a scalarized state. This is what we described as induced scalarization in Sec. II.C.4. (Strictly speaking, we do not have pure scalarization when $\alpha_0 \neq 0$, since GR is no longer a solution of the field equations.) Spontaneous scalarization was further studied by Salgado et al. (1998), who considered the problem in the Jordan frame and performed an approximate Newtonian analysis of the system. They showed that scalarization can also be associated with the fact that the effective gravitational constant in scalar-tensor theories decreases for large scalar fields. It was further argued by Whinnett and Torres (2004) that scalarization leads to violation of the weak energy condition in the inner regions of NSs, which can cause instabilities. It was later demonstrated by Salgado et al. (2004) that this is not a general feature of scalar-tensor theories and that there are subclasses of the theory where the weak energy condition is easily satisfied.

Another consequence of scalarization is that for sufficiently large $\beta_0$ the maximum allowed mass for NSs increases compared to GR, which may have various observa-

---

7 For an exception in the case of a massive scalar field, see Sec. III.A.2.

8 This happens not only in scalar-tensor theories but also in other modified gravity theories. Upper bounds on the maximum mass of NSs in GR can be found under minimal assumptions on the EOS using the approach of Rhoades and Ruffini (1974). See Hartle (1978) for an early account of applications of this method to other gravity theories.
tional consequences. This problem was studied by Sotani and Kokkotas (2017), who took the effects of various microphysics parameters into account. Empirical relations were derived for the maximum mass of scalarized NSs that are parametrized with respect to the nuclear saturation parameters and the maximum sound velocity in the core.

Until now we have discussed works that considered negative $\beta_0$. However, as we mentioned, when the trace of the energy-momentum tensor is positive, scalarization can occur for $\beta_0 > 0$ as well, and this scenario introduces qualitative differences relative to our story thus far (Mendes, 2015). In particular, for the commonly used coupling function $\alpha(\varphi) = \beta_0 \varphi$ of the DEF theory, scalarized stars are not stable for high values of $\beta_0$ ($\beta_0 \gg 1$), which is further discussed in Sec. III.B.2.

The calculation of NS parameters spanning the DEF theory parameter space provides a challenging technical task. In particular, tests of this theory against binary-pulsar observations (described in Sec. III.A.1.b) require knowledge of the scalar charges (and their derivatives with respect to the star’s mass) for a large catalog of EOSs. This calculation was performed most extensively in the works by Anderson and Yunes (2019) and Guo et al. (2021), who provide the results in tabulated form or through surrogate models. Yagi and Stepnicka (2021) [see also Horbatsch and Burgess (2011)] computed scalar charges in the DEF model analytically using a combination of perturbative weak-field expansion and Padé resummation and found excellent agreement with numerical calculations.

Spontaneous scalarization for the case of several nearby compact objects was considered by Cardoso et al. (2020). Their analytical analysis showed that, even though an isolated body might be below the threshold for scalarization, a collection of such bodies could develop a nonzero scalar field while maintaining average compactness much below the scalarization limit.

We have already discussed spontaneous scalarization of NSs, but other compact objects can also scalarize. The case of BHs is considered in Sec. IV, but as an additional nonvacuum example we now discuss the case of boson stars [cf. Liebling and Palenzuela (2023) for a review], which were also shown to scalarize (Whinnett, 2000). In such systems, one has a complex scalar field as a matter source for the boson star and an additional real scalar field responsible for scalarization similar to the DEF model. The dynamics of this process was examined by Alcubierre et al. (2010) who showed that nonlinear development of the scalar field is observed in the absence of self-interactions in the complex scalar field. Ruiz et al. (2012) studied spontaneous and induced scalarization starting with initial data corresponding to stable boson stars in GR. They showed that a strong emission of scalar radiation occurs during the scalarization process.

b. Observational constraints from binary pulsars: To date binary pulsars have set the best constraints on scalarized NSs in the DEF models (Kramer et al., 2021; Zhao et al., 2022). This is because determining different characteristics of these systems through pulsar timing can be made extremely precise by accumulating yearslong observations. One of such important observables is the rate at which the binary’s orbit decreases by energy loss through GW emission. In contrast to GR, the DEF model has an additional scalar degree of freedom that leads to a new channel of energy loss. Thus, the shrinking of the orbit should happen faster. The energy flux of the scalar-dipole radiation that gives the dominant contribution is given by (Damour and Esposito-Farèse, 1992),

$$F_{\text{dipole}}^{\text{scalar}} = A (1 + q_1 q_2)^3 (q_2 - q_1)^2,$$

where $A$ is a function depending on the properties of the binary, such as its total and reduced masses, and its eccentricity; $q_2$ and $q_1$ are the scalar charges of each binary component normalized with respect to their masses $q_i = D_i / M_i$ ($i = 1$ and 2).

Binary-pulsar observations were considered in the context of scalarized NSs for the first time by Damour and Esposito-Farèse (1996a). They calculated the gravitational form factors (also known as sensitivities) of slowly rotating NSs, which form the set of coupling constants appearing in the post-Keplerian description of the binary in scalar-tensor theory; Horbatsch and Burgess (2012) for a summary. Only a few such binary systems were known at the time, and $\beta_0$ was constrained to be greater than $-5$ using polytropic EOSs. These results were later refined to include realistic EOSs and a limit of $\beta_0 > -4.5$ was derived by Damour and Esposito-Farèse (1998). In addition, an estimate was made that it would be difficult for LIGO and VIRGO to improve $\beta_0$ bounds for most EOS possibilities [some exceptions were pointed out by Sampson et al. (2014) and Shao et al. (2017)]. The reason is that even though the merger events observed using such GW detectors can lead to much stronger scalar-dipole radiation, they are inferior in accuracy compared to the radio observations of binary pulsars, leading to weaker overall constraints. Next-generation detectors such as the Cosmic Explorer and the Einstein Telescope, though, will be able to improve the bounds on scalar-dipole radiation (Sampson et al., 2014; Shao et al., 2017).

An example of how constraints can be imposed on DEF models is presented in Fig. 2, which illustrates how the model fails the double pulsar test for specific values of the theory parameters. In this case, the failure is due to the additional energy loss from scalar GWs predicted by the DEF model, predominantly the dipolar contribution.

With the advances in observational astronomy, more pulsars in binary systems suitable for constraining the scalar-dipole radiation have been discovered, a complete and up-to-date list can be found in Freire (2022). Consequently, observational bounds on the scalar-tensor gravity
Various measured post-Keplerian parameters are shown as different curves, with the width indicating the measurement uncertainty in each parameter. For the point $\alpha_0 = 5 \times 10^{-4}$ and $\beta_0 = -4$ of the parameter space of the DEF model to be consistent with observations, all curves would have to intersect in a region of the mass-mass plane. This is not the case, as can be seen with the $\dot{\omega}$ (periastron advance) and $\dot{P}_b$ (change of orbital period) curves. For further details and the definitions of the other post-Keplerian parameters see Kramer et al. (2021). From Kramer et al. (2021).

Parameter $\beta_0$ for such systems have been widely discussed in the literature (Antoniadis et al., 2013; Chiba, 2022; Esposito-Farèse, 2004; Freire et al., 2012; Kramer et al., 2021; Shao et al., 2017; Shibata et al., 2014; Voisin et al., 2020; Wex, 2014; Zhao et al., 2022). The strongest current limit comes from Zhao et al. (2022), who practically closed the scalarization window for the original DEF model; i.e., the possibility for scalarization is ruled out in this theory. As we later discuss, though, there are a number of other well-motivated models where scalarization is still possible or cannot be constrained at all by binary-pulsar observations. These include theories with a massive scalar field, tensor-multiscalar theories, or even the standard DEF model when rapid rotation of NSs, which enhances the effect of scalarization, is considered. Furthermore, the theoretical and numerical approaches developed for the study of the DEF model are still applicable to these generalized theories in most situations. Thus, we spend considerable time here on the aspects of the DEF model despite its original form being essentially ruled out.

Constraints on scalarization with $\beta_0 > 0$ using pulsar-timing observations were investigated by Mendes and Ottoni (2019). Owing to the fact that the scalar charge

\[ \beta = 0 \]

is suppressed as $\beta_0$ increases while the range of masses allowing spontaneous scalarization decreases, it turns out that only weak constraints can be imposed by binary-pulsar observations in this part of parameter space.

### c. Rotating scalarized neutron stars

Thus far we have commented only on static NS models. All observed NSs are at least slowly rotating, and some dynamical processes such as NS mergers or stellar core collapse can produce relatively long-lived rapidly rotating supramassive protoneutron stars. Hence, the inclusion of rotation in NS physics is an inseparable part of the goal to explore their astrophysical implications.

Damour and Esposito-Farèse (1996a) were the first to study slowly rotating scalarized NSs to leading order in rotation frequency $O(\Omega)$ using the formalism of Hartle (1967) and Hartle and Thorne (1968), which allowed them to calculate the NS moment of inertia; see also Sotani (2012). It is interesting that in this case there is an exact analytical solution for the NS exterior (Damour and Esposito-Farèse, 1996a). Static and slowly rotating NSs for a wide range of realistic EOSs, including examples with hyperons or quark matter, were considered by Al-taha Motahar et al. (2017). The extension to second order in the rotational frequency $O(\Omega^2)$ was made by Pani and Berti (2014), who used the extension to calculate rotational corrections to the stellar radius and mass, and also its quadrupole moment.

Rapidly uniformly rotating scalarized stars, without approximation, were obtained by Doneva et al. (2013).
They showed that for a fixed \( \beta_0 \) the maximum deviation from GR that is achieved at the mass-shedding limit is considerably larger than in the static case, and the range of central energy densities where scalarization is possible is significantly broadened. This can be seen in Fig. 3, where sequences of static scalarized NSs are compared to NSs rotating at the Kepler limit for the same values of \( \beta_0 \).

There are a number of factors leading to differences from the static case. The first, more intuitive one is that the rotational energy of the star also acts as a source for the scalar field, and thus can change the onset and degree of scalarization. Meanwhile, the rapidly rotating models tend to be less compact, which can reduce the degree of scalarization. The large deviations from GR compared to the static case, conversely, are due mainly to the fact that scalarized stars can sustain much larger angular momentum before reaching the Kepler limit. This is a nonlinear effect that could not be normally caught in the slow-rotation approximation.

A natural consequence of the aforementioned rotation effects is that the minimum \( |\beta_0| \) where scalarization is possible changes compared to the static case. Thus, for the same EOS II (Diaz Alonso and Ibanez Cabanell, 1985) used by Damour and Esposito-Farèse (1996a), scalarization happens for \( \beta_0 < -3.9 \) in rotating stars (Doneva et al., 2013) compared to \( \beta_0 < -4.35 \) in the static case (Damour and Esposito-Farèse, 1996a). Therefore, one can conclude that, although binary-pulsar observations seem to rule out DEF scalarization for static or slowly rotating NSs (Zhao et al., 2022), there is still an observationally viable range of \( \beta_0 \) where only rapidly rotating NSs can scalarize. This is potentially relevant for binary mergers and stellar core collapse, where such rapidly rotating NSs can form.

A caveat in the previous argument is that one does not expect the star to rotate uniformly in these extremely dynamical events. Such differential rotation was first studied by Doneva et al. (2018b), who adopted a simple rotation law that can still capture some of the main properties of the merger remnants, especially a few tens of milliseconds after the merger of the binary (Bauswein and Stergioulas, 2017).

When scalarization is considered, larger values of the maximum mass as well as of the angular momentum can be achieved for supramassive NSs compared to GR. Moreover, the scalar field causes rapidly rotating models to be less quasitoroidal than their general-relativistic counterparts. This can have direct astrophysical implications, especially for binary NS mergers, where the maximum possible mass and angular momentum that a NS can sustain are crucial for determining the merger outcome and the lifetime of the merger remnant (in case a supramassive

---

9 Supramassive NSs do not have a stable static limit but are supported against collapse due to rapid rotation. Hypermassive NSs do not have a stable uniformly rotating limit but are supported against collapse due to differential rotation (Paschalidis and Stergioulas, 2017).
confined to a characteristic radius and its scalar charge is zero. If the orbital separation between the two objects in a binary-pulsar system is much larger that $\lambda_\phi$, the dynamics will not be directly altered by the scalar field and there is no significant emission of scalar-dipole radiation (Alsing et al., 2012). Since $\lambda_\phi$ is controlled by $\mu$ in the simplest case of a scalar-field potential, one can reconcile the DEF model (for arbitrarily small $\beta_0$) with binary pulsar constraints by giving the scalar field a mass $\mu \gtrsim 10^{-16}$ eV.

Scalarized NSs in massive scalar-tensor theories were first studied in the static case (Chen et al., 2015; Popchev, 2015; Ramazanoğlu and Pretorius, 2016) and later extended to slow (Yazadjiev et al., 2016) and rapid rotation (Doneva and Yazadjiev, 2016). The inclusion of a quartic self-interaction term to the potential was considered by Staykov et al. (2018). These works showed that the mass of the scalar field and the self-interaction have similar effects on the scalar field around NSs and that they both suppress scalarization. The quartic interaction by itself cannot affect the range of central energy densities where scalarized solutions exist, because it is a nonlinear contribution to the linearized scalar-field equation of motion [recall Eq. (11)]. In contrast, the mass term shrinks the domain of existence of scalarized NSs and, for large enough masses, no scalarization is possible at all. This is evident in Fig. 4, where NS mass is plotted as a function of its radius for different combinations of $\beta_0$ and $\mu$. The massive scalar-field solutions are confined between the zero scalar-field mass models (the original DEF models) and the GR ones, corresponding loosely speaking to $\mu \rightarrow \infty$. Note that for $\mu \gtrsim 10^{-16}$ eV the solutions are almost indistinguishable from the massless DEF model. For this reason the latter represents an upper limit on the possible deviations from GR in massive scalar-tensor theories.

The exponential asymptotic behavior of the massive scalar field brings computational challenges to the construction of scalarized NSs, which led to new numerical approaches (Rosca-Mead et al., 2020a) and also facilitated the construction of NSs for highly negative $\beta_0$ and large scalar-field masses. Rosca-Mead et al. (2020a) showed that for sufficiently negative $\beta_0$ qualitative changes in the strongly scalarized branch of solutions are possible. For example, the maximum of the scalar field can be located away from the stellar center. In their most extreme form, these solutions are composed of a highly compact NS model realized proper cosmic evolution, and it can also account for the cold dark matter. Chen et al. (2015) and Morisaki and Suyama (2017) focused especially on large scalar-field masses spanning several orders of magnitude and having a Compton wavelength shorter than 10 km, which is the typical size of a neutron star.

3. Incorporating further physics

New aspects of the original DEF model were recently studied with the inclusion of different physical details. For instance, Silva et al. (2015) studied the presence of anisotropic pressure of nuclear matter for both static and slowly rotating NSs. The motivation for this comes from the fact that some theoretical considerations, for instance, with magnetic fields or within the Skyrme model (a low-energy EFT of quantum chromodynamics) (Nelmes and Piette, 2012) suggest that at high densities the NS EOS might have a significant degree of anisotropy (Herrera and Santos, 1997). In such a case, the effects of scalarization increase (decrease) when the tangential pressure is larger (smaller) than the radial pressure. The threshold value of $\beta_0$ for the development of scalarization, which in the isotropic case is $\beta_0 < -4.35$, can be increased due to the presence of anisotropy, thus widening the range of parameters in which scalarization is possible.

Another astrophysically interesting extension of scalarization is to include the magnetic fields. According to observations and modeling, NS magnetic field values can span from $10^8$ to $10^{12}$ G for standard “old” pulsars, ranging from $10^{16}$ G at the surface of some magnetars to, hypothetically, as high as $10^{17} - 10^{18}$ G in the cores of newly formed protoneutron stars. Such strong magnetic fields impact the properties of scalarized NSs, including their magnetic deformability, maximum mass, and range of scalarization, as studied by Soldateschi et al. (2020). They found a magnetically induced spontaneous scalarization whose essence is the following: strong toroidal magnetic fields can support descalarized configurations and, if the star’s magnetic field decreases during some nonideal magnetohydrodynamical process, the star can undergo a rapid growth of the scalar field; i.e., it scalarizes. The magnetic quadrupolar deformations of scalarized NSs and the related GWs produced by rotating magnetars were studied by Soldateschi et al. (2021).

Another interesting extension to the standard DEF model is related to challenging the idea that the fundamental physics remains unchanged in the star’s interior, which is a common assumption when a nuclear matter EOS is constructed. This was studied by Coates et al. (2017), who considered two models in which the mass of the photon had a different value in the interior and the vicinity of a compact star compared to the mass measured by experiments performed in a weak-gravity regime. The first model is based on a Proca-like mass with an effec-
tive mass term dependent on $\phi$. The second model can be thought of as a gravitational Higgs mechanism where the Higgs potential is replaced by the scalar-gravity coupling. In both cases the scalar field undergoes spontaneous scalarization, thus acquiring a nontrivial profile if the compactness passes a certain threshold, providing a mass to the photon by coupling to it in an appropriate manner. Although the focus of Coates et al. (2017) was on the electromagnetic field as a proof of principle, these results can be extended to other fields of the standard model. The signatures of such a gravitational Higgs mechanism on the behavior of magnetic field of NSs in scalar-tensor theories was studied by Krall et al. (2020).

### B. Dynamics of scalarized neutron stars and binary mergers

The dynamics of isolated NSs can be studied by solving the full nonlinear field equations of scalar-tensor gravity, which is often a challenging task. Instead, one usually first approaches the problem by linearizing the field equations around a background solution and then analyzing the resulting linearized dynamics. The study of nonlinear dynamics is then done when necessary and feasible. We follow this sequence in the Secs. III.B.1–III.B.4. We then review what happens when NSs in scalar-tensor theories are placed in binary systems.

1. **Linearized dynamics**

   The studies of linearized dynamics concern the stability of scalarized stars and the analysis of their quasinormal mode (QNMs) spectrum. The latter involves the study of different classes of NS oscillation modes that are tied to the emission of GWs. See Kokkotas and Schmidt (1999) for a review.

   a. **Stability:** We have seen that scalarized NSs coexist with their nonscalarized counterparts as solutions within the DEF model. Which of these branches of solutions is the one realizable in nature? One way of answering this question consists of calculating the fractional binding energy $M_0/M - 1$, where $M_0$ is the star’s baryonic mass. Through this calculation, Damour and Esposito-Farèse (1993) showed that scalarized NSs are usually energetically favored over the GR ones. This is evident in Fig. 5, where we plot the binding energy as a function of the baryonic mass. At constant baryonic mass, the scalarized solutions (solid curve) have larger binding energies relative to the GR solutions (dashed curve) and are thus energetically favorable. In addition, we can see a cusp at the maximum of the mass for both the scalarized and nonscalarized NSs. This suggests a change of stability. Since the branches beyond the maximum of the mass have a lower binding energy (dotted curve in the inset), they are unstable.

   The aforementioned stability analysis relied on the bulk properties of the star. A rigorous complementary approach considers the linear perturbations of the star. The first step in this direction was taken by Harada (1997), who studied scalar-field perturbation in the background of a NS with a zero (or constant) scalar field within the DEF model. Harada (1997) studied the perturbation equations in the frequency domain and showed that the GR solution becomes unstable after a specific critical central energy density. This is the point where the scalarized solutions branch out from the GR ones; see Fig. 1. Harada (1998) reached similar conclusions but worked in the context of catastrophe theory. The radial stability of scalarized NSs was also studied by Mendes and Ortiz (2018). They considered metric and scalar-field perturbations for both signs of $\beta_0$. They found that scalarized NSs are stable against linear perturbations and that instability takes place past the point of maximum mass.

   b. **Gravitational waves from perturbed neutron stars:** Linearized perturbations are also helpful for studying the NS oscillation modes directly related to GW emission. We first note that GWs in scalar-tensor gravity can carry additional polarizations compared to GR. In particular, one can have breathing modes in addition to the standard “plus” and “cross” polarizations of GR (Will, 2018). Moreover, radial perturbations in scalar-tensor gravity can excite GWs, contrary to what happens in GR. These perturbations source monopole scalar waves that result in a nonvanishing contribution to the perturbed Jordan-frame Riemann tensor (linearized around a Minkowski
background) in the transverse-traceless gauge (Damour and Esposito-Farèse, 1992; Novak and Ibanez, 2000). For DEF-like scalar-tensor theories, this requires \( \alpha_0 \neq 0 \). This contribution is then linked to the existence of a breathing polarization mode of the GW. We can then conclude that radially oscillating scalarized NSs in scalar-tensor theories with \( \alpha_0 \neq 0 \) will emit GWs. Their amplitude is connected to \( \alpha_0 \) that controls the generation of tensorial waves from the dynamics of the scalar field. Thus, larger \( \alpha_0 \) leads to stronger coupling and stronger excitation of the breathing modes.

The first study of radial oscillations of NSs in the DEF model was performed by Sotani (2014) in the Cowling approximation (Cowling, 1941; McDermott et al., 1983). In the original GR version of this approximation, the spacetime is held fixed while only the fluid is perturbed. In scalar-tensor gravity, the scalar-field perturbations are also often neglected. Despite its limitations, the Cowling approximation can actually capture well the qualitative features of the neutron star oscillation spectrum. The full problem (i.e., when both the metric and the scalar-field perturbations are taken into account) was addressed by Mendes and Ortiz (2018). They found a new family of modes named scalar modes that have no counterpart in GR. The results show that they have distinct frequencies and damping times compared to the fluid radial oscillation modes. In Fig. 6 we show the frequencies and damping times of the fundamental fluid radial oscillation mode and its first overtone, as well as the fundamental scalar radial mode, as functions of the stellar compactness. We see in Fig. 6(c) that scalarized NSs become unstable at the maximum mass, and it is the scalar mode that is responsible for the instability. This contrasts with GR, where it is the fundamental radial fluid mode that becomes unstable at the maximum of the mass.

Not only do the radial oscillations of scalarized NSs differ significantly from those in the GR case, but the nonradial modes that are related to the tensorial gravitational wave emission can be strongly influenced by the scalar field. The behavior of nonradial oscillations modes, which are sources of the usual tensor polarizations of GWs, can also be strongly affected by the scalar field. These perturbations can be classified as “polar” or “axial” depending on how they behave under parity transformations (Regge and Wheeler, 1957; Thorne and Campolattaro, 1967).

In particular, scalar-field perturbations are of the polar type, meaning that they couple only to polar perturbations of the fluid and metric. In GR, these perturbations are the most efficient GW sources. Sotani and Kokkotas (2004) were the first to study the nonradial oscillations of scalarized NSs. They used the Cowling approximation, in which only polar-parity fluid perturbations are dynamical. This simplifies the problem considerably and Sotani and Kokkotas (2004) calculated the fundamental \( f \)-mode frequency and the pressure \( p \)-mode frequency. In a follow-up work, Sotani and Kokkotas (2005) went beyond the Cowling approximation and derived equations for both axial and polar perturbations that included metric perturbations. They analyzed only the simpler axial perturbations (discussed later). A study of polar perturbations became possible only after new techniques were developed in GR by Krüger and Kokkotas (2020a,b). These techniques were then used in scalar-tensor theory by Krüger and Doneva (2021), who also considered self-interacting massive scalar fields.\(^{10}\) Krüger and Doneva (2021) found that

---

\(^{10}\) See Blázquez-Salcedo et al. (2020d) and Staykov et al. (2015)
Numerical calculations showed that the effect of scalarization is stronger on the damping times than the effect on the frequencies, and in general the values of both are lower than their GR values. In addition, EOS-independent relations between $w$-mode properties known to exist in GR can also be obtained in scalar-tensor theory.

A class of NS modes related to the crustal torsional oscillation was studied by Silva et al. (2014) for the DEF model in the Cowling approximation. These oscillations probably follow the giant flares in soft gamma-ray repeaters (Israel et al., 2005; Strohmayer and Watts, 2005, 2006) and are associated with motions in the NS crust. Silva et al. (2014) found that, for values of $\beta_0$ consistent with binary-pulsar constraints at the time, the effect of scalarization on the torsional oscillation frequencies is smaller than the uncertainties in the microphysics modeling of the crust.

2. Nonlinear stability and collapse to a black hole

While the stability and oscillations of a scalarized NS can be studied perturbatively, the formation of scalar hair starting from a GR NS solution is a fully nonlinear process. More precisely the initial exponential growth of the scalar field starting from an unstable GR solution can be modeled by linearized dynamics, but the subsequent saturation of the scalar field to an equilibrium value is a nonlinear phenomenon.

The transition from a nonscalarized to a scalarized NS was first considered by Novak (1998a) [see also Degollado et al. (2020)], who studied the nonlinear evolution in spherical symmetry. The numerical simulations of Novak (1998a) showed that the scalar field for an unstable GR NS first grows exponentially and then saturates, with the system saturating to an equilibrium scalarized end state. A consequence of the work by Novak (1998a) was the proof of nonlinear stability of scalarized NSs. That is, the numerical simulations showed that the scalarized NSs have stable evolution if the full nonlinear system of field equations (in spherical symmetry) is considered. Although effort was not strictly dedicated to a stability analysis, a fully nonlinear evolution of scalarized NSs was performed in a series of papers that we discuss later. This provides strong support for the stability of these objects under the most general circumstances.

Novak (1998b) studied the collapse of a scalarized NS to a BH in spherical symmetry. Since BH no-hair theorems include DEF models [see Herdeiro and Radu (2015) for a review], the resulting BH will be bald and the scalar

for a discussion of a case of massive scalar-tensor theory that is mathematically equivalent to $R^2$ gravity.
field has to be radiated away during collapse.\textsuperscript{12} This takes place in the form of scalar waves, in analogy to what we discussed in the case of radial NS oscillations in Sec. III.B.1.

Our previous discussion covered the \( \beta_0 < 0 \) case. Palenzuela and Liebling (2016) and Mendes and Ortiz (2016) studied the end state of the tachyonic instability in scalar-tensor theories for representative coupling functions with \( \beta_0 > 0 \) and realistic EOSs. This was done both through an energy balance analysis of the existing equilibrium configurations and by nonlinear dynamical simulations. They found that (contrary to the \( \beta_0 < 0 \) case) the final state of the instability is highly sensitive to the details of the coupling function, varying from gravitational collapse to spontaneous scalarization. They also found that in the original DEF model [where \( \alpha(\phi) = \beta_0 \phi \) with \( \alpha_0 = 0 \)] scalarized solutions can become unstable compared to the GR ones when \( \beta_0 \gg 1 \). However, stability can be recovered for all values of \( \beta_0 \) by considering different coupling functions. This is the case for coupling functions with bounded values, such as \( \alpha(\phi \to \infty) = \alpha_\infty \) for some constant \( \alpha_\infty \). This distinction could give rise to novel astrophysical tests for determining the detailed form of the coupling.

3. Stellar core collapse

Once we know that equilibrium scalarized NS solutions exist in the DEF model and they are stable against linear perturbations, the next step is to study how they are formed. Isolated NSs can form after the collapse of the core of a massive star during a supernova explosion. A NS on the other hand can collapse to a BH if a threshold for the mass and the angular momentum is reached. All these are highly dynamical nonlinear processes that can have strong observational signatures in both the electromagnetic and GW signals.

The first study of a degenerate stellar core collapse (more specifically with white dwarf initial data) to a scalarized NS through a bounce and the formation of a shock was performed by Novak and Ibanez (2000). The simulations were done in spherical symmetry, which allowed them to calculate only the resulting gravitational monopolar radiation. They found that the emitted breathing modes can be potentially detected by LIGO or VIRGO. Notably the emitted signal will be substantially different than the collapse of a NS to a BH, allowing one to distinguish between them.

Gerosa et al. (2016) studied the problem of spherically symmetric core collapse in the DEF model in further detail. Two types of initial data were used: collapse of a stellar iron core and collapse of “realistic” NS progenitors that were obtained from computations of stellar evolution. Depending on the theory parameters three possible outcomes of the core collapse are possible – collapse to a GR NS, collapse to a scalarized NS, or collapse to a short-lived protoneutron star followed by a nonscalarized BH formation. It was in the last case that the most prominent GW signal with a clear signature from the presence of nontrivial scalar fields was observed. While the fluid dynamics during the collapse is only weakly affected by the scalar field, the converse is not true: the scalar radiation depends strongly on the specifics of the matter collapse as well as on the choice of the coupling parameters \( \alpha_0 \) and \( \beta_0 \).

The inclusion of a scalar-field mass to the core-collapse simulations has some interesting consequences (Sperhake et al., 2017). The waveform \( r_\phi \) during such events is shown in Fig. 8. The mass itself does not have a large influence on the dynamics, but it allows the theory to be reconciled with binary-pulsar observations for a much larger range of \( \alpha_0 \) and \( \beta_0 \), as discussed in Sec. III.A.2. Thus, a dramatic increase in the radiated GW signal observable even with the existing LVK detectors was predicted.

A prominent feature is that we expect to receive an inverse chirp signal from such events that will last for years, with a near monochromatic signature on timescales of \( \sim 1 \) month. In fact, the inverse chirp signal is connected to the scalar field’s mass and, consequently, to the dispersion relation. Thus, the scalar wave burst that is emitted during the collapse is effectively stretched in time with decreasing amplitude. Extensions of these results to the case of a self-interacting scalar-field potential were made by Cheong and Li (2019) and Rosca-Mead et al. (2019). Constraints on the theory based on the

\textsuperscript{12} For a discussion of collapse in theories where BHs can be endowed with a scalar field, see Sec. III.D.
scalar-field evolution in the Einstein frame were imposed by Geng et al. (2020). The problem was considered in greater detail by Rosca-Mead et al. (2020b), who found the three possible scenarios of the collapse outcome in scalar-tensor theory with sufficiently negative $\beta_0$. These constitute the formation of a BH following multiple NS stages, the multistage formation of a strongly scalarized NS, and the single-stage formation of a strongly scalarized NS. Rosca-Mead et al. (2019) found that the resulting GW signal can reach a signal-to-noise ratio of over 20 for the existing GW detectors, which has the potential to put strong constraints on the theory.

4. Dynamical scalarization and neutron star mergers

Another highly dynamical and nonlinear process that has important astrophysical, and especially GW implications is binary NS mergers. This problem can be more challenging to solve than the stellar core collapse because it is not possible by construction to apply certain approximations, such as spherical symmetry. That is why the binary merger dynamics in the DEF model was addressed shortly thereafter, first by Barausse et al. (2013) and then by Shibata et al. (2014). The overall conclusions of both works are similar and they reside in the fact that, even if one or both of the NSs are not scalarized before the merger, they can develop nonzero scalar fields during the inspiral. This was called dynamical scalarization. The main advances made by Shibata et al. (2014) constitute using several realistic EOS with consistently derived bounds on the parameter $\beta_0$, as well as developing an initial data solver for scalarized binary NSs. The overall results demonstrated a significant change of the inspiral GW signal at the moment of dynamical scalarization and afterward. The reason is that the inspiral is accelerated due to the scalar-dipole radiation and the total number of GW cycles is significantly decreased with respect to the general-relativistic case.

The actual merger and the postmerger phases were addressed only by Shibata et al. (2014), who showed that dynamical scalarization can happen not only in the inspiral phase but also during the merger since a massive compact star is formed in this process. The evolution during the NS merger of the maximum values of the rest-mass density $\rho_{\text{max}}$ and scalar field $\varphi_{\text{max}}$ for several binary NSs with total mass $M = 2.7 M_\odot$ and several values of $\beta_0$ is shown in Fig. 9. Two EOSs are considered in the two panels and the cosmological value of the scalar field is taken to be $\varphi_{0} = 10^{-5}$ in order to satisfy constraints from the binary-pulsar observations (Antoniadis et al., 2013; Freire et al., 2012). Some of the binary NSs are not scalarized at the beginning of the evolution (the models with $\varphi_{\text{max}} = 0$ at the beginning), but the values of $\beta_0$ are chosen such that all of them develop scalar fields at a certain point of the evolution. The actual merger of the two NSs is marked by the rapid increase of $\varphi_{\text{max}}$. Afterward, either a hypermassive or supramassive NS is formed or the merger remnant collapses to a bald BH and the scalar field is radiated away. In agreement with the studies of equilibrium differentially rotating NSs discussed in Sec. III A.1, the scalarized merger remnant can sustain a larger mass without collapsing to a BH (Doneva et al., 2018b). This is evident in the bottom panel of Fig. 9, where collapse to a BH is observed in pure GR, while a supramassive NS is formed after the merger in scalar-tensor gravity with small enough $\beta_0$.

The quasiperiodic oscillations of the merger remnant were examined by Shibata et al. (2014), showing a clear distinction compared to the GR case. Such oscillations were studied in a series of papers in GR (see e.g. Bauswein and Janka (2012); Bauswein et al. (2012, 2014); Clark et al. (2014); Hotokezaka et al. (2013); Maione et al. (2016); Rezzolla and Takami (2016); Takami et al. (2014, 2015)), mainly as a tool to determine the nuclear matter EOS.
from the postmerger GW signal. The observed differences with the scalarized case can potentially be used to discriminate between GR and modified gravity theories. This is especially the case because scalarized merger remnants are produced only if specific initial conditions related to the mass of the merging compact objects and the specifics of the EOS are met (Shibata et al., 2014).

Taniguchi et al. (2015) took a different approach in which they calculated quasiequilibrium sequences of binary NSs at different time instants instead of performing time evolution. This approach is well known in GR (Gourgoulhon et al., 2001; Taniguchi and Gourgoulhon, 2003; Taniguchi and Shibata, 2010), and it is assumed that the characteristic time of the system to settle to equilibrium is much smaller than the inspiral timescale. This is supposed to give a relatively accurate picture of the binary evolution even close to the merger. The results are in agreement with Shibata et al. (2014), while the small deviations relative to Barausse et al. (2013) are probably due to the fact that GR initial data were used in the latter. Taniguchi et al. (2015) showed that the absolute value of the binary binding energy is smaller than in GR. In addition, the GW cycles prior to the merger were significantly reduced compared to pure GR once scalarization kicks in, and the effect is considerably stronger than the one due to tidal interactions.

The post-Newtonian (PN) approximation has also been a valued tool to model NS inspiral in the DEF model. Initial work has focused on tensor-multiscalar theories and simple scalar-tensor theories that do not allow scalarization (Damour and Esposito-Farèse, 1992, 1996b; Lang, 2014; Mirshekari and Will, 2013). This approach cannot be immediately applied to the case of dynamical scalarization, since this is a nonperturbative effect that is absent from the weak-field regime. Palenzuela et al. (2014) were the first to address this issue. They used the equations of motion at 2.5PN order, derived in scalar-tensor gravity by Mirshekari and Will (2013), modified in such a way that the changes in the stars’ scalar charges are taken into account. More specifically, they solved a system of nonlinear algebraic equations at each step of the orbital evolution to compute the scalar charges. The resulting inspiral evolution was found to be in agreement with the numerical simulations of Barausse et al. (2013). The approach of Palenzuela et al. (2014) is computationally inexpensive, and this allowed them the to study large portions of the parameter space, including (un)equal-mass and eccentric binaries. This approach can also be used to efficiently generate inspiral GW templates in the DEF model. A further step forward taken by Sennett and Buonanno (2016) introduced a methodology called post-Dickean expansion. Their main improvement with respect to Palenzuela et al. (2014) is a 1PN extension of the feedback mechanism. The post-Dickean expansion was compared against the quasiequilibrium calculations of Taniguchi et al. (2015), and it was shown that this can accurately predict the onset and magnitude of dynamical scalarization.

Sennett et al. (2017) took a different approach in the perturbative study of the inspiral that constitutes an analytical model of dynamical scalarization using an effective action. The motivation was to cure two deficiencies in the previous PN studies. In an effective action approach, the nonlinear scalarization process is reduced to a pair of cubic equations that have a closed-form solution depending on the binary separation when dynamical scalarization happens and on the magnitude of the developed scalar charge. This simplifies the problem relative to Palenzuela et al. (2014) and Sennett and Buonanno (2016). In addition, the effective action approach allows one to construct a simple two-body Hamiltonian that can be used to compute the binary’s binding energy. Sennett et al. (2017) used this Hamiltonian, in combination with Landau’s theory of phase transitions, to interpret dynamical scalarization as a second-order phase transition. Khalil et al. (2019) extended these results to general theories admitting scalarization for either BHs or NSs and are valid for adiabatic (quasistationary) and quasicircular orbits. Khalil et al. (2022) took it a step further, where these approximations were dropped and the dynamical evolution around the phase transition to the scalarized regime was studied. The results showed that in some cases assuming a quasistationary evolution might not be accurate enough even for quasicircular binaries.

Last, Ponce et al. (2015) studied the effect of the scalar field on the electromagnetic radiation emitted in the merger of magnetized NSs. They found that deviations in the emitted electromagnetic flux due to scalarization are not negligible yet are challenging to measure. However, if combined with GW observations, constraints on scalar-tensor theory can in principle be placed, showing the usefulness of multimessenger astronomy.

C. Astrophysical implications of scalarized neutron stars in the Damour–Esposito-Farèse model

In this section, we further discuss the astrophysical implications of scalarized NSs. Many aspects have already been covered, such as the binary-pulsar observations, NS oscillations, stellar core collapse, and binary mergers, as they naturally appeared in the presentation. Here, we shed further light on the possible astrophysical implications of scalarization, trying to be as complete as possible in two main areas. We first discuss the astrophysical implications directly related to electromagnetic observables. Afterward, we address the problem of universal relations for scalarized NS models.

a. Electromagnetic observations: Scalarized NSs in the DEF model and its extensions were studied in a vari-
et al. (2014b) studied rapidly rotating NSs. The conclusion of this work is how one can disentangle uncertainties in the NS EOS and in the disk model from modifications to GR.

Another prospective way to infer information about accreting compact objects is through the shape of the Fe $K\alpha$ fluorescent line at 6.4 keV. Since it is emitted from the inner regions of the accretion disk, it carries traces of the underlying spacetime geometry. This line can be observed for both accreting NSs and BHs, and the accuracy of the observations is expected to be improved through future x-ray missions such as Athena. Bucciantini and Soldateschi (2020) calculated the shape of this line by taking the light propagation around a scalarized NS into account. They argued that the influence of GR modification both on the intensity of the low-energy tails and on the position of the high-energy edge of the line are potentially observable in the future.

Accretion onto NSs can also trigger a process called the gravitational phase transition, which was named in analogy with matter phase transitions from confined nuclear matter to deconfined quark matter (Kuan et al., 2022). This process can happen when the maximum mass of the scalarized NSs is smaller than the maximum mass of the zero scalar-field (GR) solution. The idea then is that if a scalarized NS close to this maximum mass accretes some matter, it may pass beyond the stability point. In GR this would cause a collapse to a BH. In scalar-tensor theory, the star will radiate its scalar hair and evolve toward the zero scalar-field (i.e., the GR) branch. A significant amount of GWs can be produced in this process, and they will potentially be detectable with the next generation of GW detectors.

The NS surface as an emitter of x-ray radiation is also an important probe in strong-field gravity because the electromagnetic radiation is emitted from a region with large spacetime curvature. In this regard, the simplest observable is the gravitational redshift of surface atomic lines. The redshift carries information about the NS mass, radius, spin, and, in scalar-tensor gravity, the scalar field. DeDeo and Psaltis (2003) showed that scalarization had a significant effect on the redshift, but only for negative enough values of $\beta_0$. Such values are already ruled out by binary-pulsar observations. One could entertain the idea that a small scalar-field mass could allow for large deviations in redshift relative to GR, while reconciling the theory with binary-pulsar constraints (Doneva and Yazadjiev, 2016; Popchev, 2015; Ramazano˘glu and Pretorius, 2016). A study of this problem has not yet been performed.

The x-ray pulse profiles emitted by hot spots at NS surfaces have also received considerable attention. Observations of these signals allow for a relatively clean inference of NS masses and radii (Watts et al., 2016). This potential was met with observations (Bogdanov et al., 2019a; Miller et al., 2019; Riley et al., 2019) by the Neutron Star Interior Composition Explorer (NICER) (Arzoumanian et al., 2014). Sotani (2017) and Silva and Yunes (2019b) developed pulse-profile models for NSs in a massless DEF model. Xu et al. (2020) considered the case of massive scalar fields and Hu et al. (2021) studied the positive $\beta_0$ case. Silva and Yunes (2019a) made the first study of what constraints can be placed on the DEF model with pulse-profile observations and found they can be competitive with binary-pulsar observations. However, their work is rather simple and not at the level of realism found in the analysis of real data. This remains an important avenue for future work.

At last, Tuna et al. (2022) used NS mass and radius measurements (Bogdanov et al., 2016; Ozel et al., 2016) to constrain the massive extensions of the DEF model. They obtained a weak lower bound $\beta_0 \gtrsim -20$ for scalar-field masses $\mu \lesssim 2 \times 10^{-11}$ eV; see Fig. 10. This is significant since no other bound is known for $\mu \gg 10^{-16}$ eV. These results show that large scalar masses enable agreement with observations, even for extremely negative $\beta_0$, and demonstrate the difficulty to constrain scalarization when the scalar field is massive.

b. Universal relations: One of the largest obstacles when using NS observations to test modified theories of gravity is the uncertainty in the NS EOS, which remains unknown at high densities (Baym et al., 2018; Lattimer and Prakash, 2016). In general modifications to NS properties predicted by different EOSs are degenerate with changes to the underlying gravity theory used to model these stars. One way to break this degeneracy is to consider relations between different NS properties that depend weakly on the EOS (Doneva and Pappas, 2018; Yagi and Yunes, 2017). In fact, in Sec. III.B.1.b we already met one such
example relating the QNM frequencies and damping times to the mean density of the NSs. Here we focus on other EOS-independent relations connecting various equilibrium properties of scalarized NSs.

A class of universal relations that has attracted considerable attention is the I-Love-Q relations, which connect the normalized moment of inertia, the tidal Love number, and the quadrupole moment of the NSs (Yagi and Yunes, 2013a,b). If any two quantities in the I-Love-Q trio are measured independently, one can constrain deviations from GR by determining whether these observations agree with the GR I-Love-Q relation within observational accuracy. Silva et al. (2021a) applied this strategy to constrain dynamical Chern-Simons gravity (Alexander and Yunes, 2009; Jackiw and Pi, 2003) using observational data from LVK and NICER. The application to scalarization models has not yet been done.

The I-Love-Q relations for scalarized NSs were studied in slow (Pani and Berti, 2014) and rapid rotations (Doneva et al., 2014a), including the case of massive scalar fields (Doneva and Yazadjiev, 2016; Hu et al., 2021). In Fig. 11 we show the relation between the normalized moment of inertia $I$ and the quadrupole moment $Q$ for sequences of scalarized NSs with a fixed rotational parameter $\alpha = M\Omega/2\pi$. We see that for small absolute values of $\beta$, for example, $\beta_0 = -4.5$, the deviations from GR are practically negligible for slow rotation (such as $\alpha = 0.32$) while they are enhanced by rapid rotation. As we have seen, massive scalar fields allow for much smaller values $\beta_0$ while remaining consistent with observations. Hence, these scalarization models allow for larger differences between the scalarized NS universal relations and the GR ones (Doneva and Yazadjiev, 2016).

The moment of inertia $I$ and the quadrupole moment $Q$ are among the leading-order multipole moments in the asymptotic expansion of the metric functions at infinity. There are infinitely many other higher-order multipole moments, however, and it is interesting to explore whether similar universal relations hold for them. This was first addressed in GR (Pappas and Apostolatos, 2014; Yagi et al., 2014), where such universal relations were derived for the higher multipole moments when proper normalization was applied. Pappas and Sotiriou (2015) and Pappas et al. (2019) extended these results to scalarized NSs. This required the generalization of the multipole moment formalism of Geroch (1970) and Hansen (1974) to scalar-tensor gravity. Pappas et al. (2019) showed that future observations of QPOs of low-mass x-ray binaries can in principle be used to measure different NS properties and distinguish different gravity theories.

Another class of universal relations studied in the context of the DEF model connects the NS moment of inertia and compactness. This was endorsed by Lattimer and Schutz (2005) [see also Breu and Rezzolla (2016)], who used this relation to argue that pulsar-timing observations could lead to a measurement of the moment of inertia to within 10%. These relations were generalized to scalarized NSs with massless scalar-field potential by Altaha Mo-
tahar et al. (2017), and later to massive self-interacting scalar fields by Popchev et al. (2019). As with the I-Love-Q relations, only the massive scalar-field case leads to large deviations from GR when observational constraints are taken into account.

Ofengeim (2020) proposed connecting the physical parameters of static NSs, such as mass, radius, central energy density, pressure, and sound speed, at the maximum-mass point for a given EOS. This is based on the observations that the nuclear matter EOS (without phase transitions) can be well parametrized by only two parameters (Lindblom, 2010). As a result, Ofengeim (2020) derived multiple constraints on the nuclear matter EOS. Donchev and Doneva (2021) generalized these relations to scalarized NSs. They showed how sensitive these relations can be to the underlying theory of gravity. Thus, all of the EOS restrictions derived in GR should be taken with care since the possibility for GR modifications is rarely taken into account when interpreting observational data.

D. Extended scalar-tensor theories beyond the Damour–Esposito-Farèse model

Thus far we have discussed scalarization in the DEF model or other DEF-inspired models. However, as we saw in Sec. II.B, spontaneous scalarization is also possible for other modified theories of gravity. Here we discuss the theories where scalarized NS solutions were obtained.

a. Scalar-tensor theories with disformal coupling

Minamitsuji and Silva (2016) studied NS scalarization in scalar-tensor theories with disformal coupling of the form of Eq. (17). More specifically, instead of the standard conformal transformation between the Einstein-frame metric \((g_{\mu\nu})\) and the Jordan-frame metric \((\tilde{g}_{\mu\nu})\)

\[
g_{\mu\nu} = A^2(\varphi)\tilde{g}_{\mu\nu}, \tag{33}
\]

where \(A(\varphi)\) is a function of the scalar field related to the Jordan-frame coupling between the scalar field and the Ricci scalar, we have a more general and complicated transformation that also involves scalar-field derivatives,

\[
g_{\mu\nu} = A^2(\varphi) \left[ \tilde{g}_{\mu\nu} + \Lambda B^2(\varphi) \nabla_\mu \varphi \nabla_\nu \varphi \right], \tag{34}
\]

where \(B(\varphi)\) is a function of the scalar field and \(\Lambda\) is a parameter of dimensions of [length]^{-2}. The motivation behind this modification of the conformal factor comes from Bekenstein (1993) who aimed to find the most general coupling constructed from the metric \(g_{\mu\nu}\) and the scalar field \(\varphi\) that respected causality and the WEP. More recently Andreou et al. (2019) showed that such disformal coupling is actually equivalent to a kinetic coupling with the Ricci scalar and thus is contained in the minimal action (14); see Sec. II.B.1 for a discussion.

b. Scalar-Gauss-Bonnet theory

Another well-studied modified gravity theory that allows for NS scalarization is the scalar-Gauss-Bonnet gravity discussed in Sec. II.B. Space-time curvature itself, as well as matter, can be the source of the scalar field in this theory. Scalarized NS solutions in scalar-Gauss-Bonnet gravity were studied for the first time by Doneva and Yazadjiev (2018a) and Silva et al. (2018). The solutions exhibit much different features than to the DEF model while being qualitatively similar to NS solutions in other classes of Gauss-Bonnet theories where scalarization is not possible (Kleihaus et al., 2016; Pani et al., 2011a; Safer et al., 2019). For example, for a given central energy density the stellar mass is always smaller than in GR. Moreover, there is normally only one bifurcation point at small central energy densities, and afterward the branches of solutions are terminated because of violation of the regularity conditions.

In Fig. 12, we show the branches of scalarized NS solutions for a coupling function \(f(\varphi) = -\lambda^2(2\beta)^{-1}[1 - \exp(-\beta \varphi^2/4)]\). Here \(\lambda\) is a parameter with dimensions of [length]^{-1} that controls the coupling strength in the action (25). Note that the sign of \(f(\varphi)\) is the opposite of that discussed in Sec. II.B.3, and thus cannot lead to BH spontaneous scalarization, since in vacuum \(\rho_{\text{eff}} > 0\), as defined in Eq. (27). However, owing the presence of matter, NSs can scalarize. Scalarized NS solutions were also found for the more “conventional” sign of \(f(\varphi)\), namely, \(f(\varphi) = \lambda^2(2\beta)^{-1}[1 - \exp(-\beta \varphi^2/4)]\) (Doneva and Yazadjiev, 2018a). As in the DEF model, it was demonstrated that binary pulsars strongly constrain the coupling parameters of the theory, thereby leaving a small window for scalarization (Donchev et al., 2022).

Little is currently known about the astrophysical implications of these stars. Kuan et al. (2021a) considered the spherically symmetric core collapse of a noncompact star either to a protoneutron star or to a BH in Gauss-Bonnet theory. They also proposed a realistic physical mechanism for the formation of isolated scalarized BHs and NSs. The complexity of the problem is greatly increased with respect to GR, though, and there are fundamental difficulties such as the loss of hyperbolicity of the evolution equations for certain ranges of parameters, effectively limiting the maximum possible scalar field to relatively low values; see also East and Ripley (2021a,b) and Ripley and...
They conducted a thorough exploration of the theory parameter space where the GR solution is stable against density and the MPA1 EOS (Müther et al. 1987). The white area Fig. 13 corresponds to the region of the parameter space where the GR neutron star solutions exist. Since their scalar charge is nonzero, in this theory, the effect on GW emission can be estimated only if one drops the assumption of spherical symmetry. This is something that has not yet been done in any modified gravity theories due to the complexity of the problem.

Pretorius (2020) for the case of BHs. Core-collapse simulations show that the remnant in scalar-Gauss-Bonnet theory can be rich, with (de)scalarization happening at the intermediate or final stages of the collapse, depending on the properties of the progenitor and the theory parameters. Since breathing modes are absent from this theory, the effect on GW emission can be estimated only if one drops the assumption of spherical symmetry. This is something that has not yet been done in any modified gravity theories due to the complexity of the problem.

c. Ricci-Gauss-Bonnet model: One can further modify the action in scalar-Gauss-Bonnet gravity to include additional terms. This is the case for the Ricci-Gauss-Bonnet model discussed in Sec. II.B.4. This model has advantages such as the possibility of reconciling scalarization with cosmology. Here we focus on the NS solutions within this theory that were considered by Ventagli et al. (2021). They conducted a thorough exploration of the theory parameter space in order to find the sectors where NS solutions exist. Since their scalar charge is nonzero, in these sectors one can put severe constraints on the theory based on binary-pulsar observations, as in the DEF model; see Sec. III.A.1.b.

The free parameters in the theory, as evident from the action (28), are β and α which control the coupling strength to the Ricci scalar and the Gauss-Bonnet invariant, respectively. In Fig. 13 the existence of NS solutions is shown in a two-dimensional plot spanning the parameters α and β in a broad range for a fixed central energy density and the MPA1 EOS (Müther et al., 1987). The white area Fig. 13 corresponds to the region of the parameter space where the GR solution is stable against scalar perturbations. Throughout the parameter space that is spanned by α and β, a new unstable mode appears every time that one crosses a black line. These unstable modes can be labeled by the number of scalar-field nodes (denoted by n = 0, 1, 2 and 3 in the figure). Any point in the parameter space that lies within a gray region corresponds to a configuration where the GR neutron star solution is unstable. The red (blue) area corresponds to the region where scalarized solutions with n nodes exist. It is evident that scalarized solutions exist in only part of the parameter space. In the β > 0 and small-α range (the white region in Fig. 13) the GR neutron stars are stable. This is the region that is most noteworthy from the point of view of BH scalarization and where we can reconcile scalarization with cosmology (Antoniou et al., 2021a,b); see also Sec. IV.A.2. Thus, including a Ricci coupling to the scalar-Gauss-Bonnet action seems to allow for BH scalarization while evading the binary-pulsar constraints.

Figure 13 is for only one EOS and a specific central energy density. Different EOSs and central densities can lead to significant deformations of the existence and stability regions. The general point made by Ventagli et al. (2021), however, is that the stability (white) region always survives.
d. Tensor-multiscalar theories: Another class of gravity theories where NS scalarization was considered is the tensor-multiscalar theories (TMSTs) (Damour and Esposito-Farèse, 1992), whose basics were discussed in Sec. II.B. In this theory, the gravitational interaction is mediated by the spacetime metric $g_{\mu \nu}$ and $N$ scalar fields $\varphi^a$, which take values in a coordinate patch of an $N$-dimensional Riemannian target manifold $\mathcal{E}_N$ with a positive-definite metric $\gamma_{ab}(\varphi)$ defined on it (Damour and Esposito-Farèse, 1992; Horbatsch et al., 2015).

The main features of TMST are the inclusion of more than one scalar field and a structure called the target-space metric. In that sense, there are two directions to go in order to obtain scalarized NSs. The first one is to consider a mixture of several more or less equivalent scalar fields; this was the approach followed by Horbatsch et al. (2015). They, the authors examined the case of two scalar fields in the form of a complex scalar with a maximally symmetric target-space metric $\gamma_{ab}$. Doneva and Yazadjiev (2020a) studied scalarization in tensor-multiscalar gravity in a more complicated setup, namely, when $\gamma_{ab}$ is a three-dimensional maximally symmetric space together with a nontrivial map $\varphi : \text{spacetime} \rightarrow \text{target space}$. While the solutions given by Horbatsch et al. (2015) can be viewed as a generalization of the DEF model to multiple scalar fields, the scalarized NSs given by Doneva and Yazadjiev (2020a) have some distinct properties and they can be considered more as a limiting case of the topological NSs discovered by Doneva and Yazadjiev (2020b). More specifically, the scalar field is zero in the stellar center, while only its first derivative is zero in the DEF model. The scalar charge for the scalarized stars in TMST is zero as well. This automatically reconciles this theory with the binary-pulsar observations (due to the absence of scalar-dipole radiation) while still allowing for large deviations from GR. Another interesting property is the fact that there is nonuniqueness of the scalarized solutions with respect to the central energy density, something that had never been observed before, at least for scalarized NSs. Figure 14 illustrates this property. Interestingly, Kuan et al. (2021b) showed that all of these solutions are stable until the maximum NS, mass regardless of the nonuniqueness.

e. Other scalarization models: Apart from the previously discussed models of scalarized NSs, other works have attempted to consider scalarization in more exotic theories of gravity, or even different types of scalarization.

For example, Azri and Nasri (2021) considered NS scalarization in what is called scalar-connection gravity, where gravity is mediated by a scalar field and the connection. Other types of scalarization can happen in standard scalar-tensor theories if we allow for different scalar-field potentials. Minamitsuji and Tsujikawa (2023) showed this when considering the potential $V(\varphi) = m_{\varphi}^2 f_B^2 [1 + \cos(\varphi/f_B)]$ of a pseudo-Nambu-Goldstone boson. Here $(\mu_{\varphi}, f_B)$ are constants with dimensions of length and the scalar field has a negative mass $\mu^2 = -m_{\varphi}^2$. Minamitsuji and Tsujikawa (2023) showed that in this case the scalar field sits at its vacuum expectation value far from the source, while inside a NS a symmetry restoration can take place, resulting in a new type of scalarization. This theory also has an advantage in that it avoids cosmological instabilities present in certain other scalarization models.

IV. BLACK-HOLE SCALARIZATION

In this section, we discuss the spontaneous scalarization of BHs in three parts. To begin, in Sec. IV.A we consider models in which BHs can scalarize in vacuum due to couplings between the scalar field and curvature scalars. Next, in Sec. IV.B we discuss models in which scalarization is induced by the presence of extra fields (such as gauge or matter fields) in the BH spacetime. Finally, in Sec. IV.C we review a selection of other models of BH scalarization.

A. Black-hole scalarization: Vacuum spacetimes

As we saw in Sec. II, some gravity theories admit the same vacuum BH solutions as GR, yet can give rise to new branches of solutions with scalar hair once certain
conditions are met. The prototypical example is described by scalar-Gauss-Bonnet theories, whose action in the absence of matter is given by Eq. (25), namely,

\[ S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[ R - \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi + f(\varphi) g \right]. \]

As discussed in Sec. I, the first models exhibiting BH scalarization in vacuum were proposed by Doneva and Yazadjiev (2018b) and Silva et al. (2018). These models remain the most studied BH scalarization models to date in the literature. For this reason, they will be our main focus in this section.\(^{13}\)

As explained in Sec. II.B.3, in theories described by the action (25), a no-hair theorem guarantees that the BH solutions of GR are unique to the theory, as long as \((d^2 f/d\varphi^2)_{\varphi=\varphi_0} < 0\), for some constant \(\varphi_0\) (Silva et al., 2018). If this condition is violated, scalar-field perturbations can become tachyonic unstable and the endpoint is expected to be a nonlinear, scalarized BH (Ripley and Pretorius, 2020). Scalarized BHs have been shown to form dynamically, as outcomes in the core collapse of an initially unscalarized star (Kuan et al., 2021a).

Here we first review our understanding of isolated scalarized BHs in scalar-Gauss-Bonnet theories, focusing on their properties and their stability. We then review what is understood about when they are found in binaries. Finally, we give an overview of vacuum BH solutions in models that generalize the action (25).

1. Scalarized black holes

Spontaneously scalarized BH solutions were first found by Doneva and Yazadjiev (2018b) and Silva et al. (2018). What distinguishes these works is the choice of the coupling function \(f(\varphi)\) that couples the scalar field to the Gauss-Bonnet invariant [cf. Eq. (25)],

\[
\begin{align*}
  f &= \left( \lambda^2 / 12 \right)[1 - \varepsilon \exp(-3\varphi^2 / 2)], \\
  f &= (\eta / 8) \varepsilon \varphi^2,
\end{align*}
\]

where \(\lambda^2\) and \(\eta\), positive by definition, are coupling constants with dimensions of [length]\(^2\) and \(\varepsilon = \pm 1\), which should not be confused with the energy density of Sec. III. The two coupling functions agree in the small-\(\varphi\) limit (i.e., when \(\varphi \ll 1\)) and therefore result in the same prediction for the onset of scalarization of GR BHs. This threshold can be found by searching for bound state solutions, i.e., time-independent solutions of the linearized field equation for \(\varphi\) around a GR BH, which are regular at the event horizon \(r_h\) and that vanish at spatial infinity. Hence, the determination of the scalarization threshold reduces to a boundary value problem. In the simplest case of a Schwarzschild BH of mass \(M\), the dimensionless quantity \(\eta/M^2\) plays the role of the eigenvalue and can be determined numerically with standard techniques such as the shooting method. The smallest eigenvalue gives the threshold for the formation of the “fundamental” scalarized BH solution (in the sense that the radial profile of the scalar field has \(n = 0\) nodes). The other eigenvalues give the thresholds for the formation of “excited states,” that is, solutions with one or more nodes. In Fig. 15 we show the results of such a calculation in the spacetime of a Schwarzschild BH and \(\varepsilon = 1\), as done by Silva et al. (2018). We remark that no bound states can form for \(\varepsilon = -1\), because the effective mass of scalar-field perturbations is positive, and therefore, the effective potential is positive definite. See the discussion in Sec. II.A.2.

While both models in Eqs. (35) agree in their prediction of the threshold of scalarization, they differ significantly in their prediction of the properties of the nonlinear, scalarized solution, as discussed in Sec. II.A. In Fig. 16 we show the branches of scalarized BHs, in a parameter space spanned by the dimensionless scalar charge and BH mass. The solutions were obtained by solving the full system of
field equations; see Doneva and Yazadjiev (2018b) and Julié et al. (2022) for detailed discussions. The results are for the Gaussian coupling (35a) with \( \varepsilon = 1 \). We show the fundamental, zero-node scalar-field solution (dashed red curves) and also the excited scalarized BHs (solid green and blue curves). We see that the branches of excited BHs can terminate at finite masses \( M \). This is due to a violation of an inequality at the BH horizon that must be satisfied for the scalar field to be real valued and is typical of scalar-Gauss-Bonnet theories; see e.g., Kanti et al. (1996) and Antoniou et al. (2018). The parameters in the coupling function (35a) are chosen in such a way that this disappearance of BH solutions does not affect the fundamental branch. For other couplings, such as Eq. (35b), even this branch can violate the regularity condition shortly after bifurcation; see Doneva et al. (2018a) and Silva et al. (2019). Hence, the domain of existence of scalarized solutions in the Gaussian model (35a) is larger with respect to that of the “quadratic” model (35b).

Asymptotically flat, spinning scalarized BHs in the Gaussian models (with \( \varepsilon = 1 \)) and quadratic theories were explored, respectively, by Cunha et al. (2019) and Collodel et al. (2020). In discussing these solutions, we consider a plane spanned by the dimensionless scalar charge \( Q_s/\lambda \) and dimensionless spin parameter \( j \equiv J/M^2 \), where \( J \) is the angular momentum of the BH. Figure 17 shows the domain of existence of scalarized rotating BH solutions. In particular, we focus on the inset, which corresponds to \( \varepsilon = 1 \) in Eq. (35a). We see that as the spin increases the existence domain of the solutions (the shaded region) becomes smaller. This result can be understood in terms of the Gauss-Bonnet invariant not being positive definite for the Kerr metric (Cherubini et al., 2002). More specifically, a Kerr BH develops increasingly large regions where the Gauss-Bonnet invariant is negative as the spin increases and hence suppresses scalarization (or triggers scalarization, when \( \varepsilon = -1 \), as discussed in Sec. IV.A.3.) As a consequence, the deviations relative to GR predictions on physical observables (for example, the location of the innermost stable circular orbit of massive particles or the BH shadow) are smaller for rapidly rotating BHs. See Fig. 18 for an example. Collodel et al. (2020) studied the quadratic coupling (35b) and reached similar conclusions. They obtained radially excited solutions (i.e., scalar-field profiles with nodes in the radial direction) and also notes the existence of “angularly” excited rotating BHs (i.e., scalar-field profiles with nodes in the angular directions).

2. Stability of scalarized black holes and implications for model building

Having established the existence of scalarized BHs, the natural next task is to study their stability. A first indication comes from the study of the thermodynamical properties of such BHs. Doneva and Yazadjiev (2018b) found that the scalarized solutions belonging to the fundamental branch in the Gaussian model are thermody-
This suggests the stability of the former; i.e., in a state of smaller entropy. In contrast, the quadratic dynamical process BHs in this model would evolve toward the GR mass threshold. Past the turning points (marked with blue dots), the BHs are expected to be unstable to radial perturbations.

A complementary question is whether scalarized BHs with a quadratic coupling between the scalar field and the Ricci tensor is added to the quadratic scalar-Gauss-Bonnet interaction. As in Fig. 19, we show the charge $Q = Q/\alpha^{1/2}$ and mass $M/\alpha^{1/2}$ made dimensionless by division by $\alpha^{1/2}$, for different values of scalar-Ricci coupling $\beta$. See Eq. (28) for the theory’s action. The left-bending black lines correspond to values of $\beta$ for which all scalarized BHs have masses below the GR instability mass threshold, while the right-bending red lines mark values of $\beta$ that lead to scalarized BH masses that are larger than the GR mass threshold.

This important difference between the two models can be traced back to the higher-than-quadratic terms in the scalar field in the theory’s action. More concretely Silva et al. (2019) and Minamitsuji and Ikeda (2019a) showed that the inclusion of a quartic interaction, say, $(\zeta/\eta)\phi^4$, is sufficient to yield stable BH solutions as long as $\zeta$ is negative and satisfies $\zeta/\eta < -0.8$. These two conditions are satisfied by the Gaussian model. Indeed, expanding Eqs. (35a) to order $O(\phi^4)$, we find that $\zeta/\eta = -3/2$, which helps to explain the stability of BHs in this model.

It was later shown by Macedo et al. (2019) [see also Macedo (2020)] that the original quadratic model can be made stable under radial perturbations by making the scalar field massive and self-interacting, $V(\phi) \propto \mu^2 \phi^2 + \lambda\phi^4$, as discussed in Fig. 19. From an EFT perspective, these terms are of lower order than the quartic scalar-Gauss-Bonnet interaction and hence should be included. A similar analysis was also performed for the Ricci-Gauss-Bonnet model by Antoniou et al. (2021b) and is shown in Fig. 20. See also Doneva et al. (2019) for a study of scalarization with massive scalar fields.

A complementary question is whether scalarized BHs are stable to linear perturbations. The first step in this direction was carried out by Blázquez-Salcedo et al. (2018), who considered radial scalar-field and metric perturbations. They showed that Schwarzschild BHs became unstable at the scalarization threshold and that the stability of the fundamental (i.e., the zero-node) scalarized solution depended on which model in Eq. (35) was considered: in the quadratic case all solutions were unstable, whereas stable solutions existed in the Gaussian model.
The nonlinear stability of BHs in this model was also studied in the time domain by Ripley and Pretorius (2020). They found evidence for regions in mass-coupling parameter space in which the end state of the radial instability of the Schwarzschild BH is a stable scalarized BH. For larger couplings however, regions where the time-evolution equations change character from hyperbolic to elliptic appear outside the BH horizon. This signals a regime in which the theory does not have a well-posed initial-value problem. See Hilditch (2013) and Ripley (2022) for discussions. These results are in agreement with earlier findings by Blázquez-Salcedo et al. (2018), but in the linear regime. The radial stability of scalarized BHs in the Ricci-Gauss-Bonnet model introduced in Sec. II.B.4 was analyzed by Antoniou et al. (2022). They found that although the Ricci scalar does not affect the scalarization threshold, at the nonlinear level it can stabilize the BHs in the quadratic model.

The problem of nonradial perturbations of scalarized BH solutions, which is relevant in the context of testing such theories with ringdown GW signals, was addressed by Blázquez-Salcedo et al. (2020a,b). However, the stability of rotating scalarized BHs remains an open problem, even in the slow rotating limit.

3. Spin-induced scalarization

One property of the Gauss-Bonnet invariant of the Schwarzschild metric is that it is positive definite everywhere and, as a consequence, spontaneous scalarization can occur only for positive values of the coupling constant. Does this always have to be the case? In the case of the Kerr metric, it is known that regions where $\mathcal{G} < 0$ outside the outer horizon are possible if the BH spins with $a/M \geq 0.5$ (Chernubini et al., 2002), suggesting the possibility of a spin-induced scalarization when the coupling constant is negative, i.e., $\varepsilon = -1$ in Eqs. (35).

Although conceptually simple, the problem is not straightforward to analyze because (i) the analytical form of $\mathcal{G}$ forbids the separation of variables as done in Sec. I, and thus the determination of the scalarization threshold will in principle require more sophisticated numerical methods than the Schwarzschild case, and (ii) the Kerr metric (with $a/M \geq 0.5$) has regions where $\mathcal{G}$ can be either negative or positive, where in the latter regions, for negative coupling constants, the effective mass may trigger a superradiant instability (Brito et al., 2015). To overcome these difficulties, Dima et al. (2020) reduced the $(2+1)$-dimensional nonseparable scalar-field equation to a $(1+1)$-dimensional system of equations that are coupled through spherical harmonic multipole indices $l$ and $m$. They then performed time-domain numerical integration of the scalar field borrowing methods that were previously developed to study superradiant instability (Dolan, 2013). Dima et al. (2020) showed that the tachyonic instability is the dominant one and charted the parameter space in which spin-induced scalarization would occur. This is shown in Fig. 21. Hod (2020) worked analytically in the infinitely large coupling limit ($\eta/M^2 \rightarrow \infty$) and confirmed the expectation that $a/M \gtrsim 0.5$ is the minimal necessary spin value for which spin-induced scalarization occurs. The corrections for large but finite coupling were obtained by Hod (2022).

Nonlinear spin-induced scalarized BH solutions were obtained by Herdeiro et al. (2021b) and by Berti et al. (2021). They confirmed the existence of scalarized BH solutions in the parameter space region in which linear theory predicts the spin-induced tachyonic instability. In Fig. 17 we show the domain of the spin-induced scalarized BHs in the Gaussian model (35a). We see that scalarized solutions exist only for $j \gtrsim 0.5$, and that their scalar charge increases with spin. Moreover, there are solutions that can violate the Kerr bound $j \leq 1$ on BH spins.

A approach complementary to that of Dima et al. (2020) was used by Doneva et al. (2020a) and Doneva and Yazadjiev (2021a). They evolved the scalar-field equation in 2+1 dimensions, i.e., without doing a multipolar decomposition. The influence of a nonvanishing scalar-field mass was explored by Doneva et al. (2020b). A similar timedomain study that instead used the hyperboloidal foliation method (Zenginoğlu, 2008) was performed by Zhang et al. (2020) and a model with both the Pontryagin and Gauss-Bonnet invariants was studied by Myung and Zhou.
et al. (2021). Annulli et al. (2022) studied the effect of strong magnetic fields on spin-induced scalarization. They found that the magnetic field suppressed this effect; i.e., it pushed the scalarization threshold to larger values of the dimensionless spin $j$ of the BH. The case of spin-induced scalarization has been much less studied in terms of its stability and on how additional nonlinear interactions affect the scalar charge.

4. Scalarized black holes in binary systems

Most work in BH scalarization has focused on isolated BHs, but to be able to confront these models against GW observations one has to turn to the two-body problem. What phenomenology would we expect in BH binaries? To answer this question we first consider work that explored the nonlinear regime of the late inspiral and merger of BH binaries, a regime that relies on numerical relativity. Next we consider work that focused on the inspiral alone, a regime that can be modeled with PN theory.

The first work in this area was done by Silva et al. (2021b), who studied the scalar-field dynamics (i.e., in the decoupling limit) in head-on collisions and nonspinning, quasicircular inspirals of binary BH spacetimes obtained from numerical relativity simulations. This work adapted the methods developed by Witek et al. (2019) for shift-symmetric scalar-Gauss-Bonnet theory to theories which allow for spontaneous scalarization. In particular, because they were interested at phenomenology near the scalarization threshold, Silva et al. (2021b) adopted the quadratic model (35b), with a coupling strength $\eta$ such that both (or one of them, depending on the mass ratio) binary components or the remnant BH can carry a scalar-field bound state through the simulation. They showed that BH binaries can either form a scalarized remnant or dynamically descalarize by shedding off the initial scalar hair (i.e., the scalar bound state configuration) depending on the values of the coupling constants and the mass ratio between the two BHs. An example of dynamical descalarization is shown in Fig. 22. Dynamical descalarization was also shown to occur in nonlinear scalarization models (see Sec. II.C.5) by Doneva et al. (2022) that simulated head-on BH collisions and worked in the decoupling limit.

A natural question that follows is: What would happen when $\varepsilon = -1$, i.e., the case in which spin-induced scalarization occurs? This is relevant from an observational point of view because BH remnants of binary BH coalescences have typical dimensionless spins of $j \sim 0.7$. This value comfortably meets the criteria for spin-induced scalarization to happen. This question was explored by Elley et al. (2022), who found through a suite of numerical relativity simulations that: (i) spin-induced dynamical descalarization can happen when the remnant BH spin is smaller than that of the initially spinning (and scalarized) binary components, and (ii) spin-induced dynamical scalarization in an initially nonspinning BH binary system can result in a scalarized rotating BH.

In the latter case, if the value of the coupling constant is sufficiently large the scalar field may affect the late inspiral and ringdown of the binary. For values of the coupling constant near the scalarization threshold for the remnant BH, the tachyonic instability can be delayed to some time $\approx 100M$ after the binary merger, and can thus hide the scalar field throughout the binary’s inspiral. This was termed stealth scalarization by Elley et al. (2022); Fig. 23 displays an example.

A restriction of the decoupling limit is that it does not capture the backreaction of the scalar field onto the spacetime metric, and thus does not allow one to derive the modifications to the gravitational waves in the previously mentioned binary setups. These difficulties can be partially overcome by employing a formulation of the theory’s field equations in a generalized harmonic gauge that was first developed by Kovács and Reall (2020a,b); see also Ripley (2022). This formulation in principle allows one to evolve the complete scalar-field and metric dynamics in scalar-Gauss-Bonnet gravity. However, the first binary BH coalescence simulations using this formalism in scalar-Gauss-Bonnet theories that allow for BH scalarization (East and Ripley, 2021a,b) demonstrated that evolution ceases to be well posed for larger values of the coupling. How different gauge choices and different models perform in this respect has not been well explored.

Franchini et al. (2022), motivated to go beyond the small-coupling approximation, explored the fixing-the-equations approach to perform numerical simulations, an approach inspired by work in dissipative relativistic hydrodynamics (Cayuso et al., 2017). The main idea is to modify the evolution equations of the theory such that the short wavelength modes (associated with strong coupling and the breakdown of well-posedness) with respect to some chosen scale are somewhat controlled, while longer wavelength modes become tractable. Franchini et al. (2022) were able to use the formalism to study scalar-field collapse in spherical symmetry in the scalarization model of Silva et al. (2019). This allowed them to evolve the system past situations where the original equations fail due to the appearance of regions in spacetime where the evolution changed character from hyperbolic to elliptic, which prevented evolution of the system in time (Ripley and Pretorius, 2020). Moreover, the “fixed-equation” system evolves toward the static BHs predicted by the original system of equations, thus indicating that it may provide a suitable “completion” of the original theory. Whether this approach remains valid in a situation with less symmetry, for example the case of BH binaries, remains to be explored.

Progress has also been made in modeling compact binary dynamics in the regime of wider separation and lower velocities, i.e., in the domain of validity of PN theory. In particular, Shiralilou et al. (2021, 2022) extended to a gen-
Figure 22 Dynamical descalarization in binary BH head-on collisions. Scalar-field and Gauss-Bonnet dynamics on the x-y plane of the BHs with initial mass ratio $q = m_1/m_2 = 1/2$. We show the scalar-field amplitude $\log_{10} |\Phi|$ (color map) together with the isocurvature levels of the Gauss-Bonnet invariant at the beginning of the evolution (top-left panel), during the BH approach (top-right panel), shortly before the collision (bottom-left panel) and shortly after the merger (bottom-right panel). These levels correspond to $1 M^{-4}$ (solid line), $10^{-1} M^{-4}$ (dashed line), $10^{-2} M^{-4}$ (dot-dashed line) and $10^{-3} M^{-4}$ (dotted line). From Silva et al. (2021b).

Figure 23 Snapshots of the scalar field, here labeled $\Phi$, and the Gauss-Bonnet invariant $\mathcal{G}$ in the x-z plane illustrating the stealth scalarization. The color map indicates the amplitude of the scalar field. The curves represent isocurvature contours for the following values of $|\mathcal{G}|$: $1 M^{-4}$ (solid line), $10^{-1} M^{-4}$ (dashed line), $10^{-2} M^{-4}$ (dot-dashed line), and $10^{-3} M^{-4}$ (dotted line). Black (red) lines correspond to positive (negative) values of $\mathcal{G}$. We show the inspiral (top-left panel), half an orbit before merger (top-right panel), formation of the first common apparent horizon (bottom-left panel), and at 200M after the merger. From Elley et al. (2022).

general $f(\varphi)$ coupling the calculations of Yagi et al. (2012), which applied for the shift-symmetric scalar-Gauss-Bonnet theory, i.e., $f(\varphi) \sim \varphi$. These works calculated the leading-order corrections due to curvature nonlinearities in the GW and scalar waveforms, finding that corrections due to the Gauss-Bonnet term appear at 1PN order in GWs. They also obtained the GW polarization and phasing. In addition, Julié and Berti (2019) argued that during the adiabatic inspiral of two BHs the Wald entropy of each BH is constant. This allows a precise definition of the sensitivities of BHs in scalar-Gauss-Bonnet gravity in terms of the variation of the ADM mass as a function of the ambient scalar field that the BH is embedded in at fixed Wald entropy. This parallels the notion that the baryonic mass of NSs is constant during the inspiral (Damour and Esposito-Farèse, 1996a). They also derived the two-body Lagrangian at 1PN order. Julié et al. (2022) used this result in conjunction with the numerical calculation of sensitivities of spontaneously scalarized BHs to study binaries. They found that in principle BHs can evolve toward a situation in which the inner singularity approaches the event horizon of the BH before the merger. This suggest the possibility that, depending on the model of scalar-Gauss-Bonnet gravity and the binary parameters, the field equations might simply not be able to predict a full binary evolution ranging from inspiral to merger.

B. Black-hole scalarization in the presence of matter

Stefanov et al. (2008) suggested the earliest model of BH scalarization, which consisted of the DEF model coupled to nonlinear electromagnetism (Stefanov et al., 2007a,b). They introduced a coupling between the scalar field and the Born-Infeld Lagrangian and found that BH solutions with scalar hair branch off from the general-relativistic sequence of solutions. These BHs violate the no-scalar-hair theorems of Mayo and Bekenstein (1996) and Sen and Banerjee (2001), which are valid for charged, nonrotating BHs in scalar-tensor theory. The bifurcation point occurs where the GR sequence of solutions becomes unstable to scalar perturbations (Doneva et al., 2010).

A variety of scalarization models have been studied in Einstein-Maxwell-scalar theory following Herdeiro et al. (2018), who introduced a coupling between the scalar...
field and the Lorentz invariant $F^2 = F_{\mu\nu} F^{\mu\nu}$ of the form $\exp(-\alpha \varphi^2) F^2$, where $\alpha$ is a dimensionless constant. In this model BHs can scalarize, with $F^2$ playing the role of the trace of the energy-momentum tensor in the NS scalarization in DEF-like models, or the Gauss-Bonnet invariant in BH scalarization models.

The simpler nature of the model allowed Herdeiro et al. (2018) to perform numerical relativity simulations to study the scalarization process in the time domain by expanding upon already available code bases (Sanchis-Gual et al., 2016a,b). This allowed Herdeiro et al. (2018) to show that the end point of the instability of the Reissner-Nordström BH was indeed a BH with scalar hair. Additional studies of this model were done by Astefanesei et al. (2019); Fernandes et al. (2019b); Herdeiro et al. (2021a); Luo et al. (2022); Niu et al. (2022); and Xiong et al. (2022). Within this model, Myung and Zou (2019b) studied the instability of the Reissner-Nordström solution, while Blázquez-Salcedo et al. (2021); Myung and Zou (2018, 2019c) analyzed the stability of the scalarized solutions. Fernandes et al. (2019a) and Herdeiro and Oliveira (2020) explored other variations on this theme, including scalar-axion couplings, i.e., $\propto h(\varphi) F_{\mu\nu} F_{\mu\nu}$, where $F_{\mu\nu} = (1/2) \epsilon_{\mu\nu} \alpha^2 F_{\alpha\beta}$. In addition, Ikeda et al. (2019) considered the coupling between the double dual of the Riemann tensor and $F_{\mu\nu}$.

Zhang (2021) studied an Einstein-Maxwell-scalar model, with a coupling between scalar and Maxwell fields of the form $\exp(\beta \varphi^4) F^2$. In this model, the Reissner-Nordström solution is linear stable against scalar perturbations, but it can become unstable under large, nonlinear perturbations. Zhang (2021) observed a critical phenomena separating unsualarized and scalarized BHs. This establishes an interesting connection between nonlinear scalarization with the result of critical phenomena in gravitational collapse, which was first observed by Choptuik (1993). Similar investigations were carried out in Einstein-Maxwell-scalar the Reissner-Nordström-anti-de Sitter BHs (Zhang et al., 2022) and in the Ricci-Gauss-Bonnet model (Liu et al., 2023a,b).

Cardoso et al. (2013a,b) explored a relevant scenario from an astrophysical perspective: matter in the vicinity of a BH could trigger scalarization and in turn endow the BH with scalar hair. They showed that this “matter-induced” scalarization (see Sec. II.C.3) is in principle possible for the DEF model. The viability of this proposal in realistic matter configurations, such as an accretion disk or a dark matter halo, has not yet been explored. Matter-induced scalarization seems unlikely to happen in light of the severe constraints on the DEF model. However, this might not be the case for other scalarization models, and the topic deserves further investigation.

After it was understood that in scalar-Gauss-Bonnet theories the Schwarzschild solution can scalarize, it was natural to ask if the same can happen to its charged counterpart, i.e., the Reissner-Nordström solution. Doneva et al. (2018a) analyzed this and found the existence of two bifurcation points, one at larger masses where the scalarized solutions bifurcated from the Reissner-Nordström one, and one at smaller masses where the scalar charge of the solution decreases again to zero and the branch merges again with the GR one. Scalarized charged BHs in the fundamental mode are also thermodynamically favored over the Reissner-Nordström solution. Brihaye and HARMANN (2019) showed that for the near-extremal Reissner-Nordström case, scalarization can happen for either sign of the scalar-to-Gauss-Bonnet coupling constant ($\lambda_5$). Other works vary the choice of the coupling function $f$ between the scalar field (or the axion field) and the Maxwell invariant and studying the scalarization of dyonic BHs include Fernandes et al. (2019b), Fernandes et al. (2019a), Blázquez-Salcedo et al. (2020c) and Blázquez-Salcedo et al. (2021).

Herdeiro and Radu (2019) put forward their motivation for studying these models. They noted that quantum effects can break the scale invariance and vanishing energy-momentum trace properties of electro-vacuum GR BHs. As a consequence, even the simplest nonminimally scalar-field coupling $\xi \varphi^2 R$ can result in BH scalarization when these quantum effects are taken into account. Within these models, they discussed the scalarization of the Reissner-Nördstrom solution and a noncommutative generalization of the Schwarzschild solution. They found that the resulting scalarized BHs are in general entropically favored over the GR solutions.

### C. Variations of the curvature-induced scalarization model

In light of the nontrivial effect of rotation on scalarization, it is natural to consider what happens when the Gauss-Bonnet invariant is replaced by the Pontryagin invariant $\ast RR$ as the “curvature source” to which the scalar field is coupled. The Pontryagin density is known to vanish in spherically symmetric spacetimes (such as that of a Schwarzschild BH) and becomes nonzero in nonspherically symmetric spacetimes (such as that of a Kerr BH). In this sense, all scalarized BHs in theories that replace the Gauss-Bonnet invariant with the Pontryagin density are necessarily “spin induced”; a characteristic shared with BH solutions in dynamical Chern-Simons gravity [see e.g., Komno et al. (2009) and Yunes and Pretorius (2009), and Alexander and Yunes (2009) for a review] which features a linear coupling between the scalar field and the Pontryagin invariant. The coupling between a scalar field and the Pontryagin density leads to equations that are of third order in derivatives [see Delsate et al. (2015) and Motohashi and Suyama (2012)], which is not the case for a coupling with the Gauss-Bonnet invariant. Thus, although both of these couplings can be seen as part of an EFT, they lead to distinct challenges when it comes to nonlinear evolution. Myung and Zou (2019a) studied the combined effect of Gauss-Bonnet and Pontryagin densities. In the
test field limit, the scalar-field dynamics with an effective mass proportional to $\phi^2 \cdot RR$ was also studied in a Kerr background (Doneva and Yazadjiev, 2021b; Gao et al., 2019) and in the Schwarzschild-Newman-Unti-Tamburino background (Brihaye et al., 2019).

Another extension of the original Gauss-Bonnet spontaneous scalarization model involves the consideration of $n > 1$ scalar fields $\varphi^a$, whose interaction is determined by their “target space” $\gamma_{ab}(\varphi^2)$, an $n$-dimensional Riemannian manifold (Damour and Esposito-Farèse, 1992; Horbatsch et al., 2015). These are similar to the tensor-multiscalar DEF models for NS scalarization presented in Sec. III.D. In these models, the scalar-field dynamics is determined by the quantity $\gamma_{ab}(\varphi^2)$. Doneva et al. (2020c) numerically obtained a scalarized BH for the case $n = 3$ and maximally symmetric target-space geometries. An important feature of these solutions is that the scalar fields decay asymptotically as $1/r^2$ (i.e., the BHs do not have a monopole scalar charge). This shows that BHs in these theories will not emit dipole scalar radiation when they are placed in binaries.

V. GENERALIZATIONS OF SCALARIZATION TO OTHER FIELDS AND INSTABILITIES

We identified the underlying reason for spontaneous scalarization in its various forms to be a tachyonic instability. However, the scalar nature of the field did not play a special role in the mechanism. This suggests that other fields, such as vectors, might also spontaneously develop nontrivial configurations around compact objects when they exhibit suitable couplings to curvature. In this section, we investigate this idea, which leads to the so-called spontaneous tensorization theories.

A second type of generalization of spontaneous scalarization considers new types of instabilities, as opposed to new types of fields. For instance, instead of replacing the scalar field with, say, a vector, we replace the tachyonic instability with, say, a ghost instability. We later see that a surprising key result of spontaneous tensorization is that these two paths are intimately connected. Namely, even if we intend only to have a theory of tachyons living on nonscalar fields, ghost and gradient instabilities necessarily arise in almost all models.

A. Spontaneous vectorization

What happens if we replace the scalar field of spontaneous scalarization with a vector? In scalarization, we started with the most general context of Sec. II.B.1, considering all the allowed couplings to the metric. Here we follow the reverse path, starting with individual models of vector-tensor theories and considering the more encompassing theory later. This exposition is preferable because the study of specific models that are straightforward generalizations of known scalarization models reveals some pathologies and provides guidance for further model building.

The idea and the first specific model of vectorization were introduced in analogy with a massive version of the DEF model in Eq. (18). Consider the vector-tensor theory action (Beltran Jimenez et al., 2013; Ramazanoğlu, 2017)

$$S = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - F_{\mu\nu}F^{\mu\nu} - 2\mu_2^2 X^2 \right] + S_m[\Psi_m; g_{\mu\nu}],$$

where $F_{\mu\nu} = \nabla_{\mu}X_{\nu} - \nabla_{\nu}X_{\mu}$ and $X^2 = g^{\mu\nu}X_{\mu}X_{\nu} = X_{\mu}X^{\mu}$. The vector field $X_{\mu}$ has the canonical kinetic term in this frame, and the matter degrees of freedom couple to the metric $g_{\mu\nu} = A^2(X^2)g_{\mu\nu}$, which is conformally scaled with respect to $g_{\mu\nu}$. The vector field equation is

$$\nabla_{\mu}F^{\mu\nu} = [-8\pi G_s A^4 \Omega - \mu_2^2 X^2] \Omega,$$

where $\Omega = d \ln A/dx$, $T_{\mu\nu}$ is the energy-momentum tensor in the Einstein frame, and $\Omega_T$ is the trace of the stress-energy tensor in the Jordan frame, i.e., with respect to the metric $g_{\mu\nu}$.

Equation (37) is that of a massive vector (Proca) field where the expression inside the square brackets acts as an effective mass $\mu_2^2$. In parallel with the DEF model, when $A$ has an appropriate form, for instance, $A = \exp(\beta_0 X^2/2)$ with sufficiently negative $\beta_0$, $\mu_2^2$ becomes negative. Furthermore, for dense enough objects such as NSs this occurs for order-of-unity values of $\beta_0$. The expectation is that the vector field will grow from arbitrarily small perturbations around $X_{\mu} = 0$ due to this tachyonic behavior, which can be called spontaneous vectorization, in exact analogy with spontaneous scalarization. However, we later see that there are many subtle points in this narrative.

Even though we presented spontaneous vectorization in the Einstein frame with a nonminimal coupling to matter, it can be converted to a theory of minimal matter coupling and vector-curvature couplings as in spontaneous scalarization (Ramazanoğlu, 2019a). Even more directly, there are theories of spontaneous vectorization that are purely conceived through curvature coupling, with the first example being the Hellings-Nordtvedt theory (Hellings and Nordtvedt, 1973) studied by Annuli et al. (2019)

$$S = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - F_{\mu\nu}F^{\mu\nu} + \Omega X^2 \right] R_{\mu\nu} + S_m[\Psi_m; g_{\mu\nu}],$$

where $\Omega$ and $\eta$ are coupling constants. Scalarization through coupling to the Gauss-Bonnet term also has a vector analog as in the action (Ramazanoğlu, 2019b)

$$S = \frac{1}{16\pi G} \int d^4 x \sqrt{-g} \left[ R - F_{\mu\nu}F^{\mu\nu} + f(X^2)g \right] + S_m[\Psi_m; g_{\mu\nu}],$$
which has a similar structure as Eq. (25). The action (39) leads to BH vectorization as well, unlike the other models we have seen thus far.

The models of vectorization that we have examined are merely specific examples, and any theory of scalarization can be generalized to vector fields in principle. In this regard, Brihaye and Verbin (2020) studied the coupling of a vector field to a scalar, and Oliveira and Pombo (2021) and Brihaye et al. (2022) studied the coupling between two vector fields to obtain vectorized compact objects. Kase et al. (2020) considered generalized Proca theories (Heisenberg, 2014; Tasinato, 2014) with various couplings of a massive vector field and showed that spontaneous vectorization occurred in those models as well. Ramazanoğlu (2018b) investigated models where the effective mass of the vector field was generated by the Higgs mechanism, thus preserving the gauge symmetry. On a separate front, conformal scaling of the metric in the matter action (36) can be generalized to difformal transformations, and spontaneous vectorization still occurs (Minamitsuji, 2021; Ramazanoğlu and Ünlütück, 2019). In terms of approximate solutions, Garcia-Saenz et al. (2022) recently calculated the Schwarzschild QNMs of nonminimally coupled vector fields. There have also been efforts to study all these phenomena in a unified manner in compact binaries, using more generic EFT tools (Khalil et al., 2019).

To summarize, coupling of the vector field to any nonvanishing term in the action can be considered for spontaneous vectorization (Ramazanoğlu, 2019b), and most options have been considered in at least a preliminary sense. The case of all possible couplings, an analog of the minimal action of scalarization in Sec. II.B.1, was studied by Garcia-Saenz et al. (2021). They also revealed some with the fundamental problems of vectorization, as we later discuss more broadly.

Vectorization models strongly resemble scalarization models in terms of their actions, but this can be misleading. One major difference is in the number of additional degrees of freedom. Scalar fields contribute a single new degree of freedom irrespective of whether they are part of spontaneous scalarization provided that the field equations are of second order in derivatives. For vectorization, this is not the case. The action (37) and all of the vector-tensor theory actions that we have discussed break the gauge freedom found in massless vector fields. As a result, the vectors of vectorization models carry 3 degrees of freedom instead of the 2 found in electromagnetism. This is not an immediate reason of concern, since minimally coupled massive vectors, known as Proca fields (Proca, 1936), also break the gauge symmetry, and they still provide a classical field theory with well-behaved dynamics. However, as we later see, the extra degree of freedom is far more problematic in vectorization.

Although the vectorization models are designed to have a linear tachyonic instability in analogy with scalarization, the presence of the third degree of freedom, whose dynamics is not obvious, casts doubt on the success of this design. Namely, whether vectorization can indeed be described as a tachyonic instability of the vector modes that is then quenched nonlinearly is not apparent. This concern is amplified by the fact that vectorized compact objects in the models studied thus far seem to have some striking differences with respect to their scalarized counterparts.

As an example, consider vectorized BHs in the theory described by the action (39). Entropy can be a measure of stability for BHs, where higher entropy solutions are favored in terms of stability. Entropies of spherically symmetric vectorized BHs were numerically computed by Barton et al. (2021), and GR BHs were shown to be entropically favored over vectorized ones in all cases. The verdict of stability is not definite without time evolution, but this is a strong indication for the instability of vectorized BHs. In contrast, scalarized BHs in scalar-Gauss-Bonnet theories, such as Eq. (25) can be entropically favored over those in GR for appropriate coupling functions (Doneva and Yazadjiev, 2018b), and their stability is also indicated through other methods (see Sec. IV.A.2), which makes the stability of vectorized objects even more suspect.

 Scalarized and vectorized NSs also show major differences. For example, the dependence of the strength of vectorization on the NS mass has no discernible pattern, as can be seen in Fig. 24, whereas scalarization occurs in a finite NS mass interval and vanishes at the boundary mass values; see Fig. 1. It is also known that while scalarized NSs are continuously connected to those of GR as the theory parameters are smoothly changed (for example, in the limit \( \beta \to 0 \) in Fig. 1), this is not the case in vectorization. The latter is known to be an indicator of instability in certain scalarization theories; see Mendes and Ortiz (2016). Thus, qualitative differences between scalarized and vectorized objects are also apparent for NS and, moreover, point to the instability of the vectorized ones.

These observations suggest that the vectorized compact object solutions obtained thus far may not be stable. Hence, they are not astrophysically relevant, at least in

---

14 Even though we have an intrinsic mass \( \mu_X \) for the vector field in our discussion, vectorization can occur without this term, as is the case for scalarization.

15 There is no analog of the scalar charge in Eq. (31) for massive scalarization or vectorization theories due to different asymptotic behavior at \( r \to \infty \). However, \( A - 1 \) is a monotonic function of the norm of the vector field in this case and is known to be a good indicator of the deviation of vectorized star structures from those of GR (Ramazanoğlu, 2017). Thus, we use it as a measure of vectorization strength.
Figure 24: The strength of vectorization measured by $A(r = 0) - 1$ as a function of $M$ for various values of $\mu X$ in Eq. (37). Top row: $\mu X = 1.6 \times 10^{-11}$ eV. Middle row: $\mu X = 8.0 \times 10^{-12}$ eV. Bottom row: $\mu X = 4.8 \times 10^{-12}$ eV. Left column: $\beta_0 = -4$. Middle column: $\beta_0 = -5$. Right column: $\beta_0 = -6$. The dashed and dot-dashed lines, respectively, indicate solutions where the energy density $\epsilon$ and $\Phi = -n^\mu X_{\mu}$ (with $n^\mu$ the normal vector field to spatial hypersurfaces) do not monotonically decrease with radius. From Ramazanoğlu (2017).

the case of spherically symmetric spacetimes. A close look at the instability mechanism of vectorization reveals the underlying reasons for this in Sec. V.B.

B. Vectorization and ghosts

Despite the fact that the vectorization phenomena that we have presented are directly inspired by scalarization, we have seen noteworthy differences between the two mechanisms. Even though it was puzzling for a time, this is now strongly suspected to be related to the fact that vectorization models suffer from ghost and/or gradient instabilities in addition to tachyonic ones, as demonstrated by Garcia-Saenz et al. (2021) and Silva et al. (2022). Here we first follow the latter and later outline the methodology of the former.\(^{16}\)

We begin by reviewing the basic aspects of different types of instabilities (Demirboğa et al., 2022). Recalling our discussion of the tachyon in Secs. II.A.1 and II.A.2, we consider the linearized scalar-field equation (5) in 1+1 dimensions,

\[ g^{tt} \partial_t^2 \delta \varphi + g^{xx} \partial_x^2 \delta \varphi = \mu^2 \varphi, \quad (40) \]

where we assume for simplicity that the metric is diagonal with constant components. The common case without instabilities is when $g^{tt} < 0$, $g^{xx} > 0$, and $\mu^2 > 0$. What happens when each of these three terms changes its sign while the other two stay the same?

For a plane wave mode $\delta \varphi(t, x) \propto \exp[i(\omega t - k x)]$, the dispersion relation is given by

\[ \omega(k) = \sqrt{(\mu^2 + g^{xx} k^2)/(-g^{tt})}. \quad (41) \]

We saw that for $\mu^2 < 0$ the mode behaves as a tachyon, where $\omega(k)$ becomes imaginary for sufficiently small $|k|$. This leads to exponential growth as discussed, and the fastest growing mode behaves as $\sim \exp[\sqrt{\mu^2/g^{tt}} t]$; i.e., the growth rate has an upper limit.

A ghost occurs when $g^{tt} > 0$. There is also exponential growth in this case: however, the growth rate diverges with increasing wave number as $\sim \exp[\sqrt{g^{xx} k^2}]$ and hence has no upper bound. The case $g^{xx} < 0$ also gives the same asymptotic growth rate as the ghost and is called a gradient instability.\(^{17}\) In summary, ghost and gradient instabilities as we presented them here have qualitative differences from tachyons, a with main one being the arbitrarily fast growth rates. This means that a well-defined time evolution can be problematic in such theories, and it is more difficult to devise nonlinear quenching mechanisms.

The spontaneous vectorization model of Eq. (36) modifies the “mass term” in the field equation for the vector field. Hence, the first impression might be that it leads to a tachyonic instability as in scalarization. However, this approach is naive, and a thorough analysis reveals that theories of vectorization necessarily carry ghost instabilities in addition to tachyonic ones. Following Silva et al. (2022), one way to understand this is to use the Stueckelberg trick (Ruegg and Ruiz-Altaba, 2004) where we introduce a scalar field $\psi$ into the action (36) using the substitution

\[ X_\alpha \rightarrow X_\alpha + \mu_X^{-1} \nabla_\alpha \psi. \quad (42) \]

This leads to the scalar equation of motion

\[ \Box \psi = -g^{\mu
u} \left[ \mu_X \nabla_\mu X_\nu + \frac{1}{2} (\nabla_\mu \log \xi) (\nabla_\nu \psi + \mu_X X_\nu) \right], \quad (43) \]

\(^{16}\) Esposito-Farèse et al. (2010) is an earlier similar study that specifically investigated cosmological scenarios.

\(^{17}\) For the simple equation (40), a simultaneous change of sign of two of the coefficients is equivalent to changing the sign of the third one; i.e., $g^{tt} > 0$, $g^{xx} < 0$, and $\mu^2 > 0$ is also a tachyon. Similarly, all three coefficients changing signs gives a stable theory. This is due to the fact that the equation as a whole or the action that leads to it can be multiplied by an overall factor of $-1$ without changing the physics. However, the scalar, or any investigated field that carries instabilities, is always coupled to the metric and other terms in the action, which means that the overall sign is meaningful. Hence, we use the aforementioned classification here and say that a theory with $g^{tt} > 0$, $g^{xx} < 0$, and $\mu^2 > 0$ carries both ghost and gradient instabilities according to the literature (Demirboğa et al., 2022).
where the covariant derivative $\nabla$ is compatible with the new metric

$$\bar{g}^{\mu\nu} = \hat{z} g^{\mu\nu}, \quad \hat{z} \equiv \frac{\mu^2_{\text{eff}}}{\mu^2} = 1 - \frac{8\pi G_\ast A^4 \lambda T}{\mu^2}. \quad (44)$$

The signature of $\bar{g}_{\alpha\beta}$ changes with the sign of $\hat{z}$, which controls the change of sign of the effective mass of $X_\mu$ in Eq. (37). In other words, when we try to instigate a tachyon on $X_\mu$ by having negative $\hat{z}$ in parts of the spacetime, $\psi$ carries ghost and gradient instabilities; cf. Eq. (40). This is a degree of freedom whose dynamics is governed by the effective metric $\bar{g}_{\mu\nu}$, not the spacetime metric $g_{\mu\nu}$.

It is curious to see that the Stueckelberg trick reveals the ghost, even though it is not apparent in the vector field equation (37). However, a more careful analysis can clarify this issue (Silva et al., 2022). Note first that the vector field equation (37) is not manifestly hyperbolic; i.e., its principal part, the term with the highest derivatives, is not the wave operator, since

$$\nabla_\mu F^{\mu\nu} = \nabla_\mu \nabla^{\mu} X^\nu - \nabla_\mu \nabla^\nu X^\mu. \quad (45)$$

The second term on the right-hand side of Eq. (37) vanishes in Proca theory ($A = 1$) making the principal part the wave operator. This is not the case in general due to the so-called (generalized) Lorentz condition

$$\nabla_\mu (\hat{z} X^\mu) = 0, \quad (46)$$

which is obtained by acting on both sides of Eq. (37) with $\nabla_\mu$ and recalling the antisymmetry of $F^{\mu\nu}$. We can use Eq. (46) to obtain a manifestly hyperbolic form of the linearized field equations on a fixed background of vanishing vector fields as

$$\nabla_\mu \nabla^\nu X_\nu + (\nabla_\mu \ln \hat{z}) \nabla_\nu X^\mu = M_{\nu\mu} X^\mu, \quad (47)$$

where we defined the mass-squared tensor

$$M_{\nu\mu} = \hat{z} \mu^2 X g_{\nu\mu} + R_{\nu\mu} - \nabla_\nu \nabla_\mu \ln |\hat{z}|. \quad (48)$$

In Eq. (47) and (48) the metric and $\hat{z}$ have their fixed background GR values. This is a generalized massive wave equation.

Since $\hat{z} = 1$ in the absence of matter and it becomes negative in a continuous manner inside matter for an instability to exist at all, $\hat{z}$ necessarily vanishes at some points of a NS spacetime. The $\nabla_\nu \nabla_\mu \ln \hat{z}$ term introduced by the vectorization contains powers of $\hat{z}^{-1}$, which means that the effective mass-squared tensor diverges at such points. Alternatively, if we move the $\hat{z}^{-1}$ terms to the other side of the equation, the principle part becomes $\Box X_\mu$; i.e., it changes its sign with $\hat{z}$ the same way as the $\psi$ field in Eq. (43). Hence, if we properly rewrite the vector field equation into a manifestly hyperbolic form, the ghost is apparent.

Equations (47) and (48) show that calling $\mu^2_{\text{eff}} = \hat{z} \mu^2_X$ the effective mass squared of the theory is misleading, unlike in the case of scalarization. Even though $\mu^2_{\text{eff}}$ replaces the mass term in the vector field equation of a minimally coupled Proca field, it is not actually the effective mass of physically propagating degrees of freedom. The physical mass is determined by Eq. (48), which diverges at points where $\hat{z} = 0$.

We have presented the appearance of the ghost in two separate ways, the Stueckelberg mechanism of Eq. (43) and the direct vector field formulation of Eq. (47). The degrees of freedom in the two pictures are related to each other in a nontrivial manner through Eq. (42). Hence the presence of the ghost manifests differently. Nevertheless, the criterion for the appearance of the ghost $\hat{z} = 0$ is the same in both cases. The Stueckelberg picture has an advantage in that it can straightforwardly identify the ghost as a scalar degree of freedom $\psi$ at decoupling, which is well known in the case of the Proca field.

One might at first think that the ghost and gradient instabilities can be regularized by nonlinearities. After all, the tachyon of scalarization is also an instability, but it is eventually quenched by nonlinear effects as explained in Sec. II.A. In such a scenario, one would expect the vectorized objects to be free of instabilities even though the GR solutions are unstable. However, current evidence points away from this. Even though detailed mathematical analyses of the partial differential equations have not been performed to ensure nonlinear instability, Demirboğa et al. (2022) showed that spherically symmetric vectorized NSs in the theory (36), which could be potential stable endpoints of vectorization, are also unstable to ghosts and gradient instabilities at the perturbative level.

Even if stable solutions can be found in other vectorization models, there is a more fundamental problem due to the modification to the principal part of the vector equations in equation (43) or (47). The coefficient of the wave operator vanishes at certain points in spacetime or, equivalently, the physical effective mass diverges if we move these coefficients to the other side of the equation. In either case, this leads to divergent time derivatives for arbitrarily small perturbative vector field values, which renders the initial assumptions of a perturbative approach invalid. Overall, it is not clear whether one can even define a well-posed time evolution, let alone one that can somehow suppress the exponential growth of the vector field (Silva et al., 2022). Further elucidation of these issues requires a more rigorous mathematical investigation, which has not yet been attempted.

Thus far we have investigated the ghost and gradient instabilities of the specific vectorization theory in the action (36), but similar results are known to exist for all theories of vectorization (Garcia-Saenz et al., 2021; Silva et al., 2022) due to similar mechanisms. To summarize, replacing the scalar of scalarization with a vector field results in theories where the central issue is not the nature
of the field the instability lives on, but rather the type of instabilities that arise. We cannot construct vectorization models that exhibit only the tachyonic instability commonly associated with the scalarization mechanism, which can be benign and tamed by nonlinear effects. Rather, vectorization exhibits far more threatening instabilities.

We have closely followed Silva et al. (2022), but Garcia-Saenz et al. (2021) reached similar conclusions via an alternative approach. They considered the most general action that contained a vector field and the metric (Heisenberg, 2014; Tasinato, 2014) and truncated its action at the quadratic order around GR to obtain the linearized field equations. The main qualitative difference from the scalar minimal action of Sec. II.B.1 is that one can demonstrate that if any of the physical degrees of freedom have a tachyonic instability, then there are also degrees of freedom with ghost or gradient instabilities, which are manifested by a change of sign of coefficients in the dispersion relationship. The conclusion is the same: vectorization based on tachyonic instabilities necessarily has ghost or gradient instabilities as well.

Recent work has shown that the issues with vectorization also happen in simpler theories, with an example being minimally coupled self-interacting vector field theories. For instance, for $\mathcal{A}(x) = 1$ the action (36) becomes the minimally coupled Proca theory, which is well posed. However, generalizations of the potential beyond a mass term, that is, $2\mu^2 x \to 4V(x)$ for some generic function $V$, results in ill-posedness (Clough et al., 2022; Coates and Ramazano˘glu, 2022; Mou and Zhang, 2022). The underlying reason for this is the fact that there is an analog of the generalized Lorenz condition (46) in this case as well, which ultimately leads to the same result: there is a degree of freedom in the theory whose dynamics is governed by an effective metric, an analog of $\tilde{g}_{\mu\nu}$ in Eq. (44). This new metric depends on the vector field itself and can lose its Lorentzian nature depending on how $X_\mu$ evolves in time (Coates and Ramazano˘glu, 2022).

Esposito-Far`ese et al. (2010) provided some of the earliest examples of the ill-posedness of self-interacting vector field theories that we have mentioned in a cosmological context, and it was recently demonstrated that initially healthy configurations of self-interacting vectors naturally evolve to a point where hyperbolic evolution becomes impossible (Clough et al., 2022; Coates and Ramazano˘glu, 2022, 2023; Mou and Zhang, 2022). This means that nonlinear extensions of the Proca theory are more delicate than their scalar counterparts in terms of providing physical theories. It seems that the underlying reason for the “failure” of vectorization is a much more general phenomenon in theories with conditions such as Eq. (46), at least for the models conceived thus far. In this sense, exploration of vectorization has been leading to a deeper understanding of vector fields in general, which has implications beyond theories of gravitation.

We close this section with a pertinent comment on ghost instabilities. We encountered them as an unappealing artifact in vectorization theories, but one could consider making them the driver of a spontaneous growth mechanism. This option was indeed considered prior to their discovery in vectorization models. The simplest example is another analog of the DEF model where the conformal scaling function $\mathcal{A}$ depends on derivative terms (Ramazano˘glu, 2018a)

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[ R - 2g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - 2\mu^2 \phi^2 \right] + S_m[\Psi; \mathcal{A}^2(K)g_{\mu\nu}],$$

(49)

with $K = -g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi/2$. The scalar equation of motion is

$$\nabla_\mu (\hat{\varepsilon} \nabla^\mu \phi) = \mu^2 \phi,$$

(50)

where $\hat{\varepsilon} = 1 + 4\pi G \mathcal{A}^4 (d \ln \mathcal{A}/dK)$. Unlike for the tachyonic instability where the mass term is modified, the derivative terms are modified in this case.

The nature of Eq. (50) can be better understood when we consider the linearized equation for perturbative values of the scalar around a GR background and highlight the principal part as

$$\hat{\varepsilon} \Box \delta \phi + \cdots = 0.$$  

(51)

For an appropriate choice such as $\mathcal{A} = \exp(\beta_0 K)$ with large enough $\beta_0 > 0$, the sign of $\hat{\varepsilon}$ can become negative in the presence of matter. In all spontaneous scalarization theories considered thus far, the source of the instability was the change of sign of the effective mass term, which resulted in a tachyon. In Eq. (51) the overall sign of the wave operator changes; hence, we have both ghost and gradient instabilities, recalling Eq. (40) and the subsequent discussion. Spontaneous scalarization that might arise from this new mechanism is called ghost-based spontaneous scalarization (Ramazano˘glu, 2018a). It has similar problems as vectorization in that it is hard to tame the ghost and other instabilities; hence, we do not end up with a theory with sensible dynamics.

C. Spontaneous tensorization

The basic mechanism of spontaneous scalarization can be generalized beyond vector fields, resulting in spontaneous tensorization. However, all such known theories also feature ghosts and suffer from similar problems with vectorization, aside from a single possible example of spinor-tensor theory.

The simplest mathematical generalization of scalarization beyond vectors occurs for $p$-form fields, i.e., totally antisymmetric tensors $X_{\mu_1 \ldots \mu_p}$ of rank $(0, p)$. This is due to the fact that a vector field $X_\mu$ is a 1-form field, and
the actions that we encountered in spontaneous vectorization can be readily generalized to this class of fields as (Ramazanoğlu, 2019c)

\[
S = \frac{1}{16\pi} \int d^4x \sqrt{-g} \left[ R - F_{\mu_1 \ldots \mu_p+1} F^{\mu_1 \ldots \mu_p+1} \right] + S_m \left[ \Psi_m; A^2(X^2)g_{\mu\nu} \right],
\]

(52)

where \( F_{\mu_1 \ldots \mu_p+1} = (p + 1)\nabla_{[\mu_1} X_{\mu_2 \ldots \mu_p+1]} \), with the square brackets denoting antisymmetrization and \( X^2 = X_{\mu_1 \ldots \mu_p} X^{\mu_1 \ldots \mu_p} \). It is then straightforward to show that the analog of \( \mu^2_{\text{eff}} \) becomes negative inside NSs for an appropriate choice of \( \Lambda(x) \). However, these theories carry ghost instabilities similar to those of vectors when we try to establish spontaneous growth from tachyonic instabilities since they possess an analog of the Lorenz condition (46) (Silva et al., 2022).

Spontaneous tensorization of a symmetric rank \((0,2)\) tensor field \( f_{\mu\nu} \), which is spin 2, is a natural avenue to investigate after spin-0 scalars and spin-1 vectors. Ghosts appear in this case once more since the metric in our gravity theories \( g_{\mu\nu} \) is also a spin-2 field. Note that spontaneous tensorization theories include terms where the field that tensorizes, in this case \( f_{\mu\nu} \), is coupled to the metric \( g_{\mu\nu} \). However, interacting spin-2 fields are known to generically lead to ghost degrees of freedom, rendering most such theories unphysical (Boulware and Deser, 1972; de Rham, 2014) and have led to ghost-free theories being discovered more recently in the form of massive gravity and bigravity (Hassan and Rosen, 2012; de Rham, 2014; de Rham et al., 2011). Matter can couple to one of the spin-2 fields in the novel ghost-free theories, but coupling to both metrics generically invokes a ghost again (Yamashita et al., 2014). For example, trying to mimic the DEF model by having a conformal scaling function that depends on \( f_{\mu\nu} \), i.e., changing the matter action as \( g_{\mu\nu} \rightarrow A(f_{\mu\nu}, g_{\mu\nu}) g_{\mu\nu} \), leads to ghosts. Overall, there is no known form of bigravity theory that features an analog of spontaneous scalarization.

Spinor fields present the most interesting case of generalizing scalarization, perhaps aside from vectors. The classical Lagrangian for a Dirac spinor is

\[
\mathcal{L}_\psi = \frac{1}{2} \left[ \bar{\psi} \gamma^\mu \left( \nabla_\mu \psi \right) - \nabla_\mu \bar{\psi} \gamma^\mu \psi \right] - \mu \bar{\psi} \psi,
\]

(53)

where the conventions for gamma matrices and the effect of covariant derivatives on spinors are as given by Ramazanoğlu (2018c). This action provides the usual dispersion relation \( \omega^2 = k^2 - \mu^2 \) for a plane wave of the form \( \psi \propto \exp[i(\omega t - k^0 x^0)] \). Since the mass term appears linearly, not quadratically, changing its sign results in the same dispersion relation. However, a tachyonic spinor action is still possible in the following form:

\[
\mathcal{L}_\psi^\text{tach} = \frac{1}{2} \left[ \bar{\psi} \gamma^5 \gamma^\mu \left( \nabla_\mu \psi \right) - \left( \nabla_\mu \bar{\psi} \right) \gamma^5 \gamma^\mu \psi \right] - \mu \bar{\psi} \psi,
\]

(54)

with the field equation

\[
\gamma^\mu \nabla_\mu \psi - \tilde{\gamma}^5 \psi = 0.
\]

In Eq. (55) \( \omega^2 = k^2 - \mu^2 \); that is, the dispersion relation is tachyonic (Chodos et al., 1985; Jentschura and Wundt, 2012).

Ramazanoğlu (2018c) used Eq. (54) to obtain the first spontaneous spinorization theory where spinor fields are unstable to growth around GR backgrounds of NSs. Even though the meaning of a ghost as opposed to a tachyon is a subtle issue for spinors, the equation of motion in this theory has divergent coefficients as in the vector field case (47). Hence, this form of spinorization suffers from similar problems to vectorization.

Minamitsuji (2020b) proposed an alternative model of spinorization given by the action

\[
S = \frac{1}{16\pi G_*} \int d^4x \sqrt{-g} \left\{ R + \frac{1}{2} \left[ \bar{\psi} \gamma^5 \gamma^\mu \left( \nabla_\mu \psi \right) - \left( \nabla_\mu \bar{\psi} \right) \gamma^5 \gamma^\mu \psi \right] \right\} + S_m \left[ \Psi_m, A^2(\bar{\psi} \psi)g_{\mu\nu} \right].
\]

(56)

This model has an unusual feature in that the derivative part of the spinor action is not the canonical one in Eq. (53), but rather the tachyonic one in Eq. (54). The resulting tachyonic equation of motion is

\[
\gamma^\mu \nabla_\mu \psi - \tilde{\gamma}^5 (4\pi A^4 \Lambda^\text{T}) \psi = 0,
\]

(57)

where \( \Lambda = d \ln A/d(\bar{\psi} \psi) \). The interpretation is straightforward: we have a tachyonic field with an effective mass \( \mu_{\text{eff}} = 4\pi A^4 \Lambda \); cf. Eq. (55). The most important aspect of this theory is the fact that the effective mass term and all coefficients of the field equation are regular everywhere; hence, the dynamics does not have any signs of the ill-posedness that we have found in all generalizations of scalarization thus far. In other words, the action (57) is the only known analog of spontaneous scalarization for nonscalar fields that does not suffer from ghost or gradient instabilities.

Our results on spontaneous tensorization suggest a no-go theorem for generalizing scalarization beyond scalar fields, at least if we want to avoid ghosts. It is curious that the only exception to this trend is the relatively exotic case of spinorization, which invites studies of the deeper theoretical underpinnings of the scalarization mechanism.

VI. OPEN PROBLEMS AND FUTURE PERSPECTIVES

We chose to review the literature by starting from the theoretical underpinnings of the scalarization mechanism,

---

18 Although a spinor is not a tensor in the technical sense, we classify spinorization here as an example of tensorization.
moving on to discussing NSs and BHs separately, and then returning to model building in order to discuss generalization of the mechanism to other fields. Most but not all open problems have already been mentioned in one or more of the previous sections and also discussed to some extent. Nevertheless, we revisit them in this section and cover any additional areas that require further development, opting for a concise summary of future perspectives.

A. Scalarization and cosmology

As discussed in Sec. II.B.4, one of the key challenges for scalarization is understanding whether it is compatible with cosmology. Recall that the main premise of scalarization is that there is a constant value of the scalar field $\phi_0$ for which spacetimes of GR become admissible solutions to the field equations, and it is these solutions that describe stationary objects that we expect (from an observational perspective) to carry no scalar charge. For this to be true, cosmic evolution needs to comply and drive $\phi$ to $\phi_0$ in the late Universe; otherwise, the entire Universe will in effect be “scalarized.” It was pointed out early on by Damour and Nordtvedt (1993) and more recently by Anderson et al. (2016) and Franchini and Sotiriou (2020) that reaching this preferred value for the scalar fields in the late Universe is not generic for simple models of scalarization and instead requires severe fine-tuning of initial conditions.

This could well be an artifact of not having the complete theory, and ideally one would expect the need for fine-tuned initial conditions to disappear by addressing further corrections to the model. A first step in this direction was made by Antoniou et al. (2021a) [who were inspired by Damour and Nordtvedt (1993)], who showed that the mixed model of scalarization discussed in Sec. II.B.4 has GR with a constant scalar as a late-time cosmic attractor for the right sign of the coupling between the scalar and the Ricci scalar. This demonstrated that in principle scalarization models can be made compatible with late-time cosmology and provided a recipe for doing so: include in the action terms that will dominate in late cosmology (for example, Ricci coupling) over the terms that control the onset of scalarization (for example, Gauss-Bonnet invariant) and hence impose the desired cosmological behavior. However, the specific model discussed as an example by Antoniou et al. (2021a) is by no means unique, and there have been many other attempts to address similar concerns (Chen et al., 2015; Erices et al., 2022; Minamitsuji and Tsujikawa, 2023). Perhaps more importantly, understanding the efficiency of the attractor behavior in reproducing the behavior of GR (with a cosmological constant) quantitatively, down to evolution of perturbations, structure formations, etc., certainly merits further investigation and could lead to constraints on realistic models of scalarization.

A second point of friction between scalarization and cosmology relates to the early Universe. Broadly speaking, scalarization is controlled by the coupling between the scalar-field and curvature invariants, so it is inevitable that these couplings will become increasingly important as one runs cosmic evolution backward and move to higher and higher curvatures. In particular, they will tend to dominate when the size of the Universe becomes comparable to the size of the compact objects that we want scalarization to affect today (Antoniou et al., 2021a). It was pointed out by Anson et al. (2019a,b) that quantum fluctuations could then seed a scalarization-like instability during inflation, and that it would be hard to prevent this linear instability by adding corrections to the action. It should be stressed, however, that one does not need the scalar field to necessarily remain constant or have its current value through the evolution of the Universe. Hence, one might not need to prevent such an instability, but might instead just quench it nonlinearly, exactly as it happens in scalarization itself. The scalar field could then smoothly evolve away from the vacuum that leads to this instability before it becomes relevant as one moves backward in cosmic time. The key question here is whether there is an EFT applicable to the early Universe that contains the scalarization models as late Universe limits but that can also host an inflationary scenario compatible with observations.

As a final remark in relation to cosmology, we emphasize that we have discussed here only the cosmological implications of known scalarization models and only under the assumption that the scalar field is cosmologically subdominant at late times. There are a plethora of generalized scalar-tensor theories that have been used in the context of inflation or dark energy and that exhibit couplings between the scalar and curvature invariants similar to those employed for scalarization models. In most cases, however, the nature of the couplings differs significantly from that of scalarization models. Reviewing such attempts goes well beyond the scope of this work. Exploring whether a scalar field that exhibits scalarization around compact objects could also play a role in inflation or account for dark energy would certainly be interesting, but putting together such a model (with a single scalar) seems particularly challenging. Note that in this context evading constraints on the coupling between a scalar and the Gauss-Bonnet invariant coming from the speed of GWs requires the scalar to be cosmologically subdominant (Franchini and Sotiriou, 2020).

B. Dynamical evolution

The study of the dynamical nonlinear regime of scalarization requires nonperturbative time-evolution schemes, whose availability varies among models. Finding formula-
tions of gravity theories that are amenable to numerical time evolution, i.e., numerical relativity, is nontrivial even in GR (Baker et al., 2006; Campanelli et al., 2006; Pretorius, 2005), and the issue can become even more complicated for alternative gravity theories.

Time-domain, nonlinear evolutions of DEF-like models have been made since the early work by Novak (1998a,b); and Novak and Ibanez (2000), and binary inspirals have been performed for nearly a decade (Barausse et al., 2013). This was possible thanks to the relative simplicity of the field equations, which allows one to establish their well-posedness (Salgado, 2006; Salgado et al., 2008). However, detailed and fully nonlinear numerical results in relation to GWs, a primary tool to test scalarization, are still lacking, aside from the study by Shibata et al. (2014).

The picture is much different for BH scalarization, where the coupling to the Gauss-Bonnet term can result in a complicated causal structure of the spacetime in dynamical situations. It was demonstrated at both linear (Blázquez-Salcedo et al., 2020a,b, 2018) and nonlinear (East and Ripley, 2021a,b; Ripley and Pretorius, 2020) levels that part, but not all, of the parameter space of spontaneous scalarization where scalarized BHs can be found loses hyperbolicity when the theory is evolved in time in a certain set of gauges, which means that no predictions can be obtained. The well-posedness of the initial-value problem in broader classes of modified gravity theories, including scalar-Gauss-Bonnet gravity, was studied by Kovács (2019); Kovács and Reall (2020a,b); and Papallo (2017). At the moment there is no consensus on whether this loss of hyperbolicity can be cured with a better gauge choice or if it is intrinsic to the evolution equations, which calls for further work on the issue. Furthermore, there are no detailed studies of binary evolution and merger even for the part of the parameter space for which hyperbolic time evolution has been shown to exist. The dynamics of more general models where various coupling terms are present as in Eq. (14) are not available either. The lack of results on all of these fronts presents important future research directions.

It can be the case that an EFT, obtained as a certain limit or truncation of a more fundamental theory, is ill posed while the latter is or is expected to be well posed. A typical example is relativistic hydrodynamics once viscosity has been taken into account. Indeed, it has been suggested to introduce techniques used in hydrodynamics to gravity in order to render the problematic time evolution of some theories hyperbolic by suitably modifying the equations while keeping the end point of evolution the same (Cayuso et al., 2017). This approach has seen increasing use in alternatives to GR, including the specific case of scalar-Gauss-Bonnet theories under certain symmetries (Franchini et al., 2022), and it is another avenue to explore the dynamics of scalarization in cases where current methods are inadequate. The applicability and power of such methods to general cases, for instance, fully nonlinear 3+1 evolution of a highly dynamical system, are yet to be confirmed.

C. Model building in and beyond scalarization

We have covered several models of spontaneous scalarization and their phenomenology for NSs and BHs. What they have in common is that scalarization is triggered by a tachyonic instability, whose threshold is controlled by fewer than a handful of terms in the action (Andreou et al., 2019). Yet, the properties of the final scalarized configuration depend on the nonlinearities and coupling terms that contribute beyond the linear level. The specific choices of these nonlinear interactions have been shown to crucially affect the stability (Antoniou et al., 2022; Blázquez-Salcedo et al., 2018; Silva et al., 2019) and the scalar charge (Antoniou et al., 2021b; Doneva et al., 2018a; Macedo et al., 2019; Silva et al., 2019) of scalarized solutions, as well as the cosmological behavior of the corresponding theory (Anson et al., 2019a; Antoniou et al., 2021a; Erices et al., 2022). Thus, such choices affect the viability and observability of the models. As a result, although certain models of scalarization, such as the original DEF model, have been studied fairly exhaustively, the exploration of the broader class of theories that exhibit scalarization, and their phenomenology, has only started.

Furthermore, one could consider introducing new classes of scalarization based on instabilities that are not tachyonic in nature. As discussed in Sec. V, success in this direction has thus far been limited, but it is nevertheless an interesting prospect. Alternatively, one can abandon the idea of a linear instability quenched by nonlinearities altogether, as discussed in Sec. II.C.5. Doneva and Yazadjiev (2022) demonstrated that a theory featuring a coupling to the Gauss-Bonnet invariant with a leading-order expansion with respect to the scalar field that is proportional to $\varphi^4$ can still have scalarized BHs below a certain mass threshold, although the Kerr solution is linearly stable with respect to massless scalar perturbations (Blázquez-Salcedo et al., 2022). This observation remains true in more general theories, such as Gauss-Bonnet gravity with multiple scalar fields (Staykov and Doneva, 2022). This strongly suggests the existence of a class of theories in which scalarization is a purely nonlinear phenomenon.

Scalarization models that have been considered thus far typically assume that the scalar field does not couple to matter directly in a suitable choice of variables, usually called the Jordan frame, which suffices for the theory to satisfy the WEP. However, as pointed out by Coates et al. (2017), this assumption might not be necessary because scalarization itself forces the scalar field into a trivial configuration away from a specific compact object. Therefore, scalarization models can be indistinguishable
from GR in regard to the WEP for all current observations. It was further argued by Coates et al. (2017) and Franchini et al. (2018) that particular couplings to matter that do not disrupt scalarization could introduce a Higgs-like mechanism in gravity: the scalar field changes value only near compact objects, and its coupling to matter changes the properties of the standard model in these regions. This is a largely unexplored possibility.

As discussed in Sec. V, another important open question is whether one can generalize the scalarization mechanism to other fields. Doing so in the context of a theory that is free of pathologies has thus far proven to be challenging, except perhaps in the case of spinors (Minamitsuji, 2020b). This is in part because controlling the dynamics of the extra degree of freedom is notoriously difficult in general in gravity theories with additional vector or tensor fields (de Rham, 2014), and vectorization or tensorization models are no exception.

We stress that the work on scalarization and its extensions has focused on isolated compact objects. In fact, theoretical considerations for putting together scalarization models have also been heavily influence by studies of isolated BHs and NSs. However, as previously discussed, scalarization could happen dynamically in a binary (Barausse et al., 2013; Doneva et al., 2022; Elley et al., 2022; Palenzuela et al., 2014; Shibata et al., 2014; Silva et al., 2021b; Taniguchi et al., 2015). Apart from finding ways to model this effect (Khalil et al., 2022) and potentially constraining it with GW observations, it would be particularly interesting to further explore which properties of a binary control dynamical scalarization, and how this differs among the various scalarization models.

D. Observational prospects

Perhaps the most attractive feature of scalarization is that the tachyonic instability typically leads to large scalar-field amplitudes before it is quenched, which results in nonperturbative deviations from GR in high curvature regions. At the same time, the theory mimics GR closely for weak gravity, easily satisfying existing tests. This suggests that new fundamental scalar fields that exhibit this behavior could be discovered with strong-gravity observations.

Current observations put stringent bounds on scalar radiation from pulsars that have already almost completely ruled out the original (massless) DEF model (Zhao et al., 2022). Electromagnetic observations of compact objects in binaries or surrounded by some form of matter also have the potential to extend such results to more general models of scalarization. For example, x-ray data from NS surfaces have recently been considered as a possible way to detect scalarized objects (Silva and Yunes, 2019a,b). The masses, radii, and moment of inertia of NSs can be inferred through such observations (Bogdanov et al., 2019a,b, 2021). Future observation could be used to distinguish scalarized NSs from nonscalarized ones by, for instance, using EOS-independent relations. The astrophysical signatures of most extensions of the DEF model, especially those that allow BH scalarization, have not yet been studied in detail, because they are recent.

The most interesting current development in strong gravity is the advent of GW astronomy, which is an arena for exploring spontaneous scalarization. The effects of scalarization on GWs of coalescing NSs and/or BHs have been investigated to some extent, mostly in regard to the DEF model (Khalil et al., 2022, 2019; Niu et al., 2021; Sennett and Buonanno, 2016) or scalar-Gauss-Bonnet gravity (Wong et al., 2022). The most pressing issue on this front is the development of accurate waveform models that ideally would cover the inspiral, merger, and ringdown stages of coalescing binaries. Developing such waveform models in GR has been (and remains) a formidable task that combines tools from post-Newtonian theory, numerical relativity, black-hole perturbation theory, and gravitational self-force. Progress on all these fronts is necessary for one to develop accurate scalarized waveform models. Stars that undergo core collapse are another interesting source of GWs. Albeit expected to be rare, such events may produce large bursts of scalar radiation (Kuan et al., 2022; Rosca-Mead et al., 2019; Sperhake et al., 2017) and may contribute to the stochastic GW background, which could in principle be probed by pulsar-timing arrays. In addition, some models of scalarization (but not exclusively) predict the existence of extra polarizations in GWs, which provides another way to test them.

Last, future spaceborne GW detectors such as LISA will bring another channel for testing modified gravity. Especially interesting with respect to scalarization is the case of extreme mass-ratio inspirals when the secondary object is scalarized (Maselli et al., 2020, 2022). Data from future GW detectors such as LISA or third generation ground-based detectors such as the Einstein Telescope and the Cosmic Explorer, combined with more precise electromagnetic observations such as those from improved x-ray observatories, will be crucial for detecting or ruling out scalar fields that exhibit scalarization.

ACKNOWLEDGMENTS

D.D.D. acknowledges financial support via the Emmy Noether Research Group, which was funded by the German Research Foundation (DFG) under Grant No. DO 1771/1-1. S.S.Y. thanks the University of Tübingen for the financial support and acknowledges partial support from Bulgarian NSF Grant No. KP-06-H28/7. T.P.S. acknowledges partial support from STFC Consolidated Grants No. ST/T000732/1 and No. ST/V005596/1. F.M.R acknowledges support from 2020 Bilim Akademisi GEBIP
