

Causality bounds on scalar-tensor EFTs

Dong-Yu Hong,^a Zhuo-Hui Wang^a and Shuang-Yong Zhou^{a,b}

^a*Interdisciplinary Center for Theoretical Study, University of Science and Technology of China, Hefei, Anhui 230026, China*

^b*Peng Huanwu Center for Fundamental Theory, Hefei, Anhui 230026, China*

E-mail: principle@mail.ustc.edu.cn, wzh33@mail.ustc.edu.cn, zhoushy@ustc.edu.cn

ABSTRACT: We compute the causality/positivity bounds on the Wilson coefficients of scalar-tensor effective field theories. Two-sided bounds are obtained by extracting IR information from UV physics via dispersion relations of scattering amplitudes, making use of the full crossing symmetry. The graviton t -channel pole is carefully treated in the numerical optimization, taking into account the constraints with fixed impact parameters. It is shown that the typical sizes of the Wilson coefficients can be estimated by simply inspecting the dispersion relations. We carve out sharp bounds on the leading coefficients, particularly, the scalar-Gauss-Bonnet couplings, and discuss how some bounds vary with the leading $(\partial\phi)^4$ coefficient and as well as phenomenological implications of the causality bounds.

KEYWORDS: Classical Theories of Gravity, Effective Field Theories

ARXIV EPRINT: [2304.01259](https://arxiv.org/abs/2304.01259)

Contents

1	Introduction and summary	1
2	Scalar-tensor EFT	8
3	Dispersive sum rules	12
3.1	Dispersion relations	12
3.2	Imposing st crossing symmetry	17
4	Power counting via dispersion relations	20
5	Optimization scheme	26
5.1	General strategy	26
5.2	Numerical details	28
6	Bounds on EFT coefficients and their implications	33
6.1	Scalar four-derivative term	33
6.2	Linear Gauss-Bonnet coupling	36
6.3	More generic Gauss-Bonnet couplings	43
6.4	Other six derivative terms	46
6.5	Coefficients for large α	47
6.6	Fine-tuned EFTs	50
A	Generic 4-leg amplitudes for scalar-tensor theory	55
B	Explicit sum rules with st symmetry imposed	60
C	Explicit example of optimization scheme	62

1 Introduction and summary

Causality/Positivity bounds. Relativistic causality is a foundational concept that underpins the modern construction of the fundamental models of nature. It is conjectured to imply analyticity and crossing symmetry of the S-matrix [1]. Unitarity of the quantum theory, another foundational cornerstone, also plays a vital role in restricting the forms the S-matrix can take. On the other hand, effective field theories (EFTs) are part and parcel of model building in modern particle physics and cosmology. Using merely the low energy field contents and symmetries, an EFT, arising from integrating out heavy degrees of freedom, can parametrize generic effects of possible UV completions at low energies. Interestingly, causality and unitarity, along with locality, can impose strong constraints on

the theory space, i.e., the space of the Wilson coefficients, of effective field theories, often known as causality or positivity bounds (see [2] for a concise review).

A simple and efficient way to derive these constraints on the Wilson coefficients is via the dispersion relations or dispersive sum rules, which provide a portal to connect the accessible EFT coefficients in the IR with the generic unknown physics in the UV [3]. They can be derived from analyticity, crossing symmetry and locality of the scattering amplitudes, and causality bounds are precisely the unitarity conditions on the UV amplitudes passed down to the IR via the dispersive sum rules. In the forward-limit of identical particle scattering, a simple positivity bound on the s^2 (s, t, u being the Mandelstam) coefficient can be easily seen using the textbook optical theorem [3]. The s^2 bound is usually the most accessible one phenomenologically. For 2-to-2 scattering between multiple species of particles, there are a set of s^2 coefficients since the amplitude can have different in and out states. Positivity bounds tell us that these s^2 coefficients form a convex cone, whose extremal rays (or kinks from the viewpoint of the cross section of the convex cone) correspond to tree-level UV (irrep) states, which are endowed with the projected-down versions of the UV symmetries [4, 5]. Particularly, this means that one can infer the existence of certain UV states from the causality convex cone, which helps inverse engineer the UV model from the EFT data. Furthermore, the dual cone of this amplitude cone is a spectrahedron, so the optimal causality bounds on the s^2 coefficients can also be effectively computed with semi-definite programming (SDP), even for the case of many degrees of freedom with less symmetries [6]. The Standard Model EFT (SMEFT) contains many degrees of freedom, so its parameter space is vast, especially at higher orders. Positivity bounds have been found to significantly restrict the viable space of dimension-8 operators [4, 6–19]. One may also reverse the argument and use the positivity bounds to test the fundamental principles of quantum field theory in some seemingly benign parameter regions [20–22], or inverse bootstrap the UV from the IR [23, 24].

Highly nonlinear constraints on the coefficients of higher powers of s can also be gleaned once realizing that the forward-limit dispersion relations readily define a Hausdorff moment problem [25, 26]. Away from the forward limit, a series of easily-to-use analytic bounds on both s and t derivatives of the amplitudes can be obtained using the Martin extension of analyticity [27] and the positivity of the derivatives of the Legendre polynomials [28] (see also [29–32] for related works). These bounds can be generalized to the case of massive particles with spin utilizing the transversity formalism (as opposed to the helicity formalism) for the external polarizations [33].

However, since the dispersive sum rules used to derive the above bounds are only su -symmetric, the full crossing symmetry of the S-matrix has not been used thoroughly, and they usually only constrain the coefficients from one side. Indeed, two-sided bounds can be derived for the coefficients once the full crossing symmetry is used [34, 35]. One pathway to achieve the triple crossing symmetry is simply to impose st symmetry on the su -symmetric sum rules. For the case of identical scalar scattering, the bounds on the explicitly computed coefficients are consistent with the usual dimensional analysis expectations for EFT coefficients. More importantly, this excludes the possibility that some delicate design of the UV model can lead to arbitrary disparity among different orders of Wilson coefficients

— “not everything goes for an EFT” [36]. This formalism can be easily extended to the case of multi-field theories using the generalized optical theorem for partial waves [37]. Compared to linear programming for the case of a single scalar, the optimization scheme now needs to be promoted to be a SDP problem with a continuous variable, which parametrizes the scales of the UV states. Both of them can be efficiently solved by the SDPB package [38]. Alternative methods, also based on dispersive relations, have been developed for obtaining the fully crossing symmetric causality bounds. These include directly using triple crossing symmetric dispersive relations [39], and formulating the (non-forward) dispersion relations as a double moment problem and slicing out the triple crossing bounds towards the end [40]. Triple crossing positivity bounds have also been used to constrain EFTs with spinning particles [41–47], and extra causality constraints using the upper bounds on the spectral functions can be found in [35, 45, 48]. Moreover, the powerful primal approach of S-matrix bootstrap has also been developed to chart the space of EFTs; see, e.g., [49–52] and for a review [53]. The primal approach directly parametrizes the crossing symmetric amplitudes themselves and expands viable theory space by imposing unitarity conditions. In this language, the above methods that rule out unphysical parameter regions is referred to as the dual approach, which parallels the difference between the cone and dual cone of the s^2 coefficients above.

In the presence of graviton exchanges in the scattering, a t -channel pole appears in the left hand side of the sum rules, because a spin-2 particle t -channel exchange term, different from the cases of lower spins, can survive the twice subtractions in deriving the sum rules. While we can still Taylor expand in terms of s , the existence of the t -channel pole prevents us from Taylor expanding in terms of t . Indeed, this t -channel pole must be balanced by a divergence in the dispersive integral on the right hand side as $t \rightarrow 0$. Apart from balancing the pole, the dispersive integral also gives rise to extra terms which can be negative and violate the would-be strict positivity in theories without the gravitons [24, 54, 55]. Nevertheless, each of the s -expanded sum rules can be viewed as a one-parameter (t) family of IR-UV relations, and one can effectively use them by optimizing over a set of continuous functions for the range that t can take within the EFT [56]. It turns out that the strongest constraints come from when t is far away from the forward limit and close to the cutoff. (A similar phenomenon was also seen in the earlier non-forward-limit bounds without full crossing symmetry [57, 58].) Physically, this means that some important constraints arise from when the impact parameter is small [56]. This approach has been used to constrain the Wilson coefficients of Einstein gravitational EFTs [44, 46] and Einstein-Maxwell EFTs [47].

Besides using the dispersion relations, causality bounds can also be derived from within the EFT by requiring information not propagating faster the speed of light. Although less algorithmic than the optimized dispersion relation approach, this approach is more intuitive and can sometimes produce very strong constraints with less efforts. In flat space, subluminality can usually be directly imposed on the dynamical modes of theory in a nontrivial background, which leads to conditions consistent with the positivity bounds obtained above [3]. In a gravitational EFT, the situation is more subtle, as the definition of speed is frame-dependent. So one resorts to observables such as the time delay in a

classical scattering. An often used causality condition is that the Eisenbud-Wigner time advance be not resolvable for the scattering wave, which is called asymptotic causality [59]. However, a more refined criterion for an EFT, called infrared causality, may be imposed that the time advance with the GR part subtracted should be non-resolvable for the scattering wave [60, 61]. Applications of the infrared causality can be found in [60, 62], and those of the asymptotic causality can be found in [59, 63–66]. A few other interesting applications of positivity bounds on gravitational and cosmological EFTs can be found in for example [67–82].

Scalar-tensor theory. General relativity (GR), with only the Einstein-Hilbert term, has been extensively tested in the solar system where it is relatively convenient for us to carry out gravitational experiments and where gravity is weak and velocities are small compared to the speed of light [83, 84]. The development of the Parameterized Post-Newtonian formalism has put severe constraints on possible deviations from GR in the weak gravity limit. The formalism is quite systematic, as it thoroughly parameterizes all possible deviations directly at the level of the metric. The discovery of binary pulsars has allowed us to confirm viability of GR in stronger gravity environments, with somewhat less accuracy, but those environments are still well approximated by the linearized GR. Therefore, the lesson is that, to be a viable alternative or extended gravity theory, it first needs to very precisely reduce to GR in the weak field limit.

However, this does not necessarily mean that sizable beyond GR effects have been completely ruled out in astrophysics, an intriguing possibility being that they are hidden in the highly dynamical and strong-field regimes, such as near black holes and neutron stars. Indeed, we are just starting to probe these regimes with the new observational tools such as LIGO-Virgo-KAGRA gravitational wave detectors [85] and the Event Horizon Telescope [86]. While GR can still pass the tests from these experiments to date, the accuracy is still quite low. Since interpolating between the weak gravity GR regime and the strong gravity regime with non-GR effects requires some degrees of “dynamical” nonlinearity, one of the simplest ways is to introduce new field degrees of freedom. Scalar-tensor theory is a simple extension of GR in this direction which only adds one extra field degree of freedom. Brans-Dicke theory [87], which give rises to a “variable gravitational constant”, is one of the earliest such models. It is currently tightly constrained by observations [83]. However, its extensions such as Horndeski theory/Generalized Galieon [88, 89] and Degenerate Higher-Order Scalar-Tensor theories [90] are being intensively investigated to fit astronomical and cosmological data [84]. Another motivation for scalar-tensor theory comes from string/M theory, where a dilaton naturally arises as a low energy degree of freedom from compactification [91]. The scalar degree of freedom is natural to consider also because fermions, due to the Pauli exclusion principle, can not form classical configurations, which need high occupation numbers at a range of momentum modes, while long-distance vector fields, endowed with a direction, are difficult to be compatible with the cosmological principle.

There is a growing body of research dedicated to understanding scalar-tensor theory in the strong regimes. The class of models involving the Gauss-Bonnet invariant

$\mathcal{G} = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}R^{\mu\nu} + R^2$ stand out, as they are low orders in the EFTs and can give rise to hairy black holes [92–95] and the phenomenon of (spontaneous) scalarization [96, 97]. These operators have been confronted with gravitational wave observations and beyond [98–107]. Wheeler famously coined the phrase that a black hole has no hair [108]. More precisely, due to the uniqueness theorems in GR, a (non-charged) black hole in GR can be solely described by its mass and angular momentum, and a bunch of no-hair theorems generally prevent a black hole from having other parameters/pieces of hair [108, 109]. A few exceptions include the presence of the scalar-Gauss-Bonnet couplings. In fact, assuming shift symmetry for the scalar and the equations of motion being second order, the linear scalar-Gauss-Bonnet coupling $\phi\mathcal{G}$ is necessary to sustain hairy solutions in Horndeski theory [93, 94]. Furthermore, the $\phi\mathcal{G}$ term leads to the same parametrized post-Newtonian parameters as in GR [110], and in particular it does not lead to nontrivial scalar charges for neutron stars or other extended objects [95]. Therefore, the current gravitational wave experiments are an ideal place to test this leading quadratic curvature term.

On the other hand, the Damour-Esposito-Farese model [111] is the first model of scalarization, which was proposed when the weak field gravity tests had reached an unprecedented accuracy such that viable deviations from GR was seemingly impracticable. It was also when binary pulsar observations became available, ushering in a new arena to test GR with the compact stars. In the Damour-Esposito-Farese model, the scalar field obtains a nontrivial profile once the density/curvature within the star exceeds a threshold, and this can be the case for a neutron star, resulting in strong deviations from GR, but not for the Sun. With the arrival of gravitational wave astronomy, another new window has been opened up to test GR in stronger and more dynamical gravity environments. Recently, a new class of scalarization models involving the Gauss-Bonnet invariant and black holes have been proposed, in which the black hole becomes hairy if the curvature outside the horizon exceeds a threshold [96, 97] (see [112] for a review). The underlying reason for the scalarization to happen is because in these models the strong gravity environment induces tachyonic instabilities for the unscalarized configuration. In the inspiral phase of a binary black hole coalescence, a dynamical de-scalarization can occur, which can give rise to extra scalar radiation and thus observational constraints [113].

With the arrival of the gravitational wave astronomy and advances of more traditional observational means, it is becoming increasingly accessible to test gravity, along with possible accompanied extra degrees of freedom, in the strong and dynamical regimes [84]. As we shall see, the causality bounds can strongly constrain the parameter spaces of gravitational EFTs, which may help orient current and future experiments to more theoretically favorable directions. On the flip side, one may also use the new observational data to test the fundamental principles of quantum field theory or the S-matrix theory.

Summary. In this paper, we investigate how causality bounds constrain the parameter space of scalar-tensor theory by means of dispersive sum rules of the scattering amplitudes. To fully utilize the crossing symmetry of amplitudes, we start with dispersive sum rules that are only *su*-symmetric and then impose the *st* symmetry on these sum rules. In

a multi-field theory such as scalar-tensor theory here, only a few amplitudes are truly symmetric in full permutations of s, t, u in the strict sense, some being not even strictly symmetric in s and u , so the su or st crossing symmetry is used loosely in this context, with the understanding that some crossings actually link distinct amplitudes. Nevertheless, the working mechanism of improving the bounds with crossing symmetry is exactly the same as in the single scalar case. In the presence of massless gravitons, the t -channel pole prevents us from Taylor-expanding some sum rules in the forward limit, so the decision variables for the optimization involve a set of weight functions of t , which numerically will be evaluated with a finite dimensional truncation. In this setup, some important constraint space can be effectively sampled using the impact parameter [56]. Various causality bounds without full crossing symmetry and/or neglecting the t -channel pole have previously been used to constrain scalar-tensor models [55, 72, 80, 82, 114–116].

While the Froissart-Martin bound [117, 118] for the high energy behaviors of amplitudes is rigorously established for massive particles, which suggests that only two subtractions are needed to derive the dispersive sum rules, it is more subtle for massless particles especially in the presence of gravitons. We will make the usual assumption that only two subtractions are needed when $t < 0$ and three subtractions when $t \leq 0$ [23, 44]. We will also assume that the EFT is weakly coupled in the IR so that we can use tree-level amplitudes at low energies, but we are agnostic about the attributes of the UV theory, as manifest in our exclusive use of the dispersive sum rules in deriving the bounds. We will only make use of positivity of partial wave unitarity, which leads to the semi-positive conditions on the $B_{P_X, \ell}(\mu)$ matrices (see eq. (5.10)). Nevertheless, with full crossing symmetry incorporated, we find that the Wilson coefficients projected to the gravitational coupling $1/M_P^2$ are already bounded to finite regions. This is of course except for the $(\partial\phi)^4$ coefficient (and consequently some correlated coefficients), for which the upper bound of partial wave unitarity is needed to cap from the above.

We find that a simple method can be devised to estimate the sizes/scalings of the Wilson coefficients via the dispersive sum rules, without the need for heavy numerical calculations. This proceeds by first normalizing the Mandelstam variables in the dispersive sum rules with the cutoff of the EFT. Then, from some simple sum rules that only contain the gravitational coupling $1/M_P^2$, we can establish correspondences between the UV spectral functions and the hierarchy between the cutoff and the Planck mass. A scaling correspondence can not be uniquely assigned in this way to the UV spectral function $c_{\ell, \mu}^{00 \rightarrow X}$ (the partial amplitude from two scalars to a heavy state X , cf. eq. (3.2)), for which we can either let it saturate the unitarity upper bound or assign a desired correspondence, the latter of which will lead to an *ad hoc* class of theories with reduced scalings for the relevant terms. These correspondences can then be used to infer the dimensions of the Wilson coefficients by simple inspection of available sum rules. The scalings of the coefficients extracted in this way are consistent with the sharp numerical bounds obtained by SDP.

The causality bounds on some Wilson coefficients are intimately correlated with each other, while others are quite independent. This can be often inspected from the $B_{P_X, \ell}(\mu)$ matrices that are constructed from dispersive sum rules. If the relevant quantities are in different diagonal blocks, then the corresponding coefficients are insensitive to each other.

However, even if the relevant quantities overlap in the $B_{P_X, \ell}(\mu)$ matrix, a strong correlation between the corresponding coefficients is not guaranteed. At the practical level, the bounds on certain coefficients can not be numerically optimized unless we specify the value of the coefficient of the scalar self-interaction operator $(\partial\phi)^4$. These are the coefficients that only appear in the sum rules involving the UV spectral function $c_{\ell, \mu}^{00 \rightarrow X}$.

We also derive the causality bounds on some fine-tuned EFTs. The bounds on a set of Wilson coefficients in the fine-tuned EFT can be considered as taking an appropriate crossing section in the Wilson coefficient space, while the bounds on a given set of Wilson coefficients in a generic EFT amounts to projecting the causality spectrahedron down to an appropriate subspace. We show that some phenomenological models such as the $f(\phi)\mathcal{G}$ model should not be taken at its face value, because only adding exactly $f(\phi)\mathcal{G}$ but no other terms inevitably violates causality bounds. Indeed, in a model where the operators essential for causality bounds to uphold are turned on but highly suppressed compared to the usual EFT power counting, we can see that the Wilson coefficients of concern are also highly constrained by causality bounds. We give a simple criterion to test whether a given/fine-tuned scalar-tensor model will run into contradictions with causality bounds.

Particular attention has been given to the scalar Gauss-Bonnet couplings, which can give rise to hairy black holes and scalarization and are currently undergoing intense scrutiny in astrophysics by gravitational wave and other observations. We carve out the 2D bounds on the leading order $\phi\mathcal{G}$ coefficient together with the coefficient of the Riemann cubed operator, which is independent of the coefficient α of $(\partial\phi)^4$. On the other hand, the bounds on the coefficient of $\phi^2\mathcal{G}$, which is essential for scalarization, strongly depend on α . We also compare the causality bounds with the observational bounds for the coefficients of $\phi\mathcal{G}$ and $\phi^2\mathcal{G}$, which allows us to impose bounds on the cutoffs for these EFTs and reduce the viable parameter space, thanks to the fact that for a capped α these fully crossing symmetric bounds have restricted the viable parameters to an enclosed region.

If the scalar interacts with the heavy modes weakly in the UV theory, i.e., if the UV spectral function $c_{\ell, \mu}^{00 \rightarrow X}$ is suppressed by $\mathcal{O}(\Lambda/M_P)$, the scalar will interact with the graviton with the usual gravitational strength in the low energy scalar-tensor EFT. This will lead to the scaling of eq. (4.30). For the terms involving the Gauss-Bonnet invariant, this gives rise to the usual scaling implicitly used in most literature: $\mathcal{L} \supset M_P^2 \sqrt{-g} (\frac{\mathcal{O}(1)}{\Lambda^2} \varphi \mathcal{G} + \frac{\mathcal{O}(1)}{\Lambda^2} \varphi^2 \mathcal{G})$, where $\varphi \equiv \phi/M_P$. However, for a generic UV completion, as we see in eq. (4.29), the couplings for terms like $\varphi^2\mathcal{G}$ are allowed to be much larger, without running into the trouble with causality bounds: $\mathcal{L} \supset M_P^2 \sqrt{-g} (\frac{\mathcal{O}(1)}{\Lambda^2} \varphi \mathcal{G} + \frac{\mathcal{O}(1)M_P}{\Lambda^3} \varphi^2 \mathcal{G})$. This arises when the low energy scalar interacts the heavy modes more strongly than the gravitational force, a scenario aligned with the weak gravity conjecture. Incidentally, in this scenario, the spontaneous scalarization models are natural where a vanishing $\varphi\mathcal{G}$ term is usually assumed and a sizable $\varphi^2\mathcal{G}$ is required for tachyonic instabilities to take place. We have confirmed the above scalings with the numerical causality bounds in section 6.

We have focused on the parity conserving sector in this paper. Once the parity violating operators are involved, the complexity of numerics will increase significantly, as we have to augment the dimension of the vector $\mathcal{C}_{P_X, \ell, \mu}$ and consequently the matrix $B_{P_X, \ell}(\mu)$

(see eq. (5.10)). There has also been a growing interest in examining the observational implications of parity-violating operators in scalar-tensor theories (see for example [84, 95]). We defer the extraction of causality bounds on these terms to future work [119].

The paper is organized as follows. In section 2, we present the scalar-tensor EFT both at the level of Lagrangian, with independent operators, and at the level of the amplitudes that will be needed to derive the dispersive sum rules. The sum rules will be derived in a couple of steps in section 3. In section 4, we propose a method to perform dimensional analysis of the Wilson coefficients with the dispersive sum rules. In section 5, we outline the optimization scheme to obtain the optimal bounds with positivity from unitarity, and explain its numerical implementation in details. In section 6, we present the results of the numerical causality bounds and discuss their implications. In appendix A, we show how to construct generic 4-leg EFT amplitudes from scratch. In appendix B, we explicitly list all the sum rules used to perform analyses and computations in this paper. In appendix C, we show an explicit example exhibiting how the SDP optimization is performed.

Notation and conventions. The (reduced) Planck mass is $M_P = 1/\sqrt{8\pi G_N}$. Our metric signature is $g_{\mu\nu} = \{-, +, +, +\}$. We choose all momenta to be in-going, so the Mandelstam variables are $s = -(p_1 + p_2)^2$, $t = -(p_1 + p_3)^2$, $u = -(p_1 + p_4)^2$. A generic four-point helicity amplitude is denoted as $\mathcal{M}^{1234} = \mathcal{M}(1^{h_1}2^{h_2}3^{h_3}4^{h_4})$, where h_i is the helicity for particle i , while a specific four-point helicity amplitude is denoted as, say, $\mathcal{M}^{++0-} = \mathcal{M}(1^{+2}2^{+2}3^{0}4^{-2})$. Our convention for the partial wave expansion of the four-point amplitude is $\mathcal{M}^{1234}(s, t, u) = 16\pi \sum_{\ell} (2\ell + 1) d_{h_{12}, h_{43}}^{\ell} \left(1 + \frac{2t}{s}\right) A_{\ell}^{1234}(s)$, where $A_{\ell}^{1234}(s) \equiv A_{\ell}^{h_1 h_2 h_3 h_4}(s)$ is the partial wave amplitude, $d_{h_{12}, h_{43}}^{\ell}(z)$ is the Wigner (small) d-matrices and $h_{ij} \equiv h_i - h_j$. The dimensionful scalar field ϕ is related to the dimensionless one φ by $\phi = M_P \varphi$.

2 Scalar-tensor EFT

Scalar-tensor theory is a popular extension of Einstein’s metric tensor theory. It augments gravity by coupling the massless spin-2 field to a scalar, arguably the simplest kind of fields that can form classical configurations which may affect local or large-scale gravitational physics. The scalar can minimally couple to the metric with possible potential self-interactions. However, from an EFT point of view, non-minimal and derivative interactions are generically present in the theory. For example, these couplings are also ubiquitous in EFTs from string/M theory which generally predicts existence of scalars due to compactification from higher dimensions [91]. Indeed, the effects of these non-minimal and derivative couplings have been extensively studied in astrophysics and cosmology [120].

We will be interested in 4D scalar-tensor theory where the mass of the scalar is negligible, and also assume that the theory is weakly coupled below the cutoff Λ so that we can take the tree-level approximation in the IR. We are agnostic about the UV theory, in particular, not assuming it to be weakly coupled. Up to six derivatives and including only terms that can give rise to tree-level 2-to-2 amplitudes, the lowest order terms of such a

theory are given by

$$\begin{aligned}
 S = \int d^4x \sqrt{-g} & \left(\frac{M_P^2}{2} R - \frac{1}{2} \nabla_\mu \phi \nabla^\mu \phi - \frac{\lambda_3}{3!} \phi^3 - \frac{\lambda_4}{4!} \phi^4 + \frac{\alpha}{2} (\nabla_\mu \phi \nabla^\mu \phi)^2 + \frac{\beta_1}{2!} \phi \mathcal{G} + \frac{\beta_2}{4} \phi^2 \mathcal{G} \right. \\
 & + \frac{\gamma_0}{3!} \mathcal{R}^{(3)} + \frac{\gamma_1}{3!} \phi \mathcal{R}^{(3)} + \frac{\gamma_2}{2} \nabla_\mu \phi \nabla^\mu \phi \mathcal{R}^{(2)} - \frac{4\gamma_3}{3} \nabla_\mu \phi \nabla_\rho \phi \nabla_\nu \nabla_\sigma \phi R^{\mu\nu\rho\sigma} \\
 & \left. + \frac{\gamma_4}{3} \nabla_\mu \phi \nabla^\mu \phi \nabla_\rho \nabla_\sigma \phi \nabla^\rho \nabla^\sigma \phi + \dots \right), \tag{2.1}
 \end{aligned}$$

where $M_P = 1/\sqrt{8\pi G_N}$ is the (reduced) Planck mass and we have defined $\mathcal{R}^{(2)}$, $\mathcal{R}^{(3)}$ and the Gauss-Bonnet invariant \mathcal{G} ,

$$\mathcal{R}^{(2)} = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}, \quad \mathcal{R}^{(3)} = R_{\mu\nu}{}^{\rho\sigma} R_{\rho\sigma}{}^{\alpha\beta} R_{\alpha\beta}{}^{\mu\nu}, \quad \mathcal{G} = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - 4R_{\mu\nu} R^{\mu\nu} + R^2. \tag{2.2}$$

We have focused on a scalar-tensor theory that conserves parity, so Lagrangian terms with odd numbers of the Levi-Civita tensor such as the Chern-Simons term $R_{\mu\nu\rho\sigma} \epsilon^{\mu\nu\alpha\beta} R_{\alpha\beta}{}^{\rho\sigma}$ are absent from the Lagrangian. Naively, there are several other terms that can be written down in the Lagrangian, but those terms can be reduced to the above terms by field redefinitions and integration by parts [121, 122]. This can be partially checked by explicit scattering amplitudes computed in the following, since amplitudes are free of ambiguities of field redefinitions and integration by parts.

As mentioned in the introduction, the scalar coupled quadratic curvature terms are being actively looked at phenomenologically, in search of/to rule out possible deviations from Einstein’s gravity in strong and/or dynamical gravity environments near compact stars. In principle, a couple of scalar self-interaction operators are of lower orders in terms of the EFT cutoff, but they are only minimally coupled to gravity, which by themselves would not give rise to significant modifications to the gravitational force. More practically, for the positivity bounds that will be extracted later, since we make use of the generic twice subtracted dispersion relations, the scalar potential terms are unconstrained, while, say, the scalar four-derivative self-coupling can be bounded. In fact, the coefficient of the dim-8 contact interaction being bounded to be positive in flat space has inspired the name of these bounds.

Particular attention has been paid to the operators involving the Gauss-Bonnet invariant, as these operators can give rise to hairy black holes [92–95] and the interesting phenomenon of spontaneous scalarization [96, 97], which is the reason why we have chosen to parametrize the Lagrangian terms with the Gauss-Bonnet invariant, instead of the Riemann tensor squared. The linear scalar-Gauss-Bonnet term $\phi \mathcal{G}$ [93–95] is special in the sense that it is shift-symmetric $\phi \rightarrow \phi + const$, as \mathcal{G} is famously a total derivative. Significant efforts have been put into constraining the Wilson coefficient of this operator with the gravitational wave and X-ray data from binary compact stars [98–107]. These observations capitalize on the fact that the scalar-Gauss-Bonnet coupling alters the star configurations and as well as induces significant dipole radiation in binaries, thanks to the scalar degree of freedom. In section 6, we shall use these data to infer observational bounds on the EFT cutoff. Furthermore, the $\phi^2 \mathcal{G}$ operator has also attracted a lot of interest lately, due to its

ability to generate tachyonic instabilities to make the scalar field nontrivial for black holes and neutron stars [96, 97].

Since we will be constraining the Wilson coefficients with the dispersion relations of the scattering amplitudes, we may as well parametrize the EFT at the level of amplitudes. General EFT amplitudes can be parametrize by considering little group scalings and crossing symmetries. After factoring out the helicity structures, the amplitudes can be written as scalar functions of Mandelstam variables s, t, u . Crossing symmetries dictate the symmetries of these functions, and also allow us to focus on a few independent amplitudes to extract all available information. For the lowest orders of the amplitudes with double 3-leg insertions, one can simply calculate them explicitly from the EFT Lagrangian. Contributions from the 4-leg contact interactions can be constructed based on some simple principles. For our purposes, we choose a representation for the helicity spinors to also convert the helicity structures into expressions in terms of s, t, u . After these considerations (see more details in appendix A), the independent amplitudes can thus be parametrized as follows

$$\mathcal{M}^{0000} = f_S(s, t, u) = g_{-1,1}^S \frac{x}{y} + g_{0,0}^S + g_{-1,2}^S \frac{x^2}{y} + g_{0,1}^S x + g_{1,0}^S y + g_{0,2}^S x^2 + \dots \quad (2.3)$$

$$\mathcal{M}^{++--} = ([12]\langle 34 \rangle)^4 f_{T_1}(t, u) = g_{3,-1}^{T_1} \frac{s^3}{tu} + g_{3,0}^{T_1} s^3 + g_{4,0}^{T_1} s^4 + g_{3,1}^{T_1} s^3 tu + g_{5,0}^{T_1} s^5 + \dots \quad (2.4)$$

$$\mathcal{M}^{++++} = ([12][13]\langle 14 \rangle)^4 f_{T_2}(s, t, u) = g_{1,0}^{T_2} y + g_{2,0}^{T_2} y^2 + g_{2,1}^{T_2} y^2 x + g_{3,0}^{T_2} y^3 + \dots \quad (2.5)$$

$$\mathcal{M}^{++++} = \frac{([12][34])^2}{(\langle 12 \rangle \langle 34 \rangle)^2} f_{T_3}(s, t, u) = g_{1,0}^{T_3} y + g_{0,2}^{T_3} x^2 + g_{1,1}^{T_3} yx + g_{2,0}^{T_3} y^2 + g_{0,3}^{T_3} x^3 + \dots \quad (2.6)$$

$$\mathcal{M}^{+++0} = ([12][23][31])^2 f_{M_1}(s, t, u) = g_{1,0}^{M_1} y + g_{1,1}^{M_1} yx + g_{2,0}^{M_1} y^2 + \dots \quad (2.7)$$

$$\mathcal{M}^{++0-} = [12]^6 (\langle 14 \rangle \langle 24 \rangle)^2 f_{M_2}(t, u) = g_{2,0}^{M_2} s^2 + g_{2,1}^{M_2} s^2 tu + g_{3,1}^{M_2} s^3 tu + \dots \quad (2.8)$$

$$\mathcal{M}^{++00} = [12]^4 f_{M_3}(t, u) = g_{1,0}^{M_3} s + g_{2,0}^{M_3} s^2 + g_{1,1}^{M_3} stu + g_{3,0}^{M_3} s^3 + g_{2,1}^{M_3} s^2 tu + \dots \quad (2.9)$$

$$\mathcal{M}^{+-00} = ([13]\langle 23 \rangle [14]\langle 24 \rangle)^2 f_{M_4}(t, u) = g_{-1,1}^{M_4} \frac{tu}{s} + g_{1,1}^{M_4} stu + g_{0,2}^{M_4} (tu)^2 + \dots \quad (2.10)$$

$$\mathcal{M}^{+000} = ([12]\langle 23 \rangle [31])^2 f_{M_5}(s, t, u) = g_{0,1}^{M_5} x + g_{1,0}^{M_5} y + g_{1,1}^{M_5} yx + g_{2,0}^{M_5} y^2 + \dots \quad (2.11)$$

where we have defined the shorthand for the amplitudes, say, $\mathcal{M}^{++0-} = \mathcal{M}(1^{+2}2^{+2}3^04^{-2})$ (particle 1 having helicity +2, etc.) and the basic symmetric polynomials of the Mandelstam variables

$$x = s^2 + t^2 + u^2, \quad y = stu. \quad (2.12)$$

The $f_i(t, u)$ functions are tu symmetric, while the $f_i(s, t, u)$ functions are stu symmetric. Thus, in scalar-tensor theory, a whole amplitude is either symmetric under the full permutations of s, t, u or symmetric under the exchange of two of s, t, u , accompanied by exchanges of the helicities accordingly. Explicitly, the ones with full stu permutation symmetries are given by

$$\mathcal{M}^{0000}(s, t, u) = \mathcal{M}^{0000}(u, t, s) = \mathcal{M}^{0000}(t, s, u), \quad (2.13)$$

$$\mathcal{M}^{+000}(s, t, u) = \mathcal{M}^{+000}(u, t, s) = \mathcal{M}^{+000}(t, s, u), \quad (2.14)$$

$$\mathcal{M}^{+++0}(s, t, u) = \mathcal{M}^{+++0}(u, t, s) = \mathcal{M}^{+++0}(t, s, u), \quad (2.15)$$

$$\mathcal{M}^{+++-(s, t, u) = \mathcal{M}^{+++-(u, t, s) = \mathcal{M}^{+++-(t, s, u)}, \quad (2.16)$$

$$\mathcal{M}^{++++}(s, t, u) = \mathcal{M}^{++++}(u, t, s) = \mathcal{M}^{++++}(t, s, u), \quad (2.17)$$

and the ones with only one exchange symmetry are

$$\begin{aligned} \mathcal{M}^{++00}(s, t, u) &= \mathcal{M}^{++00}(s, u, t) = \mathcal{M}^{+00+}(u, t, s) = \mathcal{M}^{+00+}(t, u, s) \\ &= \mathcal{M}^{+0+0}(t, s, u) = \mathcal{M}^{+0+0}(u, s, t), \end{aligned} \quad (2.18)$$

$$\begin{aligned} \mathcal{M}^{+-00}(s, t, u) &= \mathcal{M}^{+-00}(s, u, t) = \mathcal{M}^{+00-}(u, t, s) = \mathcal{M}^{+00-}(t, u, s) \\ &= \mathcal{M}^{+0-0}(t, s, u) = \mathcal{M}^{+0-0}(u, s, t), \end{aligned} \quad (2.19)$$

$$\begin{aligned} \mathcal{M}^{0-++}(s, t, u) &= \mathcal{M}^{0-++}(s, u, t) = \mathcal{M}^{0++-}(u, t, s) = \mathcal{M}^{0++-}(t, u, s) \\ &= \mathcal{M}^{0+-+}(t, s, u) = \mathcal{M}^{0+-+}(u, s, t), \end{aligned} \quad (2.20)$$

$$\begin{aligned} \mathcal{M}^{+--+}(s, t, u) &= \mathcal{M}^{+--+}(s, u, t) = \mathcal{M}^{+---}(u, t, s) = \mathcal{M}^{+---}(t, u, s) \\ &= \mathcal{M}^{+--+}(t, s, u) = \mathcal{M}^{+--+}(u, s, t). \end{aligned} \quad (2.21)$$

Note that for particles with spin the crossing symmetry is generally highly non-trivial except for the massless case we are considering. We see that some of the above equalities are more appropriately called crossing *relations* rather than crossing *symmetries*, as they link different amplitudes rather than reflect symmetries within an amplitude. We shall adapt the standard terminology that crossing symmetry refers to the collection of all crossing symmetries and relations. The amplitudes with the remaining helicities are not independent and can be obtained by using the relation $\mathcal{M}^{\bar{1}2\bar{3}4}(s, t, u) = (\mathcal{M}^{1234}(s^*, t^*, u^*))^*$. So we will only need to use the dispersion relations for the amplitudes above in eqs. (2.13)–(2.21) to constrain the Wilson coefficients.

By an explicit computation of the amplitudes from Lagrangian (2.1) with Feynman diagrams, we find that to the lowest orders the $g_{a,b}^i$ coefficients above are related to the Lagrangian Wilson coefficients as follows

$$\begin{aligned} \mathcal{M}^{0000} &= \lambda_3^2 \left(\frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right) - \lambda_4 + \frac{1}{M_P^2} \left(\frac{su}{t} + \frac{st}{u} + \frac{ut}{s} \right) \\ &\quad + \alpha(s^2 + t^2 + u^2) + \gamma_4 stu + \dots \end{aligned} \quad (2.22)$$

$$\mathcal{M}^{+ + - -} = \frac{1}{M_P^2} \frac{s^3}{tu} - \frac{\beta_1^2}{M_P^4} s^3 + \frac{\gamma_0^2}{M_P^6} s^3 tu + \dots \quad (2.23)$$

$$\mathcal{M}^{+ + + -} = \frac{\gamma_0}{M_P^4} stu + \dots \quad (2.24)$$

$$\mathcal{M}^{+ + + +} = \left(\frac{10\gamma_0}{M_P^4} - \frac{3\beta_1^2}{M_P^4} \right) stu + \frac{\gamma_0^2}{M_P^6} stu(s^2 + t^2 + u^2) + \dots \quad (2.25)$$

$$\mathcal{M}^{+ + + 0} = \frac{\gamma_1}{M_P^3} stu + \dots \quad (2.26)$$

$$\mathcal{M}^{+ + 0 -} = \frac{\beta_1}{M_P^3} s^2 - \frac{\gamma_0 \beta_1}{M_P^5} s^2 tu + \dots \quad (2.27)$$

$$\mathcal{M}^{+ + 0 0} = \frac{\lambda_3 \beta_1}{M_P^2} s + \frac{\beta_2}{M_P^2} s^2 + \frac{\gamma_0}{M_P^4} stu + \frac{\beta_1^2 + \gamma_2 M_P^2}{M_P^4} s^3 + \dots \quad (2.28)$$

$$\mathcal{M}^{+-00} = \frac{1}{M_P^2} \frac{tu}{s} + \frac{\beta_1^2}{M_P^4} stu + \dots \quad (2.29)$$

$$\mathcal{M}^{+000} = \frac{\beta_1}{2M_P^3} (s^2 + t^2 + u^2) + \frac{\gamma_3}{M_P} stu + \dots \quad (2.30)$$

3 Dispersive sum rules

In constructing the EFT Lagrangian or parameterizing the EFT scattering amplitudes in the last section, it would seem that the Wilson coefficients are allowed to take arbitrary values. The existence of causality/positivity bounds suggests that this would be an approach that sometimes leads to erroneous results. In particular, the consistency of the UV physics can actually impart many constraints on these EFT couplings. These UV consistency conditions include fundamental principles of S-matrix theory such as causality and unitarity, and can be utilized in the form of a series of dispersive sum rules or dispersion relations. In this section, we shall derive these dispersion relations and discuss how to effectively use them for scalar-tensor theory.

3.1 Dispersion relations

Before introducing the dispersion relations, let us first briefly recall partial wave unitarity that will be used shortly. General 2-to-2 amplitudes for particles with spin in the helicity basis can be decomposed into partial wave amplitudes in terms of the Wigner (small) d-matrices

$$\mathcal{M}^{1234}(s, t, u) = 16\pi \sum_{\ell} (2\ell + 1) d_{h_{12}, h_{43}}^{\ell} \left(1 + \frac{2t}{s}\right) A_{\ell}^{1234}(s), \quad (3.1)$$

where $A_{\ell}^{1234}(s) \equiv A_{\ell}^{h_1 h_2 h_3 h_4}(s)$ is the spin- ℓ partial wave amplitude and $d_{h_{12}, h_{43}}^{\ell}(z)$ is the Wigner (small) d-matrices with $h_{ij} \equiv h_i - h_j$ (see, e.g., appendix F of [33] for properties of the Wigner d-matrices). Note that $A_{\ell}^{1234}(s)$ is a function of s only, while $\mathcal{M}^{1234}(s, t, u) \equiv \mathcal{M}^{1234}(s, t)$ is a function of s and t because of the constraint $s + t + u = 0$. The argument of the Wigner d-matrix is $\cos \theta = 1 + 2t/s$, where scattering angle θ is the angle between the physical momenta of particle 1 and 3. Since the angular momentum is conserved in a scattering, the S-matrix is block-diagonal for different spin- ℓ , so unitarity of the \mathcal{M}^{1234} amplitudes implies that the partial wave amplitudes A_{ℓ}^{1234} are also unitary. This means that we can split the absorptive part of $A_{\ell}^{1234}(s)$ into

$$\text{Abs } A_{\ell}^{1234}(s) = \sum_X c_{\ell, s}^{12 \rightarrow X} c_{\ell, s}^{* \bar{3} \bar{4} \rightarrow X}, \quad (3.2)$$

where the sum over X is for a complete basis of the Hilbert space, $c_{\ell, s}^{12 \rightarrow X}$ denotes the partial wave amplitude from particle 1 and 2 to the intermediate state X with center of mass energy s , and $c_{\ell, s}^{* \bar{3} \bar{4} \rightarrow X} \equiv (c_{\ell, s}^{\bar{3} \bar{4} \rightarrow X})^*$ with $\bar{3}$ and $\bar{4}$ denoting that particle 3 and 4 carry helicity $-h_3$ and $-h_4$ respectively. The reason for the extra minus signs for the helicities of particle 3 and 4 is that we are using the convention where all external particles are in-going. The absorptive part of A_{ℓ}^{1234} is defined as

$$\text{Abs } A_{\ell}^{1234}(s) \equiv \frac{1}{2i} \left(A_{\ell}^{1234}(s + i\epsilon) - (A_{\ell}^{\bar{3} \bar{4} \bar{1} \bar{2}}(s + i\epsilon))^* \right) = \text{Disc } A_{\ell}^{1234}(s), \quad (3.3)$$

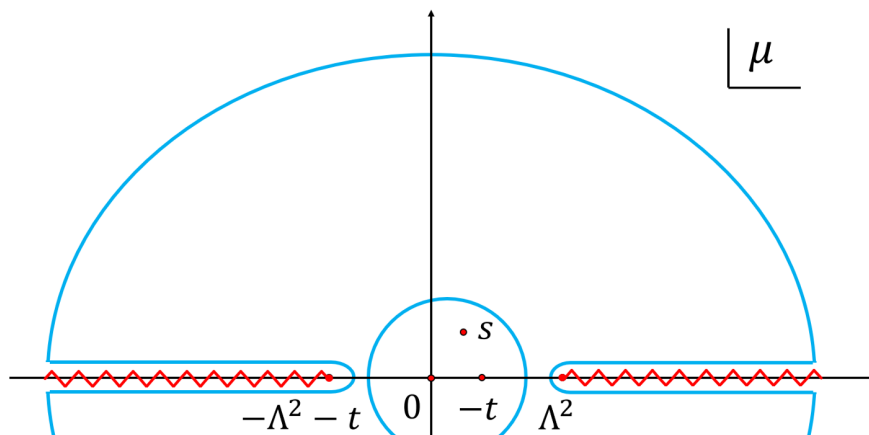


Figure 1. Analytic structure of $\mathcal{M}^{1234}(\mu, t)/(\mu - s)$ in the complex μ (center-of-mass energy squared) plane. The pole at $\mu = s$ and the low energy poles $\mathcal{M}^{1234}(\mu, t)$ are referred to as the “EFT poles”, with Λ being the EFT cutoff. The small (closed) contour is valid within the EFT, while the equivalent big (closed) contour encodes the UV information.

where the last equality is because the S-matrix is Hermitian analytic ($A_{\ell}^{\bar{3}\bar{4}\bar{1}\bar{2}}(s + i\epsilon)^* = A_{\ell}^{1234}(s - i\epsilon)$). For a time reversal invariant theory, as we are focusing on in this paper, we have $A_{\ell}^{\bar{3}\bar{4}\bar{1}\bar{2}}(s + i\epsilon) = A_{\ell}^{1234}(s + i\epsilon)$, in which case the absorptive part is simply the imaginary part: $\text{Abs } A_{\ell}^{1234}(s) = \text{Im } A_{\ell}^{1234}(s)$.

Now, let us derive the dispersion relations we will use later. The most important ingredient in deriving the dispersion relations is the analyticity of the amplitudes when s is analytically continued to the complex plane. While analyticity has not been rigorously proven, it is conjectured to be implied by causality of the UV theory (see [33] for a brief account and [123] for a recent discussion), justifying the name of causality bounds, and we shall take it as a fundamental assumption. More precisely, we will make use of the analyticity condition that for fixed t the amplitude $\mathcal{M}^{1234}(s, t)$ is analytic in the complex s plane except for singularities on the real s axis that can be readily inferred from unitarity. Additionally, we shall assume that our EFT is weakly coupled in the IR and take the leading tree level approximation below the EFT cutoff Λ . This means that we can take the approximation that the amplitudes do not have branch cuts on the real s axis in the low energy EFT region. That is, when $-t - \Lambda^2 < s < \Lambda^2$, the only singularities in the low energy amplitude \mathcal{M}^{1234} are poles from exchange diagrams calculable within the EFT. Beyond the cutoff, unknown UV poles and branch cuts can appear.

Then we can look at the quantity $\mathcal{M}^{1234}(\mu, t)/(\mu - s)$ in the complex μ plane for fixed s and t which are chosen to be in the EFT region $|s| < \Lambda^2, 0 \leq -t < \Lambda^2$. The analytic structure of this quantity is shown in figure 1, which allows us to perform the contour integration as depicted. Due to analyticity, the integration with the small (closed) contour, which is valid in the EFT, is equivalent to the big (closed) contour that goes around the UV branch cut and the infinity. We will refer to the pole at $\mu = s$ as well as other low energy poles of $\mathcal{M}^{1234}(\mu, t)$ as the “EFT poles”. For massless scalar-tensor theory we are considering, the only low energy poles of $\mathcal{M}^{1234}(\mu, t)$ for fixed t are at $\mu = 0$ and

$\mu = -t$. By the residue theorem, the big contour integral gives rise to

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s} = \int_{\Lambda^2}^{+\infty} \frac{d\mu}{\pi} \frac{\text{Abs}\mathcal{M}^{1234}(\mu, t)}{\mu - s} + \int_{-\infty}^{-t-\Lambda^2} \frac{d\mu}{\pi} \frac{\text{Abs}\mathcal{M}^{1234}(\mu, t)}{\mu - s} + \int_{C_{\infty}^{\pm}} \frac{d\mu}{2\pi i} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s}, \quad (3.4)$$

where we have made use of eq. (3.3) and C_{∞}^{\pm} denotes the upper and lower semi-circles at infinity. The second term on the right hand side can be written in a form similar to the first term by the su crossing of the amplitude and a change of the integration variable, so we get

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s} = \int_{\Lambda^2}^{+\infty} \frac{d\mu}{\pi} \left(\frac{\text{Abs}\mathcal{M}^{1234}(\mu, t)}{\mu - s} + \frac{\text{Abs}\mathcal{M}^{1432}(\mu, t)}{\mu - u} \right) + \int_{C_{\infty}^{\pm}} \frac{d\mu}{2\pi i} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s}. \quad (3.5)$$

The aforementioned equation in its current form is not particularly useful, as the two integrals on the right-hand side may not converge due to the UV behavior of the amplitude. Typically, in order to respect locality, momentum space scattering amplitudes are polynomially bounded in terms of the Mandelstam variables so that Fourier transforms to real space amplitudes are well-defined. However, the case for a theory with the massless graviton can be more delicate, as will be discussed shortly. Nevertheless, we shall assume that the UV theory is polynomially bounded such that for fixed t we have

$$\lim_{|s| \rightarrow \infty} \mathcal{M}(s, t)/s^N = 0, \quad (3.6)$$

where N is a positive integer that depends on the value of t , as will be explained shortly. To render eq. (3.5) useful, the standard remedy is to make “subtractions”. For an N subtraction, we can simply utilize the following algebraic identity

$$\frac{\text{Abs}\mathcal{M}(\mu, t)}{\mu - s} = \sum_{i=0}^N \binom{N}{i} \frac{(s - \mu_p)^{N-i} (\mu - s)^i}{(\mu - \mu_p)^N} \frac{\text{Abs}\mathcal{M}(\mu, t)}{\mu - s}, \quad (3.7)$$

where μ_p is the subtraction point that can be arbitrarily chosen and $\binom{N}{i} \equiv N!/(i!(N-i)!)$. Notice that, except for the $i = 0$ term, all the other terms in eq. (3.7) are just $(N-1)$ -th degree polynomials of s . Since the left hand side of eq. (3.5) is finite except for $t = 0$, the divergences on the right hand must cancel. So all the $i \neq 0$ terms on the right hand side eq. (3.7) must group into an $(N-1)$ -th-degree polynomial of s whose coefficients are finite functions of t , while the $i = 0$ term converges thanks to the high energy bound (3.6). Thus, eq. (3.5) can be re-written as an N -th subtracted dispersion relation:

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s} = \sum_{m=0}^{N-1} b_{(N)m}^{1234}(t) s^m + \int_{\Lambda^2}^{+\infty} \frac{d\mu}{\pi} \left(\frac{(s - \mu_s)^N}{(\mu - \mu_s)^N} \frac{\text{Abs}\mathcal{M}^{1234}(\mu, t)}{\mu - s} + \frac{(u - \mu_u)^N}{(\mu - \mu_u)^N} \frac{\text{Abs}\mathcal{M}^{1432}(\mu, t)}{\mu - u} \right), \quad (3.8)$$

where we have allowed the s and u channel subtraction points μ_s and μ_u to be different. Then, by the partial wave expansion (3.1) and the generalized optical theorem for the partial waves (3.2), we can get

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s} = \sum_{m=0}^{N-1} b_{(N)m}^{1234}(t) s^m \tag{3.9}$$

$$+ \left\langle \frac{(s - \mu_s)^N}{(\mu - \mu_s)^N} d_{h_{12}, h_{43}}^{\ell, \mu, t} c_{\ell, \mu}^{12} c_{\ell, \mu}^{*34} + \frac{(u - \mu_u)^N}{(\mu - \mu_u)^N} d_{h_{14}, h_{23}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{*32} \right\rangle,$$

where we have defined the shorthands

$$\langle \dots \rangle := 16\pi \sum_{\ell, X} (2\ell + 1) \int_{\Lambda^2}^{\infty} \frac{d\mu}{\pi} (\dots), \quad c_{\ell, s}^{12} := c_{\ell, s}^{12 \rightarrow X}, \quad d_{h_{12}, h_{43}}^{\ell, \mu, t} := d_{h_{12}, h_{43}}^{\ell} \left(1 + \frac{2t}{\mu}\right). \tag{3.10}$$

Note that each of the dispersion relations is actually a one-parameter family of relations parametrized by the momentum transfer t .

To determine the number of subtractions N , we need to have a better understanding of the Regge behavior of the amplitudes. Let us recall that for a non-gravitational massive field theory, the rigorous results of Froissart [117] and Martin [118] suggest that two subtractions are sufficient: $\lim_{|s| \rightarrow \infty} \mathcal{M}(s, t)/s^2 = 0$ for a range of physical $t \leq 0$ and even for a range of non-physical $t > 0$. For massless fields, especially when gravitons are included in the low energy spectrum, it is more subtle, not the least for the presence of the spin-2 t channel pole. Generically, one expects that for a gravitational theory the Regge behavior of the amplitude may change for different fixed t (see, e.g., [23, 124])

$$\begin{cases} \lim_{|s| \rightarrow \infty} \mathcal{M}(s, t)/s^2 = 0, & t < 0, \\ \lim_{|s| \rightarrow \infty} \mathcal{M}(s, t)/s^3 = 0, & 0 \leq t \leq \xi, \end{cases} \tag{3.11}$$

where ξ is a small positive number. While string theory gives rise to this behavior, it is believed to be generically valid for a theory with a spin-2 t -channel pole. Although the original Froissart bound does not apply for massless particles, twice subtracted dispersion relations in the physical region $t < 0$ is implied at least in the weak coupling limit by causality considerations for impact parameter amplitudes [23]. In any case, we shall assume that twice subtractions are sufficient for $t < 0$. Then, from twice-subtracted dispersion relations, say, \mathcal{M}^{++--} , in the $t \rightarrow 0^-$ limit

$$\lim_{t \rightarrow 0^-} \left(\int_{\Lambda^2}^{\infty} \frac{d\mu}{\pi} \frac{\text{Disc} \mathcal{M}^{++--}(s, t)}{\mu^2(\mu - s)} + s \leftrightarrow u \right) \sim -\frac{1}{t},$$

we can infer that the dispersive integral on the left hand side must diverge as $t \rightarrow 0^-$, since the integrand does not give rise to any negative power of t . However, a thrice subtraction eliminates the spin-2 t -channel pole s^2/t , and therefore, we have $\lim_{|s| \rightarrow \infty} \mathcal{M}(s, t)/s^3 = 0$ for $0 \leq t \leq \xi$. In this paper, we shall simply assume the Regge bounds of eq. (3.11) to hold. Since we will use the dispersion relations for the range of $t \leq 0$, N is chosen to be 2 for $t < 0$ and 3 for $t \leq 0$.

Therefore, for $t < 0$, choosing $\mu_s = \mu_u = 0$, we can get twice subtracted dispersion relations

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s} = b_{(2)0}^{1234}(t) + b_{(2)1}^{1234}(t)s + \left\langle \frac{s^2 d_{h_{12}, h_{43}}^{\ell, \mu, t} c_{\ell, \mu}^{12} c_{\ell, \mu}^{* \bar{3} \bar{4}}}{\mu^2(\mu - s)} + \frac{u^2 d_{h_{14}, h_{23}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{* \bar{3} \bar{2}}}{\mu^2(\mu - u)} \right\rangle. \quad (3.12)$$

For a su -symmetric amplitude, we additionally have $b_{(2)1}^{1234}(t) = 0$. Later, we will also use thrice subtracted dispersive relations at $t = 0$, which helps impose the st crossing symmetry of the amplitudes to get more useful dispersion relations. The use of forward-limit dispersive relations also helps harvest effective constraints numerically in the finite μ and large ℓ region. A remarkable feature of the dispersion relations (3.12) is that they link the EFT couplings in the IR (on the left hand side) to the unknown UV behaviors of the amplitudes (on the right hand side) via dispersive integrals. To see this more clearly, let us parametrize the residues of the EFT poles on the left hand side of eq. (3.12) as follows

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1234}(\mu, t)}{\mu - s} = a_{2,-1}^{1234} \frac{s^2}{t} + \sum_{k, n \geq 0} a_{k,n}^{1234} s^k t^n. \quad (3.13)$$

The $a_{k,n}^{1234}$ coefficients can be easily expressed in terms of the independent g coefficients introduced in eqs. (2.3)–(2.11) or in terms of the Lagrangian Wilson coefficients via eqs. (2.22)–(2.30). For a particular EFT amplitude, some of the a^{1234} coefficients can vanish. The term s^2/t comes from a t -channel exchange of the massless graviton. This prevents a Taylor expansion in terms of t in the forward limit $t = 0$ for the two sides of these dispersion relations. For some of the twice-subtracted dispersion relations that do not contain t -channel pole, this pathology also manifests as the fact that the expansions at $t = 0$ on the two sides can not be matched without imposing unphysical restrictions on the Wilson coefficients. (For the twice-subtracted dispersion relations listed in appendix B, those of \mathcal{M}^{0000} , \mathcal{M}^{+0-0} and \mathcal{M}^{++--} contain the s^2/t pole, while expanding those of \mathcal{M}^{+000} , \mathcal{M}^{++++} , \mathcal{M}^{+0+0} , \mathcal{M}^{+-00} , \mathcal{M}^{0-++} and \mathcal{M}^{0+-+} will impose unphysical constraints on the Wilson coefficients.) For example, if we expand the right hand side of the \mathcal{M}^{+000} dispersion relation around $t = 0$, the series of t within $\langle \rangle$ begins with t^2 because of the structure of $d_{2,0}^\ell$, which implies that the coefficient of the st term on the left hand side must be zero, i.e., $\beta_1 = 0$. This clearly is an unphysical constraint, meaning that it is invalid to expand around $t = 0$ even for those dispersion relations. Even if the two sides of a twice-subtracted dispersion relation could be matched for the expansion around $t = 0$, we might still not use its forward limit simply because of the Regge behavior eq. (3.11) of the amplitude. Nevertheless, since $\mathcal{M}^{1234}(\mu, t)$ only contains simple poles in the EFT region, the left hand side of eq. (3.12) is analytic around $s = 0$, as shown explicitly in eq. (3.13). We can Taylor-expand both sides of eq. (3.12) in the neighborhood of $s = 0$, and matching coefficients of s^k gives

$$\delta_{k,2} a_{k,-1}^{1234} \frac{1}{t} + \sum_{n=0} a_{k,n}^{1234} t^n = \left\langle \frac{\partial_s^k}{k!} \left[\frac{s^2 d_{h_{12}, h_{43}}^{\ell, \mu, t} c_{\ell, \mu}^{12} c_{\ell, \mu}^{* \bar{3} \bar{4}}}{\mu^2(\mu - s)} + \frac{(-s - t)^2 d_{h_{14}, h_{23}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{* \bar{3} \bar{2}}}{\mu^2(\mu + s + t)} \right] \Big|_{s \rightarrow 0} \right\rangle, \quad (3.14)$$

which for fixed k and n is a one-parameter (t) family of sum rules. If $\mathcal{M}^{1234}(s, t)$ is su -symmetric, eq. (3.14) is valid for $k \geq 1$, because in this case we have $b_{(2)1}^{1234}(t) = 0$; if

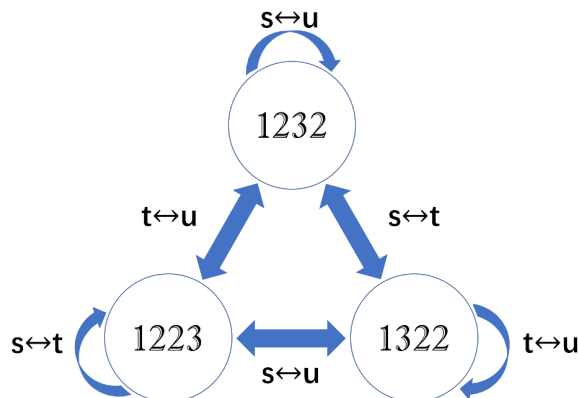


Figure 2. Crossing relations for the amplitudes with only one crossing symmetry. The su crossing symmetry for $\mathcal{M}^{1234}(s, t, u)$ is built-in in our dispersion relations.

\mathcal{M}^{1234} is not su -symmetric, eq. (3.14) is valid for $k \geq 2$, remembering that $b_{(2)1}^{1234}(t)$ is then generically nonzero and unknown. That is, for su -symmetric amplitudes, we have some extra sum rules. These extra low order sum rules are constraining in bounding the Wilson coefficients, so it is important to make use of them effectively.

3.2 Imposing st crossing symmetry

In deriving the sum rules (3.14), we have already used the su crossing symmetry of the amplitudes. However, that is not the full crossing symmetry that the amplitudes have. We also have the st crossing symmetry, whose information is not contained in the sum rules (3.14). It has been realized recently that imposing the st crossing symmetry on the su dispersion relations is very potent in improving causality bounds on the Wilson coefficients [34, 35].

As an aside, note that in the absence of gravitational interactions, dispersion relations can be expanded in the forward limit as well as around $s = 0$, and one can express individual amplitude coefficients in terms of UV dispersive integrals. In that case, the st crossing symmetry directly links different amplitude coefficients, giving rise to vanishing dispersive integrals, known as null constraints. For a theory with multiple degrees of freedom, the coefficient sum rules and the null constraints can be combined to define a SDP with one continuous decision variable [37], solvable by the powerful SDPB package. In the presence of the massless graviton, the expansion in the forward limit is invalid, and we need to be content with sum rules where the left hand sides generally contain the momentum transfer t . This will also be usually true after imposing the st crossing symmetry, as shown in appendix B.

The tu crossing symmetry is implied by the su plus st crossing symmetry, so we do not need additionally impose the tu crossing. Let us see how to implement this concretely in our case. In the massless scalar-tensor theory, there are two kinds of amplitudes: the ones that are fully stu symmetric, whose crossing symmetries have been listed in eqs. (2.13)–(2.17), and the ones with only one of the su , st and tu symmetries, whose crossing symmetries and relations have been listed in eqs. (2.18)–(2.21). For the fully crossing symmetric cases,

after imposing the st crossing symmetry, we can easily see that the tu crossing symmetry is redundant. For the amplitudes with only one crossing symmetry, there are three different types: \mathcal{M}^{1232} , \mathcal{M}^{1322} and \mathcal{M}^{1223} . Crossing then either maps one amplitude into itself or into another amplitude, see figure 2. Again, since we have used the su crossing symmetry, it is sufficient to impose the st crossing symmetry, $\mathcal{M}^{1232}(s, t, u) = \mathcal{M}^{1322}(t, s, u)$, to extract the full crossing information. We would like to remind the reader that we use the terminology that crossing symmetry refers to the collection of the crossing symmetries that map one amplitude to itself and crossing relations that map one amplitude to another.

To impose the st crossing symmetry, we first note that the amplitudes with full stu symmetry separate into 5 groups and the amplitudes with only su , st or ut crossing symmetry separate into 4 groups. The st crossing relations are imposed separately for each of these groups, which can be done by equating the following EFT coefficients in the expansion (3.13):

$$a_{k,n}^{1234} = a_{n,k}^{1324}, \quad n \geq 3, \quad (3.15)$$

where $k \geq 1$ if \mathcal{M}^{1234} is su -symmetric in the narrow sense and $k \geq 2$ if \mathcal{M}^{1234} is not su -symmetric.

Later, for technical reasons, we shall try to access dispersion relations when t is close to the cutoff $-\Lambda^2$, for which eq. (3.14) is not suitable. This is simply because the left hand side of eq. (3.14) contains an infinite number of powers of t , which all become important when t approaches $-\Lambda^2$. However, this can be overcome by combining different dispersion relations. To this end, we shall also make use of thrice subtracted dispersion relations. In eq. (3.9), we can choose the subtraction points to be $\mu_s = 0$ and $\mu_u = -t$, and get

$$\sum_{\text{EFT poles}} \text{Res} \frac{\mathcal{M}^{1324}(\mu, t)}{\mu - s} = \sum_{m=0}^2 b_{(3)m}^{1324}(t) s^m + \left\langle \frac{s^3 d_{h_{13}, h_{42}}^{\ell, \mu, t} c_{\ell, \mu}^{13} c_{\ell, \mu}^{* \bar{2}4}}{\mu^3 (\mu - s)} + \frac{(-s)^3 d_{h_{14}, h_{32}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{* \bar{2}3}}{(\mu + t)^3 (\mu - u)} \right\rangle. \quad (3.16)$$

Since these are thrice subtracted dispersive sum rules, which are free of the t -channel pole issue, we can then express both sides of eq. (3.16) as a Taylor series of t and match the expansion coefficients. The choice of $\mu_s = 0$ and $\mu_u = -t$ makes sure that the part within $\langle \dots \rangle$ only contains terms with s^3 and higher orders. This leads to

$$\sum_{k=3} a_{k,n}^{1324} s^k = \left\langle \frac{\partial_t^n}{n!} \left(\frac{s^3 d_{h_{13}, h_{42}}^{\ell, \mu, t} c_{\ell, \mu}^{13} c_{\ell, \mu}^{* \bar{2}4}}{\mu^3 (\mu - s)} + \frac{(-s)^3 d_{h_{14}, h_{32}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{* \bar{2}3}}{(\mu + t)^3 (\mu + s + t)} \right) \Big|_{t \rightarrow 0} \right\rangle, \quad (3.17)$$

where $n \geq 0$. Then, we can relabel s as t in eq. (3.17), and subtract eq. (3.14) with this s and t swapped equation. This gives the final st crossing imposed sum rules that we will use in a SDP problem to get the causality bounds

$$\delta_{k,2} a_{2,-1}^{1234} \frac{1}{t} + a_{k,0}^{1234} + a_{k,1}^{1234} t + a_{k,2}^{1234} t^2 = \left\langle F_{k,\ell}^{1234}(\mu, t) \right\rangle, \quad (3.18)$$

with $F_{k,\ell}^{1234}(\mu, t)$ defined as

$$F_{k,\ell}^{1234}(\mu, t) = \frac{\partial_s^k}{k!} \left(\frac{s^2}{\mu^2(\mu-s)} d_{h_{12}, h_{43}}^{\ell, \mu, t} c_{\ell, \mu}^{12} c_{\ell, \mu}^{*3\bar{4}} + \frac{(-s-t)^2}{\mu^2(\mu+s+t)} d_{h_{14}, h_{23}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{*3\bar{2}} \right) \Big|_{s \rightarrow 0} \quad (3.19)$$

$$- \frac{\partial_t^k}{k!} \left(\frac{s^3}{\mu^3(\mu-s)} d_{h_{13}, h_{42}}^{\ell, \mu, t} c_{\ell, \mu}^{13} c_{\ell, \mu}^{*2\bar{4}} + \frac{(-s)^3}{(\mu+t)^3(\mu+s+t)} d_{h_{14}, h_{32}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{*2\bar{3}} \right) \Big|_{t \rightarrow 0, s \rightarrow t},$$

where we have used the st crossing symmetry $a_{k,n}^{1234} = a_{n,k}^{1324}$ to cancel all the t^n terms with $n \geq 3$ and $k \geq 1$ if \mathcal{M}^{1234} is su -symmetric and $k \geq 2$ if \mathcal{M}^{1234} is not su -symmetric. These sum rules are under control even if t is close to $-\Lambda^2$. These explicit independent sum rules are listed in appendix B.

A few comments are in order. In eq. (3.15), we have only imposed st crossing relations $a_{k,n}^{1234} = a_{n,k}^{1324}$ for $n \geq 3$. In principle, we could also impose the condition $a_{2,1}^{1234} = a_{1,2}^{1324}$. However, for an stu symmetric amplitude, this is redundant, because we have already enforced $a_{2,1}^{1222} = a_{1,2}^{1222}$ when deriving the dispersion relation with the su crossing symmetry — the only su -symmetric terms at that order are stu and t^3 . For an amplitude with only one crossing symmetry, the crossing relation $a_{1,2}^{1232} = a_{2,1}^{1322}$ does provide some new information. However, since we will for our convenience use both the sum rules involving $a_{1,2}^{1232}$ and $a_{2,1}^{1322}$, it is equivalent to imposing crossing relation $a_{1,2}^{1232} = a_{2,1}^{1322}$. Using two different expressions for one Wilson coefficient is the same as using one expression for the coefficient plus one st crossing relation.

Note that sometimes the requirement of $a_{k,n}^{1234} = a_{n,k}^{1324}$ for an amplitude with stu symmetry can be redundant, since the st symmetry is occasionally guaranteed by the su symmetry already. To find redundant relations at the N -th order ($N = k + n$), we can first expand an amplitude at the N -th order as $\mathcal{M}_N = \sum_{m=0}^{\lfloor N/2 \rfloor} q_m (su)^m t^{N-2m}$, where $\lfloor \cdot \rfloor$ denotes taking the flooring integer. Further expanding u as $-s - t$, we get $\mathcal{M}_N = \sum_{k=0}^N c_{k,n} s^k t^n$, which allows us to write $c_{k,n}$ in terms of q_m . Then, requiring $c_{k,n} = c_{n,k}$ gives a set of linear equations in terms of q_m , and the redundancy of the st symmetry can be obtained by examining the linear dependence of these q_m equations. Let us take the case of scalar scattering for an example, whose amplitude is stu symmetric. When $k + n = 3$, the su symmetry requires that the terms of the amplitude must be stu or t^3 , which means that, without further imposing the st symmetry, we can already have $a_{2,1}^{0000} = a_{1,2}^{0000}$. So in this case the st symmetry is redundant. In fact, since the st symmetry results in $\lfloor (N+1)/2 \rfloor$ equations and there are only $\lfloor N/2 \rfloor + 1$ distinct values of q_m , redundancy always exists.

In principle, the sum rules in the form of eq. (3.18) are all one needs to extract the strongest causality bounds in an ideal optimization scheme. However, to have a scheme that is numerically more tractable, we find that it is beneficial to add some forward-limit sum rules, as will be discussed in section 5.2. The forward-limit sum rules can be obtained from eq. (3.18) by simply matching the coefficients in front of t^n on both sides of the equation for the cases of $k \geq 3$:

$$a_{k,0}^{1234} = \left\langle F_{k,\ell}^{1234}(\mu, 0) \right\rangle, \quad a_{k,1}^{1234} = \left\langle \partial_t F_{k,\ell}^{1234}(\mu, 0) \right\rangle, \quad a_{k,2}^{1234} = \frac{1}{2} \left\langle \partial_t^2 F_{k,\ell}^{1234}(\mu, 0) \right\rangle, \quad (3.20)$$

$$0 = \left\langle \partial_t^n F_{k,\ell}^{1234}(\mu, 0) \right\rangle, \quad \text{for } n \geq 3.$$

4 Power counting via dispersion relations

The dispersive sum rules we have derived can be used to constrain Wilson coefficients of the low energy EFT via an optimization procedure. Before doing that numerically in the next sections, we will see here that these sum rules can be used to do a dimensional analysis on the Wilson coefficients. That is, we will show how schematic estimates on the dimensions of the coefficients can be inspected from the structure of the dispersion relations.

Recall that in the absence of gravity the dimensional analysis of a scalar EFT is usually fairly simple. One just power-counts the mass dimension of an operator and suppresses it with appropriate powers of the cutoff:

$$\hat{\mathcal{O}}_\phi \sim \Lambda^4 \left[\frac{\partial}{\Lambda} \right]^{N_\partial} \left[\frac{\phi}{\Lambda} \right]^{N_\phi}, \quad (4.1)$$

where N_∂ is the number of partial derivatives and N_ϕ is the number of ϕ fields in the operator. A slightly more refined version of this analysis which takes care of loops and factors of 4π , called naive dimensional analysis, can be extended to include spin-1 and spin-1/2 fields [125]. In the presence of gravity, an extra mass scale comes in at the (reduced) Planck mass $M_P = 1/\sqrt{8\pi G_N}$. Then, an important question is how many powers of M_P there are in each of the Wilson coefficients. In the literature, there are a few seemingly plausible arguments supporting different scalings of the Wilson coefficients in terms of M_P . In the case of pure gravity that is weakly coupled in the IR, the numerical bounds from causality imply [44] that the typical scalings for generic gravitational EFT operators are given by

$$\hat{\mathcal{O}}_R \sim M_P^2 \Lambda^2 \left[\frac{\nabla}{\Lambda} \right]^{N_\nabla} \left[\frac{R}{\Lambda^2} \right]^{N_R}, \quad (4.2)$$

where N_∇ is the number of covariant derivatives, R stands for a curvature tensor and N_R is the number of curvature tensors. In the following, we shall argue that, in scalar-tensor theory, if the scaling of eq. (4.1) is recovered in the decoupling limit, the typical scalings of the EFT operators are given by

$$\hat{\mathcal{O}}_{\phi R} \sim M_P^2 \Lambda^2 \left[\frac{\nabla}{\Lambda} \right]^{N_\nabla} \left[\frac{R}{\Lambda^2} \right]^{N_R} \left[\frac{\phi}{M_P} \right]^{N_\phi} \left[\frac{M_P}{\Lambda} \right]^{\tilde{N}_\phi}, \quad (4.3)$$

where the power of the enhancement factor \tilde{N}_ϕ can be determined by counting the number of $c_{\ell,\mu}^{00}$ in the most constraining sum rule available. For the lowest orders in eq. (4.29), it happens that $\tilde{N}_\phi = \lfloor N_\phi/2 \rfloor$, where $\lfloor \cdot \rfloor$ denotes taking the flooring integer, but this has to be modified for higher orders (see section 6.5). On the other hand, for the scenario where the scalar interactions are of the gravitational strength, a typical scalar-tensor operator then has the following scaling

$$\hat{\mathcal{O}}_{\phi R}^{\text{st}} \sim M_P^2 \Lambda^2 \left[\frac{\nabla}{\Lambda} \right]^{N_\nabla} \left[\frac{R}{\Lambda^2} \right]^{N_R} \left[\frac{\phi}{M_P} \right]^{N_\phi}. \quad (4.4)$$

Of course, a caveat is that the above scalings have only been explicitly verified for EFT operators of the lowest orders with four fields in a weakly coupled EFT; see eq. (4.29) and eq. (4.30).

To see how this schematic method works, we shall first use the sum rules without the st crossing symmetry imposed, i.e., eq. (3.14), to infer the typical behaviors of the UV spectral functions $c_{\ell,\mu}^{12}$. Let us first look at the $++--$ sum rule with $k=2$, which happens to be the same as sum rule (B.55). That is, the st crossing does not alter this sum rule. Its explicit form is given by

$$-\frac{1}{M_P^2 t} = \sum_{\ell, X} 16\pi(2\ell+1) \int_{\Lambda^2}^{\infty} \frac{d\mu}{\pi} \left[\frac{d_{0,0}^{\ell,\mu,t} |c_{\ell,\mu}^{++}|^2}{\mu^3} + \frac{d_{4,4}^{\ell,\mu,t} |c_{\ell,\mu}^{+-}|^2}{(\mu+t)^3} \right]. \quad (4.5)$$

The left hand side comes from a t -channel exchange, and this sum rule is valid for a range of $|t|$ below the cutoff. When $|t|$ is small, the left hand side is large, which means that the integral over μ red and/or the sum on the right hand side converges very slowly. A quicker convergence can be achieved by choosing a large $|t|$, so for our estimates we shall choose $|t| \sim \Lambda^2$. Also, this choice does not introduce any extra scale that is not already in the problem. Introducing dimensionless variables \hat{t} and $\hat{\mu}$ and normalized $\hat{c}_{\ell,\mu}^{12}$:

$$\hat{t} = \frac{t}{\Lambda^2}, \quad \hat{\mu} = \frac{\mu}{\Lambda^2}, \quad \hat{c}_{\ell,\mu}^{12} = \sqrt{16(2\ell+1)} c_{\ell,\mu}^{12}, \quad (4.6)$$

we get

$$\frac{\Lambda^2}{M_P^2} = -\hat{t} \sum_{\ell, X} \int_1^{\infty} d\hat{\mu} \left[\frac{d_{0,0}^{\ell,\hat{\mu},\hat{t}} |\hat{c}_{\ell,\mu}^{++}|^2}{\hat{\mu}^3} + \frac{d_{4,4}^{\ell,\hat{\mu},\hat{t}} |\hat{c}_{\ell,\mu}^{+-}|^2}{(\hat{\mu}+\hat{t})^3} \right]. \quad (4.7)$$

Since the quantities on the right hand side are mostly $\mathcal{O}(1)$ numerically except for $\hat{c}_{\ell,\mu}^{++}$ and $\hat{c}_{\ell,\mu}^{+-}$, this means that $\hat{c}_{\ell,\mu}^{++}$ and $\hat{c}_{\ell,\mu}^{+-}$ must behave appropriately to make the integral and summation converge to the left hand side. That is, the spectral functions $c_{\ell,\mu}^{++}$ and $c_{\ell,\mu}^{+-}$ have to conspire to reproduce the hierarchy between Λ and M_P in the theory. Thus, we can schematically assign the following correspondences

$$\frac{\Lambda}{M_P} \Leftrightarrow \hat{c}_{\ell,\mu}^{++}, \hat{c}_{\ell,\mu}^{+-}, \hat{c}_{\ell,\mu}^{-+}, \hat{c}_{\ell,\mu}^{--}, \quad (4.8)$$

which can be used to estimate the sizes of the Wilson coefficients momentarily. Note that we have also added $\hat{c}_{\ell,\mu}^{-+}$ and $\hat{c}_{\ell,\mu}^{--}$ because they are related to $\hat{c}_{\ell,\mu}^{+-}$ and $\hat{c}_{\ell,\mu}^{++}$ by crossing or parity, and thus they must have the same scaling. In establishing the correspondences such as (4.8), the reason for not using the sum rules with the st crossing symmetry imposed is obvious: the st crossing introduces quantities that are cancelable among themselves. For example, the null sum rule (B.49) would not tell us any scaling in terms of Λ and M_P ; it only tells us that there are intricate cancellation among the terms with $\hat{c}_{\ell,\mu}^{++}$, $\hat{c}_{\ell,\mu}^{+-}$, $\hat{c}_{\ell,\mu}^{-+}$ and $\hat{c}_{\ell,\mu}^{--}$. Similarly, even though the sum rule (B.27) is not null on the left hand side, its right hand side contains terms that cancel among themselves, so it would be inappropriate to use it to estimate the behavior of $c_{\ell,\mu}^{12}$.

With these established, we can estimate the sizes of the Wilson coefficients γ_0 and β_1 via the improved sum rules in appendix B. Specifically, we can expand eq. (B.56) around

the forward limit and match the coefficients to get

$$-\frac{\gamma_0^2}{M_P^6} = \frac{1}{\Lambda^{10}} \sum_{\ell, X} \int_1^\infty d\hat{\mu} \left[\frac{\partial_{\hat{t}}^2 d_{0,0}^{\ell, \hat{\mu}, 0} |\hat{c}_{\ell, \mu}^{++}|^2}{2\hat{\mu}^4} - \frac{\partial_{\hat{t}}^2 d_{4,4}^{\ell, \hat{\mu}, 0} |\hat{c}_{\ell, \mu}^{+-}|^2}{2\hat{\mu}^4} + \frac{4\partial_{\hat{t}} d_{4,4}^{\ell, \hat{\mu}, 0} |\hat{c}_{\ell, \mu}^{+-}|^2}{\hat{\mu}^5} - \frac{10d_{4,4}^{\ell, \hat{\mu}, 0} |\hat{c}_{\ell, \mu}^{+-}|^2}{\hat{\mu}^6} \right], \quad (4.9)$$

$$-\frac{\beta_1^2}{M_P^4} = \frac{1}{\Lambda^6} \sum_{\ell, X} \int_1^\infty d\hat{\mu} \left[\frac{d_{0,0}^{\ell, \hat{\mu}, 0} |\hat{c}_{\ell, \mu}^{++}|^2}{\hat{\mu}^4} - \frac{d_{4,4}^{\ell, \hat{\mu}, 0} |\hat{c}_{\ell, \mu}^{+-}|^2}{\hat{\mu}^4} \right]. \quad (4.10)$$

Making use of the scaling correspondences (4.8), we can infer that the typical dimensional scaling of these two Wilson coefficients must be ¹

$$\gamma_0 \sim \frac{M_P^2}{\Lambda^4}, \quad \beta_1 \sim \frac{M_P}{\Lambda^2}. \quad (4.11)$$

As we will see in section 6, this is consistent with the rigorous numerical results, that is, the upper limits of the causality bounds.

One caveat is in order. Since the sum rules in appendix B are with the st crossing symmetry imposed, sometimes a coefficient's dimensional scaling from one sum rule may differ from another. In this case, one should survey all available sum rules and take the smallest dimensional scaling as the *bona fide* one. The reason for the difference from different sum rules is that these sum rules are with st crossing imposed so as to pick out a finite number of Wilson coefficients on the left hand side, but this procedure also introduces null constraints in the sum rules. That is, there are $\hat{c}_{\ell, \mu}^{12}$ terms that cancel among themselves on the right hand side of the sum rule without affecting the Wilson coefficients, and these terms may have an unusually larger scale, pessimistically overestimating the scaling of the coefficient.

To estimate the sizes of other Wilson coefficients, we also want to establish scale correspondences for the rest of the UV spectral functions $\hat{c}_{\ell, \mu}^{+0}$, $\hat{c}_{\ell, \mu}^{-0}$ and $\hat{c}_{\ell, \mu}^{00}$ that involve the scalar. For $\hat{c}_{\ell, \mu}^{+0}$, we can use the $+0-0$ sum rule of eq. (3.14) with $k = 1$, which happens to be eq. (B.36). Making use of the correspondences (4.8) and the scaling (4.11), we get

$$\frac{\Lambda^2}{M_P^2} \sim \sum_{\ell, X} \int_1^\infty d\hat{\mu} \frac{\hat{t}(2\hat{\mu} + \hat{t}) d_{2,2}^{\ell, \hat{\mu}, \hat{t}}}{\hat{\mu}^2(\hat{\mu} + \hat{t})^2} |\hat{c}_{\ell, \mu}^{+0}|^2. \quad (4.12)$$

Thus, we see that $\hat{c}_{\ell, \mu}^{+0}$ (and hence $\hat{c}_{\ell, \mu}^{-0}$) leads to the same scale correspondence as those only involving the graviton:

$$\frac{\Lambda}{M_P} \Leftrightarrow \hat{c}_{\ell, \mu}^{+0}, \hat{c}_{\ell, \mu}^{-0}, c_{\ell, \mu}^{0+}, \hat{c}_{\ell, \mu}^{0-}. \quad (4.13)$$

For $\hat{c}_{\ell, \mu}^{00}$, eq. (3.14) does not give any readily usable dispersion relation to infer its size in terms of the hierarchy between Λ and M_P . This is of course not surprising, as we should be able to define a scalar theory in the decoupling limit of the graviton where $M_P \rightarrow \infty$ and Λ is held fixed. So in principle $c_{\ell, \mu}^{00}$ should be able to reach its partial wave unitarity limit

¹By the typical scaling of, say, γ_0 , we mean that the upper bound of $|\gamma_0|$ is around $|\gamma_0| \sim M_P^2/\Lambda^4$.

$c_{\ell,\mu}^{00} \sim 1$. With a mild assumption in the spirit of lower spin dominance $c_{\ell,\mu}^{00} \sim \ell^{-1/2}$, we can then have the scaling correspondence $1 \Leftrightarrow \hat{c}_{\ell,\mu}^{00}$. This correspondence is also consistent with the pure scalar sum rules in the decoupling limit, which can be expanded in the forward limit and schematically goes like

$$a_{k,n}^{0000} \Lambda^{2k+2n} = \sum_{\ell,X} \int_1^\infty d\hat{\mu}(\dots) |\hat{c}_{\ell,\mu}^{00}|^2, \quad (4.14)$$

leading to the usual dimensional analysis in the pure scalar theory: $a_{k,n}^{0000} \sim \Lambda^{-2k-2n}$. Away from the decoupling limit, the 0000 sum rule schematically goes like

$$\frac{\Lambda^2}{M_P^2 \hat{t}} + \sum_n a_{k,n}^{0000} \Lambda^{2k+2n} \hat{t}^n = \sum_{\ell,X} \int_1^\infty d\hat{\mu}(\dots) |\hat{c}_{\ell,\mu}^{00}|^2, \quad (4.15)$$

which contains an extra subdominant $1/M_P^2$ term when $\Lambda \ll M_P$, so it is also consistent with the $1 \Leftrightarrow \hat{c}_{\ell,\mu}^{00}$ scaling. For the lowest order terms, from sum rule (B.2) or (B.3), we see that the scalar self-couplings α and γ_4 must scale as

$$\alpha \sim 1/\Lambda^4, \quad \gamma_4 \sim 1/\Lambda^6. \quad (4.16)$$

On the other hand, in scalar-tensor theory, an interesting parameter regime is when the interactions involving the scalar are comparable with those of the pure gravity, in which case one may view the scalar more as part of gravity rather than some non-minimally coupled matter field. This occurs when the first term is comparable with the rest of the terms on the left hand side of eq. (4.15), which implies a suppressed UV spectral function and the correspondence $\Lambda/M_P \Leftrightarrow \hat{c}_{\ell,\mu}^{00}$. In this case, we then have $\alpha \sim 1/(M_P^2 \Lambda^2)$ and $\gamma_4 \sim 1/(M_P^2 \Lambda^4)$. Thus, for $\hat{c}_{\ell,\mu}^{00}$, we may consider the following two scenarios

$$\begin{cases} 1 \Leftrightarrow \hat{c}_{\ell,\mu}^{00} & \implies \alpha \sim \frac{1}{\Lambda^4}, \gamma_4 \sim \frac{1}{\Lambda^6}, \\ \frac{\Lambda}{M_P} \Leftrightarrow \hat{c}_{\ell,\mu}^{00} & \implies \alpha \sim \frac{1}{M_P^2 \Lambda^2}, \gamma_4 \sim \frac{1}{M_P^2 \Lambda^4}. \end{cases} \quad (4.17)$$

While the first scenario gives the boundary of the causality bounds, the second scenario is more relevant when the scalar plays a significant role in the dynamics, which is phenomenologically more interesting. In the following, we shall discuss the typical scales of the other Wilson coefficients with both the two scenarios in mind.

Now, we are ready to deduce the dimensional scalings of the other Wilson coefficients from the scalings of $\hat{c}_{\ell,\mu}^{12}$ from the sum rules in appendix B. Let us now look at the γ_1 coefficient. From the $F_{1,\ell}^{++++0}$ sum rule (B.12) (using eq. (B.13) would be similar), we get

$$-\frac{\Lambda^6 \gamma_1}{M_P^3} = \sum_{\ell,X} \int_1^\infty d\hat{\mu} \left[\frac{(2\hat{\mu} - 3\hat{t}) d_{2,0}^{\ell,\hat{\mu},\hat{t}} \hat{c}_{\ell,\mu}^{+0} \hat{c}_{\ell,\mu}^{*-,-}}{\hat{t} \hat{\mu}^4} - \frac{\hat{t} \partial_{\hat{t}} d_{0,-2}^{\ell,\hat{\mu},0} \hat{c}_{\ell,\mu}^{++} \hat{c}_{\ell,\mu}^{*-,-0}}{\hat{\mu}^3 (\hat{\mu} - \hat{t})} + \frac{\hat{t} \partial_{\hat{t}} d_{2,0}^{\ell,\hat{\mu},0} \hat{c}_{\ell,\mu}^{+0} \hat{c}_{\ell,\mu}^{*-,-}}{\hat{\mu}^3 (\hat{\mu} + \hat{t})} \right]. \quad (4.18)$$

By the scale correspondences (4.8) and (4.13), we infer that the typical scale of γ_1 is

$$\gamma_1 \sim \frac{M_P}{\Lambda^4}. \quad (4.19)$$

Note that this is independent of the value of α , which is consistent with the numerical result in section 6. Next, we look at γ_2 , for which we can use the $F_{3,\ell}^{++00}$ sum rule, whose explicit form in the forward limit is given by

$$\frac{\Lambda^6 \beta_1^2}{M_P^4} + \frac{\Lambda^6 \gamma_2}{M_P^2} = \sum_{\ell, X} \int_1^\infty d\hat{\mu} \left[\frac{d_{0,0}^{\ell, \hat{\mu}, 0}}{\hat{\mu}^4} \hat{c}_{\ell, \hat{\mu}}^{++} \hat{c}_{\ell, \hat{\mu}}^{*,00} - \frac{d_{2,2}^{\ell, \hat{\mu}, 0}}{\hat{\mu}^4} c_{\ell, \hat{\mu}}^{+0} c_{\ell, \hat{\mu}}^{*,0-} \right]. \quad (4.20)$$

By the scale correspondences (4.8), (4.13) and (4.17), we can infer that

$$\gamma_2 \sim \frac{M_P}{\Lambda^5} \quad \text{when } \alpha \sim \frac{1}{\Lambda^4}, \quad (4.21)$$

$$\gamma_2 \sim \frac{1}{\Lambda^4} \quad \text{when } \alpha \sim \frac{1}{M_P^2 \Lambda^2}. \quad (4.22)$$

Again, this is consistent with the numerical results in the next sections, and the dependence on α is also observed there. Then, we look at the γ_3 coefficient, for which we can use the $F_{1,\ell}^{+000}$ sum rule (B.8),

$$\frac{\beta_1 \Lambda^4}{M_P^3} \hat{t} - \frac{\gamma_3 \Lambda^6}{M_P} \hat{t}^2 = \sum_{\ell, X} \int_1^\infty d\hat{\mu} \left(\frac{\hat{t}(2\hat{\mu} - 3\hat{t}) d_{2,0}^{\ell, \hat{\mu}, \hat{t}}}{\hat{\mu}^4} \hat{c}_{\ell, \hat{\mu}}^{+0} \hat{c}_{\ell, \hat{\mu}}^{*,00} + \frac{2\hat{t}^4 \partial_{\hat{t}} d_{2,0}^{\ell, \hat{\mu}, 0}}{\hat{\mu}^3 (\hat{t}^2 - \hat{\mu}^2)} \hat{c}_{\ell, \hat{\mu}}^{+0} \hat{c}_{\ell, \hat{\mu}}^{*,00} \right). \quad (4.23)$$

We already know that $\beta_1 \sim M_P/\Lambda^2$, so by the scale correspondences (4.13) and (4.17), we can infer that

$$\gamma_3 \sim \frac{1}{\Lambda^5} \quad \text{when } \alpha \sim \frac{1}{\Lambda^4}, \quad (4.24)$$

$$\gamma_3 \sim \frac{1}{M_P \Lambda^4} \quad \text{when } \alpha \sim \frac{1}{M_P^2 \Lambda^2}. \quad (4.25)$$

We also want to look at the typical size of the coefficient β_2 , which can be inferred from the $F_{2,\ell}^{++00}$ sum rule (B.24)

$$\begin{aligned} \frac{\beta_2 \Lambda^4}{M_P^2} - \frac{\gamma_0 \Lambda^6}{M_P^4} \hat{t} - g_{2,1}^{M_3} \Lambda^8 \hat{t}^2 = \sum_{\ell, X} \int_1^\infty d\hat{\mu} \left(\frac{d_{0,0}^{\ell, \hat{\mu}, \hat{t}} \hat{c}_{\ell, \hat{\mu}}^{++} \hat{c}_{\ell, \hat{\mu}}^{*,00}}{\hat{\mu}^3} + \frac{d_{2,2}^{\ell, \hat{\mu}, \hat{t}} \hat{c}_{\ell, \hat{\mu}}^{+0} \hat{c}_{\ell, \hat{\mu}}^{*,0-}}{(\hat{\mu} + \hat{t})^3} + \frac{\hat{t}^4 \partial_{\hat{t}}^2 d_{2,-2}^{\ell, \hat{\mu}, \hat{t}} \hat{c}_{\ell, \hat{\mu}}^{+0} \hat{c}_{\ell, \hat{\mu}}^{*, -0}}{2\hat{\mu}^3 (\hat{t}^2 - \hat{\mu}^2)} \right. \\ \left. - \frac{\hat{t}^3 (4\hat{\mu} + 3\hat{t}) \partial_{\hat{t}} d_{2,-2}^{\ell, \hat{\mu}, \hat{t}} \hat{c}_{\ell, \hat{\mu}}^{+0} \hat{c}_{\ell, \hat{\mu}}^{*, -0}}{\hat{\mu}^4 (\hat{t} + \hat{\mu})^2} + \frac{\hat{t}^3 (10\hat{\mu}^2 + 15\hat{\mu}\hat{t} + 6\hat{t}^2) d_{2,-2}^{\ell, \hat{\mu}, \hat{t}} \hat{c}_{\ell, \hat{\mu}}^{+0} \hat{c}_{\ell, \hat{\mu}}^{*, -0}}{\hat{\mu}^5 (\hat{t} + \hat{\mu})^3} \right). \quad (4.26) \end{aligned}$$

By the scale correspondences (4.8), (4.13) and (4.17), this gives us

$$\beta_2 \sim \frac{M_P}{\Lambda^3} \quad \text{when } \alpha \sim \frac{1}{\Lambda^4}, \quad (4.27)$$

$$\beta_2 \sim \frac{1}{\Lambda^2} \quad \text{when } \alpha \sim \frac{1}{M_P^2 \Lambda^2}. \quad (4.28)$$

As mentioned, all of these will be confirmed with the rigorous numerical results in section 6. Nevertheless, the scaling exercises above guide us to perform the numerical optimizations as they outline the rough boundaries of the causality bounds.

In summary, by simply inspecting the dispersive sum rules, one can estimate the typical sizes of the Wilson coefficients in the Lagrangian. Without imposing any *a priori* constraint on the UV spectral function $c_{\ell,\mu}^{00}$, apart from partial wave unitarity, we find that the scalar-tensor Lagrangian can be parametrized as follows

$$\begin{aligned}
 S = M_P^2 \int d^4x \sqrt{-g} & \left(\frac{1}{2} R - \frac{1}{2} \nabla_\mu \varphi \nabla^\mu \varphi + \frac{\hat{\alpha} M_P^2}{2\Lambda^4} (\nabla_\mu \varphi \nabla^\mu \varphi)^2 + \frac{\hat{\beta}_1}{2\Lambda^2} \varphi \mathcal{G} + \frac{\hat{\beta}_2 M_P}{4\Lambda^3} \varphi^2 \mathcal{G} + \frac{\hat{\gamma}_0}{3!\Lambda^4} \mathcal{R}^{(3)} \right. \\
 & + \frac{\hat{\gamma}_1}{3!\Lambda^4} \varphi \mathcal{R}^{(3)} + \frac{\hat{\gamma}_2 M_P}{2\Lambda^5} \nabla_\mu \varphi \nabla^\mu \varphi \mathcal{R}^{(2)} - \frac{4\hat{\gamma}_3 M_P}{3\Lambda^5} \nabla_\mu \varphi \nabla_\rho \varphi \nabla_\nu \nabla_\sigma \varphi R^{\mu\nu\rho\sigma} \\
 & \left. + \frac{\hat{\gamma}_4 M_P^2}{3\Lambda^6} \nabla_\mu \varphi \nabla^\mu \varphi \nabla_\rho \nabla_\sigma \varphi \nabla^\rho \nabla^\sigma \varphi + \dots \right), \tag{4.29}
 \end{aligned}$$

where we have used the dimensionless field $\varphi = \phi/M_P$ and $\hat{\alpha}, \hat{\beta}_i, \hat{\gamma}_i$ are dimensionless coefficients and are parametrically $\mathcal{O}(1)$. In this scenario, the scalar self-couplings such as α go like $\sim 1/\Lambda^p$, and these scalings remain the same in the decoupling limit of the graviton where $M_P \rightarrow \infty$ and Λ is held fixed. The scalings of the Lagrangian terms in eq. (4.29) have been summarized in eq. (4.3), which for the terms in eq. (4.29) has an intriguing integer flooring operation for the power of the M_P/Λ factor, $\tilde{N}_\phi = \lfloor N_\phi/2 \rfloor$. Having gone through the power counting with the sum rules, we can see that the flooring operation originates from the fact that, in the scaling argument above, $\hat{c}_{\ell,\mu}^{12}$ with either no or one scalar helicity corresponds to Λ/M_P (see eq. (4.8) and eq. (4.13)) while $\hat{c}_{\ell,\mu}^{12}$ with two scalar helicities corresponds to 1 (see eq. (4.17)). Also, given that each term on the right hand side of a sum rule only contains two factors of $\hat{c}_{\ell,\mu}^{12}$, there will be a $\hat{c}_{\ell,\mu}^{00}$ in the sum rule for the lowest orders as long as there are two 0 helicities in the 2-to-2 scattering (except for the case of $F_{1,\ell}^{+0-0}$, which however does not affect our argument). Thus, in these cases, the power of M_P/Λ in eq. (4.3) is determined by the number of 0 helicities in the most constraining 2-to-2 scattering amplitude, upon taking the flooring operation $\lfloor N_\phi/2 \rfloor$. We emphasize that the $\tilde{N}_\phi = \lfloor N_\phi/2 \rfloor$ rule is an coincidence, valid only for the lowest orders of the EFT operators. For higher orders, our method precisely predicts the breakdown of this rule, which will be numerically verified in section 6.5. The correct way to get \tilde{N}_ϕ for any orders is to count the number of $c_{\ell,\mu}^{00}$ in appropriate dispersion relations, as discussed through this section.

On the other hand, if the scalar interactions are constrained to be comparable with the gravitational interactions, that is, we assume the scalar UV spectral function is relatively weak and has the correspondence $\hat{c}_{\ell,\mu}^{00} \Leftrightarrow \Lambda/M_P$, then the scalar-tensor Lagrangian can be parametrized as follows

$$\begin{aligned}
 S = M_P^2 \int d^4x \sqrt{-g} & \left(\frac{1}{2} R - \frac{1}{2} \nabla_\mu \varphi \nabla^\mu \varphi + \frac{\hat{\alpha}}{2\Lambda^2} (\nabla_\mu \varphi \nabla^\mu \varphi)^2 + \frac{\hat{\beta}_1}{2\Lambda^2} \varphi \mathcal{G} + \frac{\hat{\beta}_2}{4\Lambda^2} \varphi^2 \mathcal{G} + \frac{\hat{\gamma}_0}{3!\Lambda^4} \mathcal{R}^{(3)} \right. \\
 & + \frac{\hat{\gamma}_1}{3!\Lambda^4} \varphi \mathcal{R}^{(3)} + \frac{\hat{\gamma}_2}{2\Lambda^4} \nabla_\mu \varphi \nabla^\mu \varphi \mathcal{R}^{(2)} - \frac{4\hat{\gamma}_3}{3\Lambda^4} \nabla_\mu \varphi \nabla_\rho \varphi \nabla_\nu \nabla_\sigma \varphi R^{\mu\nu\rho\sigma} \\
 & \left. + \frac{\hat{\gamma}_4}{3\Lambda^4} \nabla_\mu \varphi \nabla^\mu \varphi \nabla_\rho \nabla_\sigma \varphi \nabla^\rho \nabla^\sigma \varphi + \dots \right), \tag{4.30}
 \end{aligned}$$

where again $\hat{\alpha}, \hat{\beta}_i, \hat{\gamma}_i$ are dimensionless coefficients and are parametrically $\mathcal{O}(1)$. In this case, we have, for example, $\alpha \sim 1/(M_P^2 \Lambda^2)$. Note that the typical size of the coefficient of

$\varphi\mathcal{G}$, a leading operator that gives rise to hairy black holes, is not affected by the constraints on the scalar self-couplings. This surprising fact can be easily spotted in the dispersive sum rules. Our goal in section 6 is to use all available sum rules to numerically compute the bounds on the coefficients $\hat{\alpha}, \hat{\beta}_i, \hat{\gamma}_i$ and so on, confirming the rough estimates in this section.

5 Optimization scheme

In this section, we will set up a numerical optimization scheme that effectively utilizes the dispersive sum rules to constrain the Wilson coefficients of scalar-tensor theory in the following section. Recall that the dispersive sum rules establish a remarkable set of relations between the IR coefficients of the EFT and the amplitudes of the unknown UV completion. These relations can be fed into a semi-definite program (SDP) that can be solved numerically. This will confirm the rough estimates in the previous section and, more importantly, lead to “sharp” bounds on the coefficients in the next section. Readers uninterested in the detailed numerical setup and methods can go through section 5.1 and skip section 5.2.

5.1 General strategy

While estimating the scaling rules for the Wilson coefficients, the sum rules (3.14) are sometimes sufficient and preferred. To numerically obtain the optimal bounds, we shall always use the *st*-improved sum rules (3.18). Each of the sum rules (3.18) is actually a one-parameter family of dispersive equalities, parametrized by the momentum transfer t , connecting the Wilson coefficients and the integrals of the UV amplitudes. To effectively use all of these dispersive equalities, following the approach of [56] and [44], we integrate the dispersive sum rule against a weight function $\phi_k^{1234}(p)$ over the interval $0 \leq p \leq \Lambda$ and as well as sum over several sum rules:

$$\begin{aligned} \sum_{1234,k} \int_0^\Lambda dp \phi_k^{1234}(p) \left[\delta_{k,2} a_{k,-1}^{1234} \frac{-1}{p^2} + a_{k,0}^{1234} + a_{k,1}^{1234} (-p^2) + a_{k,2}^{1234} p^4 \right] \\ = \left\langle \sum_{1234,k} \int_0^\Lambda dp \phi_k^{1234}(p) F_{k,\ell}^{1234}(\mu, -p^2) \right\rangle, \end{aligned} \tag{5.1}$$

where we have, for later convenience, introduced a positive real number p such that

$$t := -p^2. \tag{5.2}$$

The weight functions $\phi_k^{1234}(p)$ will be the decision variables we optimize over to get the best causality bounds. (For the forward-limit sum rules that will also be used, it is suffice to use normal weight parameters; see appendix C.) By the integration and summation in eq. (5.1), we can make use of as much information as possible from the dispersive sum rules in extracting the causality bounds. If an appropriate $\phi_k^{1234}(p)$ makes the right hand side of eq. (5.1) positive, we can then obtain a condition on the Wilson coefficients

$$\sum_{1234,k} \int_0^\Lambda dp \phi_k^{1234}(p) \left(\delta_{k,2} a_{k,-1}^{1234} \frac{-1}{p^2} + a_{k,0}^{1234} + a_{k,1}^{1234} (-p^2) + a_{k,2}^{1234} (-p^2)^2 \right) \geq 0. \tag{5.3}$$

Going through all possible $\phi_k^{1234}(p)$, we can find the tightest constraints on these coefficients. The problem of finding the best bounds can be formulated as an SDP with an infinite number of constraints, enumerated by the discrete variable ℓ and the continuous variable μ . Also, the functional space of all possible $\phi_k^{1234}(p)$ is parametrized by an infinite number of parameters, so numerically we also need to approximate this functional space, which will be explained shortly in section 5.2.

To see how this optimization is implemented, notice that $F_{k,\ell}^{1234}(\mu, -p^2)$ contains an infinite number of UV partial amplitudes $c_{\ell,\mu}^{12}$ and their complex conjugates, which we are agnostic about from the point view of bootstrapping from low energies. In order to proceed, we need to eliminate them in the optimization problem, which naturally turns this into an SDP problem.

Before that, let us isolate the minimal set of $c_{\ell,\mu}^{12}$ that are necessarily involved when performing this SDP. First, note that in a theory with parity conservation, we can divide the sum over all possible intermediate states in $\langle \dots \rangle$ (see eq. (3.10)) into two parts, one being summation over parity-even X states and the other summation over parity-odd states. Denoting the parity of state X by P_X , we have the following relations for the partial wave amplitudes

$$c_{P_X,\ell,\mu}^{12} = P_X c_{P_X,\ell,\mu}^{\bar{2}\bar{1}}, \quad (5.4)$$

$$c_{P_X,\ell,\mu}^{12} = (-1)^\ell c_{P_X,\ell,\mu}^{21}. \quad (5.5)$$

Because of time reversal invariance that we assume, we have $\mathcal{M}^{1234} = \mathcal{M}^{\bar{3}\bar{4}\bar{1}\bar{2}}$, which implies that $\text{Im}(\sum_{P_X} c_{P_X,\ell,\mu}^{12} c_{P_X,\ell,\mu}^{*\bar{3}\bar{4}}) = 0$. Denoting $c_{P_X,\ell,\mu}^{12} = c_{P_X,\ell,\mu}^{12,\Re} + i c_{P_X,\ell,\mu}^{12,\Im}$, we then have

$$\sum_{P_X} c_{P_X,\ell,\mu}^{12} c_{P_X,\ell,\mu}^{*\bar{3}\bar{4}} = \sum_{\mathcal{I}=\Re,\Im} \sum_{P_X} c_{P_X,\ell,\mu}^{12,\mathcal{I}} c_{P_X,\ell,\mu}^{\bar{3}\bar{4},\mathcal{I}} = \sum_{P_X} \left(c_{P_X,\ell,\mu}^{12,\Re} c_{P_X,\ell,\mu}^{\bar{3}\bar{4},\Re} + c_{P_X,\ell,\mu}^{12,\Im} c_{P_X,\ell,\mu}^{\bar{3}\bar{4},\Im} \right). \quad (5.6)$$

So the real and imaginary parts of $c_{P_X,\ell,\mu}^{12}$ are separated and play a similar role in the dispersive sum rules. From the perspective of imposing positivity bounds, this extra summation over the real and imaginary part is essentially redundant, since, as mentioned above, we are agnostic about the values of $c_{P_X,\ell,\mu}^{12}$. Following [4, 37], we will simply absorb the summation over $\mathcal{I} = \Re, \Im$ into the summation over X and take $c_{P_X,\ell,\mu}^{12}$ as real functions in the following. Using these separations, we can express a generic quantity obtained by mixing different helicities of $F_{k,\ell}^{1234}$ and integrating over p in the following form:

$$\sum_{1234,k} \int_0^\Lambda dp \phi_k^{1234}(p) F_{k,\ell}^{1234}(\mu, -p^2) := \sum_{P_X=\pm 1} \sum_{\mathbb{A},\mathbb{B}} B_{P_X,\ell}^{\mathbb{A},\mathbb{B}}(\mu) c_{P_X,\ell,\mu}^{\mathbb{A}} c_{P_X,\ell,\mu}^{\mathbb{B}}, \quad (5.7)$$

where the summation of \mathbb{A} and \mathbb{B} is over $00, +0, ++, +-$ and $B_{P_X,\ell}^{\mathbb{A},\mathbb{B}}$ is independent of p and can be extracted from eq. (3.19). The reason why \mathbb{A} and \mathbb{B} only run over $00, +0, ++, +-$ is that we can use eqs. (5.4) and (5.5) to convert other helicities to these four. According to parity P_X and whether ℓ is odd, the summation on the right hand side of eq. (5.7) splits into four independent parts, $(P_X, \ell) = (+1, \text{even}), (+1, \text{odd}), (-1, \text{even}), (-1, \text{odd})$, each of which can be written in the following form

$$(\mathcal{C}_{P_X,\ell,\mu})^T B_{P_X,\ell}(\mu) \mathcal{C}_{P_X,\ell,\mu}, \quad (5.8)$$

where $B_{P_X, \ell}(\mu)$ is a 4×4 matrix and we have defined that

$$\mathcal{C}_{P_X, \ell, \mu} = \left(c_{P_X, \ell, \mu}^{00} \quad c_{P_X, \ell, \mu}^{+0} \quad c_{P_X, \ell, \mu}^{++} \quad c_{P_X, \ell, \mu}^{+-} \right)^T. \quad (5.9)$$

The reason why it is beneficial to separate the sum in eq. (5.7) according to parity P_X and the oddness of ℓ is that some of the $c_{P_X, \ell, \mu}^{12}$ often vanish owing to eq. (5.4) and eq. (5.5), in which case we can omit the corresponding entries of the $B_{P_X, \ell}(\mu)$ matrix in the SDP. This leads to better bounds and reduces computational costs. Again, the non-vanishing $c_{P_X, \ell, \mu}^{12}$ depend on the UV model, and for a generic bootstrap program we choose to be agnostic about them.

With these established, we see that the requirement of the right hand side of eq. (5.7) being positive is equivalent to the conditions that all the $B_{P_X, \ell}(\mu)$ matrices be positive semi-definite

$$B_{P_X, \ell}(\mu) \succeq 0, \quad \text{for } P_X = \pm 1, \text{ all possible } \ell \text{ and all } \mu \geq \Lambda^2. \quad (5.10)$$

These conditions will in turn ensure that the left hand side of eq. (5.7) is positive, giving rise to a condition for some Wilson coefficients (5.3) for a given set of $\phi_k^{1234}(p)$. To obtain the best bounds, we optimize over all possible $\phi_k^{1234}(p)$. In practice, of course, we can not impose the conditions for all ℓ and μ and go through all possible $\phi_k^{1234}(p)$, and some numerical approximations are needed. Note that the SDPB package can deal with an SDP with only one continuous parameter if the entries of the linear matrix inequalities eq. (5.10) are polynomials of this parameter, but unfortunately this is not the case here. In the following subsection, we shall outline approximations that can be used to overcome this problem, along with how to effectively truncate the $\phi_k^{1234}(p)$ functional space.

5.2 Numerical details

Having formulated the causality bounds finding as a SDP, we now get to the nitty-gritty of implementing it numerically, largely following the numerical implementation of [56] and [44]. To simplify the expressions, we shall set $\Lambda = 1$ from now on, but restore it in the final results for clarity.

As mentioned, SDPB can directly solve a SDP with a finite number of linear matrix inequalities, and the entries of these matrices can be polynomials of a continuous variable. However, for our current case, entries of $B_{P_X, \ell}(\mu)$ are more complex than polynomials of a continuous variable. To take in as many constraints as possible in the numerical program, we can divide the μ - ℓ constraint space into five regions, as shown in figure 3, and will make approximations for the five regions separately.

Finite μ and finite ℓ : in this region, we will simply discretize the continuous parameter μ . Since the UV scale $\mu \geq 1$, we can choose a discrete set of μ . We find that the point density needed to achieve convergence depends on the dimension of the truncated $\phi_k^{1234}(p)$ functional space, which is the main limiting factor to use a higher dimensional functional space. We also only make use of the partial waves up to ℓ_{\max} .

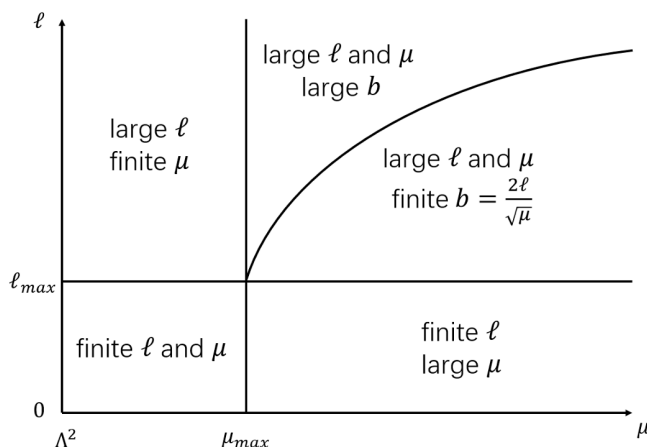


Figure 3. Various regions in the μ - ℓ constraint space. The constraints are implemented differently in different regions.

Large μ and finite ℓ : when μ is large, the entries of the $B_{P_X, \ell}(\mu)$ matrices can be expanded as a Taylor series of $1/\mu$ around $\mu \rightarrow \infty$, which allows us to approximate the entries of $B_{P_X, \ell}(\mu)$ by truncating the expansion and retaining the leading few orders. Then, we multiply all the sum rules by an appropriate power of μ to make entries of the $B_{P_X, \ell}(\mu)$ matrices polynomials of μ , and take μ as the continuous parameter in SDPB. Alternatively, when the dimension of the $\phi_k^{1234}(p)$ functional basis is not too large, we find that it is also numerically sufficient to work with the exact dependence on μ and just take a few discrete large μ points along with finite $\ell \leq \ell_{\max}$.

Finite μ and large ℓ : when ℓ is large, the Wigner d-functions (or rather the hypergeometric function) oscillate with p^2/μ and thus tend to vanish after integrating against the weight functions. This is the reason why we also seemingly redundantly add the forward-limit sum rules (3.20) in the SDP, in order to effectively use the constraints from this region. That is, in the large ℓ limit, with the forward-limit sum rules included, we can neglect the terms with the hypergeometric functions from the non-forward sum rules, since the contributions from the forward-limit sum rules dominate in this limit. In the large ℓ limit, we can approximate ℓ as a continuous variable; however, the forward-limit sum rules contain square roots of polynomials of ℓ : $\sqrt{(\ell + c_1)(\ell + c_2) \dots (\ell + c_n)}$, where c_i are real constants, which are not admissible by SDPB. To resolve this problem, we shall expand them as a Laurent series in the limit $\ell \rightarrow \infty$ and only keep a few leading terms: $\ell^{n/2} + 1/2(c_1 + c_2 + \dots + c_n)\ell^{n/2-1} + \mathcal{O}(\ell^{n/2-2})$. We then make the variable change $\ell \rightarrow (y + \sqrt{\ell_{\max}})^2$ so that it becomes a polynomial of y where $y \geq 0$. Then, we can again discretize μ , and, for a fixed μ , the entries of the $B_{P_X, \ell}(\mu)$ matrices can be viewed as polynomials of ℓ for large ℓ , the semi-positivity of $B_{P_X, \ell}(\mu)$ then becoming admissible for SDPB. Note that while the added forward limit sum rules do technically alter the SDP in this region as well as in the finite ℓ regions, they become negligible in other regions.

Large μ , ℓ and finite b : this region can be made accessible by using the asymptotic behavior of the Wigner d-functions in $F_{k,\ell}^{1234}(\mu, -p^2)$. The Wigner d-functions can be expressed in terms of the hypergeometric function, which has the following asymptotic behavior

$$\lim_{\mu, \ell \rightarrow \infty; 2\ell/\sqrt{\mu}=b} {}_2F_1(h_1 - \ell, \ell + h_1 + 1; h_1 - h_2 + 1; p^2/\mu) = \frac{\Gamma(h_1 - h_2)}{(bp/2)^{h_1 - h_2}} J_{h_1 - h_2}(bp), \quad (5.11)$$

where $J_h(x)$ is the Bessel function of the first kind and the limit is taken with fixed $b = 2\ell/\sqrt{\mu}$. That is, we sample the constraints along lines $b\sqrt{\mu} = 2\ell$ (with different b) in the region of large μ and large ℓ , and each of these lines has a natural physical interpretation of scatterings with fixed impact parameter $b = 2\ell/\sqrt{\mu}$ [44]. With these established, we can easily Taylor expand $F_{k,\ell}^{1234}(\mu, -p^2)$ around $\mu \rightarrow \infty$ with fixed b , and only retain the leading terms, namely the $1/\mu^3$ term in this case. (We do not need to expand μ in the partial wave amplitudes $c_{P_X, \ell, \mu}^{12}$, because they are limited in size by partial wave unitarity.) We find that only $F_{(1,2), \ell}^{1234}(\mu, -p^2)$ have non-vanishing $\mathcal{O}(1/\mu^3)$ terms, so only these dispersive sum rules need to be considered in the large μ and ℓ region. For example, the leading term of $F_{2,\ell}^{+0-0}(\mu, -p^2)$ in this limit is given by

$$F_{2,\ell}^{+0-0}(\mu, -p^2) = \frac{2}{\mu^3} J_0(bp) |c_{P_X, \ell, \mu}^{+0}|^2 + \mathcal{O}\left(\frac{1}{\mu^4}\right). \quad (5.12)$$

Note that in the leading order the ℓ dependence is only in $c_{P_X, \ell, \mu}^{12}$'s, which do not go into the definition of $\tilde{B}_{P_X, \tilde{\ell}}(b)$. However, $\tilde{B}_{P_X, \tilde{\ell}}(b)$ does depend on the oddness of ℓ , because we need to use $c_{P_X, \ell, \mu}^{12} = (-1)^\ell c_{P_X, \ell, \mu}^{21}$ to convert $c_{P_X, \ell, \mu}^{12}$'s to a standard independent basis. This means that the matrix $\mu^3 B_{P_X, \ell}$ only depends on b , P_X and the oddness of ℓ at leading order in the large μ and ℓ region. Let us define $\tilde{B}_{P_X, \tilde{\ell}}(b) := \mu^3 B_{P_X, \ell}(\mu)$ in this region, where $\tilde{\ell}$ means $\tilde{B}_{P_X, \tilde{\ell}}(b)$ depends on the oddness of ℓ rather than its explicit value. Therefore, for large μ and ℓ , we can simply impose the following linear matrix inequalities as a leading approximation

$$\tilde{B}_{P_X, \tilde{\ell}}(b) \succeq 0, \quad \text{for all } b > 0, P_X = \pm 1 \text{ and } \tilde{\ell} = \text{even or odd}. \quad (5.13)$$

To explicitly compute $\tilde{B}_{P_X, \tilde{\ell}}(b)$, we note the following well known integration formula

$$\frac{\Gamma(\nu)}{(b/2)^{\nu-1}} \int_0^1 dp p^{n+1-\nu} J_{\nu-1}(bp) = \frac{1}{n+1} {}_1F_2\left(\frac{n+1}{2}; \frac{n+3}{2}, \nu; -\frac{b^2}{4}\right). \quad (5.14)$$

So the entries of $\tilde{B}_{P_X, \tilde{\ell}}(b)$ are still not polynomials of b , and we need to make further approximations. For finite $b < b_{\max}$, we can discretize it into $b = \{\epsilon_b + k\delta_b | 0 \leq k < (b_{\max} - \epsilon_b)/\delta_b, k \in \mathbb{N}\}$, where ϵ_b is a very small starting point.

Large μ , ℓ and large b : for large b , by the asymptotic form of the generalized hypergeometric function, we can write $\tilde{B}_{P_X, \tilde{\ell}}(b)$ in the following form,

$$\tilde{B}_{P_X, \tilde{\ell}}(b) = f(b) + g(b) \cos(b) + h(b) \sin(b), \quad (5.15)$$

where $f(b)$, $g(b)$ and $h(b)$ are 4×4 matrices whose entries are polynomials of $1/b$, truncated to order $(1/b)^{R_b}$. For large b , it is a good approximation to replace the semi-positiveness of $\tilde{B}_{P_X, \ell}(b)$ with the following slightly stronger condition

$$b^{R_b} \begin{pmatrix} f(b) + g(b) & h(b) \\ h(b) & f(b) - g(b) \end{pmatrix} \succeq 0, \text{ for } b \geq b_{\max}, \quad (5.16)$$

where the factor b^{R_b} makes $b^{R_b}f(b)$, $b^{R_b}g(b)$ and $b^{R_b}h(b)$ polynomials of b .

Apart from the approximations in the μ - ℓ constraint space, we also need to numerically approximate the functional spaces of all possible $\phi_k^{1234}(p)$. Recall that $\phi_k^{1234}(p)$ are supposed to run over all possible functions within the interval $[0, 1]$. By the Weierstrass approximation theorem, a simple functional basis over a finite interval would be power functions p^n , and in the numerical approximation we truncate to keep the leading few orders. However, for the technical reasons to be explained below, for some $\phi_k^{1234}(p)$, we will need to choose $(1-p)^2 p^n$.

First, note that, in order to obtain the bounds on the leading order coefficients, the positivity condition (5.10) can not be satisfied without $F_{1,\ell}^{0000}$, $F_{2,\ell}^{0000}$, $F_{1,\ell}^{+0-0}$, $F_{2,\ell}^{+0-0}$ and $F_{2,\ell}^{++--}$. This is because all other leading $F_{k,\ell}^{1234}$ in the large μ and large ℓ region either lead to a non-diagonal term in $B_{P_X, \ell}$ or contribute to a term in $B_{P_X, \ell}$ that changes its sign under the parity P_X or the oddness of ℓ . For $B_{P_X, \ell}$ to be semi-positive, we need the diagonal terms to be semi-positive and we need $B_{P_X, \ell}$ to be semi-positive for both all cases of P_X and ℓ . Additionally, we aim to derive bounds projected onto $1/M_P^2$, and only the above five improved sum rules involve $1/M_P^2$.

Let us see what kinds of bases are suitable for $F_{1,\ell}^{0000}$, $F_{2,\ell}^{0000}$, $F_{1,\ell}^{+0-0}$, $F_{2,\ell}^{+0-0}$ and $F_{2,\ell}^{++--}$ for our purposes. The technical requirements come from implementing the constraints in the large ℓ and μ region. We take $F_{2,\ell}^{++--}$ as an example. In this region with fixed $b = 2\ell/\sqrt{\mu}$, a necessary condition to satisfy the positivity condition (5.13) is

$$\int_0^1 dp \phi_2^{++--}(p) J_0(bp) \geq 0, \text{ for all } b > 0. \quad (5.17)$$

This actually implies that the Fourier transform of $\phi_2^{++--}(p)/p$ is non-negative and also $\lim_{p \rightarrow 0} \phi_2^{++--}(p)/p > 0$. As a result, the basis for $\phi_2^{++--}(p)$ should start at $p^{n_{\min}}$ with $n_{\min} \leq 1$. On the other hand, this choice necessarily results in an IR divergence from integrating in the low energy region near $p = 0$. The best one can do for $F_{2,\ell}^{++--}$ is to choose $n_{\min} = 1$, which only leads to a logarithmic divergence. This IR divergence arises from how the scattering amplitudes are defined for massless particles in 4D, and may be resolved using better observables [56]. We will simply regulate it with an IR cutoff scale m_{IR} , which may be taken to be the Hubble scale as a conservative choice. The cases of $F_{1,\ell}^{0000}$, $F_{2,\ell}^{0000}$, $F_{1,\ell}^{+0-0}$ and $F_{2,\ell}^{+0-0}$ are analogous. Going through similar steps, we can see that the basis of $\phi_1^{0000}(p)$, $\phi_2^{0000}(p)$, $\phi_1^{+0-0}(p)$, $\phi_2^{+0-0}(p)$ and $\phi_2^{++--}(p)$ should be chosen to start with p^{-1} , p , p^{-1} , p and p respectively.

There is actually one additional consideration for choosing the suitable basis, namely, the requirement that $g(b)$ or $h(b)$ should not dominate in the large ℓ, μ, b region in order

ℓ_{\max}	42
b_{\max}	10001/250
ϵ_b	1/250
R_b	10
N_p	7
Discrete set of μ for finite ℓ	$\{1/(1-k/100) 0 \leq k \leq 90, k \in \mathbb{Z}\}$ $\cup \{1/(1-k/400) 361 \leq k \leq 399, k \in \mathbb{Z}\}$
Discrete set of μ for large ℓ	$\{1/2500 + 1/(1-k/100)^{1/2} 0 \leq k \leq 99, k \in \mathbb{Z}\}$
Discrete set of $b = 2\ell/\sqrt{\mu}$	$\{\epsilon_b + k/32 0 \leq k \leq 1280, k \in \mathbb{Z}\}$
Non-default SDPB parameters	<code>-precision=766</code> <code>-dualityGapThreshold=1e-11</code> <code>-maxComplementarity=1e+80</code> <code>-maxIterations=20000</code>

Table 1. Numerical parameters used in the SDPB computations.

to satisfy condition (5.16). Again, take $F_{2,\ell}^{++++}$ as an example. By eq. (5.11), we can get

$$\mu^3 \int_0^1 dp p^n F_{2,\ell}^{++++}(\mu, -p^2) \xrightarrow{\text{large } b} \left(|c_{P_X, \ell, \mu}^{++}|^2 + |c_{P_X, \ell, \mu}^{+-}|^2 \right) \quad (5.18)$$

$$\cdot \left[\frac{1}{b^{n+1}} \frac{2^n \Gamma(\frac{1+n}{2})}{\Gamma(\frac{1-n}{2})} + \frac{\sqrt{2} \cos(b - \frac{3\pi}{4})}{\sqrt{\pi} b^{\frac{3}{2}}} + \frac{\sqrt{2}(8n-5) \sin(b - \frac{3\pi}{4})}{8\sqrt{\pi} b^{\frac{5}{2}}} + \mathcal{O}\left(\frac{1}{b^{\frac{7}{2}}}\right) \right].$$

If the oscillating term $\cos(b)$ or $\sin(b)$ dominates in the large b limit, the positivity condition (5.16) can not be satisfied. However, we already require that when p goes to zero, n_{\min} in the basis $p^{n_{\min}}$ should not be less than 1 so as to avoid non-logarithmic IR divergences. To overcome this, we can multiply the corresponding weight function with a factor $(1-p)^2$, which cancel the leading oscillating terms upon integration and make $f(b)$ dominate in eq. (5.16). Again, the cases of $F_{1,\ell}^{0000}$, $F_{2,\ell}^{0000}$, $F_{1,\ell}^{+0-0}$ and $F_{2,\ell}^{+0-0}$ are analogous. Thus, the final result is that the basis of $\phi_1^{0000}(p)$, $\phi_2^{0000}(p)$, $\phi_1^{+0-0}(p)$, $\phi_2^{+0-0}(p)$ and $\phi_2^{++++}(p)$ should be chosen to start from $(1-p)^2 p^{-1}$, $(1-p)^2 p$, $(1-p)^2 p^{-1}$, $(1-p)^2 p$, $(1-p)^2 p$ respectively.

For other $F_{k,\ell}^{1234}(\mu, -p^2)$ that result in leading order contributions in the large μ, ℓ limit, the bases are chosen such that they lead to the same large b behavior in the $\tilde{B}_{P_X, \tilde{\ell}}(b)$ matrix as the above five $F_{k,\ell}^{1234}(\mu, -p^2)$. For the rest of the $F_{k,\ell}^{1234}(\mu, -p^2)$ that are sub-leading in the large μ, ℓ limit, we can simply choose their bases to be $1, p, p^2, p^3, \dots$. In our numerical calculations, it is sufficient for our purposes to choose the dimension of the functional space of $\phi_k^{1234}(p)$ to be $N_p = 7$. The numerical parameters we use to run SDPB are listed in table 1.

In general, when performing the numerical optimization to obtain bounds on a given set of Wilson coefficients, we hope to utilize as many sum rules as possible so as to derive the strongest bounds. For that, we can often include sum rules that contain Wilson coefficients not in that given set. The reason is that we can choose the weight functions $\phi_k^{1234}(p)$ properly so that the weight-function-smearred sum rules do not contain the extra Wilson coefficients.

6 Bounds on EFT coefficients and their implications

In this section, we shall constrain the Wilson coefficients of scalar-tensor theory using the dispersive sum rules obtained in section 3, via the numerical optimization procedure outlined in section 5. We will consider generic scalar-tensor EFTs as well as EFTs with some of the coefficients fine-tuned, the latter being also popular as modified gravity and cosmological models phenomenologically. In appendix C, we will give an explicit example to demonstrate how to use the optimization scheme to obtain the causality bounds. We will first derive the bound on α , the coefficient of the $(\partial\phi)^4$ term in the Lagrangian. We will show that the value of α will significantly affect the bounds on a coefficient when all the sum rules of the coefficient contain $c_{P_X, \ell, \mu}^{00}$, in agreement with the discussions in section 4. Therefore, we will compute the bounds on the other coefficients for various values of α . Particularly, we will compute the bounds on the Gauss-Bonnet couplings, which give rise to the intriguing phenomena of hairy black holes and scalarization in compact stars. These couplings are currently being intensively probed with gravitational wave and other observational means. We shall discuss the phenomenological implications of our bounds for these couplings. We will also calculate the causality bounds for large values of α , which confirms the scaling behaviors that have been estimated in section 4. We will also show that some fine-tuned EFTs can not be exact, as they will lead to inconsistencies among the sum rules, so some additional terms must exist. Moreover, some higher dimensional coefficients can significantly affect the bounds on the lower dimensional coefficients.

6.1 Scalar four-derivative term

Let us first derive the lower bound for the coefficient α , the coupling constant of the $(\partial\phi)^4 = (\nabla_\mu\phi\nabla^\mu\phi)^2$ term. In the graviton decoupling limit, the lower bound on this dim-8 coefficient is $\alpha > 0$. This was one of the earliest causality bounds [3] and gives rise to the term of “positivity bounds”, often used synonymously with “causality bounds”, as we do in this paper. In the presence of gravity, however, it has been predicted that the lower bound slightly dips below zero, the negativity being suppressed by the Planck mass squared [54]. This has been illustrated explicitly with a string theory example [55] and also numerically confirmed for generic UV completions [56].

We refer the readers to appendix C for a more detailed explanation of how to implement the numerical procedure of section 5. Here we shall simply outline the main steps of this procedure for the case of obtaining the lower bound on α .

- First, we collect relevant improved dispersive sum rules. In principle, the sum rules that do not contain α should also be included for deriving the strongest bound, because those sum rules contain the information of full crossing symmetry/null constraints. However, for this particular case, we find that the only relevant sum rules are from \mathcal{M}^{0000} , and hence we only need:

$$-\frac{1}{M_P^2} + 2\alpha t - \gamma_4 t^2 = \langle F_{1,\ell}^{0000}(\mu, t) \rangle, \tag{6.1}$$

$$-\frac{1}{M_P^2} \frac{1}{t} + 2\alpha - \gamma_4 t + 12g_{0,2}^S t^2 = \langle F_{2,\ell}^{0000}(\mu, t) \rangle. \tag{6.2}$$

- Then, we sum over the sum rules after integrating them against the weight functions and define $B_{P_X, \ell}(\mu)$ via $\sum_k \int_0^1 dp F_{k, \ell}^{0000}(\mu, -p^2) = (\mathcal{C}_{P_X, \ell, \mu})^T B_{P_X, \ell}(\mu) \mathcal{C}_{P_X, \ell, \mu}$. The weight functions are decision variables. Imposing $B_{P_X, \ell}(\mu) \succeq 0$ then gives us inequalities on the Wilson coefficients. Since here we are only concerned about the bound on α and agnostic about all the other Wilson coefficients, we can choose the weight functions such that the combinations in front of all the other Wilson coefficients vanish in the summed-over sum rules. For example, if we want to be agnostic about γ_4 , we can impose the condition on the weight functions:

$$\int_0^1 dp [\phi_1^{0000}(p)(-(-p^2)^2) + \phi_2^{0000}(p)(-(-p^2))] = 0. \quad (6.3)$$

where $\phi_1^{0000}(p)$ and $\phi_2^{0000}(p)$ are the weight functions multiplying eq. (6.1) and eq. (6.2) respectively. We must impose conditions like this such that the only coupling constants are $1/M_P^2$ and α in the summed-over sum rules, which will be used in the numerical optimization.

- Then, we impose $B_{P_X, \ell}(\mu) \succeq 0$ and conditions like eq. (6.3), which leads to inequalities like

$$\int_0^1 dp \left(-\phi_1^{0000}(p) + \phi_2^{0000}(p) \frac{1}{p^2} \right) \frac{1}{M_P^2} + 2 \int_0^1 dp \left(\phi_1^{0000}(p)(-p^2) + \phi_2^{0000}(p) \right) \alpha \geq 0. \quad (6.4)$$

To find the strongest lower bound on α , we can normalize $\int_0^1 dp \left(-\phi_1^{0000}(p) + \phi_2^{0000}(p) \frac{1}{p^2} \right) = 1$, and maximize

$$2 \int_0^1 dp \left(\phi_1^{0000}(p)(-p^2) + \phi_2^{0000}(p) \right). \quad (6.5)$$

for all possible choices of weight functions $\phi_i^{0000}(p)$. However, as discussed in section 5.2, the finite dimensional expansion of $\phi_1^{0000}(p)$ must begin with $p^{-1}(1-p)^2$, and that of $\phi_2^{0000}(p)$ must begin with $p(1-p)^2$, so the integral of the normalization condition contain a logarithmic divergence and we need to include an IR cutoff m_{IR} . It is a good approximation to only preserve the $\log(\Lambda/m_{\text{IR}})$ term in the integration, as will be explained in section 6.2. With the IR cutoff, the normalization condition becomes

$$-x_{1,-1}^{0000} + x_{2,1}^{0000} = 0, \quad (6.6)$$

where we have parameterized the weight functions as $\phi_k^{1234}(p) = \sum_{n=n_{\text{min}}} x_{k,n}^{1234} p^n (1-p)^2$.

- Finally, we solve the following SDP:

$$\begin{aligned} \text{maximize:} \quad & 2 \int_0^1 dp \left(\phi_1^{0000}(p)(-p^2) + \phi_2^{0000}(p) \right), \\ \text{subject to:} \quad & -x_{1,-1}^{0000} + x_{2,1}^{0000} = 0, \\ & \text{conditions like eq. (6.3) : } \int_0^1 dp \sum_k \phi_k^{0000}(p)(-p^2)^{n_k} = 0, \\ & \text{positivity condition: } B_{P_X, \ell}(\mu) \succeq 0, \text{ for all } P_X, \ell \text{ and } \mu. \end{aligned} \quad (6.7)$$

where Λ is set to be 1. We sum k up to $k = 4$ and n_k are chosen to eliminate higher order coefficients in the SDP. In the practical numerical calculations, since we can only optimize over a finite dimensional subspace of the infinite $\phi_k^{0000}(p)$ functional space, we supplement this SDP with some extra forward-limit sum rules. These forward-limit sum rules are redundant, but numerically they help minimize the impact of the uncertainties from implementing the constraints in the large ℓ and finite μ region (see appendix C). Thus, numerically, this SDP gives rise to the lower bound on α :

$$\alpha \geq -16.091 \frac{\log(\Lambda/m_{\text{IR}})}{\Lambda^2 M_P^2}. \tag{6.8}$$

which is consistent with the analysis in [54, 56]. However, we can not use a similar SDP to derive the upper bound on α . This is completely analogous to the pure scalar case where the corresponding α has a lower bound $\alpha \geq 0$, which can be recovered from the above bound by taking $M_P \rightarrow \infty$, but can not be bounded from above by the positivity of the spectral function. In the pure scalar case, α can be bounded from above by making use of more information from unitarity, particularly using the upper bound on the partial wave amplitude $|c_{\ell,\mu}^{00}|^2 \leq \mathcal{O}(1)$. This produces an upper bound of order $\mathcal{O}(1/\Lambda^4)$, which is very large if $M_P \gg \Lambda$. It is expected that the upper bound is something similar in the presence of gravity, which would be consistent with the estimate in section 4.

In the following subsections, we will see that the bounds on some coefficients, particularly the couplings involving the scalar, strongly depend on the value of α , while the other coefficients are insensitive to α . Specifically, we will see that the bounds on the former coefficients, projected on $1/M_P^2$, become weaker as the value of α increases. The sensitivity/insensitivity of the bounds on α originates from the fact that often being agnostic about α essentially means that we are largely agnostic about $c_{P_X,\ell,\mu}^{00}$. So, if the determination of the bounds on a coefficient requires the $B_{P_X,\ell}(\mu)$ matrices to have nontrivial $(00,*)$ or $(*,00)$ entries (cf. eqs. (5.8) and (5.9)), this coefficient will at least weakly depend on α . These coefficients include β_2 and $\gamma_{1,2,3,4}$. On the other hand, a coefficient strongly depending on α is when all of its sum rules themselves contain $c_{P_X,\ell,\mu}^{00}$; in this case, of course, the $B_{P_X,\ell}(\mu)$ matrices will have nontrivial $(00,*)$ or $(*,00)$ entries in the optimization results. For example, the γ_1 sum rules themselves do not involve $c_{P_X,\ell,\mu}^{00}$, so the bounds on γ_1 do not strongly depend on α ; nevertheless, γ_1 weakly depends on α , because the $B_{P_X,\ell}(\mu)$ matrices contain effective $(00,*)$ or $(*,00)$ entries that are nonzero when optimizing to get the bounds on γ_1 , as we shall see in section 6.5. All of these confirm the rough estimates in section 4.

More explicitly, note that, to satisfy our positivity condition $B_{P_X,\ell}(\mu) \succeq 0$, a necessary condition is that all diagonal entries of the matrix $B_{P_X,\ell}(\mu)$ must be non-negative; making use of the explicit expressions of $F_{m,\ell}^{1234}$ in appendix B and judicious choices of $\phi_k^{1234}(p)$, we find that it is easy to make the $(++,++)$, $(+-,+-)$ and $(+0,+0)$ entries of $B_{P_X,\ell}(\mu)$ positive, but often this can not be done for the $(00,00)$ entry. To see this, note that, the lowest order dispersive sum rules contributing to the $(00,00)$ entry come from amplitude

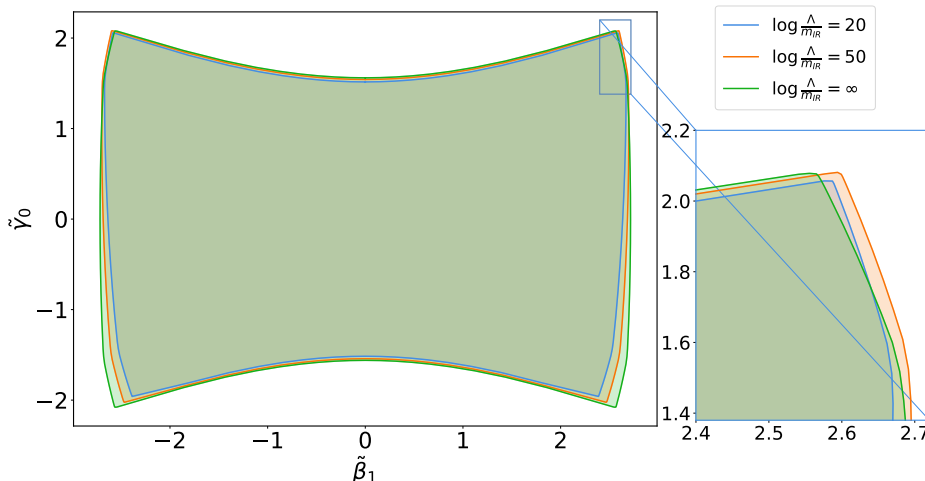


Figure 4. Causality bounds on γ_0 and β_1 . We have defined $\tilde{\beta}_1 = \beta_1 \Lambda^2 / (M_P \sqrt{\log(\Lambda/m_{\text{IR}})})$ and $\tilde{\gamma}_0 = \gamma_0 \Lambda^4 / (M_P^2 \sqrt{\log(\Lambda/m_{\text{IR}})})$. The $\log(\Lambda/m_{\text{IR}}) = \infty$ case represents the leading approximation, while the $\log(\Lambda/m_{\text{IR}}) = 50$ and $\log(\Lambda/m_{\text{IR}}) = 20$ (log being the natural logarithm) cases are computed with 2 iteration of linear improvements. The bounds are almost symmetric with respect to $\gamma_0 \rightarrow -\gamma_0$ and $\beta_1 \rightarrow -\beta_1$ because the leading approximation mostly constrains γ_0^2 and β_1^2 .

\mathcal{M}^{0000} , of which the only relevant ones, to the lowest orders, are

$$\begin{aligned}
 -\frac{1}{M_P^2} + 2\alpha t - \gamma_4 t^2 &= \left\langle F_{1,\ell}^{0000}(\mu, t) \right\rangle, \\
 -\frac{1}{M_P^2} \frac{1}{t} + 2\alpha - \gamma_4 t + 12g_{0,2}^S t^2 &= \left\langle F_{2,\ell}^{0000}(\mu, t) \right\rangle.
 \end{aligned}
 \tag{6.9}$$

If we wish to get a bound on the coefficients while being agnostic about α , we can add the above two sum rules and additionally choose more restricted $\phi_{1,2}^{0000}(p)$ to suppress the α terms on the left hand sides of the sum rules

$$\int_0^1 dp \phi_1^{0000}(p)(-p^2) + \int_0^1 dp \phi_2^{0000}(p) = 0.
 \tag{6.10}$$

However, with these extra constraints, we find that usually the (00, 00) entry of $B_{P_X, \ell}(\mu)$ can not be positive for every μ and ℓ . In the following, we shall probe how the bounds vary with the scalar dim-8 coupling α .

6.2 Linear Gauss-Bonnet coupling

The $\phi\mathcal{G}$ term has been shown to be one of the very few ways to generate black hole solutions with non-trivial/hairy solutions that are different from those of GR [93–95]. The no-hair theorems underpin many of our modern understandings of the deep nature of gravity [108, 126], and also observational confirmations of black hole solutions are important tests of Einstein’s gravity.

In this subsection, we shall compute the positivity bounds on β_1 in conjunction with the bounds on γ_0 , the coefficient of the $\mathcal{R}^{(3)}$ term,

$$\mathcal{L} \supset \sqrt{-g} \left(\frac{\beta_1}{2!} \phi\mathcal{G} + \frac{\gamma_0}{3!} \mathcal{R}^{(3)} \right).
 \tag{6.11}$$

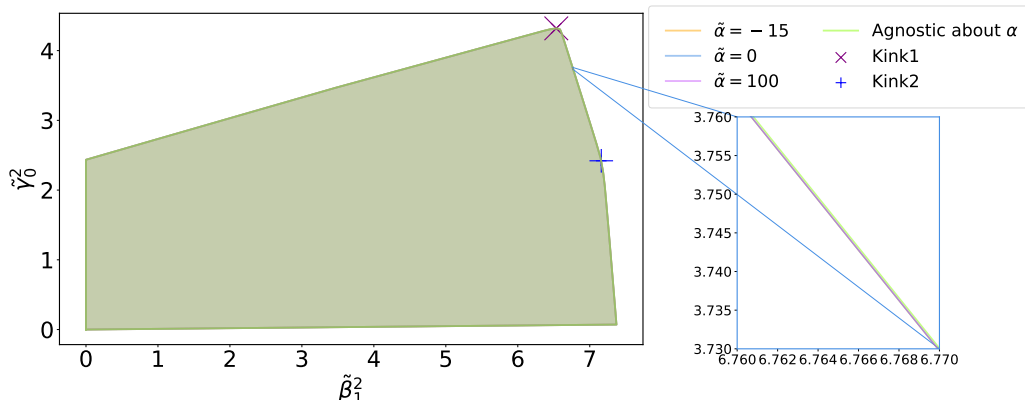


Figure 5. Bounds on γ_0^2 and β_1^2 for various α , where $\tilde{\alpha} = \alpha M_P^2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{\beta}_1^2 = \beta_1^2 \Lambda^4 / (M_P^2 \log(\Lambda/m_{\text{IR}}))$ and $\tilde{\gamma}_0^2 = \gamma_0^2 \Lambda^8 / (M_P^4 \log(\Lambda/m_{\text{IR}}))$. The main difference from figure 4 is that here we also include sum rules involving $c_{P_X, \ell, \mu}^{00}$. The four lines are almost indistinguishable from each other, meaning that the bound on γ_0^2 and β_1^2 is insensitive to α . There are kinks at $(\tilde{\gamma}_0^2, \tilde{\beta}_1^2) = (4.32, 6.53)$ and $(\tilde{\gamma}_0^2, \tilde{\beta}_1^2) = (2.42, 7.16)$ respectively.

At times, the specific structure of EFT amplitudes may lead to additional constraints in the sum rules. The case of γ_0 and β_1 provides a good example. In the tree level EFT amplitudes, some coefficients are non-negative because they are of the form of β_1^2 or γ_0^2 . These forms come from squares of 3-leg vertices in the amplitudes, as the Lagrangian terms with coefficient β_1 and γ_0 can generate 3-leg vertices. However, the corresponding right hand side terms in the sum rules do not automatically enforce such positivity. So we can take these extra constraints into account when handling the sum rules with β_1^2 and γ_0^2 . Ignoring them erroneously weakens the bounds on the other coefficients. (For the Newton’s constant $1/M_P^2$, on the other hand, there is no need to impose its positivity in our formalism, as its positivity is implied by the sum rules.)

We will be interested in bounds on γ_0 and β_1 , agnostic about α . As discussed in the previous subsection, this means that we can not use the sum rules containing $c_{P_X, \ell, \mu}^{00}$, as well as the sum rules that rely on $c_{P_X, \ell, \mu}^{00}$ to satisfy eq. (5.10). This means that we will only use improved dispersive sum rules (B.16)–(B.23), (B.49)–(B.52) and (B.55)–(B.57). Additionally, we also use some forward-limit sum rules to improve the numerical convergence in the large ℓ and finite μ region. Another *ad hoc* trick to improve the numerics in this region is to use both sum rules with helicities 1322 and 1232. This will include more null constraints in the SDP, given that our numerical implementation truncates the sum rules at a finite order of k (see eq. (3.18)).

To determine the boundary of the positivity region, we can make use of angular optimization. To this end, we parametrize γ_0 and β_1 as follows

$$\frac{\gamma_0 \Lambda^4}{M_P^2} = r \cos \theta, \quad \frac{\beta_1 \Lambda^2}{M_P} = r \sin \theta. \tag{6.12}$$

Then, for a given θ , the optimization program for bounds on γ_0 and β_1 outputs a quadratic inequality of r , which gives a bound on r ; going through sufficiently many θ , we get a 2D

bound in the γ_0 - β_1 plane. However, this is not an optimization problem directly solvable by the SDPB package, because both r and r^2 are present in the inequality. Nevertheless, for phenomenological interesting cases (for which the IR logarithm $\log(\Lambda/m_{\text{IR}})$ is not too small), we can drop the linear term in the r inequality, and then the problem becomes directly solvable by SDPB for a given θ . If we want to improve the accuracy of the bound with the linear r terms, we can use the above result as an initial background solution r_* of the quadratic r inequality and seek a linear perturbative improvement δr on top of it:

$$\frac{\gamma_0 \Lambda^4}{M_P^2} = (r_* - \delta r) \cos \theta, \quad \frac{\beta_1 \Lambda^2}{M_P} = (r_* - \delta r) \sin \theta. \quad (6.13)$$

This of course can be iterated for further improvements: set $r_* \rightarrow r_* - \delta r$ and repeat several times to a desired accuracy.

In figure 4, we compare the bounds obtained from the leading approximation and its improvements with the above iterations. The leading approximation, where the linear r terms are dropped, can be extracted by the limit $\log(\Lambda/m_{\text{IR}}) \rightarrow \infty$, while for the $\log(\Lambda/m_{\text{IR}}) = 50$ and $\log(\Lambda/m_{\text{IR}}) = 20$ cases we have performed two iterations of linear improvements. We see that, for a phenomenological interesting $\log(\Lambda/m_{\text{IR}})$, the leading approximation is actually rather good. The non-convexity of figure 4 results from the fact that the SDP is performed on quadratic functions of γ_0 and β_1 , but figure 4 is plotted for γ_0 and β_1 themselves. Also, from figure 4, we see that the allowed values of the dimensionless coefficients $\hat{\gamma}_0$ and $\hat{\beta}_1$, modulo $(\log(\Lambda/m_{\text{IR}}))^{1/2}$, are order one, which is consistent with the dimensional analysis in section 4. This is also consistent with the parametric bound on β_1 in ref. [115], estimated from requiring the absence of acausal time advances when the graviton and the scalar scatter off a heavy object in the eikonal regime. Also, the bounds on γ_0 have previously been computed in ref. [44], which can be compared with ours by setting $\beta_1 = 0$. Our bounds on γ_0 are a few percents stronger than those of ref. [44], which probably arises from the differences in using dispersion relations and approximations in the large μ, ℓ region.

To obtain the bounds on γ_0 and β_1 in figure 4, we only used sum rules that do not relate to the partial wave amplitude $c_{P_X, \ell, \mu}^{00}$. To utilize other sum rules, one needs to have some prior knowledge of α and, potentially, for a given α , the bound on γ_0 and β_1 could be significantly reduced. However, in section 4, we have estimated that this should not happen. Here, with the numerical scheme, we can confirm that the bound on γ_0 and β_1 is insensitive to the value of α ; see figure 5 for how the bound varies with α using the leading approximation. Even if the value of α has varied from near its lower bound to $\mathcal{O}(100)$, the impact on the bound on γ_0 and β_1 is only about 0.001%, consistent with a numerical error. Note that in figure 5 we only plot bounds on γ_0 and β_1 (more precisely γ_0^2 and β_1^2) in the first quadrant. The bounds on the other quadrants are almost mirror symmetric with respect to the one in the first quadrant, because the sum rules used mostly contain γ_0^2 and β_1^2 , except for one couple of them which contains $\beta_1 \gamma_0$. The effects of the sum rules with $\beta_1 \gamma_0$ turn out to be very weak in the optimization.

The $\phi \mathcal{G}$ operator itself is shift symmetric $\phi \rightarrow \phi + \text{const}$, but in plotting figure 5 we are agnostic about the coefficients of non-shift symmetric operators such as $\phi^2 \mathcal{G}$, that is, figure 5

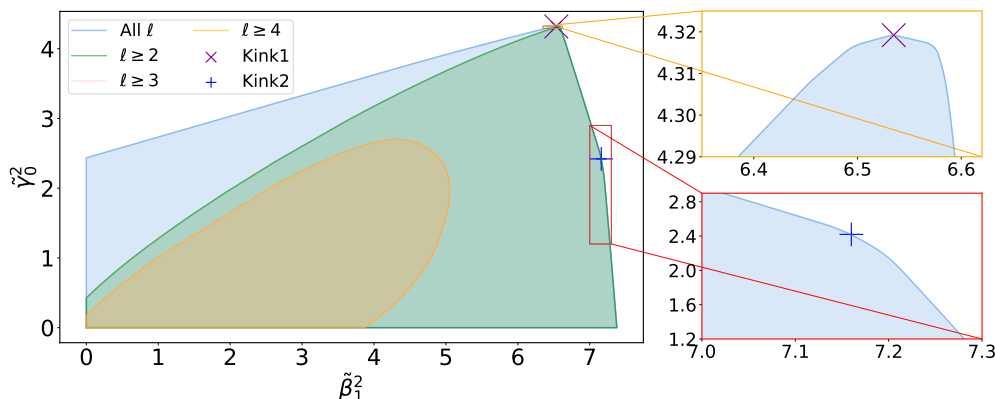


Figure 6. Bounds on γ_0^2 and β_1^2 with some low spins removed, where $\tilde{\gamma}_0^2 = \gamma_0^2 \Lambda^8 / (M_P^4 \log(\Lambda/m_{\text{IR}}))$ and $\tilde{\beta}_1^2 = \beta_1^2 \Lambda^4 / (M_P^2 \log(\Lambda/m_{\text{IR}}))$. For a line with $\ell \geq \ell_0$, we only impose positivity conditions with UV spin $\ell \geq \ell_0$. The $\ell = 1$ states decouple so that the $\ell \geq 1$ and $\ell \geq 2$ curves are the same. The $\ell \geq 3$ and $\ell \geq 4$ curves only differ slightly. The zoom-ins of the kinks are also shown.

is for a generic theory without shift symmetry. Incidentally, one may be interested in how the causality bounds differ if the theory is actually shift-symmetric. This is equivalent to setting the coefficients of terms like $\phi^2 \mathcal{G}$ to zero, which gives rise to some extra null constraints. However, our numerical results show that the bounds on γ_0 and β_1 are virtually unchanged if we impose the shift symmetry. The differences between the shift-symmetric bounds and the generic bounds are about $\mathcal{O}(10^{-5})$, which may well be numerical errors. This is not surprising from the point of view of the dispersion relations: the bounds on γ_0 and β_1 mostly come from the constraining powers of the four-graviton dispersion relations, but the latter do not contain γ_1 and β_2 at all.

In figure 5, we find that there are two kinks at $(\tilde{\gamma}_0^2, \tilde{\beta}_1^2) = (4.32, 6.53)$ and $(\tilde{\gamma}_0^2, \tilde{\beta}_1^2) = (2.42, 7.16)$, which will be referred to as “Kink 1” and “Kink 2” respectively; see figure 6 for the close-ups. Often, a kink indicates a theory with special features, as it delineates two continuous classes of EFTs. (From the perspective of the convex cone of the s^2 coefficients, already, kinks correspond to UV states that are irreps of the spacetime and internal symmetries of the EFT [4].) To probe the nature of these kinks, in figure 6, we also calculate the bounds on β_1^2 and γ_0^2 with some of the UV spin states suppressed. By $\ell \geq \ell_0$, we assume that there are no UV states with spin $\ell < \ell_0$. We do not plot the $\ell \geq 1$ case because the $\ell = 1$ states decouple from this process so that the $\ell \geq 1$ curve is the same as the $\ell \geq 2$ curve. Also, in figure 6, we also find that the $\ell \geq 3$ and $\ell \geq 4$ curves only differ slightly.

The reason why the bounds on β_1 and γ_0^2 receive no contributions from the $\ell = 1$ partial waves and are insensitive to the $\ell = 3$ partial waves is a result of spin selection rules, thanks to the fact that the EFT only includes spin-0 and spin-2 modes. To see this, note that the amplitude discontinuity in the dispersion relations schematically goes like

$$\text{Disc} \mathcal{M}^{1234}(\mu, t) \sim \sum_{\ell} d_{h_{12}, h_{43}}^{\ell}(\arccos(1 + 2t/\mu)) c_{\ell, \mu}^{12} (c_{\ell, \mu}^{\bar{3}\bar{4}})^*$$

For an odd ℓ , we have $c_{\ell, \mu}^{12} = 0$ if $1 = 2$, due to the Bose symmetry. On the other hand, for $1 \neq 2, 3 \neq 4$ and $\ell = \text{odd}$, the Wigner functions $d_{h_{12}, h_{43}}^{\ell}$ vanish if $\ell < \max\{|h_{12}|, |h_{43}|\}$. This

is the case for $\ell = 1$, because, for a scalar-tensor EFT, when $h_i \neq h_j$, we have $|h_i - h_j| = 2$ or 4. Therefore, the bounds receive no contributions from the $\ell = 1$ partial waves. For the $\ell = 3$ case, we still have $\ell < \max\{|h_{12}|, |h_{43}|\}$ if we consider pure graviton scattering in which we have $|h_i - h_j| = 4$. So, for $\ell = 3$, the Wigner functions in most of the dispersion relations vanish. Additionally, the bounds on γ_0 and β_1 turn out to be insensitive to the rest dispersion relations from the non-pure graviton scatterings. Therefore, the bounds are insensitive to the $\ell = 3$ partial waves.

Since the $\ell = 1$ states decouple, the $\ell \geq 2$ curve in figure 6 shows that the horizontal boundary of the all- ℓ bound on β_1^2 and γ_0^2 cannot be reached if we assume that there are no scalar degrees of freedom in the UV, while the vertical boundary does not have contributions from the heavy scalars.

As with the all- ℓ case, the bound on β_1^2 and γ_0^2 with $\ell \geq 2$ still mainly results from the four-graviton sum rules and is insensitive to the value of α . However, this changes if the UV theory only has higher spin states $\ell \geq 3$, and then the bounds on β_1^2 and γ_0^2 are significantly reduced in all directions, as shown in figure 6. Furthermore, for $\ell \geq 3$, we find that the dimension α is fixed to be $\mathcal{O}(1/(M_P^2 \Lambda^2))$ and can no longer reach the all- ℓ upper bound $\mathcal{O}(1/\Lambda^4)$. To understand this, we can look in the graviton decoupling limit, where all the forward limits of the dispersive sum rules can be used. Notice that the lowest order *st* null constraint $a_{3,1}^{0000} = a_{1,3}^{0000}$ gives

$$0 = 16\pi \sum_{\ell \geq 4, \text{even}; X} (2\ell + 1) \int_{\Lambda^2}^{\infty} \frac{d\mu}{\pi} \frac{\ell(\ell + 1)(\ell^2 + \ell - 8)}{2\mu^5} |c_{\ell, \mu}^{00}|^2, \quad (6.14)$$

where we have imposed $c_{\ell, \mu}^{12} = 0$ for $\ell < 3$ as intended and used the fact that $c_{\ell, \mu}^{00} = 0$ for odd ℓ . When $\ell \geq 4$, the right-hand side of eq. (6.14) is non-negative for all ℓ and μ . Therefore, eq. (6.14) implies $c_{\ell, \mu}^{00} = 0$ for all ℓ and μ , which leads to $a_{k,n}^{0000} = 0$ for $k + n \geq 3$. This means that the scalar self-interaction operators with dimension 8 or higher must vanish. Therefore, all these operators must be suppressed by appropriate powers of M_P away from the decoupling limit, which is consistent with the numerical results that $\alpha \sim \mathcal{O}(1/(M_P^2 \Lambda^2))$ in figure 6. In the language of section 4, this suggests that the correspondence $c_{\ell, \mu}^{00} \Leftrightarrow \Lambda/M_P$ is the only option. The numerical results are essentially the same if the UV theory only has higher spin states $\ell \geq 4$, as shown in figure 6.

On the other hand, a theory with only $\ell \geq 5$ does not exist. The reason is exactly the same as why a pure scalar theory with only $\ell \geq 3$ does not exist. Notice that, in the presence of gravitons, the lowest order null constraint for 2-to-2 scalar scattering in the forward limit is $0 = \langle \partial_t^4 F_{3,\ell}^{0000}(\mu, 0) \rangle$. With the assumption $\ell \geq 5$, it becomes

$$0 = 16\pi \sum_{\ell \geq 6, \text{even}; X} (2\ell + 1) \int_{\Lambda^2}^{\infty} \frac{d\mu}{\pi} \left(\frac{\ell^6}{18} + \frac{\ell^5}{6} - \frac{55\ell^4}{36} - \frac{10\ell^3}{3} + \frac{233\ell^2}{36} + \frac{49\ell}{6} \right) \frac{1}{\mu^8} |c_{\ell, \mu}^{00}|^2, \quad (6.15)$$

where we have used the fact that $c_{\ell, \mu}^{00} = 0$ for odd ℓ again. When $\ell \geq 6$, the right hand side is positive for all μ . Therefore we can infer that $c_{\ell, \mu}^{00} = 0$, which in turns implies that $1/M_P^2 = 0$. Therefore, such kind of scalar-tensor theories can not exist.

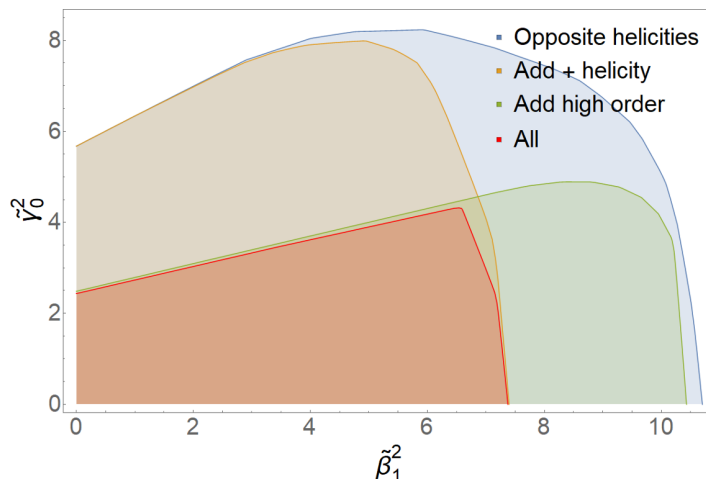


Figure 7. Bounds on β_1^2 and γ_0^2 using different sum rules. The “Opposite helicities” bound is obtained using 4 sum rules with F^{++--} and F^{+--+} , i.e., eqs. (B.49), (B.50), (B.55) and (B.56). The “Add + helicity” bound is obtained by adding the F^{++++} sum rules, i.e., eqs. (B.16) and (B.17), while the “Add high order” bound is obtained by adding 3 high order sum rules (B.51), (B.52) and (B.57), compared with the “Opposite helicities” case. The “All” bound is the one shown in figure 5. Kink 1 is located near the interaction point of “Add + helicity” and “Add high order”.

On the other hand, if the UV theory only has finite spins $\ell < \ell_M$, causality bounds will restrict β_1^2 and γ_0^2 to be zero. We can easily see this for $\ell < 4$ directly from the sum rules. To this end, note that we have $F_{1,\ell}^{++++}(\mu, t) = (\dots)d_{4,0}^{\ell,\mu,t} + (\dots)\partial_t d_{0,-4}^{\ell,\mu,0} + (\dots)\partial_t d_{4,0}^{\ell,\mu,0} = 0$ for $\ell < 4$ from their definitions. If the UV partial amplitude has no support for $\ell \geq 4$ spins, we can infer that $F_{1,\ell}^{++++}(\mu, t) = 0$ for all ℓ . Therefore, we have

$$-\frac{\gamma_0}{M_P^4}t^2 = \left\langle F_{1,\ell}^{++++}(\mu, t) \right\rangle = 0, \quad \text{for all } -\Lambda^2 < t < 0, \quad (6.16)$$

which suggest that $\gamma_0 = 0$. (The same result can also be obtained by using $F_{2,\ell}^{++++}$.) Similarly, for $F_{1,\ell}^{+0-0}$, we have $F_{1,\ell}^{+0-0}(\mu, t) = (\dots)d_{2,2}^{\ell,\mu,t} = 0$ for $\ell < 2$. If the UV partial amplitude has no support for $\ell \geq 2$ spins, we have

$$-\frac{1}{M_P^2} - \frac{\beta_1^2}{M_P^4}t^2 = \left\langle F_{1,\ell}^{+0-0}(\mu, t) \right\rangle = 0, \quad \text{for all } -\Lambda^2 < t < 0, \quad (6.17)$$

which leads to $\beta_1 = 0$. Moreover, this also leads to $M_P \rightarrow \infty$, which means that this kind of scalar-tensor theory is excluded by causality bounds. For some larger ℓ_M , we have numerically verified that β_1^2 and γ_0^2 are also forced to be zero by positivity bounds.

It is also instructive to see how presence or absence of certain sum rules impacts the bound on γ_0 and β_1 . Starting from a small set of sum rules with only the graviton scattering with opposite helicities, figure 7 shows that adding sum rules from \mathcal{M}^{++++} significantly strengthens the bound on β_1 , the coefficient of $\phi\mathcal{G}$, while adding high order sum rules with opposite helicities primarily enhances the bound on γ_0 , the coefficient of $\mathcal{R}^{(3)}$. The former is due to the fact that β_1^2 also appears in the sum rule with F^{++++} , unlike γ_0^2 . The latter is

$\Lambda(10^{-10}\text{eV})$	BHXB [98]	NS [104]	GW_{BBH} [102]	GW_{NSBH} [106]	GW_{CB} [106]
Conservative	1.5	2.2	1.6	2.1	2.4
Kink 1	1.4	2.1	1.6	2.0	2.3
Fine-tuned	0.27	0.39	0.30	0.38	0.43

Table 2. Lower bounds on the EFT cutoff Λ (in units of 10^{-10}eV) from binary compact star observations, by converting $\tilde{\beta}_1 M_P \log(\Lambda/m_{\text{IR}})/\Lambda^2 < \beta_1^{\text{obs}}$ with various choices of the dimensionless $\tilde{\beta}_1$. “Conservative”, “Kink 1” and “Fine-tuned” refer to choosing $\tilde{\beta}_1$ to be, respectively, its global upper bound, at Kink 1 in figure 5 and when $\tilde{g}_{4,0}^{T_1} = g_{4,0}^{T_1} \Lambda^6 M_P^2 / \log(\Lambda/m_{\text{IR}}) = 0.01$ and $\tilde{g}_{6,0}^{T_1} = g_{6,0}^{T_1} \Lambda^{10} M_P^2 / \log(\Lambda/m_{\text{IR}}) = 0.01$. The BHXB bound comes from a black hole low mass X-ray binary (A0620-00), while the neutron star (NS) bound is from the mass-radius measurement of pulsar J0740+6620. The other bounds are extracted from constraints from the dephasing of gravitational waves: GW_{BBH} is inferred from combining several low mass binary black hole events, GW_{NSBH} is from the best neutron star black hole binary event (GW200115) and GW_{CB} is extracted from combining several BBH and NSBH events.

because γ_0^2 can manifest in high order sum rules with opposite helicities. It can be observed that Kink 1 is roughly located at the intersection point of the two choices of adding extra sum rules in the optimization.

As mentioned previously, there has been a lot of recent interest in astrophysics to probe the $\phi\mathcal{G}$ operator in strong and dynamical gravity environments, as it is one of the leading viable scalar-curvature couplings beyond Einstein’s gravity. In some of these settings, the $\phi\mathcal{G}$ operator and the scalar kinetic term are assumed to be the only extra Lagrangian terms, which is a fine-tuned scenario we shall consider in section 6.6. From the perspective of the causality bounds, fine-tuned cases often have tighter bounds, sometimes incredibly restrictive, as we shall see. Therefore, one deduces conservative conclusions when comparing the generic causality bounds with the observational results for the fine-tuned theories. On the other hand, when constraining the β_1 coefficient from binary compact star observations, the results, in contrast to the causality bounds, are less sensitive to corrections from higher dimensional operators, as $\phi\mathcal{G}$ gives the leading contributions in the astrophysical computations. Thus, the proxy model with only $\phi\mathcal{G}$ should capture the salient astrophysical features of a generic model. With these in mind, we shall use the observational bounds to constrain the cutoff of the scalar-tensor EFT in the following.

Having established the sharp causality bounds, we can convert these experimental bounds to the bounds on the cutoff of the theory for a few specific EFTs. A specific EFT has a specific dimensionless $\tilde{\beta}_1$, and the lower bound on Λ can be extracted from

$$\frac{\tilde{\beta}_1 M_P \log(\Lambda/m_{\text{IR}})}{\Lambda^2} = \beta_1 < \beta_1^{\text{obs}}, \tag{6.18}$$

where β_1^{obs} is an observational bound and we choose $1/m_{\text{IR}}$ to be the Hubble scale. In table 2, we have surveyed three EFTs: for the “Conservative” case we take $\tilde{\beta}_1$ to be its maximum value in figure 5, which is valid regardless of values of other Wilson coefficients; “Kink 1” refers to the Kink 1 in figure 5 (the bounds on Λ for Kink 2 being almost

the same); the “Fine-tuned” case is when we take $\tilde{\beta}_1$ to be its maximum value when higher order coefficients $g_{4,0}^{T_1}$ and $g_{6,0}^{T_1}$ are set to be relatively small $\tilde{g}_{4,0}^{T_1} = 0.01$, $\tilde{g}_{6,0}^{T_1} = 0.01$, which will significantly reduce the upper bound on $|\tilde{\beta}_1|$ (see figure 13) and in turn impose much stronger bounds on the EFT cutoff. The observational constraints on β_1^{obs} in table 2 are obtained as follows. “BHXB” refers to a bound from a black hole low mass X-ray binary A0620-00 where the black hole’s companion is a K-type main-sequence star, whose matter is accreted into the black hole to produce X-rays [98]. “GW_{BBH}” [102], “GW_{NSBH}” [106] and “GW_{CB}” [106] are bounds from the newly available observational channel of gravitational waves. The BBH one is inferred from combining several most constraining low mass binary black hole events, the NSBH case is from the neutron star black hole binary event (GW200115) and the CB one is extracted from combining several BBH and NSBH events. All the bounds from these binaries are derived from the fact that the scalar Gauss-Bonnet coupling gives rise to corrections to the binary’s orbital decay rate due to extra scalar dipole radiation that is of “−1 PN” order. For X-ray observations from the BHXB, this results in corrections to the period, while for gravitational waves this leads to dephasing in the waveforms. It is also not surprising that the strongest bounds come from the lower mass compact stars, as gravity is the strongest in those environments. On the other hand, “NS” refers to a bound from the mass-radius measurement of pulsar J0740+6620, taking the most conservative case with respect to the choice of the equation of state for the neutron star [104].

To clarify, table 2 constrains the cutoffs of the models at a few special places within the causality bounds. These special models serve as benchmarks for causality-bound-compatible EFT models. Since $\tilde{\beta}_1$ is known for these specific models, one can convert the experimental bound to the lower bound on the cutoff, via eq. (6.18). This exercise is related to causality bounds because these specific models are special only from the point of view of the causality bounds. It can be viewed as a succinct survey about how the cutoff may change within the causality bounds, which serves to gauge the strength of the causality bounds in the context of observational bounds.

It is worth noting that the current experimental bounds are parametrically close to the bound from the cosmic censorship that there should be no naked singularity. An intriguing feature of the scalar-Gauss-Bonnet coupling $\phi\mathcal{G}$ is that it leads to a finite radius singularity at $r_S^4 = 48\beta_1^2$ [94]. So if we require the singularity to be cloaked by the event horizon, for a spherically symmetric black hole with its horizon at r_h , this leads to an upper bound on the coupling

$$\beta_1^2 < \frac{r_h^4}{48}, \tag{6.19}$$

(For generic $f(\varphi)\mathcal{G}$, we will have $(df(\varphi(r_h))/d\varphi)^2 < r_h^4/48$.) This will impose stronger bounds on the cutoff Λ for smaller black holes for which r_h is smaller. Taking r_h to be \sim km would lead to $\Lambda \gtrsim \mathcal{O}(1) \times 10^{-10}$ eV if $\tilde{\beta}_1$ saturates its upper bound.

6.3 More generic Gauss-Bonnet couplings

Beyond the linear coupling $\phi\mathcal{G}$, more complex couplings to the Gauss-Bonnet invariant $f(\phi)\mathcal{G}$ have also been extensively investigated, due to their ability to generate the inter-

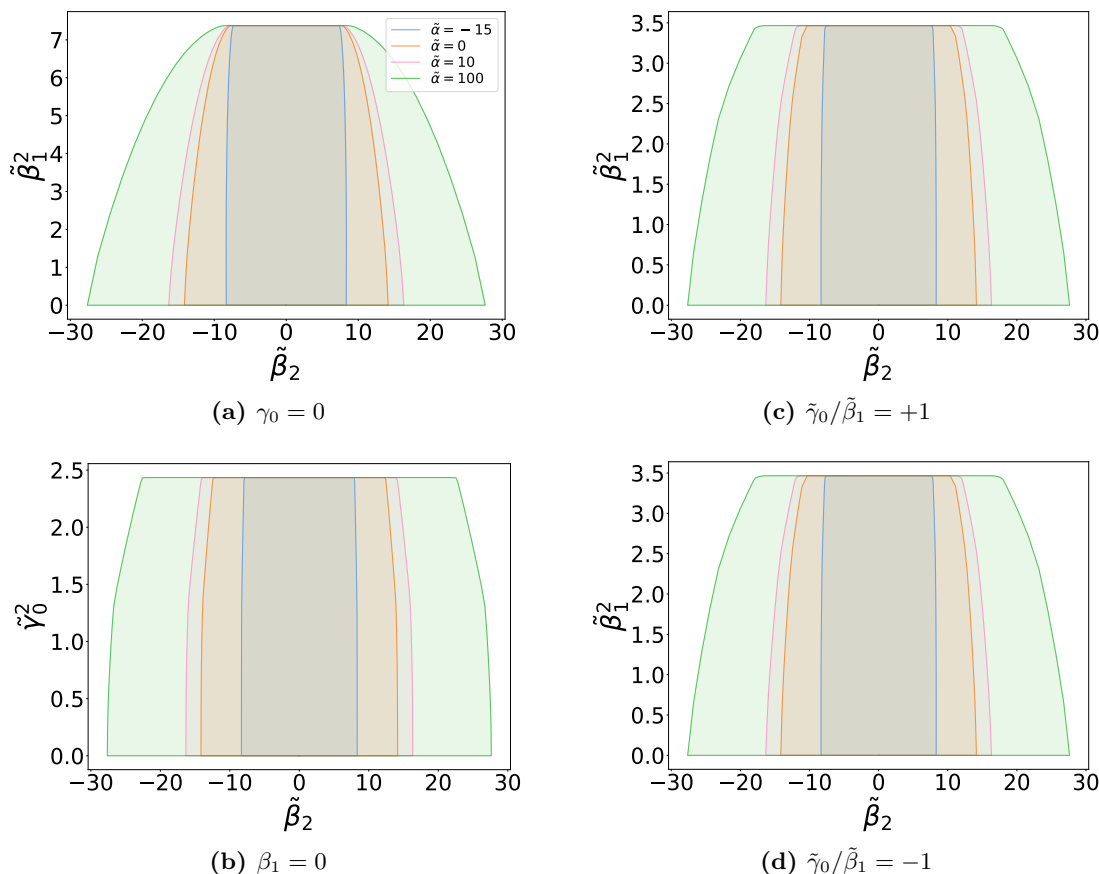


Figure 8. Bounds on β_2 , β_1^2 and γ_0^2 for various α . Four representative cross sections are chosen that pass through the β_2 axis. The dimensionless coefficients are defined as follows: $\tilde{\alpha} = \alpha M_P^2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{\beta}_2 = \beta_2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{\beta}_1^2 = \beta_1^2 \Lambda^4 / (M_P^2 \log(\Lambda/m_{\text{IR}}))$ and $\tilde{\gamma}_0^2 = \gamma_0^2 \Lambda^8 / (M_P^4 \log(\Lambda/m_{\text{IR}}))$.

esting mechanism of (spontaneous) scalarization [96, 97]. Scalarization is a phenomenon where the scalar field can develop a non-trivial profile (non-constant across the space) when the curvature near compact stars is sufficiently strong in some modified gravity models. All known scalarization mechanisms for the black hole utilize the scalar-Gauss-Bonnet coupling $f(\phi)\mathcal{G}$ [112] (for neutron stars other ways are possible). In particular, they usually rely on the $\phi^2\mathcal{G}$ term to give rise to an effective scalar mass term that is of a tachyonic nature, which is quenched by higher order terms to achieve stable scalarization. Near a compact star, the effective mass term has to be sufficiently negative to develop tachyonic instabilities, as there are positive contributions to the effective potential coming from the mass and angular momentum of the compact star. As the effective mass term is proportional to \mathcal{G} , this requires the curvature near the compact star to be sufficiently large, so smaller compact stars tend to be scalarized.

In this subsection, we shall constrain the function $f(\phi)$ to the next leading order ϕ^2 . The leading two orders of the $f(\phi)$ expansion can contribute to tree-level 4-leg amplitudes and thus can be effectively constrained with our method. More specifically, we will be

	$\beta_2 > 0$			$\beta_2 < 0$				
EoS	MS1	MPA1	WFF1	MS1	MPA1	ENG	APR4	WFF1
$\Lambda(10^{-10}\text{eV})$	1.4	2.1	3.4	2.9	4.0	4.7	5.2	5.9

Table 3. Lower bounds on the EFT cutoff Λ from the NS-WD binary J0348+0432 for various equations of state (EoS) of the neutron star.

NS-WD pair	J0348+0432	J1012+5307	J2222-0137
$\Lambda(10^{-10}\text{eV})$	4.0	3.6	3.7

Table 4. Lower bounds on the EFT cutoff Λ from three NS-WD binaries, assuming the MPA1 equation of state for the neutron star and $\beta_2 < 0$.

concerned with the coefficients in the following Lagrangian terms

$$\mathcal{L} \supset \sqrt{-g} \left(\frac{\gamma_0}{3!} \mathcal{R}^{(3)} + \frac{\beta_1}{2!} \phi \mathcal{G} + \frac{\beta_2}{4} \phi^2 \mathcal{G} + \frac{\alpha}{2} (\partial\phi)^4 \right) \quad (6.20)$$

and investigate how the bound on β_1 and β_2 varies with α and γ_0 .

Since we have seen in the last subsection that keeping only the quadratic pieces of γ_0 and β_1 in the dispersion relations quite accurately captures the bounds, we shall directly adapt that approximation here. Since β_2 is only contained in the dispersive sum rule with $F_{2,\ell}^{++00}$, in order to derive a bound on β_2 , we need to specify the value of α , as discussed in section 6.1. In this subsection, we assume $\alpha \sim \mathcal{O}(1/(M_P^2 \Lambda^2))$. (The case of $\alpha \sim \mathcal{O}(1/\Lambda^4)$, particularly the scaling of β_2 with respect to α , will be explored in section 6.5.)

A few cross sections of the 3D bounds for β_2 , β_1^2 and γ_0^2 are shown in figure 8, all of the cross sections passing through the β_2 axis. While the boundaries of β_1^2 and γ_0^2 are insensitive to the value of α , the limit of β_2 changes dramatically with it. On the other hand, the limit of β_2 is insensitive to the values of β_1^2 and γ_0^2 . We can see that β_2 is of order $\mathcal{O}(1/\Lambda^2)$ and becomes greater when α increases, consistent with the estimate in section 4. Another obvious visual pattern is that the bounds are symmetric with respect to the plane of $\beta_2 = 0$, up to about 0.01%. We have also numerically verified that, for $\tilde{\alpha} \geq -15$, the global minimum and maximum of β_2 are on the line of $\gamma_0 = 0$ and $\beta_1 = 0$. We have chosen both the cross sections of $\gamma_0 \Lambda^2 / (\beta_1 M_P) = \pm 1$ because there is a $\gamma_0 \beta_1$ term in the sum rules from \mathcal{M}^{++-0} and its crossing. However, as we can see explicitly in figure 8, the effect of this term is minimal in the results.

The reason why we choose $\tilde{\alpha}$ from $\tilde{\alpha} = -15$ to plot figure 8 is that $\tilde{\alpha} = -16.091$ is its lower bound. A peculiar feature near the lower bound when $\tilde{\alpha} = -15$ is that the cross sections in figure 8 are almost rectangular. From the 3D point of view, the bounds on β_2 , β_1 and γ_0 are basically a “plate” lying on the plane of the β_1^2 and γ_0^2 directions with a “thickness” along the β_2 direction. This means that the bounds on β_1 and γ_0 are almost independent of the bounds on β_2 , which is not at all obvious from the sum rules and we have not identified the underlying reason.

In the absence of the β_1 term, the β_2 term can result in scalarization in compact stars [96, 97]. As argued in section 4, a relatively suppressed β_1 , compared with β_2 , is

parametrically natural for a generic UV completion. We will also verify this numerically in section 6.5. The β_2 coupling has been observationally constrained with binary pulsars, as it can also give rise to dipole scalar radiation, which affects the orbital decay rate of the binaries. ref. [127] considered three neutron star-white dwarf (NS-WD) binaries, and has put some upper bounds on $|\beta_2|$ for both $\beta_2 > 0$ and $\beta_2 < 0$. Note that for black holes β_2 needs to be positive in order to have tachyonic instabilities, which is necessary for scalarization to occur, but for neutron stars both signs of β_2 are possible. Similar to the case of β_1 , we can convert these experimental constraints to bounds on the cutoff Λ by saturating $|\beta_2|$ with the causality bounds for $\alpha = 0$, $\beta_1 = 0$: $|\tilde{\beta}_2|/\Lambda^2 < |\beta_2^{\text{obs}}|$. In table 3 the uncertainties of the Λ bounds are surveyed for various different equations of state of the neutron stars, while in table 4 the Λ bounds extracted from three different NS-WD binaries are compared.

The observational bounds on β_1 and β_2 have only been established individually. Assuming the scalar interacts with gravitational strength, the β_1 term gives rise to the leading order effects in the relevant astrophysical processes, and we may take the observational upper bounds on β_1 to be valid for all reasonably valued β_2 . This is a rough approximation, which we shall be content with in the absence of any rigorous 2D astrophysical analysis so far.

Then, if we let the observational bound on β_2 saturate its causality bound (letting the right end of the thick green line, which lies on the $\tilde{\beta}_2$ axis, align with the most right end of the causality bound on $\tilde{\beta}_2$ when $\tilde{\beta}_1 = 0$), which fixes the cutoff of the EFT and in turn fixes the causality bound on β_1 and β_2 , the causality bound on β_1 and β_2 may be used to tentatively exclude parameter regions in the β_1 and β_2 space and compare with the observational bounds. See figure 9 for a comparison with a couple of choices of the observational bounds. We emphasize that this is not intended to be a rigorous comparison. Rather, it is just an attempt to estimate potential interactions between the causality bounds and the observational constraints, which should be updated when suitable astrophysical analyses become available.

6.4 Other six derivative terms

In the previous subsections, we have mainly focused on Lagrangian terms with four derivatives, except for γ_0 , which is a term with six derivatives. In this subsection, we shall compute the positivity bounds on all other six derivative terms: γ_1 , γ_2 , γ_3 and γ_4 . After all, as argued in section 4, if the scalar interactions are constrained to be comparable with the gravitational interactions (for example, when $\alpha \sim 1/(M_P^2 \Lambda^2)$), all the six derivative terms should be all suppressed by $1/\Lambda^4$ (cf. eq. (4.30)).

Numerically, for $\tilde{\alpha} = \alpha M_P^2 \Lambda^2 / \log(\Lambda/m_{\text{IR}}) \geq -15$, we find that γ_1 , γ_2 and γ_3 reach their global bounds (i.e., the loosest bounds) approximately when $\gamma_0 = 0$ and $\beta_1 = 0$, and the bounds on γ_4 are actually insensitive to the values of γ_0^2 and β_1^2 . To see how the bounds change with α , we shall present the bounds on γ_1 , γ_2 , γ_3 and γ_4 for $\tilde{\alpha} = \{-15, -10, 0, 10, 100\}$, with $\gamma_0 = 0$ and $\beta_1 = 0$, as shown in figure 10.

In section 4, we have argued that γ_1 must be $\mathcal{O}(M_P/\Lambda^4)$ and insensitive to the value of α . This is what we see with the SDP computations: in figure 10a, we see that the bounds

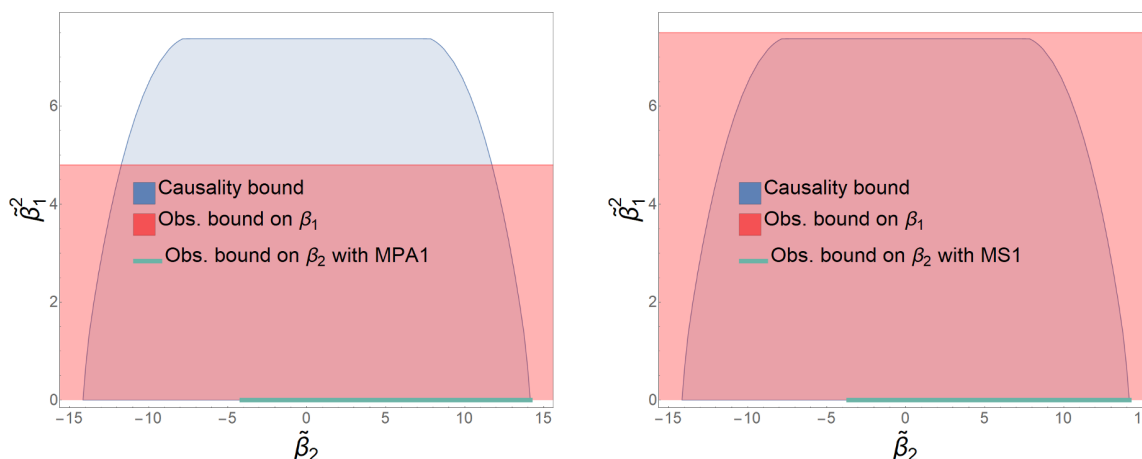


Figure 9. Comparisons of the observational bounds and causality bounds on β_1 when the observational bound on β_2 saturates its causality bound. The “Causality bound” corresponds to the theoretical bounds for the case of $\alpha = 0$ and $\gamma_0 = 0$ in figure 8. “Obs. bound on β_1 ” means the observational upper bound on β_1 coming from the most stringent gravitational wave constraint of [106], assuming $\beta_2 = 0$. “Obs. bound on β_2 with MPA1/MS1” means the observational upper bound on β_2 from the orbital decay rate measurements using the MPA1/MS1 equation of state for the neutron stars [127], assuming $\beta_1 = 0$. The observational bound on β_2 using MS1, shown on the right subfigure, is the most conservative one in [127]. We let the observational bound on β_2 saturate its causality bound, which means that the right end of the thick green line (lying on the $\tilde{\beta}_2$ axis) is at the most right end of the causality bound on $\tilde{\beta}_2$ when $\tilde{\beta}_1 = 0$.

on γ_1 depend very weakly on the value of α . Despite this, as mentioned in section 6.1, because γ_1 only appears in the sum rules of $F_{1,\ell}^{+++0}$ and $F_{2,\ell}^{+++0}$, we can not use our setup to numerically find the bounds on γ_1 without specifying α . Although hardly visible in figure 10a, the bound on γ_1 does become weaker very slowly when α increases. In fact, as will be shown in section 6.5, the value of γ_1 will be of the same order even when α is very large, for example, $\alpha \sim \mathcal{O}(1/\Lambda^4)$. On the other hand, the bounds on $\gamma_{2,3,4}$ become weaker significantly when α increases, so there is a sizable difference between the case of $\alpha \sim \mathcal{O}(1/(M_P^2\Lambda^2))$ and $\alpha \sim \mathcal{O}(1/\Lambda^4)$, which is again consistent with the analysis in section 4.

6.5 Coefficients for large α

In the previous subsections, we have seen that the bounds on some coefficients depend on the size of α . For the explicit bounds on these coefficients, we have chosen $\alpha \sim \mathcal{O}(1/(M_P^2\Lambda^2))$, in which case the scalar self-interaction, along with other interactions involving the scalar, is comparable with the spin-2 interactions. In this subsection, we shall also explore the possibility that $\alpha \sim \mathcal{O}(1/\Lambda^4)$ when the scalar self-interaction is strong, close to its upper limit. As already argued in section 4, the EFT operators scale differently in this case. In this subsection, we shall confirm these estimates with explicit calculations.

For concreteness, we will concentrate on the following three Wilson coefficients:

- γ_1 : the coefficient of the ϕR^3 term in the Lagrangian, which appears in the $F_{(1,2),\ell}^{+++0}(\mu, t)$ sum rule and thus is of order $\mathcal{O}(M_P/\Lambda^4)$ for any value of α , according to the dimensional analysis in section 4;

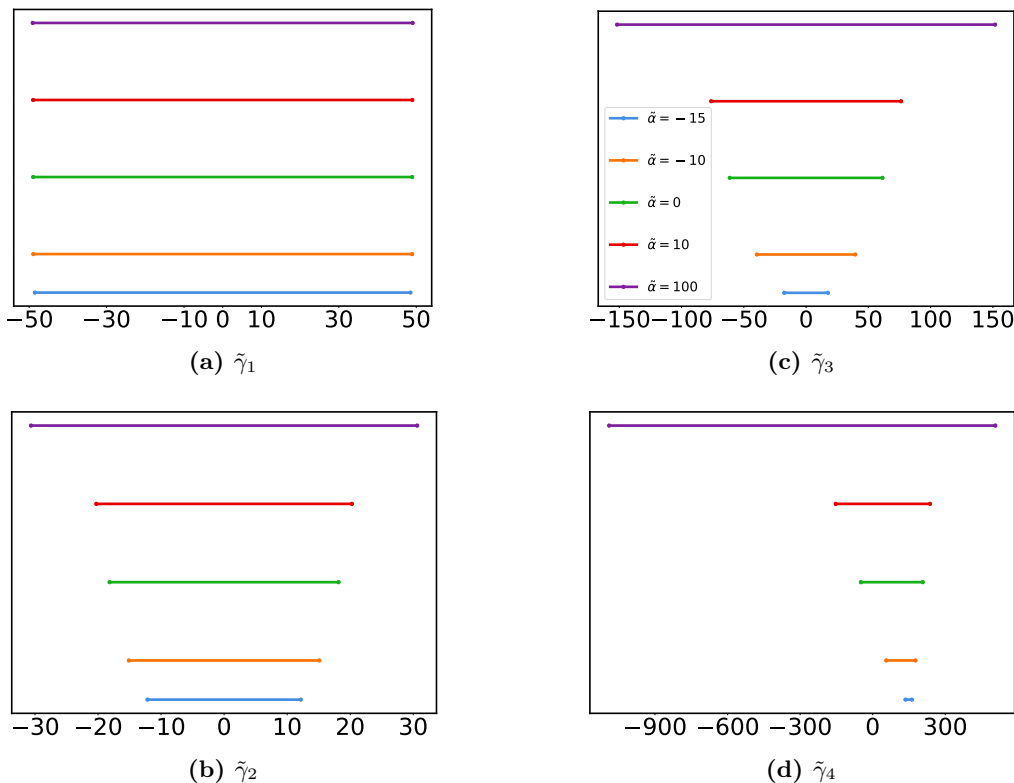


Figure 10. Bounds on $\gamma_{1,2,3,4}$ for various values of α . As expected, these bounds scale up as α increases, that is, as α transits from $\mathcal{O}(1/(M_P^2\Lambda^2))$ and $\mathcal{O}(1/\Lambda^4)$. Note that $\tilde{\alpha} = \alpha M_P^2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{\gamma}_1 = \gamma_1 \Lambda^4 / (M_P \log(\Lambda/m_{\text{IR}}))$, $\tilde{\gamma}_2 = \gamma_2 \Lambda^4 / \log(\Lambda/m_{\text{IR}})$, $\tilde{\gamma}_3 = \gamma_3 \Lambda^4 M_P / \log(\Lambda/m_{\text{IR}})$ and $\tilde{\gamma}_4 = \gamma_4 \Lambda^4 M_P^2 / \log(\Lambda/m_{\text{IR}})$.

β_2 : the coefficient of $\phi^2 \mathcal{G}$, which only appears in the sum rule with $F_{2,\ell}^{++00}$ and thus is of order $\mathcal{O}(1/\Lambda^2)$ when $\alpha \sim \mathcal{O}(1/(M_P^2\Lambda^2))$ or $\mathcal{O}(M_P/\Lambda^3)$ when $\alpha \sim \mathcal{O}(1/\Lambda^4)$;

$g_{0,2}^{M_4}$: the coefficient of the $s^2 u^2$ term in the amplitude \mathcal{M}^{+0-0} , which appears in the forward-limit sum rule $g_{0,2}^{M_4} = \langle 2|c_{\ell,\mu}^{+0}|^2/\mu^5 \rangle$ and thus should be of order $\mathcal{O}(1/(M_P^2\Lambda^6))$ and insensitive to α , meaning that the coefficient of a 8 derivative operator $\phi^2 R^2 \partial^4$ in the Lagrangian should scale as $\mathcal{O}(1/\Lambda^6)$ (this is an expected example that does not follow eq. (4.3) with $\tilde{N}_\phi = \lfloor N_\phi/2 \rfloor$, which would suggest an incorrect scaling of $\mathcal{O}(M_P/\Lambda^7)$).

We shall proceed by probing the bounds on these coefficients with a number of different hierarchies between M_P and Λ , up to a fiducial big ratio of $M_P/\Lambda = 10^{16}$. That is, we shall compute the bounds for these coefficients with α up to $\alpha \sim 10^{32} \log(\Lambda/m_{\text{IR}})/(M_P^2\Lambda^2)$. As we will see in the following, the bounds on coefficients such as γ_1 are insensitive to the changes in α , even for large α , whereas the bounds on coefficients such as β_2 increase significantly as α increases.

First, let us see how the upper bound on $|\gamma_1|$ varies for different α when $\gamma_0 = 0$ and $\beta_1 = 0$. The choice of $\gamma_0 = 0$ and $\beta_1 = 0$ makes the obtained bounds approximately the global upper bounds on $|\gamma_1|$ for all γ_0 and β_1 . As we see in figure 11a, the upper bound on

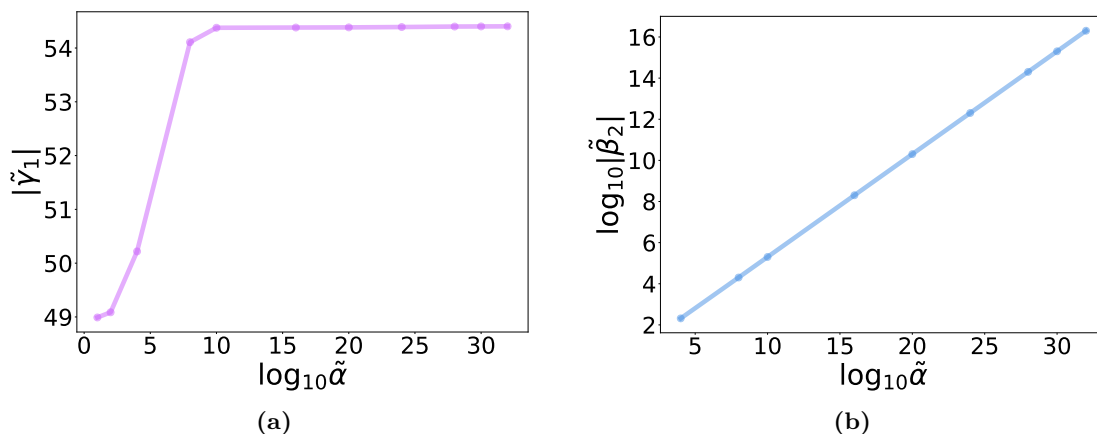


Figure 11. Upper bounds on $|\gamma_1|$ and $|\beta_2|$ for large values of α . The upper bound on $|\gamma_1|$ is insensitive to the value of α , while the upper bound on $|\beta_2|$ increases rapidly with α . The line in subfigure (b) is nearly linear with a slope of $1/2$, which is consistent with the argument in section 4. Note that $\tilde{\gamma}_1 = \gamma_1 \Lambda^4 / (M_P \log(\Lambda/m_{\text{IR}}))$, $\tilde{\beta}_2 = \beta_2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$ and $\tilde{\alpha} = \alpha \Lambda^2 M_P^2 / \log(\Lambda/m_{\text{IR}})$.

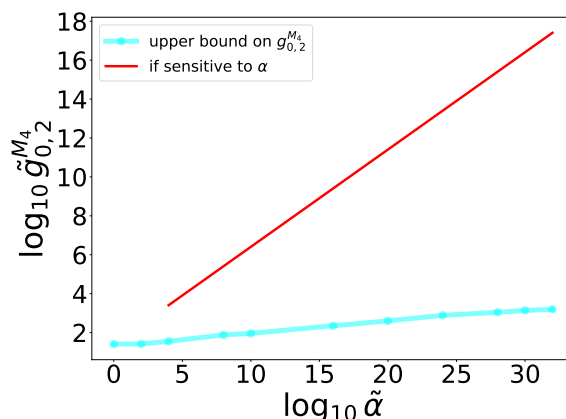


Figure 12. Upper bound on $g_{0,2}^{M_4}$ for large α . The upper bound on $\tilde{g}_{0,2}^{M_4}$ increases more rapidly than that on $\tilde{\gamma}_1$, but much slower than that on $\tilde{\beta}_2$. We have defined that $\tilde{\alpha} = \alpha \Lambda^2 M_P^2 / \log(\Lambda/m_{\text{IR}})$ and $\tilde{g}_{0,2}^{M_4} = g_{0,2}^{M_4} M_P^2 \Lambda^6 / \log(\Lambda/m_{\text{IR}})$. We see that, in contact to β_2 , its slope is much less than $1/2$ (the case of the red line), meaning that it is insensitive to α .

$|\gamma_1|$ remains stable around M_P/Λ^4 even when α has changed for many orders of magnitude, completely consistent with the dimensional analysis in section 4. Note that the dimensional analysis in section 4 suggests that γ_1 is insensitive to α , because eq. (4.18) does not contain $\hat{c}_{\ell,\mu}^{00}$. In deriving sharp bounds on γ_1 , we will use eq. (4.18) along with other dispersion relations. Although the dispersion relations containing γ_1 do not contain $c_{\ell,\mu}^{00}$, the rest dispersion relations do contain $c_{\ell,\mu}^{00}$. More specifically, γ_1 is contained in the dispersion relations with $F_{1,\ell}^{+++0}$ and $F_{2,\ell}^{+++0}$. When we add them into the numerical procedure and impose the positivity condition $B_{P_X,\ell}(\mu) \geq 0$, it is necessary for the $(+0, +0)$ element of $B_{P_X,\ell}(\mu)$ to be positive, which means that the dispersion relation with $F_{k,\ell}^{+0-0}$ must be included. For $k \geq 2$, including $F_{k,\ell}^{+0-0}$ will in turn contribute to the $(00, +0)$ and $(+0, 00)$ elements of $B_{P_X,\ell}(\mu)$. Hence, it is necessary for the $(00, 00)$ element of $B_{P_X,\ell}(\mu)$ to be

positive, so we need to include $F_{k,\ell}^{0000}$ in the SDP. $F_{k,\ell}^{0000}$ does contain $c_{\ell,\mu}^{00}$, which eventually leads to γ_1 having some dependence on α . These very indirect links also mean that the dependence of γ_1 on α is very weak, which is exactly what we see in figure 11a.

For the upper bounds on $|\beta_2|$, we again look at the direction along $\gamma_0 = 0$ and $\beta_1 = 0$, which gives approximately the global upper bounds on $|\beta_2|$. In figure 11b, we see that the upper bound on $|\beta_2|$ scales with the square root of α , accurate to several decimal places for large α ,

$$|\tilde{\beta}_2^{(\text{up})}| \propto \tilde{\alpha}^{\frac{1}{2}}, \tag{6.21}$$

precisely as what is argued in section 4. To see why this is consistent with the dimensional analysis in section 4, note that a large α of order $\alpha \sim \Lambda^{-4}$ can be viewed as originating from a large hierarchy between M_P and Λ : $\alpha \sim \Lambda^{-4} = (M_P/\Lambda)^2(M_P\Lambda)^{-2}$. So the horizontal axis in figure 11b can be viewed as depicting different values of $(M_P/\Lambda)^2$. In section 4, we argued that, switching from the $\tilde{c}_{\ell,\mu}^{00} \Leftrightarrow \Lambda/M_P$ correspondence to $\tilde{c}_{\ell,\mu}^{00} \Leftrightarrow 1$, the upper bound on $|\beta_2|$ will be boosted by an extra factor of M_P/Λ , to be of order $|\beta_2| \sim M_P/\Lambda^3 = (M_P/\Lambda)\Lambda^{-2}$, while for α the boost factor is $(M_P/\Lambda)^2$. This explains the 1/2 exponent in the fitted eq. (6.21).

The fact that β_2 increases significantly with α has interesting implications for the scalarization models. Notice that a scalarization model should accommodate non-hairy black holes, so the β_1 coefficient is usually assumed to be negligible, since a sizable $\phi\mathcal{G}$ coupling generically leads to a hairy black hole [93]. The fact that the causality bounds allow the β_2 coefficient to generically have an enhancement of a factor of up to M_P/Λ implies that the β_2 coupling can be naturally stronger than the β_1 coupling. This can be achieved by UV models where the scalar interacts with the heavy states stronger than the gravitational force.

Regarding the bounds on $g_{0,2}^{M_4}$, from the sum rule $g_{0,2}^{M_4} = \langle 2|c_{\ell,\mu}^{+0}|^2/\mu^5 \rangle$, we know that the lower bound on $g_{0,2}^{M_4}$ is 0, so let us compute its upper bound. Again, explicitly computation shows that $g_{0,2}^{M_4}$ reaches its global upper bounds when $\gamma_0 = 0$ and $\beta_1 = 0$. The dimensional analysis of the sum rule suggests that the bound should be insensitive to the value of α . Indeed, in figure 12, we see that the upper bound on $g_{0,2}^{M_4}$ only depends on α relatively weakly, although more sensitively than the case of γ_1 . This can be seen by comparing with the red line with slope 1/2, which is for the case if the upper bound were really sensitive to α . This example underlies the importance of rigorous calculations if we want to accurately capture the bounds on a specific coefficient.

6.6 Fine-tuned EFTs

Up to now, we have considered generic scalar-tensor EFTs without any *a priori* constraints on the Wilson coefficients. The bounds on them purely come from unitarity and causality of all possible UV theories, which as we have seen actually defines a power counting scheme for the higher dimensional EFT operators in the Lagrangian. However, for various reasons, one often devises models that fine-tune some of the Wilson coefficients to zero or be suppressed. These reasons may be of a UV nature, where certain UV symmetries or other mechanisms may prohibit the EFT from possessing certain operators; or, there could

be some phenomenological considerations to have certain coefficients highly suppressed or tuned to zero so as to make the model fit the observational data. Of course, many results about the leading $\phi\mathcal{G}$ coupling are insensitive to or independent of the higher order operators, as per the standard EFT power-counting. The same may not be said about other hairy black hole models or the scalarization models. In this subsection, we shall explore the consequences of fine-tunings for a couple of examples in scalar-tensor theory. Notice that *a priori* fine-tuning of the Wilson coefficients essentially gives rise to extra sum rules from the perspective of bounding the EFT. For example, restricting some coefficients to zero will lead to some null constraints, which sometimes may result in inconsistencies in the dispersion relations, as we shall see.

By this discussion, we wish to further raise the awareness of the theoretical obstacles one may face if one's model-building relies on fine-tunings of the EFT coefficients, explicitly demonstrating what can go wrong for some inconsiderate model-building with the EFT operators, illustrating the slogan “not everything goes”. That is, the main purpose of this subsection is not to promote these fine-tuned EFTs or trying to confront them with the observational constraints, rather it is to point out a caveat for potential misuses of fine-tunings in the EFT coefficients. It also acts as a reminder that even following the standard EFT power-counting may not be sufficient, and for a precision result there may be a need to also check with the causality bounds. Alternatively, one may reserve the argument and use the results here and/or the observational bounds on the lower dimensional operators to constrain the higher dimensional operators.

Let us first consider a simple example widely used in discussing hairy black holes and spontaneous scalarization, the $f(\phi)\mathcal{G}$ model. In this model, apart from the standard kinetic terms for the scalar and the graviton, all terms except the $f(\phi)\mathcal{G}$ operators are tuned to zero:

$$\mathcal{L} = \sqrt{-g} \left(\frac{M_P^2}{2} R - \frac{1}{2} (\partial\phi)^2 + \left(\frac{\beta_1}{2!} \phi + \frac{\beta_2}{4} \phi^2 + \dots \right) \mathcal{G} \right). \quad (6.22)$$

In this fine-tuned EFT, we no longer need to subtract the higher order t terms in the left hand side of the original dispersion relations, because the tree level EFT amplitude from this model only contains finite terms. Moreover, the absence of the higher order t terms gives rise to a multitude of extra null constraints. Therefore, in the model, we may simply define sum rules with

$$\tilde{F}_{k,l}^{1234}(\mu, t) := \frac{\partial_s^k}{k!} \left(\frac{s^2 d_{h_{12}, h_{43}}^{\ell, \mu, t}}{\mu^2 (\mu - s)} c_{\ell, \mu}^{12} c_{\ell, \mu}^{* \bar{3}4} + \frac{(-s - t)^2 d_{h_{14}, h_{23}}^{\ell, \mu, t}}{\mu^2 (\mu + s + t)} c_{\ell, \mu}^{14} c_{\ell, \mu}^{* \bar{3}2} \right) \Big|_{s \rightarrow 0}. \quad (6.23)$$

Taking this at its face value, it is easy to find *inconsistencies*. Let us look at the forward limit of the sum rule with $\tilde{F}_{4,\ell}^{++--}(\mu, 0)$:

$$0 = \left\langle \tilde{F}_{4,\ell}^{++--}(\mu, 0) \right\rangle = \left\langle \frac{1}{\mu^5} (|c_{\ell, \mu}^{++}|^2 + |c_{\ell, \mu}^{+-}|^2) \right\rangle. \quad (6.24)$$

Clearly, this implies that every 3-leg partial wave amplitudes $c_{\ell, \mu}^{++}$ and $c_{\ell, \mu}^{+-}$ must vanish for all μ . However, this clearly contradicts with the sum rule from $\tilde{F}_{2,\ell}^{++--}(\mu, t)$:

$$-\frac{1}{M_P^2} \frac{1}{t} = \left\langle \frac{1}{\mu^3} d_{0,0}^\ell (1 + 2t/\mu) |c_{\ell, \mu}^{++}|^2 + \frac{1}{(\mu + t)^3} d_{4,4}^\ell (1 + 2t/\mu) |c_{\ell, \mu}^{+-}|^2 \right\rangle = 0, \quad (6.25)$$

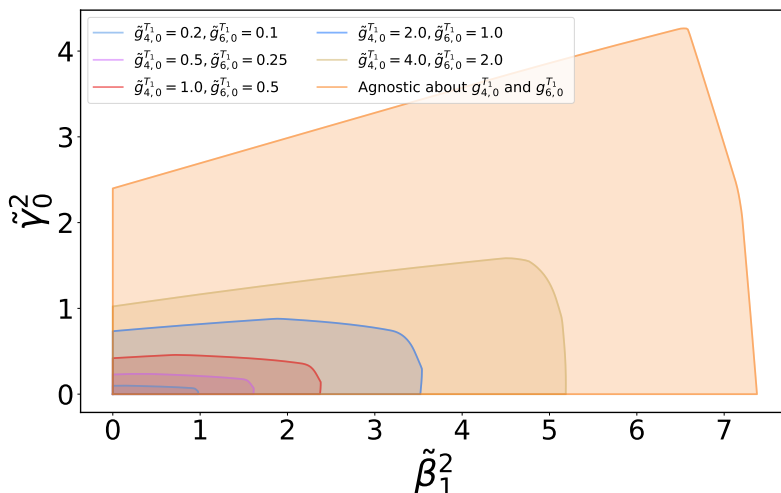


Figure 13. Bounds on γ_0^2 and β_1^2 for various $g_{4,0}^{T_1}$ and $g_{6,0}^{T_1}$. Causality bounds require $g_{4,0}^{T_1}$ and $g_{2,0}^{T_2}$ to be nonzero for γ_0 and β_1 to have non-vanishing values. We have defined that $\tilde{\beta}_1^2 = \beta_1^2 \Lambda^4 / (M_P^2 \log(\Lambda/m_{\text{IR}}))$, $\tilde{\gamma}_0^2 = \gamma_0^2 \Lambda^8 / (M_P^4 \log(\Lambda/m_{\text{IR}}))$, $\tilde{g}_{4,0}^{T_1} = g_{4,0}^{T_1} \Lambda^6 M_P^2 / \log(\Lambda/m_{\text{IR}})$ and $\tilde{g}_{6,0}^{T_1} = g_{6,0}^{T_1} \Lambda^{10} M_P^2 / \log(\Lambda/m_{\text{IR}})$.

because it forces M_P to be infinitely large or the Newton’s gravitational constant to vanish, which prevents the existence of a coupled scalar-tensor theory. This tells us that the $f(\phi)\mathcal{G}$ model (6.22) should not be taken at its face value. Instead, we should allow some nonzero values for the higher order operators.

Generically, our formalism provides us with the following criterion to test whether a fine-tuned scalar-tensor model is compatible with unitarity and causality: if we can deduce either $c_{\ell,\mu}^{00} = 0$ or $c_{\ell,\mu}^{+0} = 0$ or $c_{\ell,\mu}^{++} = c_{\ell,\mu}^{+-} = 0$ in the sum rules, then the scalar-tensor theory is inconsistent, in the sense that its Planck mass is forced to be infinitely large. To see this, notice that we have the following sum rules

$$-\frac{1}{M_P^2} + 2\alpha t - \gamma_4 t^2 = \left\langle F_{1,\ell}^{0000}(\mu, t) \right\rangle, \tag{6.26}$$

$$-\frac{1}{M_P^2} - \frac{\beta_1^2}{M_P^4} t^2 = \left\langle F_{1,\ell}^{+0-0}(\mu, t) \right\rangle, \tag{6.27}$$

$$-\frac{1}{M_P^2} \frac{1}{t} = \left\langle F_{1,\ell}^{++--}(\mu, t) \right\rangle. \tag{6.28}$$

Using the explicit expressions of the Wigner d-functions, we can see that $F_{1,\ell}^{0000}$ only contains $|c_{\ell,\mu}^{00}|^2$, $F_{1,\ell}^{+0-0}$ only contains $|c_{\ell,\mu}^{+0}|^2$ and $F_{2,\ell}^{++--}$ only contains a sign-definite combination of $|c_{\ell,\mu}^{++}|^2$ and $|c_{\ell,\mu}^{+-}|^2$. Thus, we can infer that $1/M_P^2$ must go to zero if $c_{\ell,\mu}^{00} = 0$ or $c_{\ell,\mu}^{+0} = 0$ or $c_{\ell,\mu}^{++} = c_{\ell,\mu}^{+-} = 0$.

Therefore, care should be taken to completely switch off coefficients that are allowed by the symmetries of the EFT. For the $f(\phi)\mathcal{G}$ model to be consistent with the causality bounds, we need to abandon the rigid definition of (6.22) and switch back on some other operators in the Lagrangian, for example, the γ_0 term or some other higher dimensional

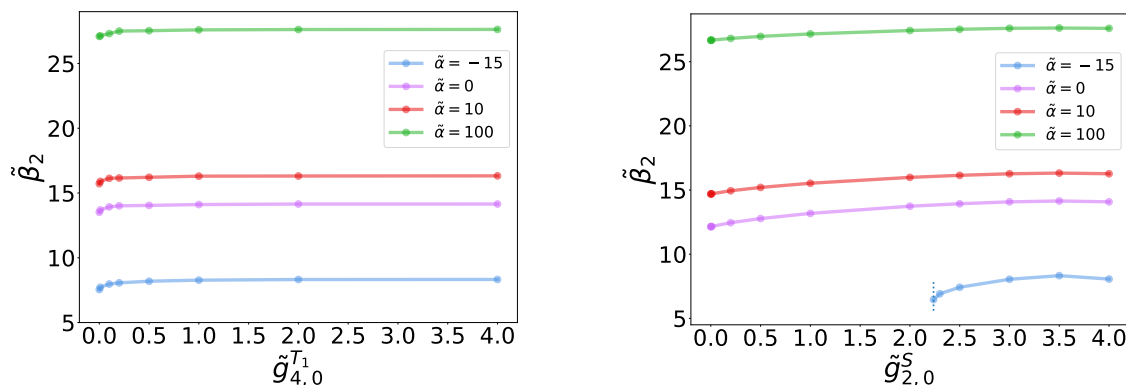


Figure 14. Insensitivity of the upper bound of $\tilde{\beta}_2$ to higher order coefficients. The short dotted line in the right subfigure denotes the lower bound on $\tilde{g}_{2,0}^S$ when $\tilde{\alpha} = -15$. We have defined that $\tilde{\alpha} = \alpha M_P^2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{\beta}_2 = \beta_2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{g}_{4,0}^{T_1} = g_{4,0}^{T_1} M_P^2 \Lambda^6 / \log(\Lambda/m_{\text{IR}})$ and $\tilde{g}_{2,0}^S = g_{2,0}^S M_P^2 \Lambda^6 / \log(\Lambda/m_{\text{IR}})$. Note that $\tilde{\alpha} = -15$ is almost the lower causality bound of $\tilde{\alpha}$.

terms. To determine how large the extra coefficients need to be in order to be consistent with causality and unitarity, we can run our numerical programs. We will see that the bounds on the coefficients of $f(\phi)$ shrink as we tune the higher dimensional coefficients to be smaller. For example, in figure 13, we can see how the bounds on γ_0^2 and β_1^2 reduce as $g_{4,0}^{T_1}$ and $g_{6,0}^{T_1}$ go toward zero along the surface $g_{4,0}^{T_1} = 2\Lambda^4 g_{6,0}^{T_1}$. It is interesting to see that these higher order terms in the Lagrangian can have such dramatic effects on the bounds on the lower order Wilson coefficients, merely assuming that there exists an analytic UV model, even though the higher orders may be negligible phenomenologically.

On the other hand, the bound on the β_2 coupling is insensitive to the higher order Wilson coefficients; see figure 14. Note that as $g_{4,0}^{T_1}$ and $g_{2,0}^S$ approach zero, the upper bound on β_2 decreases only slightly. Not visible in figure 14, we have also verified this numerically as both $g_{4,0}^{T_1}$ and $g_{2,0}^S$ approach zero.

We can understand the difference in sensitivity for β_1 and β_2 without actually solving the SDP. Let us look at the example of how $g_{4,0}^{T_1}$ affects the bounds on β_1 and β_2 . For $g_{4,0}^{T_1}$, its forward-limit sum rule is positive definite on the right hand side: $g_{4,0}^{T_1} = \langle \tilde{F}_{4,\ell}^{++--}(\mu, 0) \rangle = \langle (|c_{\ell,\mu}^{++}|^2 + |c_{\ell,\mu}^{+-}|^2) / \mu^5 \rangle \geq 0$. Note that this positive structure is important for the arguments below and, fortunately, this kind of forward-limit sum rules come by quite often. In the SDP, the $g_{4,0}^{T_1}$ coupling enters $B_{P_X,\ell}$ in eq. (5.10) as

$$B_{P_X,\ell} \sim (\dots) + y_* \mathcal{O}(\mu^{-n_*}) + y_{4,0}^{T_1} \mathcal{O}(\mu^{-5}) \geq 0 \tag{6.29}$$

where $y_{4,0}^{T_1}$ is the decision variable associated with $g_{4,0}^{T_1}$ in the optimization process and y_* is the decision variable associated with a Wilson coefficient β_* that we are concerned with. For every viable set of decision variables, acting $\langle \dots \rangle$ on eq. (6.29), we get a condition on the Wilson coefficients

$$[\dots] + y_* \beta_* + y_{4,0}^{T_1} g_{4,0}^{T_1} \geq 0 \tag{6.30}$$

First, suppose that $g_{4,0}^{T_1}$ is suppressed and becomes smaller, and let us see how it affects the bounds on the coefficient β_* . Owing to the smallness of $g_{4,0}^{T_1}$, $y_{4,0}^{T_1}$ can be very large and

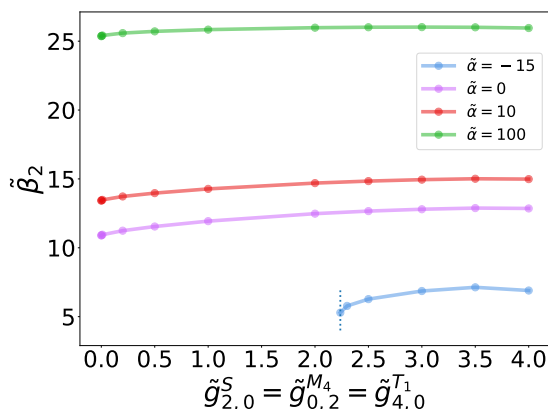


Figure 15. Entanglement in $B_{P_X, \ell}$ does not necessarily leads to strong correlations between the coefficients. We take $g_{2,0}^S$, $g_{0,2}^{M_4}$ and $g_{4,0}^{T_1}$ to approach zero along $\tilde{g}_{2,0}^S = \tilde{g}_{0,2}^{M_4} = \tilde{g}_{4,0}^{T_1}$. The short dotted line denotes the lower bound on $\tilde{g}_{2,0}^S$ when $\tilde{\alpha} = -15$. Note that $\tilde{\beta}_2 = \beta_2 \Lambda^2 / \log(\Lambda/m_{\text{IR}})$, $\tilde{g}_{4,0}^{T_1} = g_{4,0}^{T_1} M_P^2 \Lambda^6 / \log(\Lambda/m_{\text{IR}})$, $\tilde{g}_{0,2}^{M_4} = g_{0,2}^{M_4} M_P^2 \Lambda^6 / \log(\Lambda/m_{\text{IR}})$ and $\tilde{g}_{2,0}^S = g_{2,0}^S M_P^2 \Lambda^6 / \log(\Lambda/m_{\text{IR}})$.

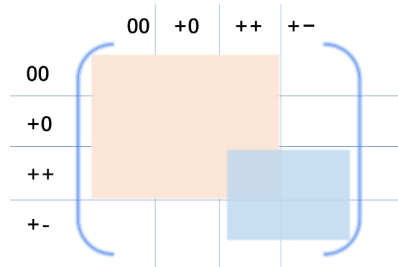
still does not significantly affect the inequality (6.30), and $g_{4,0}^{T_1}$ becoming smaller will allow $y_{4,0}^{T_1}$ to be larger. In the small $g_{4,0}^{T_1}$ limit, eq. (6.30) can be approximated by $[\dots] + y_* \beta_* \geq 0$, which leads to the upper bound on β_* : $\beta_* \leq |[\dots]/y_*|$. Then, thanks to the positivity of the $g_{4,0}^{T_1}$ sum rule, when $g_{4,0}^{T_1}$ becomes smaller, linear matrix inequality (6.29) will allow y_* to take more values, which in turn means that the bounds on β_* will become tighter. This is what we have seen for both β_1 and β_2 in figure 13 and figure 14, albeit for β_2 the effect is very small.

The reason why β_1 is sensitive to the value of $g_{4,0}^{T_1}$ is linked to the fact that $n_* = 4$ for β_1 . To see this, let us first consider the large μ region of linear matrix inequality (6.29). In this region, both the y_* and $y_{4,0}^{T_1}$ terms (i.e., the β_1 and $g_{4,0}^{T_1}$ terms) are negligible, compared to the leading $\mathcal{O}(\mu^{-3})$ term in $B_{P_X, \ell}$. So the large μ region does not significantly constrain y_* . On the other hand, in the small μ region, the positive $y_{4,0}^{T_1}$ term can be significant due to the $\mathcal{O}(\mu^{-5})$ scaling, which also leads to weak constraints on y_* . Therefore, a loosely constrained y_* results in a strong bound on β_1 . Furthermore, as $g_{4,0}^{T_1}$ becomes smaller, $y_{4,0}^{T_1}$ is allowed to take larger values, which leads to stronger bounds on β_1 . This is what we saw in figure 13.

On the other hand, for β_2 , we have $n_* = 3$ in eq. (6.29). In this case, the argument for the small μ region is similar to that of β_1 . However, in the large μ region, the y_* term has the leading $\mathcal{O}(\mu^{-3})$ scaling, so linear matrix inequality (6.29) now does impose significant constraints on y_* , which leads to loose bounds on β_2 . This explains why β_2 is insensitive to $g_{4,0}^{T_1}$.

One may be tempted to suggest that the difference between β_1 and β_2 in sensitivity to $g_{4,0}^{T_1}$ is due to the distribution of relevant terms in the $B_{P_X, \ell}$ matrix. One may observe that the $g_{4,0}^{T_1}$ sum rules give rise to nonzero contributions to the blue region in the lower

right corner of the $B_{P_X,\ell}$ matrix, as depicted below:



The main β_1 sum rules occupy the same region in $B_{P_X,\ell}$, while the β_2 sum rule occupies the brown region in the upper left corner, which only slightly overlaps with the $g_{4,0}^{T_1}$ block. This means that β_1 and $g_{4,0}^{T_1}$ are more entangled in the $B_{P_X,\ell}$ matrix, which might suggest that the mixing in $B_{P_X,\ell}$ is the main reason for β_1 to be more sensitive to $g_{4,0}^{T_1}$. However, this might not be the case here. We find that, while being separated in $B_{P_X,\ell}$ generally leads to insensitivity between the coefficients, being mixed in $B_{P_X,\ell}$ does not necessarily leads to strong correlations between the coefficients. For example, for the case of figure 15, the three higher order coefficients occupy the whole $B_{P_X,\ell}$ matrix, and yet we still find that β_2 is insensitive to these coefficients.

Acknowledgments

We would like to thank Yue-Zhou Li, Ning Su, Shi-Lin Wan, Hao Xu and Yang Zhang for helpful discussions. SYZ acknowledges support from the Fundamental Research Funds for the Central Universities under grant No. WK2030000036, from the National Natural Science Foundation of China under grant No. 12075233 and 12247103, and from the National Key R&D Program of China under grant No. 2022YFC220010. We acknowledge use of the GNU parallel package [128, 129].

A Generic 4-leg amplitudes for scalar-tensor theory

In this appendix, we shall derive the generic forms of the tree-level amplitudes for scalar-tensor theory. The amplitudes can be written as functions of s, t, u , with certain symmetries among the Mandelstam variables, and also need to satisfy the helicity structure of the scattering particles. For a tree-level amplitude, there are only two types of contributions, one being two 3-leg vertices connected by a propagator and the other type being a 4-leg contact vertex.

For massless particles, the on-shell 3-leg amplitudes, with the momenta extended to be complex, are uniquely fixed by the momentum conservation and the little group scaling up to an overall constant [130]

$$\mathcal{M}(1^{h_1} 2^{h_2} 3^{h_3}) \propto \begin{cases} \langle 12 \rangle^{h_3-h_1-h_2} \langle 23 \rangle^{h_1-h_2-h_3} \langle 31 \rangle^{h_2-h_3-h_1}, & h \leq 0, \\ [12]^{h_1+h_2-h_3} [23]^{h_2+h_3-h_1} [31]^{h_3+h_1-h_2}, & h \geq 0, \end{cases} \quad (\text{A.1})$$

where h_i is the helicity of particle i and $h \equiv h_1 + h_2 + h_3$. A 4-leg amplitude can be obtained by glueing one leg of a 3-leg amplitude with one leg of another 3-leg amplitude with opposite helicity. Alternatively, we can simply enumerate the Lagrangian terms with lowest few mass dimensions and compute the leading few orders of amplitudes from those terms. This allows us to enumerate all possible pole contributions to the amplitudes from double 3-leg vertex insertions. Then, the rest of the terms can be enumerated in a fashion similar to how eq. (A.1) is obtained, as we shall see shortly. For the latter approach, note that the Lagrangian terms that can give rise to 3-leg vertices are given by

$$\mathcal{L} \supset \sqrt{-g} \left(\frac{M_P^2}{2} R - \frac{1}{2} \nabla_\mu \phi \nabla^\mu \phi - \frac{\lambda_3}{3!} \phi^3 + \frac{\beta_1}{2!} \phi \mathcal{G} + \frac{\gamma_0}{3!} \mathcal{R}^{(3)} \right). \quad (\text{A.2})$$

Computing the relevant amplitudes with these terms only, the independent 2-to-2 amplitudes are given by

$$\mathcal{M}_{(3)}^{0000} = \lambda_3^2 \left(\frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right) + \frac{1}{M_P^2} \left(\frac{su}{t} + \frac{st}{u} + \frac{ut}{s} \right), \quad (\text{A.3})$$

$$\mathcal{M}_{(3)}^{++--} = \frac{1}{M_P^2} \frac{s^3}{tu} - \frac{\beta_1^2}{M_P^4} s^3 + \frac{\gamma_0^2}{M_P^6} s^3 tu, \quad (\text{A.4})$$

$$\mathcal{M}_{(3)}^{++++} = \frac{\gamma_0}{M_P^4} stu, \quad (\text{A.5})$$

$$\mathcal{M}_{(3)}^{++++} = \frac{10\gamma_0}{M_P^4} stu - \frac{3\beta_1^2}{M_P^4} stu + \frac{\gamma_0^2}{M_P^6} stu(s^2 + t^2 + u^2), \quad (\text{A.6})$$

$$\mathcal{M}_{(3)}^{+++0} = 0, \quad (\text{A.7})$$

$$\mathcal{M}_{(3)}^{++0-} = \frac{\beta_1}{M_P^3} s^2 - \frac{\gamma_0 \beta_1}{M_P^5} s^2 tu, \quad (\text{A.8})$$

$$\mathcal{M}_{(3)}^{++00} = \frac{\lambda_3 \beta_1}{M_P^2} s + \frac{\gamma_0}{M_P^4} stu + \frac{\beta_1^2}{M_P^4} s^3, \quad (\text{A.9})$$

$$\mathcal{M}_{(3)}^{+-00} = \frac{1}{M_P^2} \frac{tu}{s} + \frac{\beta_1^2}{M_P^4} stu, \quad (\text{A.10})$$

$$\mathcal{M}_{(3)}^{+000} = \frac{\beta_1}{2M_P^3} (s^2 + t^2 + u^2), \quad (\text{A.11})$$

where we have included contributions from both the amplitudes with double 3-leg insertions and those from contact 4-leg vertices. The subscript (3) indicates that these contributions are from the above Lagrangian terms containing 3-leg vertices. Note that most of the terms in eq. (A.2) can generate both 3-leg and 4-leg vertices, and it is only when both of them are included can the Ward identities be satisfied. Amplitudes with other helicities can be obtained from the above ones via crossing, using eq. (2.13) to eq. (2.21).

Having found all terms including the 3-leg vertices, we now turn to those purely from the contact vertices. Possible such terms can be constructed by considering restrictions from little group scaling, momentum conservation, locality and crossing symmetry. To see how this works, first note that a 4-momentum goes like $p \sim ||\langle |$ and a polarization vector goes like $\epsilon \sim ||\langle | / \rangle$ or $||\langle | / \rangle$. So, by Lorentz symmetry, a 4-leg amplitude term from a

contact term must be a product of powers of $[ij]$ and $\langle ij \rangle$ going like

$$\mathcal{M}_m^{1234} \propto [12]^{a_{12}} [13]^{a_{13}} [14]^{a_{14}} [23]^{a_{23}} [24]^{a_{24}} [34]^{a_{34}} \langle 12 \rangle^{b_{12}} \langle 13 \rangle^{b_{13}} \langle 14 \rangle^{b_{14}} \langle 23 \rangle^{b_{23}} \langle 24 \rangle^{b_{24}} \langle 34 \rangle^{b_{34}}, \quad (\text{A.12})$$

where m denotes the number of partial derivatives in the contact vertex and a_{ij} and b_{ij} are integers. From little group scaling and locality, we can infer some constraints on a_{ij} and b_{ij} .

Let us see what these constraints are. Since a momentum in the contact vertex is of dimension 1 and polarization vectors or tensors are dimensionless, we must have

$$a_{12} + a_{13} + a_{14} + a_{23} + a_{24} + a_{34} + b_{12} + b_{13} + b_{14} + b_{23} + b_{24} + b_{34} = m. \quad (\text{A.13})$$

Also, since $|i]$ and $|i\rangle$ scale as $|i] \rightarrow t_i|i]$, $|i\rangle \rightarrow t_i^{-1}|i\rangle$ and polarization tensors scales $\epsilon_{h_i} \rightarrow t_i^{2h_i}\epsilon_{h_i}$ under little group scaling, we have

$$\begin{cases} a_{12} + a_{13} + a_{14} - b_{12} - b_{13} - b_{14} = 2h_1, \\ a_{12} + a_{23} + a_{24} - b_{12} - b_{23} - b_{24} = 2h_2, \\ a_{13} + a_{23} + a_{34} - b_{13} - b_{23} - b_{34} = 2h_3, \\ a_{14} + a_{24} + a_{34} - b_{14} - b_{24} - b_{34} = 2h_4. \end{cases} \quad (\text{A.14})$$

Furthermore, since we are considering a theory that is local and the momenta from the partial derivatives in the local EFT operators contribute non-negatively to the amplitude's dimension, it must be that the minima of m for the following helicity amplitudes are as follows

$$\begin{cases} m \geq 0, & \text{for } \mathcal{M}_m^{0000}, \\ m \geq 2, & \text{for } \mathcal{M}_m^{+000}, \\ m \geq 4, & \text{for } \mathcal{M}_m^{++00}, \mathcal{M}_m^{+-00}, \mathcal{M}_m^{++--}, \\ m \geq 6, & \text{for } \mathcal{M}_m^{+++0}, \mathcal{M}_m^{++-0}, \\ m \geq 8, & \text{for } \mathcal{M}_m^{++++}, \mathcal{M}_m^{+++ -}. \end{cases} \quad (\text{A.15})$$

The reason why the lower m terms vanish is similar to the well-known fact that in non-Abelian gauge theory non-MHV 2-to-2 amplitudes must vanish. The easiest way to see this for cases other than \mathcal{M}_m^{++--} , as in the case of non-Abelian gauge theory, is to appropriately choose reference momenta to make all the contractions between the polarizations vanish. So the nonzero \mathcal{M}_m^{1234} (except for \mathcal{M}_m^{++--}) are those where every Lorentz index in the polarizations is accompanied by one contracting momentum from the vertex partial derivative. As for \mathcal{M}_m^{++--} , we can set the reference momentum in the polarization tensor of the 1st and 2nd particle to be p_4 and that of the 3rd and 4th particle to be p_1 , which will only leave the contraction of polarizations between particle 2 and 3 nonzero, and then the polarizations of particle 1 and 4 must contract with momenta from the vertex to give rise to nonzero terms, leading to $m \geq 4$ for \mathcal{M}_m^{++--} . We emphasize that although it is easy to see it by choosing some special reference momenta, the constraint (A.15) obviously holds

independent of the choice of reference momenta. So we still have the freedom to choose the reference momenta.

Indeed, if we choose the reference momentum in the polarization tensor of the 1st and 2nd particle to be p_4 and that of the 3rd and 4th particle to be p_1 for all the helicity amplitudes, then we further have the following constraints

$$a_{13} \geq \min[0, h_3], \quad a_{14} \geq \min[0, h_1] + \min[0, h_4], \quad a_{24} \geq \min[0, h_2], \quad (\text{A.16})$$

$$b_{13} \geq \min[0, -h_3], \quad b_{14} \geq \min[0, -h_1] + \min[0, -h_4], \quad b_{24} \geq \min[0, -h_2], \quad (\text{A.17})$$

and all the other a_{ij} and b_{ij} are non-negative. (A caveat is that one should find appropriate reference momenta in the above construction; otherwise there can be spurious terms in the final amplitude. This can be done by going through a few choices of the reference momenta and pick up the most constraining one.) Furthermore, the contact vertices do not give rise to poles of s, t, u in the amplitude, so we also have

$$\begin{cases} a_{12} + a_{34} + b_{12} + b_{34} \geq 0, \\ a_{13} + a_{24} + b_{13} + b_{24} \geq 0, \\ a_{14} + a_{23} + b_{14} + b_{23} \geq 0. \end{cases} \quad (\text{A.18})$$

With all these constraints established, we can solve these constraints for a_{ij} and b_{ij} . Typically, these constraint equations lead to multiple (in fact, many) solutions. For example, for the case of $m = 4, h_1 = h_2 = +2, h_3 = h_4 = 0$, there are 9 solutions for a_{ij} and b_{ij} , while we have 4570 solutions for $m = 10, h_1 = h_2 = h_3 = h_4 = +2$. However, they all collapse to a small number of cases after converting to expressions in terms of s, t, u .

To convert to an expression in terms of s, t, u , we can use an explicit choice for the momenta (all momenta chosen as ingoing and thus related to physical ones by $p_3^{\text{physical}} = -p_3$ and $p_4^{\text{physical}} = -p_4$) and the spinors

$$\begin{aligned} p_1 &= \omega(1, 0, 0, 1), & p_2 &= \omega(1, 0, 0, -1), & p_3 &= -\omega(1, \sin \theta, 0, \cos \theta), & p_4 &= -\omega(1, -\sin \theta, 0, -\cos \theta), \\ |1\rangle &= \sqrt{2\omega} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & |2\rangle &= \sqrt{2\omega} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & |3\rangle &= i\sqrt{2\omega} \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix}, & |4\rangle &= i\sqrt{2\omega} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}, \end{aligned} \quad (\text{A.19})$$

where particle 1, 2, 3 and 4 are moving in the direction of $(0, 0), (\pi, \pi), (\theta, \phi)$ and $(\pi - \theta, \phi + \pi)$ with $\phi = 0$ respectively. Here θ and ϕ are the polar and azimuthal angles. There is an extra i in the $|3\rangle$ and $|4\rangle$ expression because we need to analytically continue $\sqrt{\omega}$ to $i\sqrt{\omega}$ to account for unphysical p_3 and p_4 .² Also, we have $s = -(p_1 + p_2)^2 = -(p_3 + p_4)^2 =$

²For a massless particle, a generic momentum is given by $p^\mu = \omega(1, \sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$, and a generic helicity-spinor can be written as

$$|p\rangle^{\dot{a}} = \tilde{\lambda}^{\dot{a}} = \sqrt{2\omega} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{pmatrix}, \quad |p\rangle_a = \lambda_a = \sqrt{2\omega} \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\phi} \\ +\cos \frac{\theta}{2} \end{pmatrix}. \quad (\text{A.20})$$

Note that $\varepsilon^{12} = \varepsilon_{21} = +1, \varepsilon^{21} = \varepsilon_{12} = -1, \varepsilon^{ij} = \varepsilon^{i\dot{j}}, \varepsilon_{ij} = \varepsilon_{i\dot{j}}$, We have $p^{\text{physical}} = \varepsilon_p p$, where $\varepsilon_p = -1$ if the direction of the physical momentum goes against the assumed direction (otherwise $\varepsilon_p = 1$), and also $\langle pk \rangle^* = \varepsilon_p \varepsilon_k [kp]$ because there is an extra i in $|j\rangle$ and $|j\rangle$.

$(2\omega)^2$, $t = -s(1 - \cos\theta)/2 = -s \sin^2(\theta/2)$, $u = -s(1 + \cos\theta)/2 = -s \cos^2(\theta/2)$. We can cast ω and θ in terms of s, t, u

$$\omega = \frac{1}{2}\sqrt{s}, \quad \sqrt{2\omega} = s^{\frac{1}{4}}, \quad \sin\frac{\theta}{2} = \sqrt{\frac{-t}{s}}, \quad \cos\frac{\theta}{2} = \sqrt{\frac{-u}{s}}, \quad (\text{A.21})$$

from which we can find that

$$\langle 12 \rangle = \sqrt{s}, \quad \langle 13 \rangle = -i\sqrt{-t}, \quad \langle 14 \rangle = i\sqrt{-u}, \quad \langle 23 \rangle = -i\sqrt{-u}, \quad \langle 24 \rangle = -i\sqrt{-t}, \quad \langle 34 \rangle = -\sqrt{s}, \quad (\text{A.22})$$

$$[12] = -\sqrt{s}, \quad [13] = i\sqrt{-t}, \quad [14] = -i\sqrt{-u}, \quad [23] = i\sqrt{-u}, \quad [24] = i\sqrt{-t}, \quad [34] = \sqrt{s}. \quad (\text{A.23})$$

Substituting these replacements into the large numbers of expressions in terms of $[ij]$ and $\langle ij \rangle$ and imposing appropriate crossing symmetries, we can see that they collapse to a small number of functions of s, t, u .

In the end, we find that the results are consistent with simply taking the $f_i(s, t, u)$ functions in eqs. (2.3)–(2.11) to be generic polynomials of s, t, u that share the symmetries of the corresponding amplitudes, except for \mathcal{M}^{++++} . In the \mathcal{M}^{++++} case, letting $f_{T_3}(s, t, u)$ be generic polynomials of x, y would give rise to a couple of spurious terms, which should vanish according to the analysis above. Therefore, including the contributions from the 3-leg vertices (eq. (A.3) to eq. (A.11)), we can parametrize the 4-leg amplitudes as follows:

$$\mathcal{M}^{0000} = \lambda_3^2 \left(\frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right) - \lambda_4 + \frac{1}{M_P^2} \left(\frac{su}{t} + \frac{st}{u} + \frac{ut}{s} \right) + \sum_{n \geq 0, m \geq 0} g_{m,n}^S x^n y^m, \quad (\text{A.24})$$

$$\mathcal{M}^{++--} = \frac{1}{M_P^2} \frac{s^3}{tu} - \frac{\beta_1^2}{M_P^4} s^3 + \frac{\gamma_0^2}{M_P^6} s^3 tu + \sum_{n \geq 4, m \geq 0} g_{n,m}^{T_1} s^n (tu)^m, \quad (\text{A.25})$$

$$\mathcal{M}^{++++} = \frac{\gamma_0}{M_P^4} y + \sum_{n \geq 0, m \geq 2} g_{m,n}^{T_2} x^n y^m \quad (\text{A.26})$$

$$\mathcal{M}^{++++} = \left(\frac{10\gamma_0}{M_P^4} - \frac{3\beta_1^2}{M_P^4} \right) y + \sum_{\substack{n \geq 0, m \geq 0, \\ m+n \geq 2}} g_{m,n}^{T_3} x^n y^m, \quad (\text{A.27})$$

$$\mathcal{M}^{+++0} = \sum_{n \geq 0, m \geq 1} g_{m,n}^{M_1} x^n y^m, \quad (\text{A.28})$$

$$\mathcal{M}^{++0-} = \frac{\beta_1}{M_P^3} s^2 - \frac{\gamma_0 \beta_1}{M_P^5} s^2 tu + \sum_{n \geq 3, m \geq 1} g_{n,m}^{M_2} s^n (tu)^m, \quad (\text{A.29})$$

$$\mathcal{M}^{++00} = \frac{\lambda_3 \beta_1}{M_P^2} s + \frac{\gamma_0}{M_P^4} stu + \sum_{n \geq 2, m \geq 0} g_{n,m}^{M_3} s^n (tu)^m, \quad (\text{A.30})$$

$$\mathcal{M}^{+-00} = \frac{1}{M_P^2} \frac{tu}{s} + \frac{\beta_1^2}{M_P^4} stu + \sum_{n \geq 0, m \geq 2} g_{n,m}^{M_4} s^n (tu)^m, \quad (\text{A.31})$$

$$\mathcal{M}^{+000} = \frac{\beta_1}{2M_P^3} x + \sum_{n \geq 0, m \geq 1} g_{m,n}^{M_5} x^n y^m, \quad (\text{A.32})$$

where $x = s^2 + t^2 + u^2$, $y = stu$. All other amplitudes can be obtained by crossing.

B Explicit sum rules with st symmetry imposed

Here we explicitly list all the st -symmetry imposed sum rules that we use in this paper, for a quick reference. The definition of $F_{k,\ell}^{1234}(\mu, t)$ is given by

$$F_{k,\ell}^{1234}(\mu, t) = \frac{\partial_s^k}{k!} \left(\frac{s^2}{\mu^2(\mu-s)} d_{h_{12}, h_{43}}^{\ell, \mu, t} c_{\ell, \mu}^{12} c_{\ell, \mu}^{*\bar{3}\bar{4}} + \frac{(-s-t)^2}{\mu^2(\mu+s+t)} d_{h_{14}, h_{23}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{*\bar{3}\bar{2}} \right) \Big|_{s \rightarrow 0} \quad (\text{B.1})$$

$$- \frac{\partial_t^k}{k!} \left(\frac{s^3}{\mu^3(\mu-s)} d_{h_{13}, h_{42}}^{\ell, \mu, t} c_{\ell, \mu}^{13} c_{\ell, \mu}^{*\bar{2}\bar{4}} + \frac{(-s)^3}{(\mu+t)^3(\mu+s+t)} d_{h_{14}, h_{32}}^{\ell, \mu, t} c_{\ell, \mu}^{14} c_{\ell, \mu}^{*\bar{2}\bar{3}} \right) \Big|_{t \rightarrow 0, s \rightarrow t}.$$

The st -symmetry imposed sum rules derived from the stu -symmetric amplitudes are:

$$-\frac{1}{M_P^2} + 2\alpha t - \gamma_4 t^2 = \langle F_{1,\ell}^{0000}(\mu, t) \rangle \quad (\text{B.2}) \quad -2g_{1,1}^{M_1} t + g_{2,0}^{M_1} t^2 = \langle F_{4,\ell}^{++++0}(\mu, t) \rangle \quad (\text{B.15})$$

$$-\frac{1}{M_P^2} \frac{1}{t} + 2\alpha - \gamma_4 t + 12g_{0,2}^S t^2 = \langle F_{2,\ell}^{0000}(\mu, t) \rangle \quad (\text{B.3}) \quad -\frac{\gamma_0}{M_P^4} t^2 = \langle F_{1,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.16})$$

$$8g_{0,2}^S t - 4g_{1,1}^S t^2 = \langle F_{3,\ell}^{0000}(\mu, t) \rangle \quad (\text{B.4}) \quad -\frac{\gamma_0}{M_P^4} t = \langle F_{2,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.17})$$

$$4g_{0,2}^S - 2g_{1,1}^S t + (g_{2,0}^S + 48g_{3,0}^S) t^2 = \langle F_{4,\ell}^{0000}(\mu, t) \rangle \quad (\text{B.5}) \quad 0 = \langle F_{3,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.18})$$

$$24g_{0,3}^S t - 12g_{1,2}^S t^2 = \langle F_{5,\ell}^{0000}(\mu, t) \rangle \quad (\text{B.6}) \quad g_{2,0}^{T_2} t^2 = \langle F_{4,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.19})$$

$$8g_{0,3}^S - 4g_{1,2}^S t + (2g_{2,1}^S + 160g_{0,4}^S) t^2 = \langle F_{6,\ell}^{0000}(\mu, t) \rangle \quad (\text{B.7}) \quad \left(-\frac{10\gamma_0}{M_P^4} + \frac{3\beta_1^2}{M_P^4} \right) t^2 = \langle F_{1,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.20})$$

$$\frac{\beta_1}{M_P^3} t - \frac{\gamma_3}{M_P} t^2 = \langle F_{1,\ell}^{+000}(\mu, t) \rangle \quad (\text{B.8})$$

$$\frac{\beta_1}{M_P^3} - \frac{\gamma_3}{M_P} t = \langle F_{2,\ell}^{+000}(\mu, t) \rangle \quad (\text{B.9}) \quad \left(-\frac{10\gamma_0}{M_P^4} + \frac{3\beta_1^2}{M_P^4} \right) t + 12g_{0,2}^{T_3} t^2 = \langle F_{2,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.21})$$

$$-4g_{1,1}^{M_5} t^2 = \langle F_{3,\ell}^{+000}(\mu, t) \rangle \quad (\text{B.10})$$

$$-2g_{1,1}^{M_5} t + g_{2,0}^{M_5} t^2 = \langle F_{4,\ell}^{+000}(\mu, t) \rangle \quad (\text{B.11}) \quad 8g_{0,2}^{T_3} t - 4g_{1,1}^{T_3} t^2 = \langle F_{3,\ell}^{++++}(\mu, t) \rangle \quad (\text{B.22})$$

$$-\frac{\gamma_1}{M_P^3} t^2 = \langle F_{1,\ell}^{++++0}(\mu, t) \rangle \quad (\text{B.12})$$

$$-\frac{\gamma_1}{M_P^3} t = \langle F_{2,\ell}^{++++0}(\mu, t) \rangle \quad (\text{B.13}) \quad 4g_{0,2}^{T_3} - 2g_{1,1}^{T_3} t + (g_{2,0}^{T_3} + 48g_{0,3}^{T_3}) t^2 = \langle F_{4,\ell}^{++++}(\mu, t) \rangle. \quad (\text{B.23})$$

$$-4g_{1,1}^{M_1} t^2 = \langle F_{3,\ell}^{++++0}(\mu, t) \rangle \quad (\text{B.14})$$

The st -symmetry imposed sum rules derived from amplitudes with su , st or ut symmetry are:

$$\frac{\beta_2}{M_P^2} - \frac{\gamma_0}{M_P^4} t - g_{2,1}^{M_3} t^2 = \langle F_{2,\ell}^{++00}(\mu, t) \rangle \quad (\text{B.24})$$

$$\frac{\gamma_2}{M_P^2} + \frac{\beta_1^2}{M_P^4} - g_{2,1}^{M_3} t - g_{3,1}^{M_3} t^2 = \langle F_{3,\ell}^{++00}(\mu, t) \rangle \quad (\text{B.25})$$

$$g_{4,0}^{M_3} - g_{3,1}^{M_3} t + (g_{2,2}^{M_3} - g_{4,1}^{M_3}) t^2 = \langle F_{4,\ell}^{++00}(\mu, t) \rangle \quad (\text{B.26})$$

$$-\frac{\gamma_0}{M_P^4} t^2 = \langle F_{1,\ell}^{+0+0}(\mu, t) \rangle \quad (\text{B.27})$$

$$-\frac{\gamma_0}{M_P^4} t - g_{2,1}^{M_3} t^2 = \langle F_{2,\ell}^{+0+0}(\mu, t) \rangle \quad (\text{B.28})$$

$$0 = \langle F_{3,\ell}^{+0+0}(\mu, t) \rangle \quad (\text{B.29})$$

$$g_{2,2}^{M_3} t^2 = \langle F_{4,\ell}^{+0+0}(\mu, t) \rangle \quad (\text{B.30})$$

$$-\frac{\beta_1^2}{M_P^4} t + g_{0,2}^{M_4} t^2 = \langle F_{2,\ell}^{+-00}(\mu, t) \rangle \quad (\text{B.31})$$

$$g_{1,2}^{M_4} t^2 = \langle F_{3,\ell}^{+-00}(\mu, t) \rangle \quad (\text{B.32})$$

$$g_{2,2}^{M_4} t^2 = \langle F_{4,\ell}^{+-00}(\mu, t) \rangle \quad (\text{B.33})$$

$$g_{3,2}^{M_4} t^2 = \langle F_{5,\ell}^{+-00}(\mu, t) \rangle \quad (\text{B.34})$$

$$g_{4,2}^{M_4} t^2 = \langle F_{6,\ell}^{+-00}(\mu, t) \rangle \quad (\text{B.35})$$

$$-\frac{1}{M_P^2} - \frac{\beta_1^2}{M_P^4} t^2 = \langle F_{1,\ell}^{+0-0}(\mu, t) \rangle \quad (\text{B.36})$$

$$-\frac{1}{M_P^2} \frac{1}{t} - \frac{\beta_1^2}{M_P^4} t + g_{0,2}^{M_4} t^2 = \langle F_{2,\ell}^{+0-0}(\mu, t) \rangle \quad (\text{B.37})$$

$$2g_{0,2}^{M_4} t + 2g_{1,2}^{M_4} t^2 = \langle F_{3,\ell}^{+0-0}(\mu, t) \rangle \quad (\text{B.38})$$

$$g_{0,2}^{M_4} + g_{1,2}^{M_4} t + (g_{2,2}^{M_4} - 3g_{0,3}^{M_4}) t^2 = \langle F_{4,\ell}^{+0-0}(\mu, t) \rangle \quad (\text{B.39})$$

$$-3g_{0,3}^{M_4} t - 3g_{1,3}^{M_4} t^2 = \langle F_{5,\ell}^{+0-0}(\mu, t) \rangle \quad (\text{B.40})$$

$$-g_{0,3}^{M_4} - g_{1,3}^{M_4} t + (-g_{2,3}^{M_4} + 6g_{0,4}^{M_4}) t^2 = \langle F_{6,\ell}^{+0-0}(\mu, t) \rangle \quad (\text{B.41})$$

$$\frac{\beta_1}{M_P^3} + \frac{\beta_1 \gamma_0}{M_P^5} t^2 = \langle F_{2,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.42})$$

$$\frac{\gamma_0 \beta_1}{M_P^5} t - g_{3,1}^{M_2} t^2 = \langle F_{3,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.43})$$

$$-g_{3,1}^{M_2} t - g_{4,1}^{M_2} t^2 = \langle F_{4,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.44})$$

$$0 = \langle F_{1,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.45})$$

$$\frac{\gamma_0 \beta_1}{M_P^5} t^2 = \langle F_{2,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.46})$$

$$0 = \langle F_{3,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.47})$$

$$0 = \langle F_{4,\ell}^{0-++}(\mu, t) \rangle \quad (\text{B.48})$$

$$0 = \langle F_{1,\ell}^{+-+-}(\mu, t) \rangle \quad (\text{B.49})$$

$$0 = \langle F_{2,\ell}^{+-+-}(\mu, t) \rangle \quad (\text{B.50})$$

$$0 = \langle F_{3,\ell}^{+-+-}(\mu, t) \rangle \quad (\text{B.51})$$

$$0 = \langle F_{4,\ell}^{+-+-}(\mu, t) \rangle \quad (\text{B.52})$$

$$0 = \langle F_{5,\ell}^{+-+-}(\mu, t) \rangle \quad (\text{B.53})$$

$$0 = \langle F_{6,\ell}^{+-+-}(\mu, t) \rangle \quad (\text{B.54})$$

$$-\frac{1}{M_P^2} \frac{1}{t} = \langle F_{2,\ell}^{++--}(\mu, t) \rangle \quad (\text{B.55})$$

$$-\frac{\beta_1^2}{M_P^4} - \frac{\gamma_0^2}{M_P^6} t^2 = \langle F_{3,\ell}^{++--}(\mu, t) \rangle \quad (\text{B.56})$$

$$g_{4,0}^{T_1} - \frac{\gamma_0^2}{M_P^6} t - g_{4,1}^{T_1} t^2 = \langle F_{4,\ell}^{++--}(\mu, t) \rangle \quad (\text{B.57})$$

$$g_{5,0}^{T_1} - g_{4,1}^{T_1} t - g_{5,1}^{T_1} t^2 = \langle F_{5,\ell}^{++--}(\mu, t) \rangle \quad (\text{B.58})$$

$$g_{6,0}^{T_1} - g_{5,1}^{T_1} t + (g_{4,2}^{T_1} - g_{6,1}^{T_1}) t^2 = \langle F_{6,\ell}^{++--}(\mu, t) \rangle. \quad (\text{B.59})$$

Note that one can take the forward limit of the sum rules $F_{k,\ell}^{1234}(\mu, t)$ with $k \geq 3$, which are valuable to extract the useful constraints in the finite μ and large ℓ region.

C Explicit example of optimization scheme

The purpose of this appendix is to provide the nitty-gritty of using the dispersive sum rules to obtain causality bounds on the Wilson coefficients. We shall demonstrate these by the explicit example of deriving the bound on β_1^2 and γ_0^2 (projected to $1/M_P^2$), using only the dispersive sum rules from four graviton scattering for the sake of simplicity. This example provides a representative illustration of the essential computational steps and subtleties involved, while in the actual figures in section 6.2 all available dispersive sum rules are used.

Even if we restrict to four graviton scattering, there are already quite a few dispersive sum rules available to constrain β_1 and γ_0 , which are given by

$$\begin{aligned}
 -\frac{\beta_1^2}{M_P^4} - \frac{\gamma_0^2}{M_P^6} p^4 &= \langle F_{3,\ell}^{++--}(\mu, -p^2) \rangle, & \frac{1}{M_P^2} \frac{1}{p^2} &= \langle F_{2,\ell}^{++--}(\mu, -p^2) \rangle, \\
 g_{4,0}^{T_1} + \frac{\gamma_0^2}{M_P^6} p^2 - g_{4,1}^{T_1} p^4 &= \langle F_{4,\ell}^{++--}(\mu, -p^2) \rangle, \\
 0 &= \langle F_{1,\ell}^{+--+}(\mu, -p^2) \rangle, & 0 &= \langle F_{2,\ell}^{+--+}(\mu, -p^2) \rangle, \\
 0 &= \langle F_{3,\ell}^{+--+}(\mu, -p^2) \rangle, & 0 &= \langle F_{4,\ell}^{+--+}(\mu, -p^2) \rangle, \\
 -\frac{\gamma_0}{M_P^4} p^4 &= \langle F_{1,\ell}^{++++}(\mu, -p^2) \rangle, & \frac{\gamma_0}{M_P^4} p^2 &= \langle F_{2,\ell}^{++++}(\mu, -p^2) \rangle, \\
 0 &= \langle F_{3,\ell}^{++++}(\mu, -p^2) \rangle, & g_{2,0}^{T_2} p^4 &= \langle F_{4,\ell}^{++++}(\mu, -p^2) \rangle, \\
 -\left(\frac{10\gamma_0}{M_P^4} - \frac{3\beta_1^2}{M_P^4}\right) p^4 &= \langle F_{1,\ell}^{++++}(\mu, -p^2) \rangle, \\
 \left(\frac{10\gamma_0}{M_P^4} - \frac{3\beta_1^2}{M_P^4}\right) p^2 + 12g_{0,2}^{T_3} p^4 &= \langle F_{2,\ell}^{++++}(\mu, -p^2) \rangle, \\
 -8g_{0,2}^{T_3} p^2 - 4g_{1,1}^{T_3} p^4 &= \langle F_{3,\ell}^{++++}(\mu, -p^2) \rangle, \\
 4g_{0,2}^{T_3} + 2g_{1,1}^{T_3} p^2 + (g_{2,0}^{T_3} + 48g_{0,3}^{T_3}) p^4 &= \langle F_{4,\ell}^{++++}(\mu, -p^2) \rangle. \tag{C.1}
 \end{aligned}$$

where the EFT cutoff has been set to $\Lambda = 1$. As mentioned in section 5.2, for technical reasons, we also add some forward-limit sum rules (3.20) to effectively make use of the constraints in the finite μ and large ℓ region:

$$\begin{aligned}
 -\frac{2\gamma_0^2}{M_P^6} &= \langle \partial_t^2 F_{3,\ell}^{++--}(\mu, 0) \rangle, & -\frac{\gamma_0^2}{M_P^6} &= \langle \partial_t F_{4,\ell}^{++--}(\mu, 0) \rangle, & -\frac{\beta_1^2}{M_P^4} &= \langle F_{3,\ell}^{++--}(\mu, 0) \rangle, \\
 8g_{0,2}^{T_3} &= \langle \partial_t F_{3,\ell}^{++++}(\mu, 0) \rangle, & 0 &= \langle \partial_t^3 F_{3,\ell}^{++--}(\mu, 0) \rangle, \tag{C.2}
 \end{aligned}$$

where for illustration purposes only one forward-limit null constraint is included. It is usually beneficial to include more forward-limit sum rules, as we do for the plots in section 6. The reason for using forward-limit sum rules, as mentioned, is that the Wigner d-functions, when integrated over p against the weight functions, tend to vanish in this region, and the added forward-limit sum rules provide terms that are polynomials of ℓ in the large ℓ limit

and that are dominant in this region. Ultimately, the reason why adding forward-limit sum rules is helpful is that we only parametrize each of the weight functions with a few parameters, while the actual functional spaces are infinite dimensional.

Similar considerations also apply when choosing the forward-limit sum rules. The two dispersion relations with γ_0 in eq. (C.2) are formally independent, but they are actually linked by the st crossing. Nevertheless, we use both of them, as the enforcing of the st crossing is not complete due to the finite dimensional truncation of the weight functions. On the other hand, we do not use the forward-limit sum rule $g_{2,0}^{T_3} = \langle F_{4,\ell}^{++++}(\mu, 0)/4 \rangle$ because it is formally the same as the sum rule with $\partial_t F_{3,\ell}^{++++}(\mu, 0)$, already guaranteed by the su symmetry of the dispersion relations.

Our goal is to extract as much information as possible from these sum rules. To that end, we integrate both sides of the sum rules (C.1) over various weight functions $\phi_k^{1234}(p)$ and sum both sides of the sum rules (C.2) over weight parameters $z_{k,n}^{1234}$, which leads to

$$\begin{aligned}
 & \left\{ \int_0^1 dp \phi_2^{++--}(p) \frac{1}{p^2} \right\} \frac{1}{M_P^2} \\
 & + \left\{ \int_0^1 dp \left(-\phi_1^{++++}(p)p^4 + \phi_2^{++++}(p)p^2 - 10\phi_1^{++++}(p)p^4 + 10\phi_2^{++++}(p)p^2 \right) \right\} \frac{\gamma_0}{M_P^4} \\
 & + \left\{ \int_0^1 dp \left(-\phi_3^{++--}(p)p^4 + \phi_4^{++--}(p)p^2 \right) - 2z_{3,2}^{++--} - z_{4,1}^{++--} \right\} \frac{\gamma_0^2}{M_P^6} \\
 & + \left\{ \int_0^1 dp \left(-\phi_3^{++--}(p) + 3\phi_1^{++++}(p)p^4 - 3\phi_2^{++++}(p)p^2 \right) - z_{3,0}^{++--} \right\} \frac{\beta_1^2}{M_P^4} \\
 & + \left\{ \int_0^1 dp \phi_4^{++--}(p) \right\} g_{4,0}^{T_1} + \left\{ \int_0^1 dp \phi_4^{++--}(p)p^4 \right\} g_{4,1}^{T_1} + \left\{ \int_0^1 dp \phi_4^{++++}(p)p^4 \right\} g_{2,0}^{T_2} \\
 & + \left\{ \int_0^1 dp \left(12\phi_2^{++++}(p)p^4 - 8\phi_3^{++++}(p)p^2 + 4\phi_4^{++++}(p) \right) + 8z_{3,1}^{++++} \right\} g_{0,2}^{T_3} \\
 & + \left\{ \int_0^1 dp \left(-4\phi_3^{++++}(p)p^4 + 2\phi_4^{++++}(p)p^2 \right) \right\} g_{1,1}^{T_3} + \left\{ \int_0^1 dp \left(\phi_4^{++++}(p)p^4 \right) \right\} (g_{2,0}^{T_3} + 48g_{0,3}^{T_3}) \\
 & = \left\langle \int_0^1 dp \left(\sum_{1234,k} \phi_k^{1234}(p) F_{k,\ell}^{1234}(\mu, -p^2) \right) + \sum_{1234,k,n} z_{k,n}^{1234} \partial_t^n F_{k,\ell}^{1234}(\mu) \right\rangle \quad (C.3) \\
 & := \left\langle (\mathcal{C}_{P_X,\ell,\mu})^T B_{P_X,\ell}(\mu) \mathcal{C}_{P_X,\ell,\mu} \right\rangle, \quad (C.4)
 \end{aligned}$$

where the last equality implicitly defines the $B_{P_X,\ell}(\mu)$ matrices mentioned in section 5.1. The weight functions and parameters are so-called decision variables in the optimization problem. For some appropriate chosen decision variables, these matrices can be made semi-positive:

$$B_{P_X,\ell}(\mu) \succeq 0, \quad \text{for } P_X = \pm 1, \text{ all possible } \ell \text{ and all } \mu \geq \Lambda^2, \quad (C.5)$$

which in turn results in the right hand side of eq. (C.3) being semi-positive. This gives rise to a bound on the Wilson coefficients appearing on the left hand side of eq. (C.3). However, our goal here is more specific: we want to derive a bound on β_1^2 and γ_0^2 , projected onto $1/M_P^2$. So we do not want other Wilson coefficients to be involved on the left hand

side of eq. (C.3). We can achieve this by imposing the following constraints on the weight functions:

$$\begin{aligned} \int_0^1 dp \left(\phi_4^{++--}(p) \right) &= 0, & \int_0^1 dp \left(\phi_4^{++--}(p)p^4 \right) &= 0, & \int_0^1 dp \left(\phi_4^{++++}(p)p^4 \right) &= 0, \\ \int_0^1 dp \left(12\phi_2^{++++}(p)p^4 - 8\phi_3^{++++}(p)p^2 + 4\phi_4^{++++}(p) \right) + 8z_{3,1}^{++++} &= 0, \\ \int_0^1 dp \left(-4\phi_3^{++++}(p)p^4 + 2\phi_4^{++++}(p)p^2 \right) &= 0, & \int_0^1 dp \left(\phi_4^{++++}(p)p^4 \right) &= 0. \end{aligned} \tag{C.6}$$

That is, we are deriving bounds on β_1 and γ_0 while being agnostic about all other Wilson coefficients, except for $1/M_P^2$. (This is in contrast to the bounds with some other Wilson coefficients fixed, which can often be stronger.) Then, the causality bounds that we can derive are given by

$$\begin{aligned} \int_0^1 dp \left(\phi_2^{++--} \frac{1}{p^2} \right) \frac{1}{M_P^2} + \int_0^1 dp \left(-\phi_1^{++++} p^4 + \phi_2^{++++} p^2 - 10\phi_1^{++++} p^4 + 10\phi_2^{++++} p^2 \right) \frac{\gamma_0}{M_P^4} \\ + \left\{ \int_0^1 dp \left(-\phi_3^{++--} p^4 + \phi_4^{++--} p^2 \right) - 2z_{3,2}^{++--} - z_{4,1}^{++--} \right\} \frac{\gamma_0^2}{M_P^6} \\ + \left\{ \int_0^1 dp \left(-\phi_3^{++--} + 3\phi_1^{++++} p^4 - 3\phi_2^{++++} p^2 \right) - z_{3,0}^{++--} \right\} \frac{\beta_1^2}{M_P^4} \geq 0, \end{aligned} \tag{C.7}$$

for all sets of weight functions $\phi_k^{1234}(p)$ and parameters $z_{k,n}^{1234}$ satisfying eq. (C.5) and eq. (C.6). If a set of weight functions $\phi_k^{1234}(p)$ and parameters $z_{k,n}^{1234}$ satisfy condition eq. (C.5) and eq. (C.6), so do the scaled set of $\lambda\phi_k^{1234}(p)$ and $\lambda z_{k,n}^{1234}$ with $\lambda > 0$, leading to an arbitrary normalization of eq. (C.7). Considering that we project all of our bounds onto $1/M_P^2$, one is tempted to fix the normalization of eq. (C.7) by setting

$$\int_0^1 dp \phi_2^{++--}(p) \frac{1}{p^2} = 1. \tag{C.8}$$

However, this is only a formal/schematic assignment, which can not be implemented numerically. As discussed in section 5.2, the integration on the left hand side of eq. (C.8) is actually divergent, so we need to introduce an IR cutoff m_{IR} to regulate it, that is, we actually choose to integrate from m_{IR} to 1 against all the weight functions in the numerical implementation. As we parametrize $\phi_2^{++--}(p)$ by

$$\phi_2^{++--}(p) = (1-p)^2 \sum_{i=1} x_{2,i}^{++--} p^i \tag{C.9}$$

and the $x_{2,1}^{++--}$ term, i.e., the t -channel contribution, must be present to yield positivity bounds, the leading term on the left hand side, which is logarithmic divergent as $m_{\text{IR}} \rightarrow 0$, comes from the $x_{2,1}^{++--}$ term. For phenomenological interesting cases, the $\log(\Lambda/m_{\text{IR}})$ term is usually $\mathcal{O}(10^2)$, which dominates the left hand side of eq. (C.8).

Thus, in the numerical implementation, we can choose the normalization to be

$$x_{2,1}^{++--} = 1. \tag{C.10}$$

Also, because of the large logarithmic term, it is a good approximation to neglect the linear term γ_0/M_P^4 in the sum rules. (These two approximations are justified numerically in more details in section 6.2.) After these considerations, the inequality (C.7) becomes

$$\begin{aligned} \frac{\log \Lambda/m_{\text{IR}}}{M_P^2} + \left\{ \int_0^1 dp \left(-\phi_3^{++--} p^4 + \phi_4^{++--} p^2 \right) - 2z_{3,2}^{++--} - z_{4,1}^{++--} \right\} \frac{\gamma_0^2}{M_P^6} \\ + \left\{ \int_0^1 dp \left(-\phi_3^{++--} + 3\phi_1^{++++} p^4 - 3\phi_2^{++++} p^2 \right) - z_{3,0}^{++--} \right\} \frac{\beta_1^2}{M_P^4} \geq 0 \end{aligned} \quad (\text{C.11})$$

where the decision variables $\phi_k^{1234}(p)$ and $z_{k,n}^{1234}$ must satisfy linear conditions eq. (C.5), eq. (C.6) and eq. (C.10).

To carve out the boundary of the causality bound in a 2D parameter space, we choose a fixed point within the convex bound region, use the optimization scheme to find the end points of a ray at a fixed angle from the fixed point, and scan over all angles to get the boundary. Although only γ_0^2 and β_1^2 appear in the inequality (C.11), to use the same parametrization as section 6.2, we parametrize γ_0 and β_1 as

$$\frac{\gamma_0}{M_P^2} = r \cos \theta, \quad \frac{\beta_1}{M_P} = r \sin \theta. \quad (\text{C.12})$$

Then the inequality (C.11) becomes

$$\log \frac{\Lambda}{m_{\text{IR}}} \geq \left\{ \dots \right\} \frac{\gamma_0^2}{M_P^4} + \left\{ \dots \right\} \frac{\beta_1^2}{M_P^2} = \left(\left\{ \dots \right\} \cos^2 \theta + \left\{ \dots \right\} \sin^2 \theta \right) r^2. \quad (\text{C.13})$$

Therefore, for every fixed θ , we solve the following SDP problem

$$\begin{aligned} \text{maximize: } & \left\{ \int_0^1 dp \left(\phi_3^{++--}(p) p^4 - \phi_4^{++--}(p) p^2 \right) + 2z_{3,2}^{++--} + z_{4,1}^{++--} \right\} \cos^2 \theta \quad (\text{C.14}) \\ & + \left\{ \int_0^1 dp \left(\phi_3^{++--}(p) - 3\phi_1^{++++}(p) p^4 + 3\phi_2^{++++}(p) p^2 \right) + z_{3,0}^{++--} \right\} \sin^2 \theta, \end{aligned} \quad (\text{C.15})$$

$$\text{subject to: eqs. (C.5), (C.6), (C.10) for all functions } \phi_k^{1234}(p) \text{ and parameters } z_{k,n}^{1234}, \quad (\text{C.16})$$

to get the lowest upper bound on r^2 at the given θ , which can be implemented by the SDPB package. Scanning θ from 0 to $\pi/2$ for sufficiently many angles, the optimal results on r^2 from different angles will depict the boundary of the bounds on β_1^2 and γ_0^2 .

Before ending, we would like to comment on whether more constraints can be added to get more information, using the current example. This seems to be possible at first glance but actually unachievable. For example, we know from amplitudes \mathcal{M}^{+-+} and \mathcal{M}^{++--} that the coefficients $a_{n,2}^{+-+} = 0$, $a_{2,n}^{++--} = 0$ with $n \geq 3$. Thus, besides the st null constraints $a_{2,n}^{++--} = a_{n,2}^{+-+}$ with $n \geq 3$, it seems that we can use naively stronger constraints $a_{n,2}^{+-+} = a_{2,n}^{++--} = 0$. However, the information $a_{n,2}^{+-+} = 0$ is actually already contained in the dispersive sum rules and thus does not give rise to extra null constraints. This can be seen from eq. (3.17):

$$\sum_{k=3} a_{k,2}^{+-+} s^k = \left\langle \frac{\partial_t^2}{2!} \left(\frac{s^3 d_{4,-4}^{\ell,\mu,t} c_{\ell,\mu}^{+-} c_{\ell,\mu}^{*-++}}{\mu^3(\mu-s)} + \frac{(-s)^3 d_{4,-4}^{\ell,\mu,t} c_{\ell,\mu}^{+-} c_{\ell,\mu}^{*-++}}{(\mu+t)^3(\mu+s+t)} \right) \Big|_{t \rightarrow 0} \right\rangle = \langle 0 \rangle, \quad (\text{C.17})$$

That is, the dispersion relations automatically enforce this extra information. Another example is that, as a result of the structure of 3-leg vertices in the theory, β_1^2 appears in both the sum rules from \mathcal{M}^{++--} and \mathcal{M}^{++++} , which leads to a constraint when equating the expressions for β_1^2 . This constraint does not come from crossing symmetry. However, again, we do not need to explicitly impose this constraint, as we have used the dispersion relations from both \mathcal{M}^{++--} and \mathcal{M}^{++++} in our optimization programs. Using several dispersion relations involving a Wilson coefficient will give the same result as using one of these dispersion relations and the constraints from these dispersion relations.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne, *The analytic S-matrix*, Cambridge Univ. Press, Cambridge (1966) [[INSPIRE](#)].
- [2] C. de Rham et al., *Snowmass White Paper: UV Constraints on IR Physics*, in the proceedings of the *Snowmass 2021*, (2022) [[arXiv:2203.06805](#)] [[INSPIRE](#)].
- [3] A. Adams et al., *Causality, analyticity and an IR obstruction to UV completion*, *JHEP* **10** (2006) 014 [[hep-th/0602178](#)] [[INSPIRE](#)].
- [4] C. Zhang and S.-Y. Zhou, *Convex Geometry Perspective on the (Standard Model) Effective Field Theory Space*, *Phys. Rev. Lett.* **125** (2020) 201601 [[arXiv:2005.03047](#)] [[INSPIRE](#)].
- [5] B. Bellazzini, L. Martucci and R. Torre, *Symmetries, Sum Rules and Constraints on Effective Field Theories*, *JHEP* **09** (2014) 100 [[arXiv:1405.2960](#)] [[INSPIRE](#)].
- [6] X. Li et al., *Positivity in Multifield Effective Field Theories*, *Phys. Rev. Lett.* **127** (2021) 121601 [[arXiv:2101.01191](#)] [[INSPIRE](#)].
- [7] C. Zhang and S.-Y. Zhou, *Positivity bounds on vector boson scattering at the LHC*, *Phys. Rev. D* **100** (2019) 095003 [[arXiv:1808.00010](#)] [[INSPIRE](#)].
- [8] Q. Bi, C. Zhang and S.-Y. Zhou, *Positivity constraints on aQGC: carving out the physical parameter space*, *JHEP* **06** (2019) 137 [[arXiv:1902.08977](#)] [[INSPIRE](#)].
- [9] B. Bellazzini and F. Riva, *New phenomenological and theoretical perspective on anomalous ZZ and Z γ processes*, *Phys. Rev. D* **98** (2018) 095021 [[arXiv:1806.09640](#)] [[INSPIRE](#)].
- [10] G.N. Remmen and N.L. Rodd, *Consistency of the Standard Model Effective Field Theory*, *JHEP* **12** (2019) 032 [[arXiv:1908.09845](#)] [[INSPIRE](#)].
- [11] K. Yamashita, C. Zhang and S.-Y. Zhou, *Elastic positivity vs extremal positivity bounds in SMEFT: a case study in transversal electroweak gauge-boson scatterings*, *JHEP* **01** (2021) 095 [[arXiv:2009.04490](#)] [[INSPIRE](#)].
- [12] T. Trott, *Causality, unitarity and symmetry in effective field theory*, *JHEP* **07** (2021) 143 [[arXiv:2011.10058](#)] [[INSPIRE](#)].
- [13] G.N. Remmen and N.L. Rodd, *Flavor Constraints from Unitarity and Analyticity*, *Phys. Rev. Lett.* **125** (2020) 081601 [*Erratum ibid.* **127** (2021) 149901] [[arXiv:2004.02885](#)] [[INSPIRE](#)].

- [14] Q. Bonnefoy, E. Gendy and C. Grojean, *Positivity bounds on Minimal Flavor Violation*, *JHEP* **04** (2021) 115 [[arXiv:2011.12855](#)] [[INSPIRE](#)].
- [15] J. Davighi, S. Melville and T. You, *Natural selection rules: new positivity bounds for massive spinning particles*, *JHEP* **02** (2022) 167 [[arXiv:2108.06334](#)] [[INSPIRE](#)].
- [16] M. Chala and J. Santiago, *Positivity bounds in the standard model effective field theory beyond tree level*, *Phys. Rev. D* **105** (2022) L111901 [[arXiv:2110.01624](#)] [[INSPIRE](#)].
- [17] X. Li and S. Zhou, *Origin of neutrino masses on the convex cone of positivity bounds*, *Phys. Rev. D* **107** (2023) L031902 [[arXiv:2202.12907](#)] [[INSPIRE](#)].
- [18] D. Ghosh, R. Sharma and F. Ullah, *Amplitude's positivity vs. subluminality: causality and unitarity constraints on dimension 6 & 8 gluonic operators in the SMEFT*, *JHEP* **02** (2023) 199 [[arXiv:2211.01322](#)] [[INSPIRE](#)].
- [19] G.N. Remmen and N.L. Rodd, *Spinning sum rules for the dimension-six SMEFT*, *JHEP* **09** (2022) 030 [[arXiv:2206.13524](#)] [[INSPIRE](#)].
- [20] B. Fuks, Y. Liu, C. Zhang and S.-Y. Zhou, *Positivity in electron-positron scattering: testing the axiomatic quantum field theory principles and probing the existence of UV states*, *Chin. Phys. C* **45** (2021) 023108 [[arXiv:2009.02212](#)] [[INSPIRE](#)].
- [21] J. Gu, L.-T. Wang and C. Zhang, *Unambiguously Testing Positivity at Lepton Colliders*, *Phys. Rev. Lett.* **129** (2022) 011805 [[arXiv:2011.03055](#)] [[INSPIRE](#)].
- [22] X. Li et al., *Moments for positivity: using Drell-Yan data to test positivity bounds and reverse-engineer new physics*, *JHEP* **10** (2022) 107 [[arXiv:2204.13121](#)] [[INSPIRE](#)].
- [23] L. Alberte, C. de Rham, S. Jaitly and A.J. Tolley, *Reverse Bootstrapping: IR Lessons for UV Physics*, *Phys. Rev. Lett.* **128** (2022) 051602 [[arXiv:2111.09226](#)] [[INSPIRE](#)].
- [24] L. Alberte, C. de Rham, S. Jaitly and A.J. Tolley, *QED positivity bounds*, *Phys. Rev. D* **103** (2021) 125020 [[arXiv:2012.05798](#)] [[INSPIRE](#)].
- [25] N. Arkani-Hamed, T.-C. Huang and Y.-T. Huang, *The EFT-Hedron*, *JHEP* **05** (2021) 259 [[arXiv:2012.15849](#)] [[INSPIRE](#)].
- [26] B. Bellazzini et al., *Positive moments for scattering amplitudes*, *Phys. Rev. D* **104** (2021) 036006 [[arXiv:2011.00037](#)] [[INSPIRE](#)].
- [27] A. Martin, *Extension of the axiomatic analyticity domain of scattering amplitudes by unitarity. 1.*, *Nuovo Cim. A* **42** (1965) 930 [[INSPIRE](#)].
- [28] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *Positivity bounds for scalar field theories*, *Phys. Rev. D* **96** (2017) 081702 [[arXiv:1702.06134](#)] [[INSPIRE](#)].
- [29] A.V. Manohar and V. Mateu, *Dispersion Relation Bounds for $\pi\pi$ Scattering*, *Phys. Rev. D* **77** (2008) 094019 [[arXiv:0801.3222](#)] [[INSPIRE](#)].
- [30] A. Nicolis, R. Rattazzi and E. Trincherini, *Energy's and amplitudes' positivity*, *JHEP* **05** (2010) 095 [*Erratum ibid.* **11** (2011) 128] [[arXiv:0912.4258](#)] [[INSPIRE](#)].
- [31] B. Bellazzini, *Softness and amplitudes' positivity for spinning particles*, *JHEP* **02** (2017) 034 [[arXiv:1605.06111](#)] [[INSPIRE](#)].
- [32] Y.-J. Wang, F.-K. Guo, C. Zhang and S.-Y. Zhou, *Generalized positivity bounds on chiral perturbation theory*, *JHEP* **07** (2020) 214 [[arXiv:2004.03992](#)] [[INSPIRE](#)].

- [33] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *UV complete me: Positivity Bounds for Particles with Spin*, *JHEP* **03** (2018) 011 [[arXiv:1706.02712](#)] [[INSPIRE](#)].
- [34] A.J. Tolley, Z.-Y. Wang and S.-Y. Zhou, *New positivity bounds from full crossing symmetry*, *JHEP* **05** (2021) 255 [[arXiv:2011.02400](#)] [[INSPIRE](#)].
- [35] S. Caron-Huot and V. Van Duong, *Extremal Effective Field Theories*, *JHEP* **05** (2021) 280 [[arXiv:2011.02957](#)] [[INSPIRE](#)].
- [36] C. Vafa, *The string landscape and the swampland*, [hep-th/0509212](#) [[INSPIRE](#)].
- [37] Z.-Z. Du, C. Zhang and S.-Y. Zhou, *Triple crossing positivity bounds for multi-field theories*, *JHEP* **12** (2021) 115 [[arXiv:2111.01169](#)] [[INSPIRE](#)].
- [38] D. Simmons-Duffin, *A Semidefinite Program Solver for the Conformal Bootstrap*, *JHEP* **06** (2015) 174 [[arXiv:1502.02033](#)] [[INSPIRE](#)].
- [39] A. Sinha and A. Zahed, *Crossing Symmetric Dispersion Relations in Quantum Field Theories*, *Phys. Rev. Lett.* **126** (2021) 181601 [[arXiv:2012.04877](#)] [[INSPIRE](#)].
- [40] L.-Y. Chiang et al., *Into the EFThedron and UV constraints from IR consistency*, *JHEP* **03** (2022) 063 [[arXiv:2105.02862](#)] [[INSPIRE](#)].
- [41] Z. Bern, D. Kosmopoulos and A. Zhiboedov, *Gravitational effective field theory islands, low-spin dominance, and the four-graviton amplitude*, *J. Phys. A* **54** (2021) 344002 [[arXiv:2103.12728](#)] [[INSPIRE](#)].
- [42] J. Henriksson, B. McPeak, F. Russo and A. Vichi, *Rigorous bounds on light-by-light scattering*, *JHEP* **06** (2022) 158 [[arXiv:2107.13009](#)] [[INSPIRE](#)].
- [43] S.D. Chowdhury et al., *Crossing Symmetric Spinning S-matrix Bootstrap: EFT bounds*, *SciPost Phys.* **13** (2022) 051 [[arXiv:2112.11755](#)] [[INSPIRE](#)].
- [44] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez and D. Simmons-Duffin, *Causality constraints on corrections to Einstein gravity*, *JHEP* **05** (2023) 122 [[arXiv:2201.06602](#)] [[INSPIRE](#)].
- [45] L.-Y. Chiang et al., *(Non)-projective bounds on gravitational EFT*, [arXiv:2201.07177](#) [[INSPIRE](#)].
- [46] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez and D. Simmons-Duffin, *Graviton partial waves and causality in higher dimensions*, *Phys. Rev. D* **108** (2023) 026007 [[arXiv:2205.01495](#)] [[INSPIRE](#)].
- [47] J. Henriksson, B. McPeak, F. Russo and A. Vichi, *Bounding violations of the weak gravity conjecture*, *JHEP* **08** (2022) 184 [[arXiv:2203.08164](#)] [[INSPIRE](#)].
- [48] L.-Y. Chiang, Y.-T. Huang, L. Rodina and H.-C. Weng, *De-projecting the EFThedron*, [arXiv:2204.07140](#) [[INSPIRE](#)].
- [49] A.L. Guerrieri, J. Penedones and P. Vieira, *S-matrix bootstrap for effective field theories: massless pions*, *JHEP* **06** (2021) 088 [[arXiv:2011.02802](#)] [[INSPIRE](#)].
- [50] A. Guerrieri, J. Penedones and P. Vieira, *Where Is String Theory in the Space of Scattering Amplitudes?*, *Phys. Rev. Lett.* **127** (2021) 081601 [[arXiv:2102.02847](#)] [[INSPIRE](#)].
- [51] J. Elias Miro, A. Guerrieri and M.A. Gumus, *Bridging positivity and S-matrix bootstrap bounds*, *JHEP* **05** (2023) 001 [[arXiv:2210.01502](#)] [[INSPIRE](#)].
- [52] K. Häring et al., *Bounds on photon scattering*, [arXiv:2211.05795](#) [[INSPIRE](#)].

- [53] M. Kruczenski, J. Penedones and B.C. van Rees, *Snowmass White Paper: S-matrix Bootstrap*, [arXiv:2203.02421](#) [[INSPIRE](#)].
- [54] L. Alberte, C. de Rham, S. Jaitly and A.J. Tolley, *Positivity Bounds and the Massless Spin-2 Pole*, *Phys. Rev. D* **102** (2020) 125023 [[arXiv:2007.12667](#)] [[INSPIRE](#)].
- [55] J. Tokuda, K. Aoki and S. Hirano, *Gravitational positivity bounds*, *JHEP* **11** (2020) 054 [[arXiv:2007.15009](#)] [[INSPIRE](#)].
- [56] S. Caron-Huot, D. Mazac, L. Rastelli and D. Simmons-Duffin, *Sharp boundaries for the swampland*, *JHEP* **07** (2021) 110 [[arXiv:2102.08951](#)] [[INSPIRE](#)].
- [57] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *Massive Galileon Positivity Bounds*, *JHEP* **09** (2017) 072 [[arXiv:1702.08577](#)] [[INSPIRE](#)].
- [58] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *Positivity Bounds for Massive Spin-1 and Spin-2 Fields*, *JHEP* **03** (2019) 182 [[arXiv:1804.10624](#)] [[INSPIRE](#)].
- [59] X.O. Camanho, G. Lucena Gómez and R. Rahman, *Causality Constraints on Massive Gravity*, *Phys. Rev. D* **96** (2017) 084007 [[arXiv:1610.02033](#)] [[INSPIRE](#)].
- [60] C.Y.-R. Chen, C. de Rham, A. Margalit and A.J. Tolley, *A cautionary case of casual causality*, *JHEP* **03** (2022) 025 [[arXiv:2112.05031](#)] [[INSPIRE](#)].
- [61] C. de Rham and A.J. Tolley, *Causality in curved spacetimes: The speed of light and gravity*, *Phys. Rev. D* **102** (2020) 084048 [[arXiv:2007.01847](#)] [[INSPIRE](#)].
- [62] C. de Rham, A.J. Tolley and J. Zhang, *Causality Constraints on Gravitational Effective Field Theories*, *Phys. Rev. Lett.* **128** (2022) 131102 [[arXiv:2112.05054](#)] [[INSPIRE](#)].
- [63] G. Goon and K. Hinterbichler, *Superluminality, black holes and EFT*, *JHEP* **02** (2017) 134 [[arXiv:1609.00723](#)] [[INSPIRE](#)].
- [64] K. Hinterbichler, A. Joyce and R.A. Rosen, *Massive Spin-2 Scattering and Asymptotic Superluminality*, *JHEP* **03** (2018) 051 [[arXiv:1708.05716](#)] [[INSPIRE](#)].
- [65] M. Accattulli Huber, A. Brandhuber, S. De Angelis and G. Travaglini, *Eikonal phase matrix, deflection angle and time delay in effective field theories of gravity*, *Phys. Rev. D* **102** (2020) 046014 [[arXiv:2006.02375](#)] [[INSPIRE](#)].
- [66] B. Bellazzini, G. Isabella, M. Lewandowski and F. Sgarlata, *Gravitational causality and the self-stress of photons*, *JHEP* **05** (2022) 154 [[arXiv:2108.05896](#)] [[INSPIRE](#)].
- [67] B. Bellazzini, C. Cheung and G.N. Remmen, *Quantum Gravity Constraints from Unitarity and Analyticity*, *Phys. Rev. D* **93** (2016) 064076 [[arXiv:1509.00851](#)] [[INSPIRE](#)].
- [68] C. Cheung and G.N. Remmen, *Positive Signs in Massive Gravity*, *JHEP* **04** (2016) 002 [[arXiv:1601.04068](#)] [[INSPIRE](#)].
- [69] J. Bonifacio, K. Hinterbichler and R.A. Rosen, *Positivity constraints for pseudolinear massive spin-2 and vector Galileons*, *Phys. Rev. D* **94** (2016) 104001 [[arXiv:1607.06084](#)] [[INSPIRE](#)].
- [70] B. Bellazzini, F. Riva, J. Serra and F. Sgarlata, *Beyond Positivity Bounds and the Fate of Massive Gravity*, *Phys. Rev. Lett.* **120** (2018) 161101 [[arXiv:1710.02539](#)] [[INSPIRE](#)].
- [71] J. Bonifacio and K. Hinterbichler, *Bounds on Amplitudes in Effective Theories with Massive Spinning Particles*, *Phys. Rev. D* **98** (2018) 045003 [[arXiv:1804.08686](#)] [[INSPIRE](#)].

- [72] S. Melville and J. Noller, *Positivity in the Sky: Constraining dark energy and modified gravity from the UV*, *Phys. Rev. D* **101** (2020) 021502 [Erratum *ibid.* **102** (2020) 049902] [[arXiv:1904.05874](#)] [[INSPIRE](#)].
- [73] C. de Rham and A.J. Tolley, *Speed of gravity*, *Phys. Rev. D* **101** (2020) 063518 [[arXiv:1909.00881](#)] [[INSPIRE](#)].
- [74] L. Alberte et al., *Positivity Constraints on Interacting Spin-2 Fields*, *JHEP* **03** (2020) 097 [[arXiv:1910.11799](#)] [[INSPIRE](#)].
- [75] W.-M. Chen, Y.-T. Huang, T. Noumi and C. Wen, *Unitarity bounds on charged/neutral state mass ratios*, *Phys. Rev. D* **100** (2019) 025016 [[arXiv:1901.11480](#)] [[INSPIRE](#)].
- [76] Y.-T. Huang, J.-Y. Liu, L. Rodina and Y. Wang, *Carving out the Space of Open-String S-matrix*, *JHEP* **04** (2021) 195 [[arXiv:2008.02293](#)] [[INSPIRE](#)].
- [77] Z.-Y. Wang, C. Zhang and S.-Y. Zhou, *Generalized elastic positivity bounds on interacting massive spin-2 theories*, *JHEP* **04** (2021) 217 [[arXiv:2011.05190](#)] [[INSPIRE](#)].
- [78] M. Herrero-Valea, I. Timiryasov and A. Tokareva, *To Positivity and Beyond, where Higgs-Dilaton Inflation has never gone before*, *JCAP* **11** (2019) 042 [[arXiv:1905.08816](#)] [[INSPIRE](#)].
- [79] M. Herrero-Valea, R. Santos-Garcia and A. Tokareva, *Massless positivity in graviton exchange*, *Phys. Rev. D* **104** (2021) 085022 [[arXiv:2011.11652](#)] [[INSPIRE](#)].
- [80] C. de Rham, S. Melville and J. Noller, *Positivity bounds on dark energy: when matter matters*, *JCAP* **08** (2021) 018 [[arXiv:2103.06855](#)] [[INSPIRE](#)].
- [81] N. Arkani-Hamed, Y.-T. Huang, J.-Y. Liu and G.N. Remmen, *Causality, unitarity, and the weak gravity conjecture*, *JHEP* **03** (2022) 083 [[arXiv:2109.13937](#)] [[INSPIRE](#)].
- [82] B. Bellazzini, G. Isabella and M.M. Riva, *Classical vs. quantum eikonal scattering and its causal structure*, *JHEP* **04** (2023) 023 [[arXiv:2211.00085](#)] [[INSPIRE](#)].
- [83] C.M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Rel.* **17** (2014) 4 [[arXiv:1403.7377](#)] [[INSPIRE](#)].
- [84] E. Berti et al., *Testing General Relativity with Present and Future Astrophysical Observations*, *Class. Quant. Grav.* **32** (2015) 243001 [[arXiv:1501.07274](#)] [[INSPIRE](#)].
- [85] LIGO SCIENTIFIC and VIRGO collaborations, *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102 [[arXiv:1602.03837](#)] [[INSPIRE](#)].
- [86] EVENT HORIZON TELESCOPE collaboration, *First M87 Event Horizon Telescope Results. I. The shadow of the Supermassive Black Hole*, *Astrophys. J. Lett.* **875** (2019) L1 [[arXiv:1906.11238](#)] [[INSPIRE](#)].
- [87] C. Brans and R.H. Dicke, *Mach's principle and a relativistic theory of gravitation*, *Phys. Rev.* **124** (1961) 925 [[INSPIRE](#)].
- [88] G.W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int. J. Theor. Phys.* **10** (1974) 363 [[INSPIRE](#)].
- [89] C. Deffayet, X. Gao, D.A. Steer and G. Zahariade, *From k-essence to generalised Galileons*, *Phys. Rev. D* **84** (2011) 064039 [[arXiv:1103.3260](#)] [[INSPIRE](#)].
- [90] D. Langlois and K. Noui, *Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability*, *JCAP* **02** (2016) 034 [[arXiv:1510.06930](#)] [[INSPIRE](#)].

- [91] M.B. Green, J.H. Schwarz and E. Witten, *Superstring Theory Vol. 1: 25th Anniversary Edition*, Cambridge University Press (2012) [[DOI:10.1017/CB09781139248563](#)] [[INSPIRE](#)].
- [92] P. Kanti et al., *Dilatonic black holes in higher curvature string gravity*, *Phys. Rev. D* **54** (1996) 5049 [[hep-th/9511071](#)] [[INSPIRE](#)].
- [93] T.P. Sotiriou and S.-Y. Zhou, *Black hole hair in generalized scalar-tensor gravity*, *Phys. Rev. Lett.* **112** (2014) 251102 [[arXiv:1312.3622](#)] [[INSPIRE](#)].
- [94] T.P. Sotiriou and S.-Y. Zhou, *Black hole hair in generalized scalar-tensor gravity: An explicit example*, *Phys. Rev. D* **90** (2014) 124063 [[arXiv:1408.1698](#)] [[INSPIRE](#)].
- [95] K. Yagi, L.C. Stein, N. Yunes and T. Tanaka, *Post-Newtonian, Quasi-Circular Binary Inspirals in Quadratic Modified Gravity*, *Phys. Rev. D* **85** (2012) 064022 [Erratum *ibid.* **93** (2016) 029902] [[arXiv:1110.5950](#)] [[INSPIRE](#)].
- [96] H.O. Silva et al., *Spontaneous scalarization of black holes and compact stars from a Gauss-Bonnet coupling*, *Phys. Rev. Lett.* **120** (2018) 131104 [[arXiv:1711.02080](#)] [[INSPIRE](#)].
- [97] D.D. Doneva and S.S. Yazadjiev, *New Gauss-Bonnet Black Holes with Curvature-Induced Scalarization in Extended Scalar-Tensor Theories*, *Phys. Rev. Lett.* **120** (2018) 131103 [[arXiv:1711.01187](#)] [[INSPIRE](#)].
- [98] K. Yagi, *A new constraint on scalar Gauss-Bonnet gravity and a possible explanation for the excess of the orbital decay rate in a low-mass X-ray binary*, *Phys. Rev. D* **86** (2012) 081504 [[arXiv:1204.4524](#)] [[INSPIRE](#)].
- [99] H. Witek, L. Gualtieri, P. Pani and T.P. Sotiriou, *Black holes and binary mergers in scalar Gauss-Bonnet gravity: scalar field dynamics*, *Phys. Rev. D* **99** (2019) 064035 [[arXiv:1810.05177](#)] [[INSPIRE](#)].
- [100] Z. Carson, B.C. Seymour and K. Yagi, *Future prospects for probing scalar–tensor theories with gravitational waves from mixed binaries*, *Class. Quant. Grav.* **37** (2020) 065008 [[arXiv:1907.03897](#)] [[INSPIRE](#)].
- [101] H.-T. Wang et al., *Tight constraints on Einstein-dilation-Gauss-Bonnet gravity from GW190412 and GW190814*, *Phys. Rev. D* **104** (2021) 024015 [[arXiv:2104.07590](#)] [[INSPIRE](#)].
- [102] S.E. Perkins, R. Nair, H.O. Silva and N. Yunes, *Improved gravitational-wave constraints on higher-order curvature theories of gravity*, *Phys. Rev. D* **104** (2021) 024060 [[arXiv:2104.11189](#)] [[INSPIRE](#)].
- [103] P. Pani, E. Berti, V. Cardoso and J. Read, *Compact stars in alternative theories of gravity. Einstein-Dilaton-Gauss-Bonnet gravity*, *Phys. Rev. D* **84** (2011) 104035 [[arXiv:1109.0928](#)] [[INSPIRE](#)].
- [104] A. Saffer and K. Yagi, *Tidal deformabilities of neutron stars in scalar-Gauss-Bonnet gravity and their applications to multimessenger tests of gravity*, *Phys. Rev. D* **104** (2021) 124052 [[arXiv:2110.02997](#)] [[INSPIRE](#)].
- [105] G. Antoniou, A. Papageorgiou and P. Kanti, *Probing Modified Gravity Theories with Scalar Fields Using Black-Hole Images*, *Universe* **9** (2023) 147 [[arXiv:2210.17533](#)] [[INSPIRE](#)].
- [106] Z. Lyu, N. Jiang and K. Yagi, *Constraints on Einstein-dilation-Gauss-Bonnet gravity from black hole-neutron star gravitational wave events*, *Phys. Rev. D* **105** (2022) 064001 [Erratum *ibid.* **106** (2022) 069901] [[arXiv:2201.02543](#)] [[INSPIRE](#)].

- [107] L.K. Wong, C.A.R. Herdeiro and E. Radu, *Constraining spontaneous black hole scalarization in scalar-tensor-Gauss-Bonnet theories with current gravitational-wave data*, *Phys. Rev. D* **106** (2022) 024008 [[arXiv:2204.09038](#)] [[INSPIRE](#)].
- [108] J.D. Bekenstein, *Black hole hair: 25 - years after*, in the proceedings of the *2nd International Sakharov Conference on Physics*, (1996), p. 216–219 [[gr-qc/9605059](#)] [[INSPIRE](#)].
- [109] C.A.R. Herdeiro and E. Radu, *Asymptotically flat black holes with scalar hair: a review*, *Int. J. Mod. Phys. D* **24** (2015) 1542014 [[arXiv:1504.08209](#)] [[INSPIRE](#)].
- [110] T.P. Sotiriou and E. Barausse, *Post-Newtonian expansion for Gauss-Bonnet gravity*, *Phys. Rev. D* **75** (2007) 084007 [[gr-qc/0612065](#)] [[INSPIRE](#)].
- [111] T. Damour and G. Esposito-Farese, *Nonperturbative strong field effects in tensor - scalar theories of gravitation*, *Phys. Rev. Lett.* **70** (1993) 2220 [[INSPIRE](#)].
- [112] D.D. Doneva et al., *Scalarization*, [arXiv:2211.01766](#) [[INSPIRE](#)].
- [113] H.O. Silva, H. Witek, M. Elley and N. Yunes, *Dynamical Descalarization in Binary Black Hole Mergers*, *Phys. Rev. Lett.* **127** (2021) 031101 [[arXiv:2012.10436](#)] [[INSPIRE](#)].
- [114] M. Herrero-Valea, *The shape of scalar Gauss-Bonnet gravity*, *JHEP* **03** (2022) 075 [[arXiv:2106.08344](#)] [[INSPIRE](#)].
- [115] F. Serra, J. Serra, E. Trincherini and L.G. Trombetta, *Causality constraints on black holes beyond GR*, *JHEP* **08** (2022) 157 [[arXiv:2205.08551](#)] [[INSPIRE](#)].
- [116] M.P. Hertzberg, J.A. Litterer and N. Shah, *Causal modifications of gravity and their observational bounds*, *Phys. Rev. D* **107** (2023) 024037 [[arXiv:2209.07525](#)] [[INSPIRE](#)].
- [117] M. Froissart, *Asymptotic behavior and subtractions in the Mandelstam representation*, *Phys. Rev.* **123** (1961) 1053 [[INSPIRE](#)].
- [118] A. Martin, *Unitarity and high-energy behavior of scattering amplitudes*, *Phys. Rev.* **129** (1963) 1432 [[INSPIRE](#)].
- [119] D.-Y. Hong, Z.-H. Wang, H. Xu and S.-Y. Zhou, *Causality bounds on scalar-tensor EFTs II*, in preparation.
- [120] T. Clifton, P.G. Ferreira, A. Padilla and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1 [[arXiv:1106.2476](#)] [[INSPIRE](#)].
- [121] A.R. Solomon and M. Trodden, *Higher-derivative operators and effective field theory for general scalar-tensor theories*, *JCAP* **02** (2018) 031 [[arXiv:1709.09695](#)] [[INSPIRE](#)].
- [122] M. Ruhdorfer, J. Serra and A. Weiler, *Effective Field Theory of Gravity to All Orders*, *JHEP* **05** (2020) 083 [[arXiv:1908.08050](#)] [[INSPIRE](#)].
- [123] S. Mizera, *Bounds on Crossing Symmetry*, *Phys. Rev. D* **103** (2021) 081701 [[arXiv:2101.08266](#)] [[INSPIRE](#)].
- [124] M. Herrero-Valea, A.S. Koshelev and A. Tokareva, *UV graviton scattering and positivity bounds from IR dispersion relations*, *Phys. Rev. D* **106** (2022) 105002 [[arXiv:2205.13332](#)] [[INSPIRE](#)].
- [125] B.M. Gavela, E.E. Jenkins, A.V. Manohar and L. Merlo, *Analysis of General Power Counting Rules in Effective Field Theory*, *Eur. Phys. J. C* **76** (2016) 485 [[arXiv:1601.07551](#)] [[INSPIRE](#)].

- [126] J.D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333 [INSPIRE].
- [127] V.I. Danchev, D.D. Doneva and S.S. Yazadjiev, *Constraining scalarization in scalar-Gauss-Bonnet gravity through binary pulsars*, *Phys. Rev. D* **106** (2022) 124001 [[arXiv:2112.03869](https://arxiv.org/abs/2112.03869)] [INSPIRE].
- [128] O. Tange, *GNU Parallel 20150322 ('Hellwig')*, <http://www.gnu.org/s/parallel> [DOI:10.5281/ZENODO.16303].
- [129] O. Tange, *GNU Parallel - The Command-Line Power Tool*, *login:* **36** (2011) 42.
- [130] P. Benincasa and F. Cachazo, *Consistency Conditions on the S-Matrix of Massless Particles*, [arXiv:0705.4305](https://arxiv.org/abs/0705.4305) [INSPIRE].